Dondena Working Papers
Carlo F. Dondena Centre for Research
on Social Dynamics and Public Policy

Taxes and the Division of Social Status

S. Bastani
T. Blumkin
L. Micheletto

Working Paper No. 156
February 2023

Università Bocconi • The Dondena Centre

Via Guglielmo Röntgen 1, 20136 Milan, Italy
http://www.dondena.unibocconi.it

ISSN-2035-2034

Taxes and the Division of Social Status

Spencer Bastani[*] Tomer Blumkin [†] Luca Micheletto[‡]

February 4, 2023

Abstract

It is widely acknowledged that the quest for social status can result in an inefficient consumption "rat-race" and the existing literature has discussed how taxes can mitigate the associated externalities. We suggest a new reason to tax conspicuous consumption. Our paper highlights that taxing status goods can achieve a more equitable distribution of welfare by compressing the status distribution. By curbing the conspicuous consumption of the wealthy, the government renders signaling less informative and increases the share of the social status surplus derived by the less wealthy. This "status channel" serves as a complement to traditional monetary channels of redistribution.

Keywords: optimal taxation, signaling, status

JEL classification: H21, D63, D82

[*]Institute for Evaluation of Labour Market and Education Policy (IFAU); Research Institute of Industrial Economics (IFN); UCFS, UCLS, CESIfo, Germany. E-mail: spencer.bastani@ifau.uu.se.

[†]Department of Economics, Ben Gurion University; CESifo, Germany; IZA. E-mail: tomerblu@bgu.ac.il.

[‡]Department of Law, University of Milan, and Dondena Centre for Research on Social Dynamics and Public Policy, Bocconi University; UCFS; CESifo, Germany. E-mail: luca.micheletto@unimi.it.

1 Introduction

> ”...in the present times, through the greater part of Europe, a creditable day-laborer would be ashamed to appear in public without a linen shirt, the want of which would be supposed to denote that disgraceful degree of poverty which, it is presumed, nobody can well fall into without extreme bad conduct. Custom, in the same manner, has rendered leather shoes a necessary of life in England. The poorest creditable person of either sex would be ashamed to appear in public without them.”[1]

The idea that consumers use visible expenditure on consumption goods to signal their wealth and acquire social status has a long history in the social sciences. The quote from Adam Smith above dates back to 1776, but the important role of visibility in consumption expenditures was recognized already by Plato.[2] In the late 19th century, Thorstein Veblen (Veblen 1899) coined the term conspicuous consumption, referring to lavish spending on visible signaling goods and Hirsch (1976) later introduced the term ’positional good’ referring to goods that are valued for their relative rather than their absolute properties. The importance of positional goods and the welfare costs of the associated negative externalities were emphasized by Robert Frank in a series of papers that appeared in the late 90’s and early 2000’s (see, e.g., Frank 1997, 2005) and the theory was further developed by important contributions such as Bagwell and Bernheim (1996) and Hopkins and Kornienko (2004).

The empirical relevance of consumption visibility has been demonstrated in numerous empirical studies. In his path-breaking work, Heffetz (2011) constructed a measure of expenditure visibility based on US survey data and showed that it could explain as much as one third of the observed variation in income elasticities across consumption categories. The importance of visibility has also been confirmed in recent experimental work. Examples include Bursztyn et al. (2017) who found compelling evidence on the relevance of social status signaling in the context of Platinum credit cards in Indonesia, and Butera et al. (2022) who found an almost universal positive willingness to pay for public recognition in the context of donations to the Red Cross.

Given the the salience of conspicuous consumption in modern societies and the centrality of status motives in shaping consumption patterns, an important question is what the implications are for the design of consumption taxation. Most research has focused on corrective (Pigouvian) commodity taxes and their role in combating positional externalities. The literature has also highlighted that efficient corrective taxes are useful from an equity perspective as such taxes are progressive (since conspicuous consumption goods are disproportionately consumed by the wealthy) and can be used to finance transfers to the poor.[3] However, the literature has not paid

[1] See Smith (1776), Book V, Chapter 2. p.148.

[2] Plato writes in Book II of The Republic: ”Since then, as philosophers prove, appearance tyrannizes over truth and is lord of happiness, to appearance I must devote myself”, see e.g. the 1888 translation by Benjamin Jowett, Jowett (1888).

[3] Examples of papers addressing the policy implications of conspicuous consumption include Ng (1987), Corneo and Jeanne (1997), Ireland (2001), Bilancini and Boncinelli (2012), Truyts (2012), and Friedrichsen et al. (2020).

attention to the fact that taxes on conspicuous consumption affect the information transmission among consumers, and, ultimately, the division of a potentially huge social status surplus.

In this paper, we view conspicuous consumption as a form of rent-seeking activity aimed at capturing a part of the social status surplus. Following the literature on noisy signaling, we consider a variation of the Spence (1973) signaling game. We assume the visibility of people's status consumption, and thereby the likelihood of being perceived to be of high social status, is determined by the variety of conspicuous consumption goods consumed. The game structure implies that, under an efficient separating regime, the resulting allocation is not fully revealing due to the imperfect visibility of the consumption signals. In this context, we illustrate that commodity taxation serves a new purpose. By making wealthy individuals consume a lesser variety of consumption goods, commodity taxation renders status signaling less informative. This, in turn, increases the share of the social status surplus derived by the less wealthy, providing equity gains. However, this new status redistribution channel must be weighed against traditional monetary channels for redistribution, and we offer a tractable optimal commodity taxation framework that allows us to capture this trade-off.

To illustrate the trade-off facing the government, think about the notion of enforcing mandatory uniform dress codes in schools. Pupils coming from higher socio-economic backgrounds, may, in the absence of regulation, show up at school wearing expensive brand-name clothing. In the resulting separating equilibrium, expensive visible clothing enables wealthy students to signal a higher social status at the expense of less wealthy students (a zero-sum status contest). Regulation can be used to address this issue by enforcing a uniform dress code, implying an equitable allocation of social status in the resulting pooling equilibrium. An alternative to regulation is to allow students to buy their way out of the dress code, by paying a fee. The revenues collected could finance extra tutoring or the purchase of school materials for the general use of pupils at school. Under this policy regime, wealthy students would choose to distinguish themselves, albeit to a lesser extent than in the laissez-faire (no regulation) regime, but the welfare of students with a less wealthy background could still be improved through the proposed fine/compensation-scheme.

In the broader context of redistributive policy designed by the government, the optimal policy ranges from relying exclusively on monetary channels of redistribution by heavily exploiting conspicuous consumption as a source of tax revenue, to full suppression of status signaling, where no tax revenue from conspicuous consumption is raised. The general insight of our analysis is that the equity gains of high commodity tax rates should not be judged solely on the basis of their revenue-raising effects, as they can promote a more equitable distribution of welfare even when they contribute very little to the revenue collected by the government.

The paper is organized as follows. Section 2 presents the basic ingredients of our model, section 3 presents the two-stage signaling game (between the government and private agents), and section 4 offers concluding remarks.

2 Model

Consider an economy with two types of agents, denoted by $j = 1, 2$, who differ in their wealth endowment. We let w^2 denote the 'wealthy' type and w^1 denote the 'poor' type, where $w^2 > w^1 > 0$. The measure of each type in the economy is normalized to one. The individual's wealth endowment is assumed to be private information observed by neither the government nor by other agents. Each agent spends his/her endowment on $n + 1$ consumption goods: (i) y, defined as the numéraire good which provides intrinsic utility and is observed by neither the government nor by other agents; (ii) $x_i, i = 1, \ldots, n$, $x_i \in \{0, 1\}$, a set of n binary pure signaling goods that serve for status-signaling purposes (as will be explained below). The utility of a type-j agent is given by:

$$U^j(y^j, x_1^1, \ldots, x_n^1, x_1^2, \ldots, x_n^2) = y^j + \mathbf{P}[\tilde{w}^j = w^2 | x_1^1, \ldots, x_n^1, x_1^2, \ldots, x_n^2] \cdot B. \quad (1)$$

The first term on the right hand side captures the intrinsic utility from consuming the numéraire good y, whereas the second term captures the utility from status. The latter is given by the product of $\mathbf{P}$, the probability that type j is perceived to have a high wealth endowment conditional on the consumption choices of the two type of individuals (with $\tilde{w}^j$ being the perceived wealth endowment), and $B > 0$, the gain from social status associated with being perceived as a high type. The vector x_i^j denotes the consumption choices made by types $j = 1, 2$ with respect to the pure signaling goods $x_i, i = 1, \ldots, n$.[4]

The formulation of the utility function in (1) reflects our choice to confine attention to pure strategies when characterizing the perfect Bayesian equilibrium of the signaling game, assuming all agents of the same type will choose the same consumption bundle on the equilibrium path. Perceptions are formed by Bayesian updating conditional on the vector x chosen by both types on this path.

Notice that in our context, a low level of utility derived by poor agents is not merely driven by being deprived (and thereby consuming a lower amount of the numéraire good) but also through the status channel, by being perceived to be poor. Poor type-1 agents therefore have an incentive to mimic their wealthy counterparts; namely, behaving as if they were type-2 agents. This, in turn, induces wealthy type-2 agents, to credibly signal the fact that they possess a larger endowment by spending on conspicuous consumption, from which they derive no intrinsic utility.

The budget constraints faced by the two types of agents are given by:

$$y^j + \sum_{i=1}^{n} p_i^j \cdot x_i^j = w^j, j = 1, 2, \quad (2)$$

[4] These are pure signaling goods in the sense that neither type derives any direct utility from consuming them. Assuming, alternatively, that consumers derive some intrinsic utility from x would not change the qualitative features of the analysis. Wasteful signaling would then be simply measured relative to a reference level which would be bounded away from zero (and could possibly differ across types). Our simplifying assumption is made for tractability.

where p_i^j denotes the price incurred by type $j = 1, 2$ when purchasing a unit of good $i = 1, \ldots, n$. For each $i = 1, \ldots, n$, we assume that $p_i^2 = \theta/n$ and $p_i^1 = 1/n$, where $0 < \theta < 1$.

The fact that the (per-unit) prices incurred by type 2 (when purchasing x) are lower than those incurred by type 1 is consistent with the following three interpretations. First, dwelling on the canonical signaling model of Spence (1973), the cost of acquiring the signal may be lower for type-2. Second, the lower costs incurred by type-2 could reflect, in reduced form, the diminishing marginal utility from consumption of the numéraire good y (assumed to be constant for tractability), implying that the effective (unit) prices of x are decreasing with respect to the wealth endowment. Third, the lower costs incurred by type-2 may reflect heterogeneity in preferences, where type-2 derives some direct utility (or a higher direct utility) from consumption of x. The latter implies again that the effective price incurred by type-2 would be lower than that incurred by type-1.[5]

We turn next to discuss the role of visibility in the status-generation process. Dwelling on the literature on noisy signaling (e.g., Matthews and Mirman 1983), we plausibly assume that engaging in conspicuous consumption is imperfectly observed by the target population.[6] This fundamental feature of our model will play an important role in the policy analysis which will follow (see subsection 3.2 below). In particular, under a separating equilibrium regime, each individual will be faced with a trade-off between spending more on the numéraire good, y, or investing more in signaling, thereby enhancing the likelihood of being perceived as a high type and gaining a higher level of (expected) social status.[7] The above trade-off will enable the government to re-distribute via the status channel, by controlling the level of noise (which is determined endogenously in equilibrium), along with redistributing through the traditional income channel.

For tractability, we choose a fairly simple structure of the stochastic process which determines the level of noise. The qualitative features of our analysis are robust to using other specifications of the stochastic process. Formally, we assume that the visibility of each $x_i, i = 1, \ldots, n$ is determined by a binary random variable, z, which takes the value of 1 ("visible"), with probability $0 < q/n < 1$, and, the value of 0 ("non-visible"), with the complementary probability $0 < 1 - q/n < 1$. The realization of the random variable z is assumed to be independent across the various x-goods and perfectly correlated across all agents that choose to spend on a given signaling good. Visibility is therefore an (ex-post) attribute of each signaling good. Our

[5]To see this, suppose for simplicity that $n = 1$ (and let $x_1^j \equiv x^j; j = 1, 2$). Let the utility derived by type 2 be given by: $u^2(y^2, x^1, x^2) = y^2 + \frac{1}{n} \cdot (1-\theta)x^2 + \mathbf{P}[\tilde{w}^2 = w^2 \mid x^1, x^2] \cdot B$. Assuming that both types incur the same prices (per unit of x) given by $1/n$ implies that the budget constraint faced by type-2 is given by: $w^2 = y^2 + \frac{1}{n} \cdot x^2$. Substituting for y^2 from the budget constraint into the utility yields: $u^2 = w^2 - \frac{1}{n} \cdot \theta x^2 + \mathbf{P}[\tilde{w}^2 = w^2 \mid x^1, x^2] \cdot B$. It is straightforward to verify that, assuming $n = 1$, substituting for y^2 from the budget constraint in (2) into the utility in (1) yields an identical expression.

[6]Matthews and Mirman (1983) is considered the seminal paper in this strand of the literature, introducing noisy signals in a classical model of limit pricing. More recently, de Haan et al. (2011) analyzed theoretically the behavioral implications of varying levels of noise and provided supporting experimental evidence.

[7]When signals are perfectly observed, no such trade-off arises under a separating equilibrium which is fully revealing.

assumptions imply that, ex-post, a fraction q/n of the signaling goods will be visible and a complementary fraction of $1-q/n$ will be non-visible. Consumption choices are assumed to be taken prior to the realizations of z. The induced (ex-ante) imperfect visibility of each signaling good implies that consumers may enhance their exposure, thereby increasing their acquired (expected) social status, by spending on a larger subset of the signaling goods.[8]

A simple interpretation of the stochastic process which generates the noise could be as follows. When agents ex-ante choose the variety of goods they purchase for signaling purposes, they don't know on which set of goods the attention of the individuals in the target population will be focused on ex-post. The perfect correlation assumption implies that all individuals in the target population will confine attention to the same goods (which could potentially reflect herding patterns). One could alternatively assume that signaling takes the form of informative advertising (following Butters 1977 and Grossman and Shapiro 1984) viewing the consumption of each x-good by an individual as an advertisement which is sent to the target population (and is received with some probability). Rendering the calculations more complicated, the qualitative features of our analysis would remain unscathed under the alternative specification.

3 The two-stage game

We consider a two-stage game. The government levies a uniform ad-valorem tax, $t \geq 0$, on all signaling goods so that the after-tax prices of each good, faced by types $j = 1, 2$, respectively, are given by $p^1 = (1+t)\cdot\frac{1}{n}$ and $p^2 = (1+t)\cdot\frac{\theta}{n}$, with $p^1 > p^2$. Tax revenues serve to finance a universal lump-sum transfer T, maintaining government budget balance. In the first stage, the government is setting its tax instruments (t and T) so as to maximize social welfare subject to a revenue constraint. In the second stage, a signaling game forms (given the tax instruments in place), and each agent is choosing how to spend his/her wealth endowment on the consumption goods. We will start by analyzing the second stage and then proceed to the first stage.

3.1 Stage II: The signaling game

Applying standard refinement considerations (the intuitive criterion) we characterize a separating equilibrium in which type 2 agents spend on $0 \leq m \leq n$ of the signaling goods, whereas type 1 agents spend their entire wealth endowment on the numéraire good, y.[9] The separating equilibrium is defined as the solution to the following constrained maximization program:

$$\max_{0\leq m\leq n} w^2 - \theta\cdot(1+t)\cdot\frac{m}{n} + T + \left\{\left[1-(1-\frac{q}{n})^m\right] + \frac{1}{2}\cdot(1-\frac{q}{n})^m\right\}\cdot B \qquad (3)$$

[8] The symmetry across goods is invoked for simplicity. We revisit this assumption in section 3.3.

[9] We assume that off-equilibrium, when beliefs cannot be formed in a Bayesian fashion, an agent is perceived to be a low type with probability 1.

subject to

$$w^1 - (1+t)\cdot\frac{m}{n} + T + \left\{\left[1-(1-\frac{q}{n})^m\right] + \frac{1}{2}\cdot(1-\frac{q}{n})^m\right\}\cdot B \leq w^1 + T + \frac{1}{2}\cdot(1-\frac{q}{n})^m\cdot B. \tag{4}$$

Notice that when none of the signals are being observed, which, given the stochastic process described above, occurs with probability $(1-\frac{q}{n})^m$, the status surplus is evenly divided between the two types (based on the prior symmetric distribution). Otherwise, when at least one of the signals is being observed, which occurs with probability $1-(1-\frac{q}{n})^m$, a Bayesian update yields a posterior distribution which supports full separation. Accordingly, the entire status surplus is derived by the high-type. Partial observability (noisy signaling), though, implies that low-type agents enjoy 'the benefit of the doubt' and derive a positive fraction of the surplus under the separating equilibrium. Notice that this qualitative feature of the model is robust to perturbations of the specific stochastic process that generates the noise. Further notice that this 'benefit of doubt' is diminishing with respect to m, which measures the intensity of signaling chosen by the high type. The constraint (4) reflects a standard (no-mimicking) incentive constraint which may (or may not) bind in the optimal solution and states that the low-type weakly prefers to refrain from spending on the signaling goods.

Denote by $\alpha \equiv m/n$ the fraction of signaling goods on which the high types spend. Further assume that both m and n are large. Then, using the fact that $e = \lim_{h\to\infty}\left(1+\frac{1}{h}\right)^h$, it follows that the constrained maximization in (3)-(4) can be re-formulated as follows

Problem $\mathcal{P}1$

$$\max_{0\leq\alpha\leq1} V(\alpha) \equiv w^2 - \theta\cdot(1+t)\cdot\alpha + T + \left(1-\frac{1}{2}\cdot e^{-q\alpha}\right)\cdot B \tag{5}$$

subject to

$$(IC) \quad (1-e^{-q\alpha})\cdot B \leq (1+t)\cdot\alpha. \tag{6}$$

Before turning to the formal analysis of the constrained maximization program in $\mathcal{P}1$, two important observations are called for. First notice that the reformulated constrained maximization in (5)-(6) is implicitly inter-changing the 'max' and 'limit' operators; that is, doing the maximization over the limiting expression rather than taking the limit of the maximized expression. These procedures are equivalent if and only if the convergence of the limiting expression is uniform (rather than point-wise) and is satisfied if α is chosen from a fixed finite partition of the unit interval $[0,1]$. However, as we can fix the finite grid to be arbitrarily fine, we stick to the continuum approximation in the analysis that follows. Second, notice that in $\mathcal{P}1$, we implicitly assume that the only choice faced by a type-1 agent is whether to refrain from spending on the x-goods or mimicking his/her type-2 counterparts (by choosing the same α). In principle, type-1 agents may choose to spend on a smaller subset of the signaling goods ($0<\tilde{\alpha}<\alpha$). However, one can show that the constrained maximization program solved by type-1, namely, maximizing his/her expected utility by choosing $\tilde{\alpha} \in [0, \alpha]$, given the separating equilibrium strategies stated above, is strictly convex. Thus, one can confine attention to the two corner solutions: $\tilde{\alpha} = 0$ and $\tilde{\alpha} = \alpha$.[10] The incentive compatibility constraint (6) is hence well defined. The formal details are provided in Appendix A.1.

Assumption 1 below presents a set of parametric assumptions that we impose in the subsequent analysis.

Assumption 1. *We impose the following parametric assumptions*

$$e^{-q} < \frac{2\theta}{qB} < 1, \tag{7}$$

$$\ln qB - \ln 2\theta < qB - 2\theta, \tag{8}$$

$$(1 - e^{-q})B < 1. \tag{9}$$

The first part of assumption (7), $qBe^{-q} < 2\theta$, guarantees that, for $t = 0$, the unconstrained value of α (i.e., the value for α implied by Problem $\mathcal{P}1$ when neglecting the IC constraint (6)) is strictly smaller than one; the second part of assumption (7), $2\theta < qB$, guarantees that, for $t = 0$, the unconstrained value of α is strictly positive. Assumption (8) guarantees that, for $t = 0$, the unconstrained value of α violates the IC constraint (6). Thus, assumption (8) guarantees that under laissez-faire the IC constraint (6) is binding. Assumption (9) guarantees that, for $t = 0$, the smallest positive value for α that satisfies the IC constraint (6) is strictly lower than one.

The solution to problem $\mathcal{P}1$ is characterized by the following Proposition.

Proposition 1. *When $\theta < 1/2$, there exists a positive threshold value for t, which we will denote by t^*, such that the optimal solution to $\mathcal{P}1$ is:*

$$\alpha(t) = \begin{cases} \alpha_2(t), & 0 \leq t < t^* \\ \frac{1}{q} \ln \frac{qB}{2\theta(1+t)}, & t^* \leq t < \frac{qB}{2\theta} - 1 \\ 0, & t \geq \frac{qB}{2\theta} - 1, \end{cases}$$

where $\alpha_2(t)$ is implicitly given by the strictly positive (interior) solution to the binding (IC) constraint (6) and $\frac{1}{q} \ln \frac{qB}{2\theta(1+t)}$ is the optimal solution to the unconstrained maximization of (5). In contrast, when $\theta \geq 1/2$, we have that $\alpha(t) = 0$ for all $t \geq 0$.

Proof See Appendix A.2. □

Based on the characterization of the optimal solution to problem $\mathcal{P}1$, it is straightforward to verify that $\frac{d\alpha}{dt} \equiv \alpha'(t) < 0$ for all $t < \frac{qB}{2\theta} - 1$.[11] Thus, as anticipated, levying higher tax rates on

[10]Notice that spending on a larger subset, $\alpha' > \alpha$, would be clearly sub-optimal, by virtue of our assumption on off-equilibrium beliefs. See footnote 9.

[11]See appendix A.2. For the case where (IC) is slack, this follows immediately from (A8). When (IC) binds,

the set of signaling goods induces type 2 individuals to spend their income on a smaller subset of the signaling goods (reducing thereby the intensity of signaling). Proposition 1 highlights that status signaling serves two purposes for type-2 agents. First, by increasing the number of status goods purchased, they increase their expected utility from status. This property holds for sufficiently high tax rates for which the IC constraint is slack, implying that the threat of mimicking by type-1 agents is not a relevant concern. Hence, α is in this case chosen to achieve the optimal trade-off between non-status and status goods. This is a non-standard property driven by the presence of noisy signaling. Second, spending on signaling goods serves to deter mimicking by low types and thereby enables high types to distinguish themselves from their less wealthy counterparts. This property holds for sufficiently low tax rates for which the IC constraint is binding.

We turn next to analyze the first stage of the game in which the government is setting its tax instruments. To render our analysis non-trivial, we will focus on the case where $\theta < 1/2$.

3.2 Stage I: Government problem

We now formulate the government program and characterize the optimal redistributive policy. The (binding) revenue constraint is given by:

$$\theta \cdot \alpha(t) \cdot t = 2T, \tag{10}$$

where $\alpha(t)$ denotes the optimal fraction of signaling goods on which type-2 spends in equilibrium, and is characterized by Proposition 1. In a separating equilibrium, type-1 agents refrain from engaging in signaling and spend their entire wealth endowment on the numéraire good, y. Thus, in equilibrium, utility is given by:

$$u^1 = w^1 + T + \frac{1}{2} \cdot e^{-q\alpha(t)} \cdot B = w^1 + \frac{1}{2} \cdot \theta \cdot \alpha(t) \cdot t + \frac{1}{2} \cdot e^{-q\alpha(t)} \cdot B, \tag{11}$$

where the second equality follows by substituting for T from the revenue constraint in (10). We assume an egalitarian government is seeking to maximize the well-being of type-1 agents. The social welfare measure is given by:

$$W = \delta \cdot \left[w^1 + \frac{1}{2} \cdot \theta \cdot \alpha(t) \cdot t\right] + (1-\delta) \cdot \left[\frac{1}{2} \cdot e^{-q\alpha(t)} \cdot B\right], \tag{12}$$

where $\delta \in [0.5, 1]$ denotes the weight assigned to consumption of the $num\acute{e}raire$ good and $(1-\delta)$ denotes the weight assigned to social status.[12] Differentiation of (12) w.r.t. t to obtain the

it follows from Figure 4. Notice that we are focusing on the interior solution. Hence, when the red curve shifts upwards, the intersection point shifts to the left.

[12]Notice that for $\delta = 0.5$, the welfare measure is non-paternalistic and coincides with the utility derived by type-1, whereas for $\delta > 0.5$ the welfare measure is paternalistic and exhibits a bias towards utility from consumption of the numeraire good. In particular, for $\delta = 1$, the welfare measure reflects a preference for 'income maintenance', utterly laundering out status utility from the social calculus.

FOC for the government maximization program yields

$$\frac{\delta}{2}\left[\alpha(t)+t\alpha'(t)\right]\theta-\frac{1-\delta}{2}qB\alpha'(t)e^{-q\alpha(t)}=0, \tag{13}$$

or equivalently,

$$\delta\theta\left(t+\frac{\alpha(t)}{\alpha'(t)}\right)-(1-\delta)\,qBe^{-q\alpha(t)}=0. \tag{14}$$

Denoting by η the elasticity of α with respect to the after-tax price $(1+t)$, i.e. $\eta\equiv\frac{\alpha'(t)\cdot(1+t)}{\alpha(t)}$, eq. (14) can be restated as follows:

$$\frac{t}{1+t}=\frac{1-\delta}{\delta\theta}\frac{qBe^{-q\alpha(t)}}{1+t}+\frac{1}{|\eta|}. \tag{15}$$

Eq. (15) shows that, except for the limiting case when $\delta=1$, the optimal tax rate exceeds the Laffer rate $t/\,(1+t)=|\eta|^{-1}$. When $\delta=1$ the government attaches zero weight to social status and the optimal tax rate is given by the inverse-elasticity rule; in this case revenue collected from taxing signaling goods (which are only purchased by the high-type) is maximized and redistribution is exclusively accomplished through the income channel (maximizing the value of the demogrant T). For $\delta\in[0.5,1)$, i.e. when the government is either non-paternalistic ($\delta=0.5$) or, while being paternalistic, it does not fully launder out status utility from social calculus ($\delta\in(0.5,1)$), the optimal tax rate is above the Laffer rate. This is because tax revenue considerations are mitigated by status-distribution considerations , i.e. the incentive to use t also as an instrument to promote a more egalitarian distribution of status.

To provide further insights on the optimal tax system, we will reformulate (15). Denoting by λ the Lagrange multiplier attached to the IC constraint (6), we can write the first order condition for an optimal choice of α by the high-type agent as

$$-(1+t)\,\theta+\frac{qB}{2}e^{-q\alpha}+\left(1+t-qBe^{-q\alpha}\right)\lambda=0, \tag{16}$$

from which we obtain

$$e^{-q\alpha}=\frac{2\,(1+t)\,(\theta-\lambda)}{(1-2\lambda)\,qB}. \tag{17}$$

Substituting for $e^{-q\alpha(t)}$ in (15) the value provided by (17) allows re-expressing eq. (15) as follows:

$$\frac{t}{1+t}=2\frac{1-\delta}{\delta}\frac{1-\lambda/\theta}{1-2\lambda}+\frac{1}{|\eta|}. \tag{18}$$

Remembering that t^{*} is the threshold (positive) value for t that separates the region where the IC constraint is binding ($t<t^{*}$) from the region where it is slack ($t>t^{*}$), Proposition 2 below characterizes the relationship between the elasticity $|\eta|$ and the tax rate.

Proposition 2. *The absolute size of the elasticity $|\eta|$ is endogenous to the tax rate t and is characterized as follows:*

(i) For $t \in [0, t^*)$, we have that

$$|\eta| = \frac{1-2\lambda}{1-2\theta},$$

which is monotonically increasing in t, given that λ is decreasing in t.

(ii) As t approaches t^* from the left, $|\eta|$ drops discontinuously.

(iii) For $t \in [t^*, \frac{qB}{2\theta} - 1]$, we have that

$$|\eta| = \left(\ln \frac{qB}{2\theta\,(1+t)}\right)^{-1},$$

which is monotonically increasing in t and tends to infinity at $t = t^{max} \equiv \frac{qB}{2\theta} - 1$.

Proof See Appendix A.3 □

Notice that according to (iii), the inverse elasticity is equal to $\ln \frac{qB}{2\theta(1+t)}$ and reflects the return on status signaling, given by the ratio between the expected benefit and the cost of signaling.[13] The features of the inverse elasticity $|\eta|^{-1}$ as a function of t are described graphically in Figure 1. The figure also shows that two different tax rates can be consistent with the same elasticity. For instance, defining $\hat{t} = -1 + \frac{qB}{2\theta} e^{2\theta-1}$, we have that $\lim_{t \to t^{*-}} |\eta(t)| = |\eta(\hat{t})| = \frac{1}{1-2\theta}$.

[13] To see this, consider a single (discrete) signaling good x visible with probability q with an associated cost θ and let the gains from status be denoted by B. If x is not purchased, the status surplus is split evenly across the two agents and type 2 derives an expected net benefit of $B/2$. Alternatively, if type 2 purchases a unit of x which costs θ , the social status derived by type-2 is given $qB + (1-q)B/2$, as x is only visible with probability q. The net benefit associated with spending on x is hence $[(qB + (1-q)B/2) - B/2] = qB/2$. Dividing by the cost θ yields $qB/2\theta$.

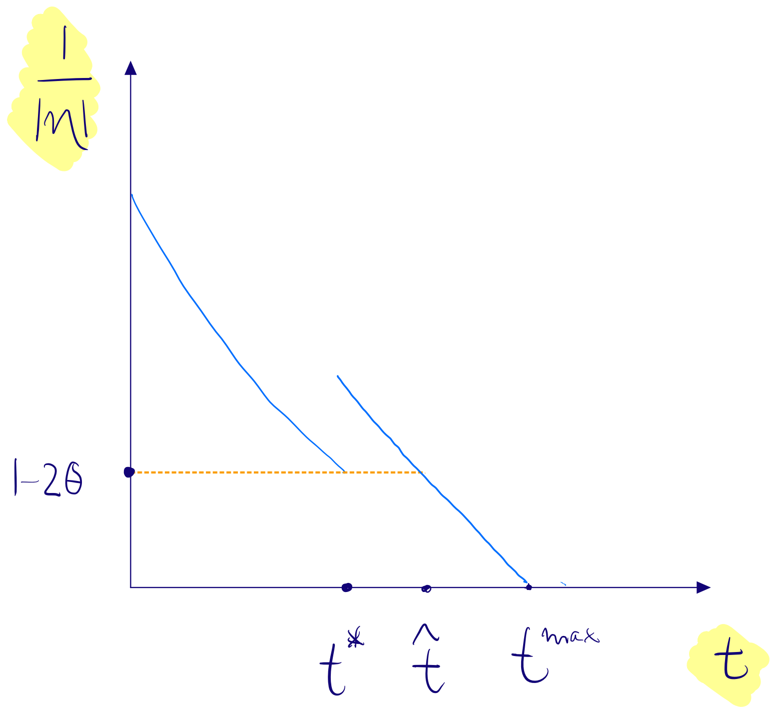

Figure 1: The shape of $|\eta|^{-1}$.

The reason why there is a discrete drop in $|\eta|$ at $t = t^*$ is that for $t < t^*$ the value of α (as a function of t) is dictated by the binding IC constraint $(1+t)\,\alpha = (1 - e^{-q\alpha})\,B$, whereas for $t^* \leq t \leq t^{\max}$ we have that α is given by the unconstrained demand function $\alpha\,(t) = \frac{1}{q}\ln\frac{qB}{2\theta(1+t)}$. Although this does not disturb the continuity of $\alpha(t)$ at $t = t^*$, it implies a discontinuous drop in $\alpha'(t)$ at $t = t^*$.[14] The intuition is that when the threat of mimicking by type-1 ceases to be a concern for type-2, the demand for α becomes less sensitive to changes in t.

Having completed the analysis of the shape of $|\eta|^{-1}$ as a function of t, in Proposition 3 we provide a characterization of the optimal tax policy. This characterization uses the first-order condition (18) to the government's problem, taking into account the potential multiple solutions to this equation, and distinguishing between local and global optimal solutions.

Proposition 3. *Letting* $\hat{\delta} \equiv \frac{2qB}{3qB-2\theta}$ *and* $\widehat{\hat{\delta}} \equiv \frac{1}{1+\theta}$*, the optimal tax policy is characterized as:*

i) For $\delta \in [\frac{1}{2}, \hat{\delta}]$*, the optimal tax policy fully suppresses signaling, no tax revenues are raised, and re-distribution is exclusively carried out by promoting an egalitarian distribution of status; any tax rate weakly larger than* $\frac{qB}{2\theta} - 1$ *is optimal.*

(ii) For $\delta \in (\hat{\delta}, 1]$*, signaling is not fully suppressed, and the optimal tax rate* t^{opt} *is monotoni-*

[14] To see this, notice that, given that the existence of t^* requires that $1 + t < qB$, we have that $\lim_{t\to t^{*-}} \alpha'(t) = -\frac{\alpha}{(1+t)(1+q\alpha)-qB} < -\frac{1}{(1+t)q} = \lim_{t\to t^{*+}} \alpha'(t)$ (i.e., $\lim_{t\to t^{*-}} |\alpha'(t)| = \frac{\alpha}{(1+t)(1+q\alpha)-qB} > \frac{1}{(1+t)q} = \lim_{t\to t^{*+}} |\alpha'(t)|$).

cally decreasing in δ; moreover, t^{opt} satisfies eq. (18), namely

$$\frac{t^{opt}}{1+t^{opt}} = 2\frac{1-\delta}{\delta}\frac{1-\lambda/\theta}{1-2\lambda} + \frac{1}{|\eta|}, \tag{19}$$

where λ is necessarily equal to zero for $\delta \in (\hat{\delta}, \widehat{\hat{\delta}}]$.

Proof See Appendix A.4 □

Proposition 3 is confined to cases where the government attaches at least as large a weight on consumption as it does on status ($\delta \geq 1/2$), with the non-paternalistic case being represented by $\delta = 1/2$. It illustrates that, when δ is sufficiently large, the government does not suppress conspicuous consumption as it serves as a source of tax revenue that can be used to achieve an egalitarian distribution of consumption.[15] However, as δ is gradually lowered, the Proposition shows that redistribution is best achieved by suppressing signaling and achieving redistribution through the status channel. Notice in particular that this is indeed the optimal policy in the non-paternalistic case where $\delta = 1/2$.

In sum, a key insight which emerges from Proposition 3 is the novel and potentially significant role played by redistribution via the signaling channel, along with the traditional income channel, when individuals exhibit social status concerns and engage in conspicuous consumption to signal their wealth.

According to (19) and for given values of δ and η, the upward adjustment on the Laffer rate, that is called for by status-redistribution purposes, is smaller when the IC constraint is binding ($\lambda \neq 0$).[16] Intuitively, the reason is that a binding IC constraint implies that α is upward distorted (compared to the choice that would have been made by a type-2 agent in the absence of a mimicking threat by type-1). On one hand this implies that a marginal reduction in α delivers smaller gains in terms of status redistribution (since the status-redistribution effects of a marginal variation in α become smaller as α gets larger); on the other hand, it implies a base-broadening effect that makes more effective to achieve redistributive goals through the traditional income channel.

Figures 2-3 illustrate the two possible profiles of the optimal tax function $t(\delta)$.

[15] The fact that the non-paternalistic optimum implies redistribution exclusively via the status channel is driven by the linearity of the cost of signaling. In a more general setting, the optimum would combine redistribution through the two channels. However, the status channel would always imply that the optimal tax rate exceeds the Laffer rate.

[16] This is because $\partial\left(\frac{1-\lambda/\theta}{1-2\lambda}\right)/\partial\lambda = \frac{2-1/\theta}{(1-2\lambda)^2} < 0$ (for $0 < \theta < 1/2$).

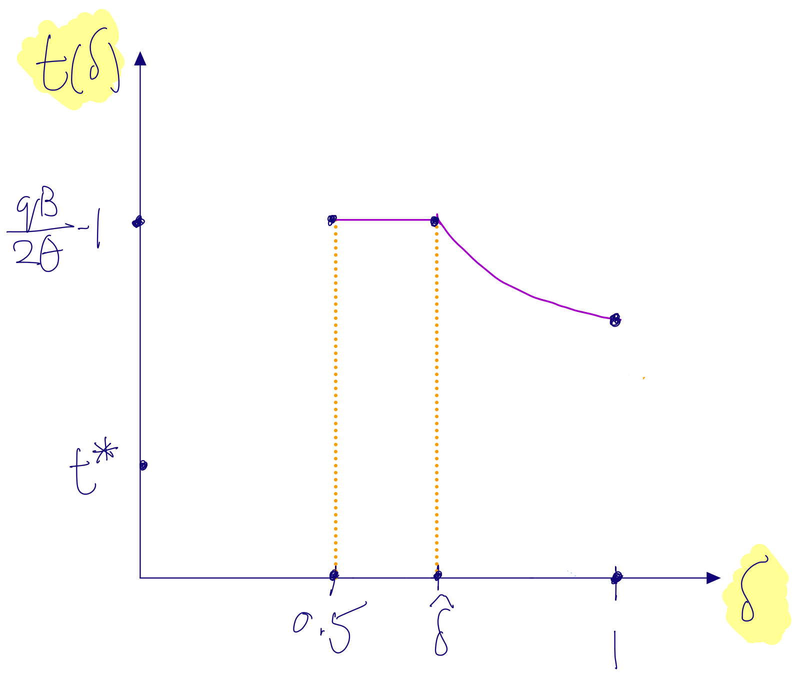

Figure 2: Illustration of the shape of $t(\delta)$ when $t(1)$ is larger than t^*.

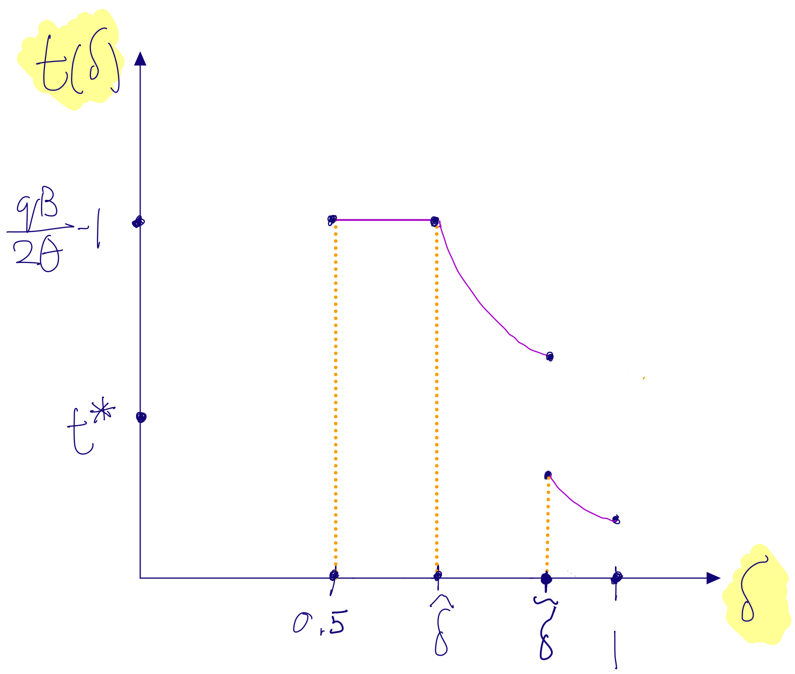

Figure 3: Illustration of the shape of $t(\delta)$ when $t(1)$ is smaller than t^*.

The crucial difference between the two figures is that $t(1) > t^*$ in Figure 2 whereas $t(1) < t^*$ in Figure 3. When the Laffer rate is higher than t^* (i.e., $t(1) > t^*$), the function $t(\delta)$ is continuous within its domain $[0.5, 1]$. There is first a range where the optimal tax rate is constant and given by $\frac{qB}{2\theta} - 1$ (for $\delta \in [\frac{1}{2}, \hat{\delta}]$, where $\hat{\delta} \equiv \frac{2qB}{3qB-2\theta}$), and then $t(\delta)$ decreases monotonically and continuously until it reaches its minimum for $\delta = 1$.

Instead, when the Laffer rate is lower than t^* (i.e., $t(1) < t^*$), the function $t(\delta)$ has a point of discontinuity. There is again an initial range where the optimal tax rate is constant and then, from $\delta = \hat{\delta}$, a region where $t(\delta)$ decreases monotonically and continuously. However, this region does not extend until $\delta = 1$; at some threshold value for δ, denoted by $\widetilde{\delta}$ in Figure 3, the function $t(\delta)$ jumps from a value that is strictly larger than t^* to a value that is strictly smaller than t^*, and from there the function resumes its continuously decreasing profile.

The presence of a discontinuous jump in the function $t(\delta)$ implies that, as societies become over time more sensitive to differences in status, it is possible to reach a point where a marginal reduction in δ (i.e., a marginal increase in the weight assigned by the government to the status component of utility) triggers a policy-regime change: from a low-tax regime, where the bulk of redistribution is accomplished through the income channel, to a high-tax regime, where the bulk of redistribution is shifted to the status-channel.

3.3 The status production function and optimal tax differentiation

In the analysis thus far, in order to simplify the exposition, we have assumed that all signaling goods were symmetric (in terms of acquisition costs, visibility and benefits from status). One could in principle generalize the model and assume some asymmetries. One simple way to do this in a tractable manner is to think about a status-production technology that exhibits perfect substitutability. In particular, suppose that there are two categories of signaling goods, x_{ki}, $i = 1, \dots, n_k, k = 1, 2$, with respective visibility parameters q_k/n_k and unit costs θ_k/n_k, which are respectively targeted at two reference groups of agents (say, category $k = 1$ focuses on colleagues at work whereas category $k = 2$ targets friends or relatives). The benefits from signaling to each group may vary (intrinsically or due to differences in size). Denote the benefit associated with group k by B_k.

The separable technology implies that the status derived by type-2 agents is given by:

$$Status^2(\alpha_1, \alpha_2) = \sum_{k=1}^{2} [1 - \frac{1}{2} \cdot e^{-\alpha_k q_k}] \cdot B_k, \tag{20}$$

where α_k denotes the fraction of signaling goods in category k on which type-2 agents spend their wealth; whereas the status derived by type-1 agents is given by:

$$Status^1(\alpha_1, \alpha_2) = \sum_{k=1}^{2} [\frac{1}{2} \cdot e^{-\alpha_k q_k}] \cdot B_k \tag{21}$$

The separability assumption implies zero cross tax elasticities between the two categories and one can generalize our analysis by allowing for differentiated commodity tax rates across the two categories of consumption goods, $t_k, k = 1, 2$.

An immediate extension of our analysis implies a generalization of formula (A24), which characterizes the optimal tax formula for the baseline case with a single category of consumption goods when the (IC) constraint is slack (see the proof of Proposition 3 in Appendix A.4). Assuming that the (IC) constraint is slack (which holds for an intermediate range of values of δ, as shown above) it follows that:

$$\frac{t_k}{1+t_k} + \ln(1+t_k) = 2\frac{1-\delta}{\delta} + \ln\frac{q_k B_k}{2\theta_k}. \tag{22}$$

In the separable case, tax differentiation draws on the variation in the returns on status signaling across categories of consumption. Recalling that the expression on the left-hand side of (22) is increasing with t_k, it follows that a higher tax rate is levied on a good which exhibits a higher return on signaling.[17]

3.4 The efficiency-enhancing role of commodity taxation

So far we have focused on equity considerations. We now briefly discuss efficiency aspects. As is typically the case with pure signaling, the resulting allocation in equilibrium is inefficient. In our setup, the status surplus is fixed at the size of B and engaging in signaling by type-2 agents is essentially a form of rent seeking. The only Pareto efficient allocation is one where type-2 agents refrain from signaling, and set $\alpha = 0$. This implies that a Pareto improvement can be attained by a reduction in the intensity of signaling (a decrease in α) supplemented by a proper transfer from type-1 to type-2 agents (in units of the numéraire good, y, by virtue of the quasi-linearity of the utility function). In general, this can be achieved via a system of non-linear commodity taxes.

Confining attention to a linear regime (ad-valorem taxes levied on the signaling goods accompanied by a universal lump-sum transfer, as we have done) could still potentially attain a Pareto improvement. To see this, notice that based on our parametric assumptions, as stated in Proposition 1, in the laissez faire equilibrium, namely for $t = 0$, the incentive constraint is binding. Thus, with a linear tax system in place, although tax revenues are split between the two types, and hence cross subsidization goes in the 'wrong' direction from type-2 to type-1 agents, the former may still become better-off due to the reduction in the extent of (excessive) signaling, which is desirable in light of the binding incentive constraint.[18]. In general, the feasibility of attaining a Pareto improvement depends on the magnitude of the distortion associated with

[17] One could in principle assume, alternatively, full complementarity between the categories of consumption goods in the production of social status. That is, all agents observe both categories of consumption and form their beliefs about the type of agent. The benefit from status is given by B as in the case with one category. The full benefit is obtained when at least one signal (of either category) is observed. Otherwise, the surplus is evenly split between the two types. Such a technology would lead to the status measures $Status^2(\alpha_1, \alpha_2) = [1 - \frac{1}{2} \cdot e^{-\sum_{k=1}^{2} \alpha_k q_k}] \cdot B$ and $Status^1(\alpha_1, \alpha_2) = [\frac{1}{2} \cdot e^{-\sum_{k=1}^{2} \alpha_k q_k}] \cdot B$. The tax formula in the case of full complementarity would account for cross tax elasticities.

[18] The argument bears similarity to the role played by a binding parental leave mandate in a labor market marred by adverse selection (see e.g., Bastani et al. 2019)

excessive consumption of the signaling goods (due to the binding IC constraint) and the degree of cross subsidization needed to maintain the information rent associated with the low-type agents. With a sufficiently large distortion, and a relatively moderate extent of cross subsidization, a Pareto improvement becomes feasible.[19]

It is important to notice the difference from the standard argument which supports a Pareto improvement in the presence of wasteful signaling when linear instruments are in place. In the traditional context, the information content is fixed and a separating equilibrium requires that high types spend a sufficient amount of resources on the signal to induce no-mimicking. The expenditure could either be driven by 'burning money' which is wasteful, or, be associated with higher tax payments, that could be diverted to consumption (through transfers). Thus, given that the tax parameters are common knowledge, paying taxes could serve as an instrumental signal. Hence, it is always desirable to tax signals. In our context, in contrast, signals are not wasteful in the sense that consumption needs to be visible to acquire status. Thus, status is driven by the number of units (or variety) of signaling goods that are purchased rather than the resources spent on these goods.

4 Concluding remarks

In most societies, status-seeking through conspicuous consumption is prevalent. The research literature has widely discussed how taxes levied on such consumption may serve to promote both efficiency, by mitigating the extent of wasteful signaling, and equity, by using the resulting tax revenues for redistribution. In the current study, we have demonstrated a new channel through which taxes on conspicuous consumption can be welfare-enhancing. The novel insight is that redistribution can be achieved by promoting a more equitable distribution of social status. This 'status channel' of redistribution, therefore, has to be balanced against traditional motives to tax conspicuous consumption.

A key feature of our framework is noisy signaling, reflected by the imperfect visibility of consumption used to signal social status. This feature implies that, under a separating equilibrium, unlike in the standard signaling setup with perfectly visible signals, the identity of the agents engaging in signaling is not perfectly revealed. Thus, by investing in a larger variety of conspicuous consumption goods, agents with a high wealth endowment may separate themselves from their less wealthy counterparts. This increases their likelihood of being perceived to be of high social status, and thereby also increases their (expected) share of the total social status surplus. By curbing the conspicuous consumption of the wealthy (say, through taxation), the government can render signaling less informative and thereby increase the share of the social status surplus derived by the less wealthy.

[19] In general, with linear instruments in place, the unregulated market equilibrium may turn out to be second-best Pareto efficient. Notice the difference from the standard setup with noise-less signaling, in which the first best can be attained with linear instruments by setting a confiscatory tax rate on the pure signaling/diamond good (see Ng 1987 and Truyts 2012).

Considering a Rawlsian welfare function, we have shown that when the weight assigned to consumption in the welfare function is relatively low, the optimal tax fully suppresses signaling and no tax revenues are raised. In this case, re-distribution is exclusively carried out by promoting an egalitarian distribution of status. In contrast, when the weight assigned to consumption is relatively high, the optimal tax balances the revenue-raising motive to tax conspicuous consumption, which depends negatively on the elasticity of conspicuous consumption, against the new motive to promote an egalitarian status distribution, which depends positively on the elasticity of conspicuous consumption.

More generally, our analysis highlights that the equity gains of a high marginal tax should not be solely judged on the basis of its revenue-raising effect. As with the case of high marginal income tax rates, which can exert beneficial pre-distributive effects (Bozio et al. 2020), high commodity tax rates can promote a more equitable distribution of welfare even when they contribute very little to the revenue collected by the government.

References

Bagwell, L. S. and Bernheim, B. D. (1996). Veblen effects in a theory of conspicuous consumption. *The American Economic Review*, 86(3):349–373.

Bastani, S., Blumkin, T., and Micheletto, L. (2019). The welfare-enhancing role of parental leave mandates. *The Journal of Law, Economics, and Organization*, 35(1):77–126.

Bilancini, E. and Boncinelli, L. (2012). Redistribution and the notion of social status. *Journal of Public Economics*, 96(9):651–657.

Bozio, A., Garbinti, B., Goupille-Lebret, J., Guillot, M., and Piketty, T. (2020). Predistribution vs. Redistribution: Evidence from France and the U.S. World Inequality Lab, Working Paper No 2020/22.

Bursztyn, L., Ferman, B., Fiorin, S., Kanz, M., and Rao, G. (2017). Status Goods: Experimental Evidence from Platinum Credit Cards*. *The Quarterly Journal of Economics*, 133(3):1561–1595.

Butera, L., Metcalfe, R., Morrison, W., and Taubinsky, D. (2022). Measuring the welfare effects of shame and pride. *American Economic Review*, 112(1):122–68.

Butters, G. R. (1977). Equilibrium distributions of sales and advertising prices. *The Review of Economic Studies*, 44(3):465–491.

Corneo, G. and Jeanne, O. (1997). Conspicuous consumption, snobbism and conformism. *Journal of Public Economics*, 66(1):55–71.

de Haan, T., Offerman, T., and Sloof, R. (2011). Noisy signaling: Theory and experiment. *Games and Economic Behavior*, 73(2):402–428.

Frank, R. H. (1997). The frame of reference as a public good. *The Economic Journal*, 107(445):1832–1847.

Frank, R. H. (2005). Positional externalities cause large and preventable welfare losses. *The American Economic Review*, 95(2):137–141.

Friedrichsen, J., König, T., and Lausen, T. (2020). Social Status Concerns and the Political Economy of Publicly Provided Private Goods. *The Economic Journal*, 131(633):220–246.

Grossman, G. M. and Shapiro, C. (1984). Informative advertising with differentiated products. *The Review of Economic Studies*, 51(1):63–81.

Heffetz, O. (2011). A test of conspicuous consumption: Visibility and income elasticities. *The Review of Economics and Statistics*, 93(4):1101–1117.

Hirsch, F. (1976). *Social Limits to Growth*. Cambridge, Mass.: Harvard University Press.

Hopkins, E. and Kornienko, T. (2004). Running to keep in the same place: Consumer choice as a game of status. *American Economic Review*, 94(4):1085–1107.

Ireland, N. (2001). Optimal income tax in the presence of status effects. *Journal of Public Economics*, 81(2):193–212.

Jowett, B. (1888). *The Republic of Plato, translated into English with Introduction, Analysis, Marginal Analysis, and Index*. Oxford: Clarendon Press.

Matthews, S. A. and Mirman, L. J. (1983). Equilibrium limit pricing: The effects of private information and stochastic demand. *Econometrica*, 51(4):981–996.

Ng, Y.-K. (1987). Diamonds are a government's best friend: Burden-free taxes on goods valued for their values. *The American Economic Review*, 77(1):186–191.

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. McMaster University Archive for the History of Economic Thought.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374.

Truyts, T. (2012). Signaling and indirect taxation. *Journal of Public Economics*, 96(3):331–340.

Veblen, T. (1899). *The Theory of the Leisure Class*. New York, NY: Macmillan.

A Appendix

A.1 Strict convexity of type-1 optimization problem

Formulating the constrained maximization program faced by type-1 yields:

$$\max_{0\leq\tilde{\alpha}\leq\alpha} J(\tilde{\alpha}) \equiv \left[w^1 - (1+t)\cdot\tilde{\alpha} + T\right] + e^{-q(\alpha-\tilde{\alpha})}\cdot\left[1 - \frac{1}{2}\cdot e^{-q\tilde{\alpha}}\right]\cdot B \tag{A1}$$

In words, type-1, given the separating equilibrium profile of strategies, is choosing to spend on a subset of α, so as to maximize his expected utility. Notice that $e^{-q(\alpha-\tilde{\alpha})}$ measures the probability that none of the signals on which type-2 spends but type-1 refrains from spending on, $(\alpha-\tilde{\alpha})$, is visible. If at least one of these signals is visible, no surplus is derived by type-1, as, in equilibrium, all type-2 agents spend on these signals, which serve to distinguish them from their lower-type counterparts. If none of these signals is visible, the surplus derived by type-1 is given by the last term in brackets in (A1). Notably, this term is identical in structure to the second term in brackets of the objective (5), with the exception that $\tilde{\alpha}$ replaces α. That is, the relevant subset of signaling goods that identify type-2 agents is given by $\tilde{\alpha}$. Differentiating $J(\tilde{\alpha})$ with respect to $\tilde{\alpha}$ yields:

$$\frac{\partial J}{\partial \tilde{\alpha}} = -(1+t) + qB \cdot e^{-q(\alpha - \tilde{\alpha})} \tag{A2}$$

Taking the derivative one more time yields:

$$\frac{\partial^2 J}{\partial \tilde{\alpha}^2} = q^2 \cdot B \cdot e^{-q(\alpha - \tilde{\alpha})} > 0 \tag{A3}$$

Thus, $J(\tilde{\alpha})$ is strictly convex with respect to $\tilde{\alpha}$. The optimum for the maximization in (A1) is hence attained by either one of the two corner solutions: $\tilde{\alpha} = 0$ or $\tilde{\alpha} = \alpha$. We conclude that constraint (IC) in program $\mathcal{P}1$ is well defined.

A.2 Proof of Proposition 1

We begin by assuming that (IC) is slack in the optimal solution to $\mathcal{P}1$. Then one can formulate the first-order condition:

$$-\theta \cdot (1+t) + \frac{B}{2} \cdot e^{-q\alpha} \cdot q = 0. \tag{A4}$$

It is straightforward to verify that the second-order condition is satisfied. Denoting by $\alpha(t)$ the optimal choice of the high type (as a function of t) given by the solution to (A4), an interior solution $0 < \alpha(t) < 1$ exists, by virtue of (7), when $1 + t < \frac{qB}{2\theta}$. When $1 + t \geq \frac{qB}{2\theta}$, a corner solution in which the high-type refrains from spending on the signaling goods emerges, namely, $\alpha(t) = 0$. The IC constraint (6) is not necessarily slack, however. Whether (6) is binding or not depends on parametric conditions. We separate between different cases.

Case I: $1 + t \geq \frac{qB}{2\theta}$ As shown above, in this case, assuming (IC) is not violated, the optimal choice is $\alpha(t) = 0$. It is straightforward to verify that for $\alpha = 0$, the IC constraint is trivially satisfied. Thus, this forms indeed the optimal solution. Levying a sufficiently high tax on the x goods, hence, induces the high-type to refrain from engaging in any signaling. Clearly, in such a case, no tax revenues are being collected and $T = 0$. Thus, redistribution is exclusively confined to the status channel, ensuring that the low-type derives the largest possible share of the social status surplus (an expected surplus of $B/2$).

Case II: $1 + t < \frac{qB}{2\theta}$ We will separate this case into two sub-cases. We first assume that $\theta \geq 1/2$. That is, engaging in signaling is fairly costly for the high-type. As shown in Figure 4 below, which represents condition (IC) under the invoked parametric assumptions, for each t, there are two values of α for which (IC) is satisfied as equality: $\alpha_1(t) = 0$ and $0 < \alpha_2(t) < 1$.

To see that there are two values, notice that the relevant range we are considering (when IC is binding) is $1 + t < \frac{qB}{2\theta}$ where $\theta \geq 1/2$. Thus, we have that:

$$qB > 1 + t, \tag{A5}$$

Consider Figure 4. Differentiating the left-hand-side of (IC) with respect to α and taking the limit when $\alpha \to 0$ yields:

$$\lim_{\alpha \to 0} \frac{\partial}{\partial \alpha} \left(1 - e^{-q\alpha}\right) \cdot B = qB > 1 + t. \tag{A6}$$

Thus, by virtue of (A6), as (IC) is satisfied as equality for $\alpha = 0$, by invoking a first-order approximation, it follows that for sufficiently small $\alpha > 0$ the left hand side expression of (IC) is strictly exceeding the RHS and hence (IC) is violated. Taking the limit as $\alpha \to 1$ implies that the LHS of (IC) is given by:

$$\lim_{\alpha \to 1} \left(1 - e^{-q\alpha}\right) \cdot B = \left(1 - e^{-q}\right) \cdot B < 1 < 1 + t. \tag{A7}$$

where the first inequality follows from (9). Thus, for sufficiently high $\alpha > 0$, the RHS of (IC) is strictly exceeding the LHS and, hence, (IC) is satisfied as a strict inequality. By virtue of the intermediate value theorem, hence, there exists some $0 < \alpha < 1$ for which (IC) is satisfied as an equality. The strict concavity (with respect to α) of the left-hand side expression of (IC), which can be readily verified, along with the linearity of the RHS expression, imply that this value of α is unique.

Let us now go back to Figure 4. The red line represents the RHS of (6), whereas, the blue curve represents the LHS. For $\alpha > \alpha_2(t)$, (IC) is satisfied as a strict inequality, and for $0 < \alpha < \alpha_2(t)$, (IC) is violated. We next show that the interior (unconstrained) solution $\alpha(t)$ to the first order condition (A4) violates (IC), that is, $0 < \alpha(t) < \alpha_2(t)$, which implies that the optimal solution is either given by $\alpha_1(t)$ or $\alpha_2(t)$.

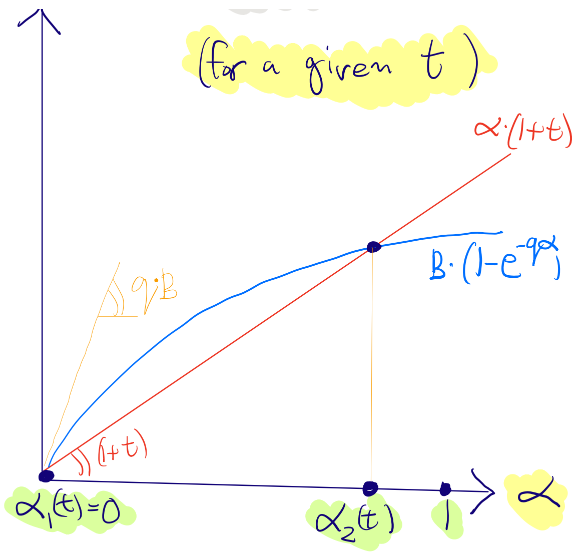

Figure 4: Illustration of the IC constraint (6) .

To show that $0 < \alpha(t) < \alpha_2(t)$, we exploit (A4) and re-arrange to obtain:

$$q \cdot \alpha(t) = -\ln \frac{2\theta \cdot (1+t)}{qB}. \tag{A8}$$

Insertion of (A8) into (IC) given by (6) allows us to express (IC) as follows:

$$H(t) \equiv (1+t) \ln \frac{2\theta \cdot (1+t)}{qB} + qB - 2\theta \cdot (1+t) \leq 0. \tag{A9}$$

It follows immediately that for $1 + t = \frac{qB}{2\theta}$, $H(t) = 0$. Thus, to show that (IC) is violated for $1 + t < \frac{qB}{2\theta}$, it suffices to show that $\frac{dH(t)}{dt} < 0$ in this range. We have that:

$$\frac{dH(t)}{dt} = \ln \frac{2\theta \cdot (1+t)}{qB} + 1 - 2\theta < 0, \tag{A10}$$

where the inequality follows as $1 + t < \frac{qB}{2\theta}$ and the assumption that $\theta \geq \frac{1}{2}$. Thus, $H(t) > 0$ and (IC) is violated in the unconstrained optimum for $1 + t < \frac{qB}{2\theta}$. It follows that in the optimal solution, (IC) is binding and the optimal solution is either given by $\alpha_1(t) = 0$ or $0 < \alpha_2(t) < 1$. We can compare the two candidates for the optimal solution by plugging them into the objective function (5):

$$V(\alpha_1) = \bar{w} + B/2 \tag{A11}$$

$$V(\alpha_2) = \bar{w} + B \cdot (1 - \theta) + B \cdot e^{-q\alpha_2} \cdot (\theta - 1/2). \tag{A12}$$

For $\theta = 1/2$, we have that $V(\alpha_1) = V(\alpha_2)$. Differentiating $V(\alpha_2)$ with respect to θ yields:

$$\frac{\partial V(\alpha_2)}{\partial \theta} = (e^{-q\alpha_2} - 1) \cdot B < 0, \quad \text{as} \quad \alpha_2 > 0. \tag{A13}$$

It follows that $V(\alpha_1) > V(\alpha_2)$ for $\theta > 1/2$. Hence, for any $t < \frac{qB}{2\theta} - 1$ and $\theta \geq 1/2$, the optimal solution is given by: $\alpha(t) = 0$.

We next consider $H(t)$ for the case $\theta < 1/2$. We make the following observations:

- $H(0) > 0$ due to Assumption 1 (by virtue of condition 8)
- $H(t) = 0$ when $1 + t = \frac{qB}{2\theta}$ (as above, by virtue of equation A9)
- $\frac{dH(t)}{dt} > 0$ when $1 + t = \frac{qB}{2\theta}$ since $\theta < 1/2$ (by virtue of equation A10)
- $\frac{d^2H(t)}{dt^2} = \frac{1}{1+t} > 0$ for all $1 + t \leq \frac{qB}{2\theta}$ implying that $H(t)$ is strictly convex.

The properties of $H(t)$ imply that there exists a unique $t^* \in (0, \frac{qB}{2\theta} - 1)$, such that $H(t) \leq 0$ (and hence IC is satisfied) for $t \in [t^*, \frac{qB}{2\theta} - 1)$, whereas $H(t) > 0$ (and hence IC is violated) for $t \in [0, t^*)$. To see this formally, notice that as $H(t) = 0$ and $\frac{dH(t)}{dt} > 0$ when $1 + t = \frac{qB}{2\theta}$, by applying a first-order approximation, it follows that for t smaller than but sufficiently close to $\frac{qB}{2\theta} - 1$, $H(t) < 0$. As $H(0) > 0$, it follows by the Intermediate Value Theorem that t^* exists. Uniqueness follows from the strict convexity of $H(t)$. $H(t)$ is illustrated in Figure 5 below.

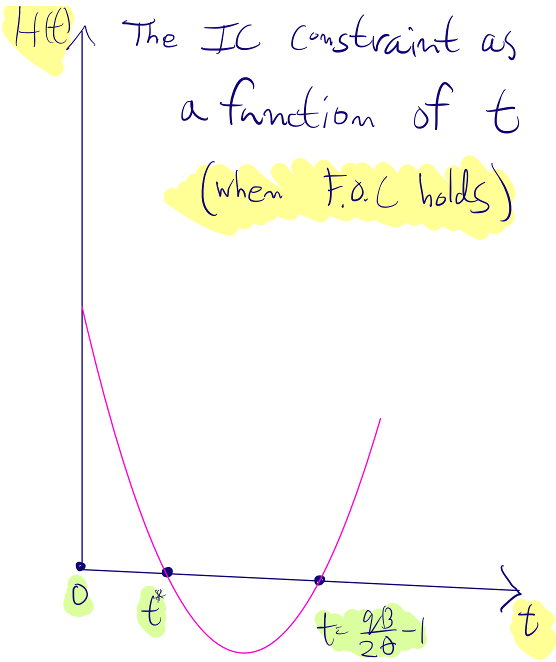

Figure 5: An illustration of $H(t)$ when $\theta < 1/2$.

We conclude that for $t \in [t^*, \frac{qB}{2\theta} - 1)$, the optimal solution is given by condition (A4), which upon re-arrangement yields $\alpha(t) = -\frac{1}{q} \ln \frac{2\theta \cdot (1+t)}{qB}$, whereas for $t \in [0, t^*)$, the optimum is given by a solution to the binding incentive constraint (6). As before, there are two possibilities for this constraint to bind, either $\alpha_1(t) = 0$ or $0 < \alpha_2(t) < 1$. To see this, notice that the relevant range we are considering (when IC is binding) is $t \in [0, t^*)$. By definition of t^* (see Figure 5) we have that:

$$H(t^*) \equiv (1 + t^*) \ln \frac{2\theta \cdot (1 + t^*)}{qB} + qB - 2\theta \cdot (1 + t^*) = 0. \tag{A14}$$

Moreover,

$$\left. \frac{dH(t)}{dt} \right|_{t=t^*} = \ln \frac{2\theta \cdot (1 + t^*)}{qB} + 1 - 2\theta < 0. \tag{A15}$$

Substituting for $\ln \frac{2\theta \cdot (1+t^*)}{qB}$ from (A14) into (A15) yields upon re-arrangement:

$$qB > 1 + t^* \Longrightarrow qB > 1 + t \quad \text{for} \quad t < t^*. \tag{A16}$$

Having established this, we can proceed in an identical fashion as for the case $\theta \geq 1/2$.

The level of the objective in each case is again given by equations (A11) and (A12). The difference now is that $\theta < 1/2$. Exploiting (A13), we conclude that $V(\alpha_2) > V(\alpha_1)$. Hence, for $\theta < 1/2$, the optimal solution is given by: $\alpha(t) = \alpha_2(t) > 0$ when $0 \leq t < t^*$.

A.3 Proof of Proposition 2

By Proposition 1, for $t \in [0, t^*)$, $\alpha(t)$ is dictated by the binding IC constraint $(1 + t)\,\alpha = (1 - e^{-q\alpha})\,B$ and $\alpha'(t) = -\frac{\alpha}{1+t-qBe^{-q\alpha}}$. This implies that

$$\frac{1}{|\eta|} = 1 - \frac{qBe^{-q\alpha}}{1 + t},$$

or equivalently, exploiting (17),

$$\frac{1}{|\eta|} = 1 - 2\frac{\theta - \lambda}{1 - 2\lambda} = \frac{1 - 2\theta}{1 - 2\lambda}. \tag{A17}$$

For $t \in [t^*, \frac{qB}{2\theta} - 1]$, the IC constraint is slack, and $\alpha(t)$ is given by the unconstrained demand function $\alpha\,(t) = \frac{1}{q} \ln \frac{qB}{2\theta(1+t)}$, we have that $\alpha'(t) = -\left[q\,(1 + t)\right]^{-1}$ and:

$$\frac{1}{|\eta|} = \ln \frac{qB}{2\theta\,(1 + t)}. \tag{A18}$$

Recall that t^* is the threshold (positive) value for t that separates the region where the IC constraint is binding ($t < t^*$) from the region where it is slack ($t > t^*$), and satisfies the following equations:

$$\begin{aligned}(1+t)\,\alpha &= \left(1-e^{-q\alpha}\right)B,\\ \alpha &= \frac{1}{q}\ln\frac{qB}{2\theta\,(1+t)},\end{aligned}$$

where the first equation states the IC constraint as an equality and the second equation provides the unconstrained optimal choice for α by a high-type individual. Combining these equations, we thus have that t^* is implicitly given by the following condition:

$$(1+t)\left[2\theta+\ln\frac{qB}{2\theta}-\ln\left(1+t\right)\right]=qB. \tag{A19}$$

As we have shown in the Proof of Proposition 1, for $\theta<1/2$ there exists a unique t^* in the range $t\in\left(0,\frac{qB}{2\theta}-1\right)$. Moreover, for $t\in(0,t^*)$ the LHS of (A19) is smaller than its RHS, and vice versa for $t\in\left(t^*,\frac{qB}{2\theta}-1\right)$. Consider now the value of $\frac{1}{|\eta|}$ when t approaches t^* from the left (IC constraint is binding) and from the right (IC constraint is slack). When t approaches t^* from the left, we have that (see expression (A17)) $\frac{1}{|\eta|}=1-2\theta$ (i.e., $\lim\limits_{t\to t^{*-}}\frac{1}{|\eta|}=\lim\limits_{\lambda\to 0}\frac{1-2\theta}{1-2\lambda}=1-2\theta$). Now consider (A18) and evaluate at which level for t we get that $\frac{1}{|\eta|}=1-2\theta$. Solving the equation

$$\ln\frac{qB}{2\theta}-\ln\left(1+t\right)=1-2\theta,$$

we get

$$t=-1+\frac{qB}{2\theta}e^{2\theta-1}\equiv\hat{t}. \tag{A20}$$

Notice that, inserting into (A19) the value for t provided by (A20), the LHS of (A19) boils down to $\frac{qB}{2\theta}e^{2\theta-1}$, which is a decreasing function of θ (under our assumption that $0<\theta<1/2$). Given that $\lim\limits_{\theta\to 1/2}\frac{qB}{2\theta}e^{2\theta-1}=qB$, we have that, for $0<\theta<1/2$, the LHS of (A19) is larger than its RHS when $t=-1+\frac{qB}{2\theta}e^{2\theta-1}$. Thus, the IC constraint is slack for $t=-1+\frac{qB}{2\theta}e^{2\theta-1}$, which implies that $-1+\frac{qB}{2\theta}e^{2\theta-1}>t^*$. Moreover, given that the RHS of (A18) is a decreasing function of t, it also follows that $\lim\limits_{t\to t^{*+}}\frac{1}{|\eta|}>1-2\theta$. We can then conclude that

$$\lim_{t\to t^{*-}}\frac{1}{|\eta|}=1-2\theta<\lim_{t\to t^{*+}}\frac{1}{|\eta|},$$

i.e., the function $\frac{1}{|\eta|}$ is discontinuous at $t=t^*$. Regarding the shape of $\frac{1}{|\eta|}$ for $t\in(0,t^*)$, i.e. when the IC constraint is binding, rearranging (16) we have that

$$\lambda=\frac{(1+t)\,(q\alpha+2\theta)-qB}{2\left[(1+t)\,(1+q\alpha)-qB\right]}, \tag{A21}$$

from which we obtain that

$$\frac{\partial\lambda}{\partial t} = \frac{2\left[(1+t)(1+q\alpha)-qB\right]\left[q\alpha+2\theta+(1+t)\,q\frac{\partial\alpha}{\partial t}\right]}{4\left[(1+t)(1+q\alpha)-qB\right]^2} - \frac{2\left[(1+t)(q\alpha+2\theta)-qB\right]\left[1+q\alpha+(1+t)\,q\frac{\partial\alpha}{\partial t}\right]}{4\left[(1+t)(1+q\alpha)-qB\right]^2}. \tag{A22}$$

Taking into account that $\alpha'(t) = -\frac{\alpha}{1+t-qBe^{-q\alpha}}$ for $t \in (0, t^*)$, eq. (A22) can be simplified to obtain

$$\frac{\partial\lambda}{\partial t} = \frac{(1+t-qB)\left[\alpha(1+t)-B\right]}{2\left[(1+t)(1+q\alpha)-qB\right]^2}\frac{(1-2\theta)\,q}{qB-(1+q\alpha)(1+t)} = \frac{(1-2\theta)(1+t-qB)\left[\alpha(1+t)-B\right]q}{2\left[qB-(1+t)(1+q\alpha)\right]^3}. \tag{A23}$$

Given that $1-2\theta > 0$ (by assumption), $1+t-qB < 0$ (a necessary condition for the IC constraint to be binding), $\alpha(1+t)-B < 0$ (since a binding IC constraint requires that $(1+t)\,\alpha - B = -Be^{-q\alpha}$), and $qB-(1+q\alpha)(1+t) < 0$ (since $\alpha'(t) = \frac{\alpha}{qB-(1+q\alpha)(1+t)}$ and we know that $\alpha'(t) < 0$), it follows from (A23) that $\frac{\partial\lambda}{\partial t} < 0$. Therefore, it also follows (see (A17)) that $|\eta|^{-1}$ is a monotonically decreasing function for $t \in (0, t^*)$.

A.4 Proof of Proposition 3

When the IC constraint is slack, the first order condition (18) can be rewritten as (taking into account Proposition 2)

$$\frac{t}{1+t} = 2\frac{1-\delta}{\delta} + \ln\frac{qB}{2\theta(1+t)} \tag{A24}$$

Given that the LHS of (A24) takes value 0 at $t = 0$ and is monotonically increasing in t, and that the RHS is monotonically decreasing in t, for given δ there is at most one value for t in the range $t \in \left(t^*, \frac{qB}{2\theta} - 1\right)$ that satisfies condition (A24).

When the IC constraint is binding, the first order condition (18) can be rewritten as (taking into account Proposition 2)

$$\frac{t}{1+t} = 2\frac{1-\delta}{\delta}\frac{1-\lambda/\theta}{1-2\lambda} + \frac{1-2\theta}{1-2\lambda}. \tag{A25}$$

Next, we show that the RHS of (A25) is either monotonically decreasing in t or monotonically increasing in t depending on whether, respectively, $\delta > (1+\theta)^{-1}$ or $\delta < (1+\theta)^{-1}$. We also show that for $\delta \leq (1+\theta)^{-1}$ the optimal value for t must be necessarily greater than t^*. Therefore, given that the LHS of (A25) takes value 0 at $t = 0$ and is monotonically increasing in t, for given δ there is at most one value for t in the range $t \in (0, t^*)$ that satisfies condition (A25).

We have that

$$\begin{aligned}\frac{\partial\left(2\frac{1-\delta}{\delta}\frac{1-\lambda/\theta}{1-2\lambda}+\frac{1-2\theta}{1-2\lambda}\right)}{\partial t} &= \left[2\frac{1-\delta}{\delta}\frac{-(1-2\lambda)\frac{1}{\theta}+2(1-\lambda/\theta)}{(1-2\lambda)^2}+\frac{2(1-2\theta)}{(1-2\lambda)^2}\right]\frac{\partial\lambda}{\partial t}\\ &= \left[-2\frac{1-\delta}{\delta\theta}\frac{1-2\theta}{(1-2\lambda)^2}+2\frac{1-2\theta}{(1-2\lambda)^2}\right]\frac{\partial\lambda}{\partial t}\\ &= 2\frac{1-2\theta}{(1-2\lambda)^2}\left(1-\frac{1-\delta}{\delta\theta}\right)\frac{\partial\lambda}{\partial t}\\ &= -2\frac{1-2\theta}{(1-2\lambda)^2}\frac{1-\delta-\delta\theta}{\delta\theta}\frac{\partial\lambda}{\partial t}.\end{aligned}$$

We know that $1-2\theta>0$ (by assumption) and that $\frac{\partial\lambda}{\partial t}<0$; thus, we have that

$$sign\left\{\frac{\partial\left(2\frac{1-\delta}{\delta}\frac{1-\lambda/\theta}{1-2\lambda}+\frac{1-2\theta}{1-2\lambda}\right)}{\partial t}\right\}=sign\left\{1-\delta-\delta\theta\right\}. \tag{A26}$$

From the first order condition (14) of the government's problem we have that

$$t=\frac{1-\delta}{\delta}\frac{qB}{\theta}e^{-q\alpha}-\frac{\alpha}{\alpha'}. \tag{A27}$$

Assuming that the IC constraint is binding we have that $qBe^{-q\alpha}=qB-(1+t)\,q\alpha$ and $\alpha/\alpha'=qB-(1+q\alpha)(1+t)$, and therefore we can rewrite (A27) as

$$t=\frac{1-\delta}{\delta\theta}\left[-(1+t)\,q\alpha+qB\right]-qB+(1+q\alpha)(1+t),$$

from which we obtain (after some algebraic manipulations)

$$t=\frac{\delta\theta}{(1-\delta-\delta\theta)\,q\alpha}+\frac{B}{\alpha}-1, \tag{A28}$$

i.e.,

$$\alpha(1+t)=\frac{\delta\theta}{(1-\delta-\delta\theta)\,q}+B.$$

Given that a binding IC constraint requires that $\alpha(1+t)=(1-e^{-q\alpha})\,B$, we can rewrite the equation above as

$$-Be^{-q\alpha}=\frac{\delta\theta}{(1-\delta-\delta\theta)\,q},$$

from which we obtain

$$\alpha=\frac{1}{q}\ln\frac{(-1+\delta+\delta\theta)\,qB}{\delta\theta}, \tag{A29}$$

and therefore, substituting in (A28) the value for α provided by (A29)), we get that

$$t = \frac{\delta\theta}{(1-\delta-\delta\theta)\ln\frac{(-1+\delta+\delta\theta)qB}{\delta\theta}} + \frac{qB}{\ln\frac{(-1+\delta+\delta\theta)qB}{\delta\theta}} - 1. \tag{A30}$$

The equation above gives the optimal value of t as a function of the various parameters when α is implicitly given by the equation $\alpha(1+t) = \left(1-e^{-q\alpha}\right)B$. Notice that a necessary condition for α, as defined by (A29), to be positive is that $1-\delta-\delta\theta < 0$, i.e. $\delta > (1+\theta)^{-1}$. Notice also that, even for a given value of δ that is greater than $(1+\theta)^{-1}$, the value for t provided by (A30) might be larger than t^*, in which case one should conclude that the first order condition of the government's problem cannot be satisfied within the range of values for t that make the IC constraint binding. Thus, a necessary (but not sufficient) condition for the first order condition of the government's problem to be satisfied for $t \in (0, t^*)$ is that $1-\delta-\delta\theta < 0$. Taking into account (A26), this means that a necessary (but not sufficient) condition for the first order condition of the government's problem to be satisfied for $t \in (0, t^*)$ is that $\partial\left(2\frac{1-\delta}{\delta}\frac{1-\lambda/\theta}{1-2\lambda} + \frac{1-2\theta}{1-2\lambda}\right)/\partial t < 0$.

Figures 6-9 below illustrate the various possibilities that can arise when there is at least one value for t that satisfies condition (18). Each figure plots two functions. The increasing function starting at the origin of the axes represents the function $t/(1+t)$, i.e. the LHS of (18). The other function represents the profile of the RHS of (18), plotted for a given value of δ; this function is characterized by a discontinuity point (at $t = t^*$), which reflects the discontinuity of $1/|\eta|$. The passage from one figure to the next can be interpreted as illustrating the effect of lowering the underlying value of δ. This is because a reduction in δ shifts up the function describing the RHS of (18), albeit in a different way for $t < t^*$ (where the shifting is non-parallel) and $t > t^*$ (where the shifting is parallel).[20] Notice also that in drawing figures 6-8 we implicitly assume that $\delta > (1+\theta)^{-1}$ so that, except at $t = t^*$, the RHS of (18) is decreasing in t. In Figure 9 we instead assume that $\delta < (1+\theta)^{-1}$ so that the RHS of (18) is increasing in t for $t \in [0, t^*)$.

Finally, to interpret the figures notice that, if at a given value for t the function $t/(1+t)$ lies below (resp.: above) the other function, it is socially desirable to marginally raise (resp.: lower) t. Moreover, the values for t at which the two functions intersect represent values for t that satisfy the first order condition (18).

In Figure 6 there is only one value for t, lower than t^* and denoted by t^{opt}, that satisfies the first order condition (18).[21] The single value for t that satisfies the first order condition is in this case the optimal tax rate: for all values of t smaller than t^{opt} social welfare increases when t is marginally raised, and for all values of t larger than t^{opt} social welfare decreases when t is

[20] For $t > t^*$ a marginal variation in δ changes the RHS of (18) by $-2\delta^{-2}d\delta$; for $t < t^*$ a marginal variation in δ changes the RHS of (18) by $-2\delta^{-2}\frac{1-\lambda/\theta}{1-2\lambda}d\delta$. Thus, while for $t > t^*$ a variation in δ shifts (up or down) in a parallel way the function describing the RHS of (18), this is not the case for $t < t^*$. However, given that $\lambda \to 0$ when t approaches t^* from the left, we have that $\lim_{t\to t^{*-}} 2\frac{1-\delta}{\delta}\frac{1-\lambda/\theta}{1-2\lambda} = 2\frac{1-\delta}{\delta}$. This implies that the magnitude of the jump at t^* does not vary with δ.

[21] For this to happen, a necessary condition is that the value of δ is sufficiently large (i.e., sufficiently close to 1). The condition is however not sufficient. The reason is that there can be cases when the optimal tax rate is larger than t^* also for $\delta = 1$. If this happens, the optimal tax rate will be larger than t^* for all values of $\delta \in [0.5, 1]$.

marginally raised.

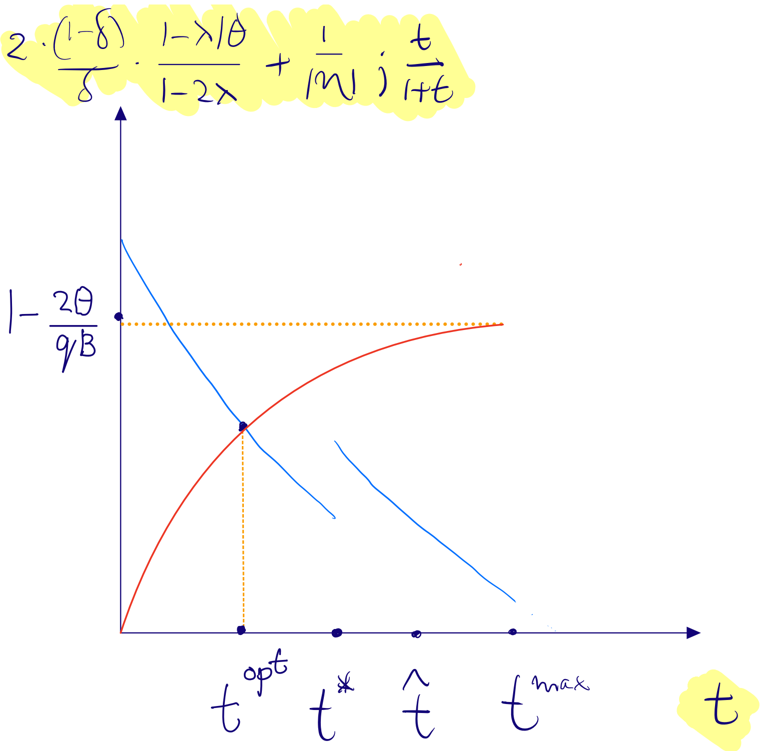

Figure 6: $\delta > (1+\theta)^{-1}$ and $t^{opt} < t^{*}$.

Figure 7 shows a case where a reduction in δ implies that there are two values of t, one smaller than t^{*} (denoted in the figure by $\underline{t}$) and the other larger than t^{*} (denoted in the figure by $\bar{t}$), that satisfy condition (18). For values of t smaller than $\underline{t}$ social welfare increases when t is marginally raised, for $t \in (\underline{t}, t^{*})$ social welfare is decreasing in t, but for $t \in \left(t^{*}, \bar{t}\right)$ social welfare is again increasing in t thanks to the upward jump in $1/|\eta|$. In this case, whether the optimal tax rate is given by $\underline{t}$ or by $\bar{t}$ will depend on whether or not the losses that are accumulated raising t from $\underline{t}$ to t^{*} are more than offset by the gains that accrue raising t from t^{*} to $\bar{t}$.[22]

[22]In principle, it is also possible that social welfare is the same at $\underline{t}$ and $\bar{t}$, in which case the optimal tax rate would not be unique.

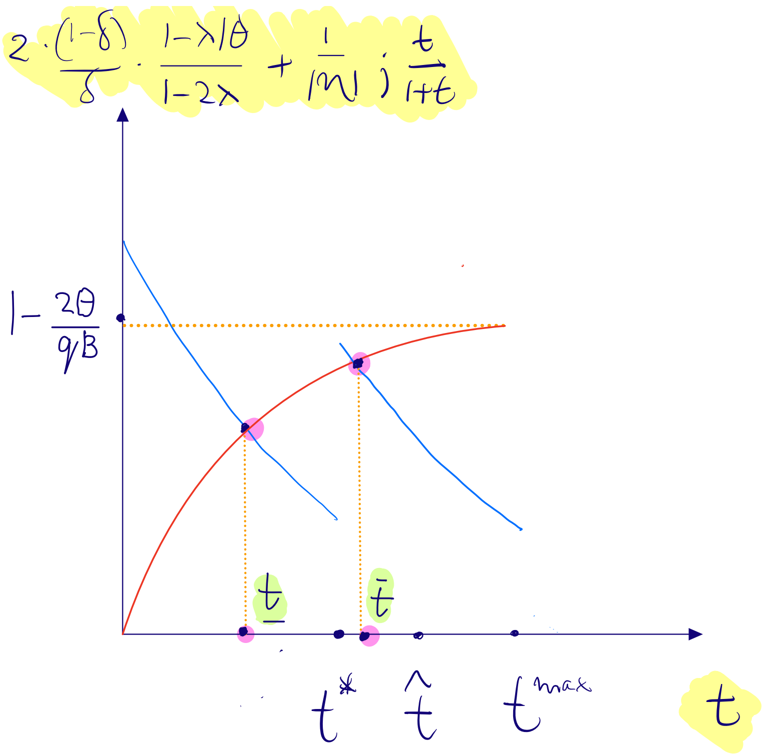

Figure 7: $\delta > (1+\theta)^{-1}$ and two locally optimal solutions.

Figures 8-9 show two cases where, further lowering δ, there is only one value for t, larger than t^* and denoted by t^{opt}, that satisfies condition (18). Once again, the single value for t that satisfies the first order condition is in this case the optimal tax rate.

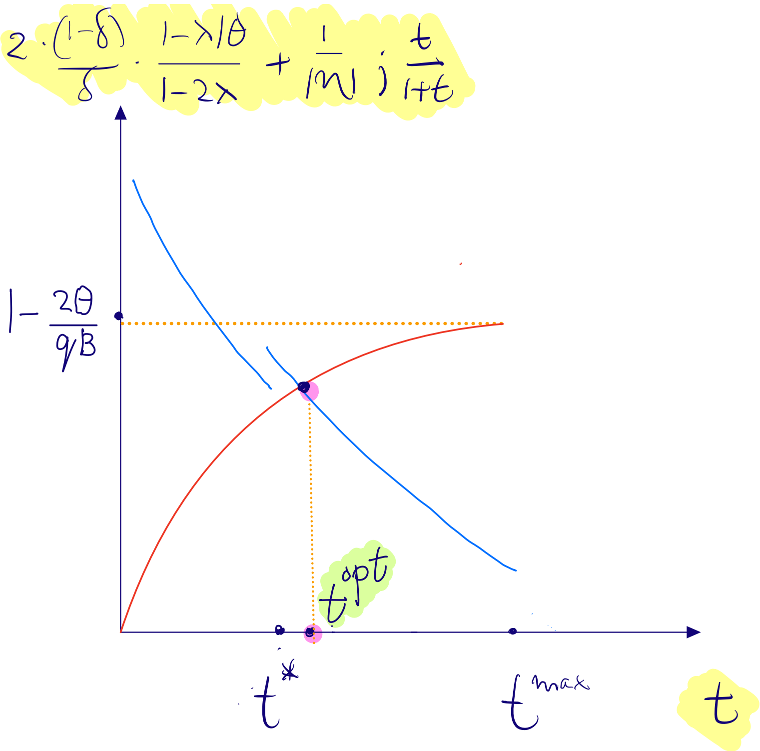

Figure 8: $\delta > (1+\theta)^{-1}$ and $t^{opt} > t^{*}$

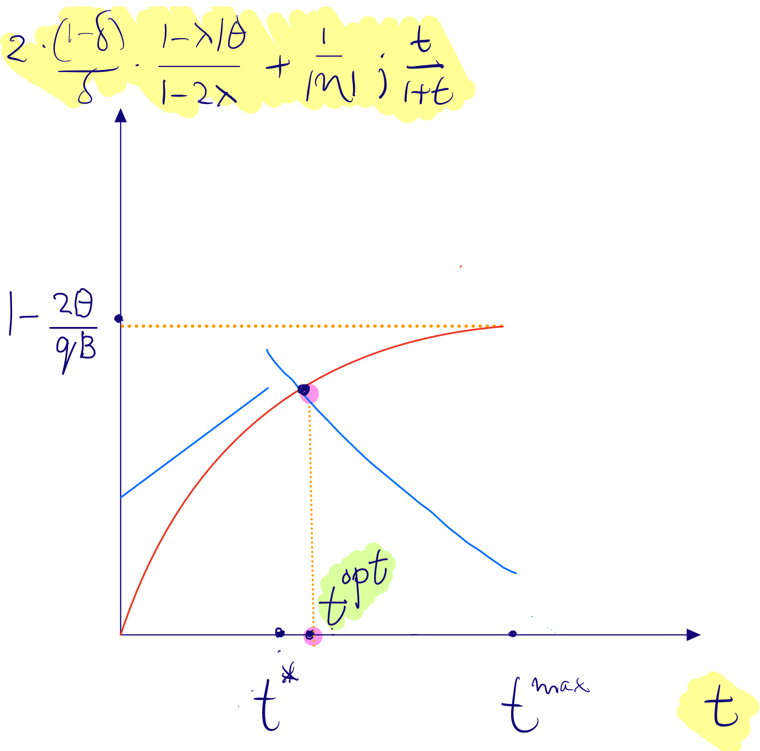

Figure 9: $\delta < (1+\theta)^{-1}$

Apart from the cases when the optimal tax rate fulfills the first order condition (18), there is however the possibility that, for δ sufficiently low (i.e., close to 0.5), the optimal tax policy is given by the corner solution $t = \frac{qB}{2\theta} - 1$ (which induces $\alpha = 0$). To see this, denote by $t^{opt}(\delta)$ the optimal value for t as a function of δ, and remember that we have shown that, for $\delta \leq (1+\theta)^{-1}$, $t^{opt}(\delta)$ necessarily belongs to the set $(t^{*}, \frac{qB}{2\theta} - 1]$. Given our assumption that $0 < \theta < 1/2$, we have that $\min_{0<\theta<1/2} (1+\theta)^{-1} > 2/3$, and therefore $t^{opt}(\delta) \in (t^{*}, \frac{qB}{2\theta} - 1]$ for δ sufficiently close to its lower bound 0.5. When $t^{opt}(\delta) \in \left(t^{*}, \frac{qB}{2\theta} - 1\right)$, it will necessarily satisfy the first order condition (obtained from (18) by using Proposition 2)

$$\frac{t}{1+t} - 2\frac{1-\delta}{\delta} - \ln \frac{qB}{2\theta(1+t)} = 0. \tag{A31}$$

When instead $t^{opt}(\delta) = \frac{qB}{2\theta} - 1$, it will either be the case that $t^{opt}(\delta)$ satisfies the first order condition $\frac{t}{1+t} = 2\frac{1-\delta}{\delta} + \ln \frac{qB}{2\theta(1+t)}$ or that it represents a corner solution (since $\alpha = 0$ for $t \geq \frac{qB}{2\theta} - 1$). Evaluating the LHS of (A31) for $t \to \frac{qB}{2\theta} - 1$, we have that

$$\lim_{t \to \frac{qB}{2\theta} - 1} \frac{t}{1+t} - 2\frac{1-\delta}{\delta} - \ln \frac{qB}{2\theta(1+t)} = 1 - \frac{2\theta}{qB} - 2\frac{1-\delta}{\delta}. \tag{A32}$$

Thus, given that $1 - \frac{2\theta}{qB} - 2\frac{1-\delta}{\delta} \leq 0$ for $\delta \leq \frac{2qB}{3qB-2\theta}$, and since $\frac{1}{2} < \frac{2qB}{3qB-2\theta} < 1$ (under the assumption that $0 < \theta < 1/2$ and $qB > 2\theta$), it follows that for $\delta \in \left[\frac{1}{2}, \frac{2qB}{3qB-2\theta}\right]$ the optimal tax policy will entail $t = \frac{qB}{2\theta} - 1$.