Rethinking Commodity Taxation: The New Status Redistribution Channel

Spencer Bastani, Tomer Blumkin, Luca Micheletto

Discussion Paper No. 24-12

June 17, 2024

Monaster Center for
Economic Research
Ben-Gurion University of the Negev
P.O. Box 653
Beer Sheva, Israel

Fax: 972-8-6472941
Tel: 972-8-6472286

Rethinking Commodity Taxation: The New Status Redistribution Channel

Spencer Bastani* Tomer Blumkin † Luca Micheletto‡

June 17, 2024

Abstract

This paper presents a novel perspective on the taxation of conspicuous consumption, highlighting its potential to achieve a more equitable distribution of welfare by compressing the status distribution. By curbing the conspicuous consumption of the affluent, the government reduces the informativeness of status signaling, leading to an increased share of the social status surplus for the less wealthy. This "status channel" serves as a complement to traditional monetary channels of redistribution. The findings emphasize the importance of incorporating the status dimension in policy design and shed new light on the benefits of taxing conspicuous consumption in pursuit of societal equity.

Keywords: optimal taxation, signaling, status, redistribution

JEL classification: H21, D63, D82

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

†Department of Economics, Ben Gurion University of the Negev; 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

Consumers often allocate their spending to visible goods in order to signal their wealth and social status. This idea has a long history in the social sciences, dating back to ancient times. Plato, for example, eloquently expressed this idea in The Republic: "Since then, as philosophers prove, appearance tyrannizes over truth and is lord of happiness, to appearance I must devote myself" (Plato, Rep. 3.414b [Jowett 1888]). Adam Smith, in his seminal work of 1776, also recognized this concept and provided an illustration of how consumption norms are influenced by social expectations.[1]

The term "conspicuous consumption" was coined by Veblen (1899) to describe the extravagant spending on visible goods that serve as status symbols. Hirsch (1976) also introduced the concept of "positional goods," referring to goods valued for their relative rather than absolute qualities. Building on these basic insights, recent economic literature has developed the theory of conspicuous consumption. In particular, Frank (1997, 2005) emphasized the importance of positional goods and the welfare costs associated with the negative externalities they generate. Bagwell and Bernheim (1996) and Hopkins and Kornienko (2004) made influential contributions to the analysis of signaling games involving conspicuous consumption. Other studies have examined the determinants, effects, and implications of conspicuous consumption in various economic domains, including income inequality, social welfare, taxation, savings behavior, and environmental sustainability.

Empirical studies have provided compelling evidence of the importance of consumption visibility and its impact on consumer behavior. Heffetz (2011), using survey data from the United States, constructed a visibility index for various consumption categories and found that it can account for as much as one-third of the variation in income elasticities across these categories. In addition, experimental evidence supports the role of visibility in signaling social status. For example, Bursztyn et al. (2017) showed that consumers in Indonesia were willing to pay a premium for platinum credit cards, which are more visibly prestigious than gold cards. Butera et al. (2022) showed that donors to the Red Cross exhibited a positive and widespread desire for public recognition of their contributions, indicating a demand for recognition.[2]

Given the prevalence and importance of conspicuous consumption in modern societies, a natural question is how to design optimal consumption taxes that take into account consumers' status motives. Most of the existing literature has focused on corrective (Pigouvian) taxes, which can internalize the positional externalities generated by conspicuous consumption. The literature has also shown that such taxes can improve equity because they are progressive (since status goods are mainly consumed by the rich) and can finance transfers

[1] Smith remarked that "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" (Smith 1776, 5.2.148).

[2] Notably, when charitable acts are driven by status concerns rather than altruism, they may be perceived as a form of conspicuous consumption for signaling purposes, as argued by Glazer and Konrad (1996).

to the poor. [3] However, a novel aspect that has been overlooked in the literature is that conspicuous consumption taxes affect the transmission of information among consumers and, ultimately, the allocation of a potentially large social status surplus.

To further illustrate this point, consider the scenario of implementing mandatory uniform dress codes in schools. In the absence of regulation, affluent students may engage in conspicuous consumption by wearing expensive brand-name clothing as a means of signaling their higher social status at the expense of their less affluent peers in a zero-sum competition for status. To address this problem, regulations can be introduced to enforce a uniform dress code, which would result in an equitable distribution of social status within a pooling equilibrium. Alternatively, the government could allow students to opt out of the dress code by paying a fee. The tax revenue generated by this fee, under the resulting separating equilibrium, can still be used to provide additional tutoring or school supplies that benefit all students, while maintaining an inequitable distribution of social status. Under this policy framework, wealthy students would still have the ability to differentiate, albeit to a lesser extent than under the laissez-faire regime, while the proposed fee/compensation scheme could potentially increase the welfare of less wealthy students.

Another example of status competition is the choice of cars by consumers. Without regulation, wealthy consumers may buy expensive and luxurious cars to signal their high income and gain higher social status at the expense of less wealthy consumers. Regulation can address this problem by imposing a uniform car standard, which leads to an equal distribution of social status in a pooling equilibrium. For example, in East Germany, the car market was dominated by two models, the Trabant and the Wartburg, which were produced by state-owned enterprises. Alternatively, the government can allow consumers to buy more expensive cars by paying a luxury tax. The tax revenue can be used to finance public goods or transfers that benefit all consumers. Under this policy regime, wealthy consumers would still choose to buy more expensive cars, but less so than under the laissez-faire regime, and the welfare of less wealthy consumers could be improved by the proposed tax/transfer scheme.

In both examples, the proposed tax/transfer schemes serve a dual purpose: they reduce income inequality between the wealthy and the less wealthy, while at the same time they reduce status inequality by reducing the informativeness of signaling. As a result, a novel channel of status redistribution emerges that complements the conventional monetary redistribution channel. By implementing these policies, the government not only addresses the economic imbalance, but also reduces the effectiveness of conspicuous consumption as a means of signaling social status. This approach allows for a broader redistribution of both wealth and status, ultimately contributing to a more equitable society.

This paper develops a simple framework for analyzing the role of commodity taxation in pursuing redistributive goals, trading off between the two aforementioned monetary and status channels of redistribution. We view conspicuous consumption as a rent-seeking ac-

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

tivity aimed at extracting a higher share of the social surplus from status. The effectiveness of status signaling, and thus the likelihood of being recognized as high status, depends on the variety of conspicuous goods consumed. We use an extension of Spence (1973) classic signaling game that allows for noisy signals. The game structure implies that in a separating equilibrium, consumption signals are not fully informative due to their limited visibility. Commodity taxation reduces the variety of goods consumed by the rich, making status signals less reliable. This in turn raises the relative status of the poor, who receive a larger share of the social surplus from status. However, this novel mechanism of redistribution through status must be balanced against traditional mechanisms of redistribution through the income channel.

We address the challenge of formulating an optimal government policy aimed at redistributing welfare through a combination of monetary and status-based mechanisms. We explore a range of potential strategies, from relying exclusively on monetary redistribution by using conspicuous consumption as a tax base, to completely suppressing the signaling of social status, thereby foregoing any tax revenue derived from conspicuous consumption. Our main contribution is the recognition of the status channel and that the evaluation of the equity benefits of a high marginal tax should go beyond its revenue raising capacity alone. Drawing parallels to the effects of high marginal income tax rates, which can have favorable predistributive effects (Stiglitz 2018, Bozio et al. 2020), we show that the imposition of higher commodity tax rates can facilitate a more equitable distribution of welfare, even if their contribution to government revenue is relatively modest.

The rest of the paper proceeds as follows. Section 2 presents the main elements of our model, and section 3 describes the two-stage signaling game (between the government and private agents). Finally, section 4 provides a concluding discussion.

2 Model

In the economy under consideration, there are two types of agents, denoted by $j = 1, 2$, characterized by their different wealth endowments. The rich type is represented by the endowment w^2, while the poor type has an endowment of w^1, where $w^2 > w^1 > 0$. The population size of each type is normalized to one for simplicity. It is important to note that individual wealth endowments remain private information, undisclosed to the government and other agents.

Each agent allocates its endowment among $n + 1$ consumption goods. The numeraire good, denoted by y, generates intrinsic utility from consumption and remains unobserved by the government and other agents. In addition, there are n binary signaling goods, denoted by x_i with $i = 1, \ldots, n$, which are used for status signaling purposes. These signaling goods are binary in nature, meaning that an agent can choose to either consume or not consume each signaling good ($x_i \in 0, 1$).

The utility of an agent of type j is given by the following expression:

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

An agent's utility consists of two main factors. The first factor accounts for the intrinsic satisfaction derived from consuming the numeraire good, denoted by y. The second factor captures the utility derived from social status, which is determined by two components. The first component, $\mathbf{P}$, represents the conditional probability that an agent of type j will be perceived as having a high wealth endowment, given the consumption choices of both types of individuals. The perceived wealth endowment is denoted by $\tilde{w}^j$. The second component, $B > 0$, represents the additional social status satisfaction associated with being perceived as a wealthy type.

The vector x_i^j represents the consumption choices made by agents of type $j = 1, 2$ with respect to the pure signaling goods x_i ($i = 1, \dots, n$). It is important to note that these signaling goods are considered "pure" because neither type derives direct utility from consuming them.[4] The formulation of the utility function in equation (1) is chosen to focus specifically on pure strategies when analyzing the perfect Bayesian equilibrium of the signaling game. It is assumed that all agents of the same type will choose the same consumption bundle along the equilibrium path. The formation of perceptions is done by Bayesian updating, conditioned on the vector x chosen by both types along this path.

It is important to note that, in our context, a low level of utility experienced by poor agents is not only a result of being deprived and consuming a smaller quantity of the numeraire good. It is also influenced by the status channel, specifically by being perceived as poor. Consequently, type-1 agents with lower wealth endowments have an incentive to imitate their wealthier counterparts by behaving as if they were type-2 agents. This, in turn, induces wealthy type-2 agents to credibly signal their greater endowments by allocating their resources to conspicuous consumption, which does not provide them with intrinsic utility.

The budget constraints faced by the two types of agents are given by equation (2):

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

where p_i^j represents the price paid by type $j = 1, 2$ when purchasing a unit of good $i = 1, \dots, n$. Specifically, we assume that $p_i^2 = \theta/n$ and $p_i^1 = 1/n$, where $0 < \theta < 1$.

The fact that the unit prices of type 2 are lower than those of type 1 can be interpreted in several ways. First, it is consistent with the canonical signaling model of Spence (1973), where the cost of acquiring the signal may be lower for type 2. Second, the lower cost of type 2 may reflect, in a simplified form, the diminishing marginal utility of consuming the

[4] Our assumption simplifies the analysis for tractability, but the qualitative features of the analysis would remain unchanged even if intrinsic utility were derived from these signaling goods, as long as wasteful signaling is measured relative to a nonzero reference level (which may differ across types).

numeraire good y, which is assumed to be constant for tractability. This implies that the effective prices of x decrease with respect to the wealth endowment. Third, the lower costs incurred by type 2 may reflect heterogeneity in preferences, where type 2 derives either direct utility or higher direct utility from the consumption of x. In this case, the effective price incurred by type 2 would be lower than that incurred by type 1.

To illustrate this, consider the scenario where $n = 1$ (and let $x_1^j \equiv x^j$ for simplicity). If the utility derived by type 2 is given by $u^2(y^2, x^1, x^2) = y^2 + \frac{1}{n}(1-\theta)x^2 + \mathbf{P}[\tilde{w}^2 = w^2|x^1, x^2] \cdot B$, Assuming that both types incur the same price per unit of x (given by $1/n$) implies that the budget constraint faced by type 2 is $w^2 = y^2 + \frac{1}{n}x^2$. By substituting y^2 from the budget constraint into the utility function, we obtain an expression identical to the utility function in equation (1), after substituting for y^2 from equation (2).

We now consider the role of visibility in the status generation process, drawing on the literature on noisy signaling (e.g., Matthews and Mirman 1983).[5] We plausibly assume that conspicuous consumption is not perfectly observed by the target population, introducing an element of noise into the signaling mechanism. This feature is fundamental to our model and will have important implications for subsequent policy analysis. In particular, in a separating equilibrium, each individual faces a trade-off between allocating more resources to the numeraire good y or investing more in signaling activities. The latter increases the probability of being perceived as a high type, leading to a higher expected level of social status. This trade-off arises because of the imperfect observability of signals. With perfectly observed signals, there would be no such trade-off in a fully revealing separating equilibrium. The trade-off provides the government with an opportunity to redistribute resources through the status channel. By controlling the level of noise, which is endogenously determined in equilibrium, the government can influence the signaling process and thereby engage in redistribution through the status channel. This complements the traditional income channel of redistribution, allowing for multiple ways of achieving desired distributional outcomes.

For the sake of tractability, we adopt a relatively simple structure for the stochastic process that determines the level of noise in our model. It is important to note, however, that the qualitative results of our analysis remain robust to alternative specifications of the stochastic process.

Specifically, we assume that from an observer's point of view, the visibility of each signaling good x_i ($i = 1, \ldots, n$) purchased by a signaling agent is determined by a binary random variable denoted z. The random variable z takes the value 1 ("visible") with the probability $0 < q/n < 1$ and the value 0 ("not visible") with the complementary probability $0 < 1 - q/n < 1$. The realization of z is assumed to be independent across different signaling goods. Note that for simplicity, we invoke symmetry across goods in our analysis. However,

[5] The paper by Matthews and Mirman (1983) is widely regarded as a seminal paper in the noisy signaling literature. It introduced the concept of noisy signals in the context of a classical model of limit pricing. More recently, de Haan et al. (2011) provided theoretical analysis and experimental evidence on the behavioral implications of different levels of noise in signaling games.

we discuss a generalization that allows for asymmetry across goods in Online Appendix B.

A straightforward interpretation of the stochastic process generating the noise draws on the literature on informative advertising, following the seminal studies of Butters (1977) and Grossman and Shapiro (1984). According to this approach, an individual's consumption of each signaling good is viewed as an advertisement that is sent to the target population and received with some probability. For example, suppose there are n social events over a period of time (say, a year). The number of individuals attending a single social event is given by $q/n{\cdot}B$ in the target population of size B. That is, a fraction q/n of the target population is exposed to each event. Suppose that the target population is randomly assigned to the events, and further suppose that the probability that any individual in the target population participates in a given event, q/n, is independent of the probability of participating in any other event.

Now consider an individual who engages in signaling by renting a luxury car for the event he attends. By showing up with a luxury car, the individual is signaling that he or she is a wealthy individual. The individual may choose to rent a car for a subset of events (possibly none or all as special cases). Willing to "impress" as many individuals from the target population as possible, the independence property implies that different individuals would attend different events, so there is an incentive to rent the car for many events to increase exposure. Assuming that the signaling agent rents a car and attends m events (not for the intrinsic benefit of attending the event, but for a signaling motive), the probability of being perceived as a wealthy type increases with the likelihood of exposure, i.e., the likelihood that an individual in the target population will attend at least one of these events. The probability of exposure increases with the number of events attended by the signaling agent.

The benefit from social status can be captured simply by the observer's perception, namely the probability that the signaling agent is perceived by the observer as the rich type. If the observer is exposed, he believes with probability 1 that the agent is rich. Otherwise, if the observer is not exposed, he assigns a probability of 1/2, based on the prior distribution of types in the population. The gain from status associated with this observer is thus given by applying the law of total probability, taking into account the probability of exposure. Aggregating over the entire target population (of observers) yields the total social status utility.

3 The two-stage game

We analyze a two-stage game. In the first stage, the government imposes a uniform ad valorem tax, denoted $t \geq 0$, on all signaling goods. This tax affects the after-tax prices of each signaling good for types $j = 1, 2$. Specifically, the after-tax prices faced by types 1 and 2 are given by $p^1 = (1+t) \cdot \frac{1}{n}$ and $p^2 = (1+t) \cdot \frac{\theta}{n}$, respectively, with $p^1 > p^2$. The tax revenue generated is used to finance a universal lump-sum transfer, denoted by T. In the second stage, given the tax instruments implemented by the government, a signaling game unfolds. Each agent decides, on the basis of his wealth endowment, how to allocate his resources among consumption goods. The tax instruments (t and T) set by the government in the first stage are chosen to maximize social welfare, subject to a balanced revenue constraint and taking into account the optimal choices of consumers in the second stage of the game.

Next, we analyze the second stage and then solve the government constrained maximization program.

3.1 Stage II: The signaling game

Consider a separating equilibrium in which type 2 agents allocate their spending between the numeraire good y and a set of signaling goods ($0 \leq m \leq n$), while type 1 agents spend their entire wealth endowment solely on the numeraire good.[6]

Formally, the separating equilibrium can be characterized as the solution to the following constrained maximization program:

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

subject to

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

It is important to note that in cases where none of the signals are observed, which occurs with a probability of $(1 - \frac{q}{n})^m$ based on the stochastic process outlined earlier, the social status surplus is evenly distributed between the two types according to the prior symmetric distribution. However, in situations where at least one signal is observed that occurs with a probability of $1 - (1 - \frac{q}{n})^m$, a Bayesian update leads to a posterior distribution that supports full separation. Consequently, the entire social status surplus is received by the high-type agents.

However, due to the presence of noisy signaling and partial observability, the low-type agents benefit from the 'benefit of the doubt' and are able to derive a positive fraction of the surplus under the separating equilibrium. Furthermore, this 'benefit of the doubt' derived by the low-type agents diminishes as the intensity of the signaling chosen by the high-type agents, m, increases. These two qualitative features of the model are generic and do not depend on the specific stochastic process invoked that generates the noise.

The constraint (4) captures a standard incentive constraint known as the no-mimicking constraint. This constraint may or may not be binding in the optimal solution, and states that low-type agents have a weak preference for not spending on the signaling goods.

Let $\alpha \equiv \frac{m}{n}$ denote the fraction of signaling goods on which the high types allocate their

[6] In non-equilibrium situations, where agents' beliefs cannot be formed based on Bayesian updating, an agent is assumed to be perceived as a low type with probability 1. In other words, if agents do not adhere to the equilibrium strategies and their behavior deviates from the expected pattern, they are assumed to be perceived as low types by default.

spending. Suppose also that both m and n are large. In this case, we can use the fact that $e = \lim_{h\to\infty}\left(1+\frac{1}{h}\right)^h$ to reformulate the constrained maximization problem in (3)–(4) as follows:

Problem $\mathcal{P}1$

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

subject to

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

Before proceeding with the formal analysis of the constrained maximization program in $\mathcal{P}1$, two important observations must be made.

First, note that the reformulated constrained maximization problem in (5)–(6) implicitly exchanges the 'max' and 'limit' operators. In other words, it maximizes the limiting expression instead of 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 pointwise. In our analysis, we assume that α is chosen from a fixed finite partition of the unit interval $[0,1]$. Under this assumption, uniform convergence is guaranteed. However, since we can set the finite grid to be arbitrarily fine, we adopt the continuum approximation for the following analysis.

Second, within $\mathcal{P}1$ we implicitly assume that type-1 agents have two choices: either to refrain from spending on the signaling goods, or to mimic the spending behavior of type-2 agents by choosing the same α. In principle, type-1 agents could choose to spend on a smaller subset of signaling goods ($0<\tilde{\alpha}<\alpha$). However, it can be shown that the constrained maximization program solved by type-1 agents, which involves maximizing their expected utility by choosing $\tilde{\alpha}\in[0,\alpha]$, is strictly convex given the separating equilibrium strategies outlined above. As a result, we can focus on the two corner solutions: $\tilde{\alpha}=0$ and $\tilde{\alpha}=\alpha$.[7] The incentive compatibility constraint (6) is thus well defined. The formal details are given in Online Appendix A.1.

Before proceeding with the analysis, we invoke Assumption 1, which guarantees an interior solution for α in the laissez-faire ($t=0$).

Assumption 1

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

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

Proposition 1 *For $\theta < 1/2$, there exists a threshold value, denoted by $0 \leq t^* < \frac{qB}{2\theta}$-1, such that the*

[7] It is worth noting that spending on a larger subset, $\alpha' > \alpha$, would be suboptimal because of our assumption about off-equilibrium beliefs (see footnote 6) and because the cost of acquiring the signal is higher for the low type.

optimal solution to $\mathcal{P}1$ can be characterized as follows:

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

where $\alpha_2(t)$ is the strictly positive interior solution to the binding incentive compatibility constraint (6), and $\frac{1}{q} \ln \frac{qB}{2\theta(1+t)}$ represents the optimal solution to the unconstrained maximization of (5). Conversely, for $\theta \geq 1/2$, we have $\alpha(t) = 0$ for all $t \geq 0$.

Proof. See Online Appendix A.2. ∎

Proposition 1 highlights that status signaling serves two purposes for type-2 agents. First, by increasing the number of status goods purchased, they can increase their expected utility from status. This property holds for sufficiently high tax rates, where the incentive compatibility constraint is not binding and the threat of mimicking by type-1 agents is not a concern. In this case, the choice of α achieves the optimal tradeoff between status and non-status goods, given the noisy signaling environment. 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 allows high types to distinguish themselves from their less wealthy counterparts. This property may hold for sufficiently low tax rates where the incentive compatibility constraint is binding. By spending on signaling goods, high types can effectively distinguish themselves from low types and maintain their higher perceived social status. [8]

Based on the characterization of the optimal solution to problem $\mathcal{P}1$, we can observe that $\frac{d\alpha}{dt} \equiv \alpha'(t) < 0$ for all $t < \frac{qB}{2\theta} - 1$.[9] This means that as tax rates on signaling goods increase, type 2 individuals choose to spend their income on a smaller subset of signaling goods, reducing the intensity of signaling.

We now turn to an analysis of the first stage of the game, in which the government sets its tax instruments. To make 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 redistribution policy. The (binding) revenue constraint is formulated as follows:

$$\theta \cdot \alpha(t) \cdot t = 2T, \tag{8}$$

[8] Formally, one can show that the threshold, t^*, is strictly positive when θ approaches $1/2$ from below. In this case the IC constraint binds for all $0 \leq t < t^*$.

[9] See online appendix A.2. For the case where (IC) is slack, this follows immediately from (A8). If (IC) is binding, it follows from Figure 4. Note that we are focusing on the interior solution. Thus, as the red curve shifts upward, the intersection shifts to the left.

where $\alpha(t)$ denotes the optimal fraction of signaling goods that type-2 spends on in equilibrium, and is characterized by Proposition 1.

In a separating equilibrium, where type-1 agents refrain from signaling and spend all their wealth on the numeraire good y, their equilibrium utility is given by:

$$u^1 = w^1 + T + \frac{1}{2} \cdot e^{-q\alpha(t)} \cdot B = w^1 + \frac{1}{2} \cdot \theta \cdot \alpha(t) \cdot t + \frac{1}{2} \cdot e^{-q\alpha(t)} \cdot B. \tag{9}$$

The second equality follows by substituting the expression for T from the revenue constraint in equation (8). We consider an egalitarian government that aims to maximize the welfare of type-1 agents. The social welfare measure is given by

$$W = \delta \cdot \left[w^1 + \frac{1}{2} \cdot \theta \cdot \alpha(t) \cdot t\right] + (1-\delta) \cdot \left[\frac{1}{2} \cdot e^{-q\alpha(t)} \cdot B\right]. \tag{10}$$

The weight $\delta \in [0.5, 1]$ represents the importance assigned to consumption of the numeraire good, while $(1-\delta)$ represents the weight assigned to social status.[10] By differentiating equation (10) with respect to t, we obtain the first-order condition (FOC) for the government maximization program:

$$\frac{\delta}{2}\left[\alpha(t) + t\alpha'(t)\right]\theta - \frac{1-\delta}{2} qB\alpha'(t)e^{-q\alpha(t)} = 0, \tag{11}$$

or equivalently,

$$\delta\theta\left(t + \frac{\alpha(t)}{\alpha'(t)}\right) - (1-\delta)\, qBe^{-q\alpha(t)} = 0. \tag{12}$$

By denoting η as the elasticity of α with respect to the after-tax price $(1+t)$, defined as $\eta \equiv \frac{\alpha'(t)\cdot(1+t)}{\alpha(t)}$, equation (12) can be rewritten as follows:

$$\frac{t}{1+t} = \frac{1-\delta}{\delta\theta}\frac{qBe^{-q\alpha(t)}}{1+t} + \frac{1}{|\eta|}. \tag{13}$$

Equation (13) shows that, except in the limiting case when $\delta = 1$, the optimal tax rate exceeds the Laffer rate $t/(1+t) = |\eta|^{-1}$. When $\delta = 1$, meaning that the government ignores social status, the optimal tax rate follows the inverse elasticity rule. In this scenario, revenue is maximized by taxing signaling goods (purchased only by high types), and redistribution is achieved exclusively through the income channel by maximizing the demogrant value T. For $\delta \in [0.5, 1)$, which includes both non-paternalistic government ($\delta = 0.5$) and partially paternalistic government (with δ between 0.5 and 1), the optimal tax rate exceeds the Laffer rate. This divergence arises because revenue considerations are tempered by the desire to promote a more equitable distribution of social status using t as an instrument.

[10]For $\delta = 0.5$, the welfare measure is non-paternalistic, consistent with the utility derived by type-1. However, for $\delta > 0.5$, the welfare measure becomes paternalistic, exhibiting a bias toward the utility derived from consumption of the numeraire good. In particular, when $\delta = 1$, the welfare measure reflects a preference for "income maintenance," ignoring the utility derived from social status. For further discussion of the concept of status laundering in social welfare functions, see Aronsson and Johansson-Stenman (2018).

To gain further insight into the optimal tax system, we can reformulate equation (13). By introducing λ as the Lagrange multiplier associated with the IC constraint (6), we can express the first-order condition for the high-type agent's optimal choice of α as follows:

$$-\left(1+t\right)\theta + \frac{qB}{2}e^{-q\alpha} + \left(1+t-qBe^{-q\alpha}\right)\lambda = 0. \tag{14}$$

Solving for $e^{-q\alpha}$ in (14), we obtain the expression:

$$e^{-q\alpha} = \frac{2\left(1+t\right)\left(\theta-\lambda\right)}{\left(1-2\lambda\right)qB}. \tag{15}$$

Substituting the value of $e^{-q\alpha(t)}$ from (15) into (13), we can restate (13) as follows:

$$\frac{t}{1+t} = 2\frac{1-\delta}{\delta}\frac{1-\lambda/\theta}{1-2\lambda} + \frac{1}{|\eta|}. \tag{16}$$

Recalling that t^{*}, when bounded away from zero, represents the threshold for t that distinguishes the region where the IC constraint is binding from the region where it is slack, the following proposition characterizes the relationship between the elasticity $|\eta|$ and the tax rate.

Proposition 2 *The absolute value 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 monotonically increases in t since λ decreases 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

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

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

Proof. See Online Appendix A.3 ■

Note that according to (iii), the inverse elasticity is equal to $\ln \frac{qB}{2\theta(1+t)}$ and reflects the return to status signaling, which is given by the ratio of the expected benefit to the cost of signaling.[11] The characteristics of the inverse elasticity $|\eta|^{-1}$ as a function of t are illustrated

[11] To see this, consider a single signaling good x that is visible with probability q and has an associated cost θ, while the gains from status are denoted by B. If x is not purchased, the status surplus is divided equally between the two agents, and type 2 derives an expected net benefit of $B/2$. Alternatively, if type 2 purchases a unit of x that costs θ, the social status derived by type 2 is $qB + (1-q)B/2$, since x is only visible with probability q. The net benefit associated with spending on x is therefore $[(qB + (1-q)B/2) - B/2] = qB/2$. Dividing by the cost θ gives $qB/2\theta$.

graphically in Figure 1, based on the presumption that $t^*>0$. The figure also shows that two different tax rates can be consistent with the same elasticity. For example, if we define $\hat{t} = -1 + \frac{qB}{2\theta}e^{2\theta-1}$, we have that $\lim_{t\to t^{*-}} |\eta(t)| = |\eta(\hat{t})| = \frac{1}{1-2\theta}$.

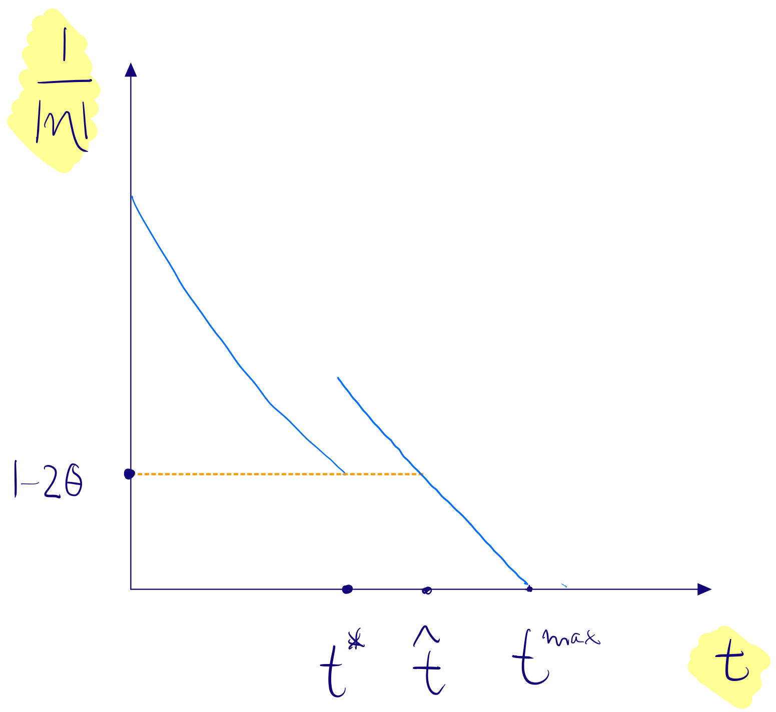

Figure 1: The shape of $|\eta|^{-1}$.

The reason why there is a discrete decrease 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 $(1+t)\,\alpha = (1-e^{-q\alpha})\,B$, while 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 break the continuity of $\alpha(t)$ at $t = t^*$, it does imply a discontinuous decrease in $\alpha'(t)$ at $t = t^*$. 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.[12] Clearly, the discontinuity property disappears when the IC constraint is slack throughout the entire range (i.e., $t^* = 0$). As will be explored in the analysis which follows, the (potential) kink in the demand for status goods would allow for multiple local optima.

In Proposition **??** we provide a characterization of the optimal tax policy based on the analysis of the inverse elasticity $|\eta|^{-1}$ as a function of t. The characterization takes into account the first-order condition (16) for the government problem, considers the potential existence of multiple solutions, and distinguishes between local and global optima.

Proposition 3 *There exist threshold levels , $0.5 < \hat{\delta} < \tilde{\delta} \leq 1$ such that the optimal solution for the government problem is given by:*

[12] To see this, note that since 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., $\lim_{t\to t^{*-}} |\alpha'(t)| = \frac{\alpha}{(1+t)(1+q\alpha)-qB} > \frac{1}{(1+t)q} = \lim_{t\to t^{*+}} |\alpha'(t)|$).

$$t(\delta)=\begin{cases} qB/2\theta-1 & 0.5\leq\delta\leq\hat{\delta} \\ \bar{t}(\delta) & \hat{\delta}<\delta\leq\tilde{\delta} \\ \underline{t}(\delta) & \tilde{\delta}<\delta\leq 1 \end{cases}$$

where:

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

and

$\frac{\underline{t}(\delta)}{1+\underline{t}(\delta)}=\frac{2(1-\delta)}{\delta}\cdot\frac{1-\lambda/\theta}{1-2\lambda}+\frac{1-2\theta}{1-2\lambda}$

Moreover, $t(\delta)$ is strictly decreasing with respect to δ for all $\delta>\hat{\delta}$.

Proof. See Online Appendix A.4 ∎

Proposition 3 focuses on cases where the government places at least as much weight on consumption as it does on social status ($\delta\geq 1/2$), with the non-paternalistic case represented by $\delta=1/2$. The proposition shows that when δ is sufficiently large, the government does not suppress conspicuous consumption because it serves as a source of tax revenue that can be used to achieve an egalitarian distribution of consumption. However, as δ decreases, the proposition shows that the optimal approach is to suppress signaling and achieve redistribution primarily through the status channel. In particular, this optimal policy is consistent with the non-paternalistic case ($\delta=1/2$). [13]

In summary, Proposition 3 highlights the novel and potentially significant role of redistribution via the signaling channel, in addition to the traditional income channel, when individuals are concerned about social status and engage in conspicuous consumption to signal their wealth.

According to equation (16) and for given values of δ and η, the upward adjustment to the Laffer rate required for status redistribution is smaller when the IC constraint is binding ($\lambda\neq 0$). This is because a binding IC constraint implies an upward bias in the choice of α relative to the choice made by a type-2 agent in the absence of a type-1 mimicking threat. The smaller adjustment is due to two factors. First, a marginal reduction in α yields smaller gains in terms of status redistribution, as the effects diminish with larger values of α. Second, the base-broadening effect of the binding IC constraint makes it more effective to achieve redistributive goals through the traditional income channel.[14]

Figures 2 and 3 provide graphical illustrations of the two possible profiles of the optimal tax function $t(\delta)$, demonstrating the different adjustments of the Laffer rate depending on the binding or slack IC constraint. The two figures focus on the case where $t^*>0$, hence the IC constraint (associated with type-2 individual optimization program) is binding for small enough values of t. Notice that the case where $t^*=0$, in which the IC constraint is slack

[13]The non-paternalistic optimum, where redistribution occurs exclusively through the status channel, is a result of the linearity of the signaling cost. In a more general setting, the optimal policy would involve redistribution through both channels, but the status channel would always require a tax rate higher than the Laffer rate.

[14]Formally, the partial derivative $\frac{\partial}{\partial\lambda}\left(\frac{1-\lambda/\theta}{1-2\lambda}\right)=\frac{2-1/\theta}{(1-2\lambda)^2}<0$ (for $0<\theta<1/2$). The negative derivative implies that as λ increases from zero to positive values (corresponding to the IC constraint turning binding), the adjustment to the Laffer rate decreases.

for all values of t, is qualitatively similar to the scenario presented in Figure 2 and is hence omitted.

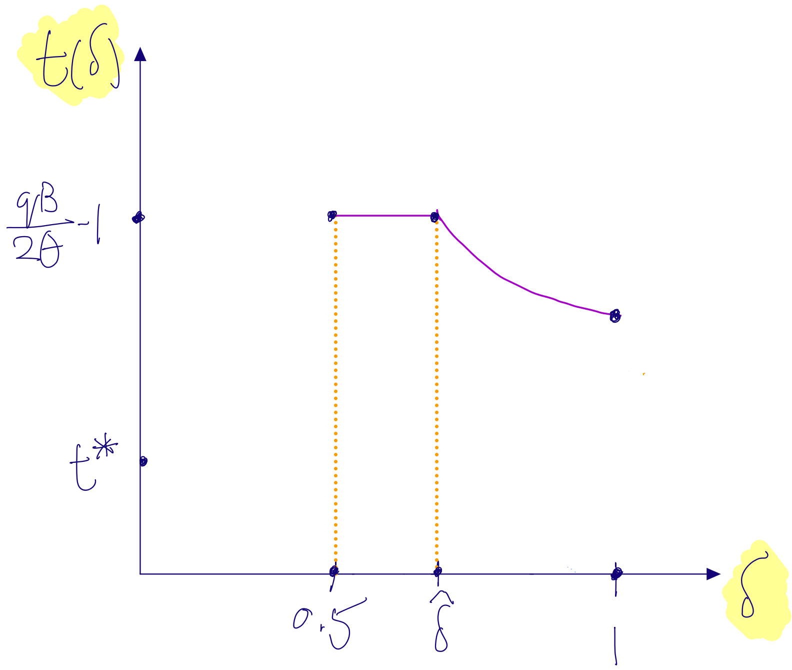

Figure 2: Illustration of the shape of $t(\delta)$ when $t(1)$ is larger than t^*.

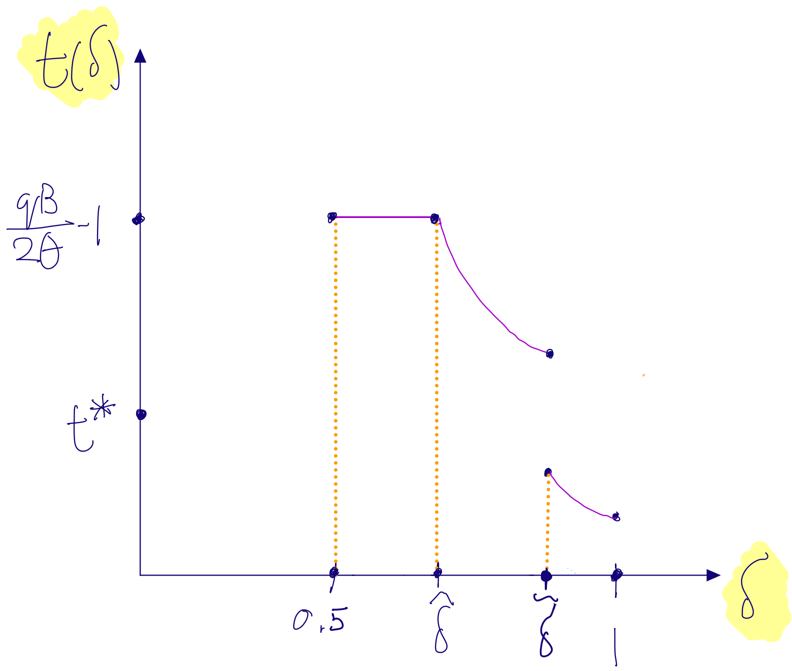

Figure 3: Illustration of the shape of $t(\delta)$ when $t(1)$ is smaller than t^*.

The two figures, Figure 2 and Figure 3, highlight a crucial difference in the behavior of the optimal tax function $t(\delta)$. In Figure 2, where $t(1) > t^*$, the function $t(\delta)$ is continuous within its range $[0.5, 1]$. It starts with a constant optimal tax rate of $\frac{qB}{2\theta} - 1$ for δ in the range

$[0.5, \hat{\delta}]$. Then $t(\delta)$ decreases monotonically and continuously until it reaches its minimum for $\delta = 1$. Notice that this scenario trivially holds when $t^* = 0$, namely, the IC constraint is slack for all values of t.

In contrast, in Figure 3, where $t(1) < t^*$, the optimal tax function $t(\delta)$ has a point of discontinuity. Similar to Figure 2, there is an initial range where the optimal tax rate is constant. However, starting at $\delta = \hat{\delta}$, there is a continuous and monotonically decreasing region. But instead of extending all the way to $\delta = 1$, there is a threshold, denoted $\widetilde{\delta}$, where a jump occurs. At this threshold, $t(\delta)$ jumps from a value strictly greater than t^* to a value strictly less than t^*. From there, the function follows a continuous decrease.

The presence of a discontinuous jump in the optimal tax function $t(\delta)$ indicates that, as societies become more sensitive to status differences over time, a marginal decrease in δ (i.e., an increase in the weight assigned to the status component by the government) can trigger a regime change in tax policy. This change shifts from a low-tax regime, where most redistribution occurs through the income channel, to a high-tax regime, where the focus of redistribution shifts to the status channel. This observation suggests that similar countries (in terms of demographics, preferences and technology) may differ substantially in their tax regimes

3.3 The efficiency-enhancing role of commodity taxation

So far, we have focused on equity considerations. However, it is important to briefly discuss the efficiency aspects of our model. As is typically the case with pure signaling, the resulting allocation in equilibrium is inefficient. In our setup, the status surplus is fixed at B, and signaling by type-2 agents is essentially a form of rent-seeking. The only Pareto efficient allocation is one in which type-2 agents refrain from signaling and set $\alpha = 0$. This implies that a Pareto improvement can be achieved by reducing the intensity of signaling (a decrease in α), complemented by an appropriate transfer from type-1 to type-2 agents (in units of the numeraire good, y, due to the quasi-linearity of the utility function). In general, this can be achieved by a system of non-linear commodity taxes.

However, even in the linear regime considered in our model (ad valorem taxes on the signaling goods accompanied by a universal lump-sum transfer), it is still potentially possible to achieve a Pareto improvement, in case the incentive constraint is binding when $t = 0$. Thus, although the tax revenue is shared between the two types, and thus the cross-subsidy goes in the "wrong" direction from type-2 to type-1 agents, the former may still be better off due to the reduction in the extent of (excessive) signaling, which is desirable in light of the binding incentive constraint.[15] The feasibility of achieving a Pareto improvement depends on the size of the distortion associated with over-consumption of signaling goods (due to the binding incentive constraint) and the degree of cross-subsidization required to maintain the information rent associated with the low-type agents. With a sufficiently large distor-

[15] The argument is similar to the role played by a binding parental leave mandate in a labor market plagued by adverse selection (see, e.g., Bastani et al. 2019).

tion and a relatively moderate amount of cross-subsidization, Pareto improvement becomes feasible.

It is important to note the difference from the standard argument supporting Pareto improvement in the presence of wasteful signaling when linear instruments are present. 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, which could be diverted to consumption (via transfers). Thus, given that the tax parameters are common knowledge, paying taxes could serve as an instrumental signal. Therefore, it is always desirable to tax signals. In our context, however, signals are not wasteful in the sense that consumption must be visible in order to gain status. Thus, status is determined by the number of units (or variety) of signaling goods purchased rather than by the resources spent on those goods. It may therefore be the case that the laissez-faire equilibrium is second-best efficient, i.e. linear instruments need not yield a Pareto improvement.[16]

4 Concluding discussion

In this paper, we identify a novel "status" channel of redistribution. We consider a model in which individuals value social status, which is realistically determined by spending on acquiring noisy consumption signals (conspicuous consumption). In this context, we show that commodity taxes can be a useful tool to reduce the level of conspicuous consumption of the wealthy individuals, and thereby the visibility/informativeness of their acquired signals. The latter serves to raise the welfare of the poor individuals by generating a more egalitarian distribution of status.

The optimal policy trades off between attaining redistribution via the income channel and promoting redistribution through the status channel. The social optimum depends on the extent to which the government recognizes status concerns as a determinant of individual welfare. Assuming that the government's goal is to maximize the welfare of the poor, we find that if the government places a relatively high weight on status utility relative to normal consumption, the optimal tax strategy involves the complete suppression of signaling, resulting in no tax revenues being raised. In this scenario, redistribution is achieved solely by promoting an egalitarian distribution of social status. In contrast, when the government places a relatively low weight on status utility, the optimal tax strikes a balance between the motive to tax conspicuous consumption in-order-to raise revenues serving to fund transfers, and the newly discovered motive to promote an egalitarian distribution of social status. Status-driven signaling has been already well embedded in the optimal tax framework. The literature typically considers noise-less signals and focuses on a separating tax equilibrium, in which types are fully revealed, and hence the distribution of social

[16] Ng (1987) and Truyts (2012) discuss the use of linear instruments, including confiscatory tax rates on pure signaling goods, to achieve Pareto improvements.

status is maintained fixed. The government goal is, therefore, confined to enhancing redistribution via the income channel. Our model, in contrast, considers a setup with noisy consumption signals, in which types are never fully revealed in equilibrium, and the extent to which they are depends on the level of expenditure on the consumption signals. It enables us to characterize the optimal mix of redistribution via the income and status channels, by examining a continuum of (partially) separating tax equilibria. The equilibria differ in the extent to which signals are informative and range from a scenario where agents spend significantly on signaling, in which substantial tax revenues are being raised and serve to finance generous transfers, whereas the distribution of status in highly inequitable; to a scenario in which agents utterly refrain from engaging in signaling (a pooling equilibrium) in which no revenues are being raised and the distribution of status is perfectly equitable (poor and wealthy individuals are indistinguishable). From a methodological perspective, the combination of a two-type setup with a noisy continuous signal enables us to 'replicate' a much more complex framework with many types and a noise-less signal, in which we would compare the social desirability of a bunch of tax equilibria (separating, hybrid, pooling). Our analysis demonstrates the role of the tax system in selecting the equilibrium configuration, which determines the degree of redistribution via the status channel, in addition to optimally choosing the tax instruments to determine the extent of redistribution through the income channel per given equilibrium configuration. The equilibrium selection component was entirely missing from the literature on status and optimal taxation. Our noisy signaling framework further implies that the incentive compatibility constraint associated with the mimicking poor individuals (aiming to gain social status by pretending to be wealthy) need not necessarily bind in equilibrium, in sharp contrast to the canonical noise-less signaling setup. The latter implies that the laissez-faire equilibrium in the presence of signaling may be second-best efficient in contrast to the existing literature. The possibility of having either a binding or a slack incentive compatibility constraint implies that the demand for conspicuous consumption has a kink. This implies the possibility of having multiple local tax optima and the potential discontinuity of the optimal tax function.

We characterize the modified Ramsey Rule for the optimal taxation of conspicuous consumption goods which serve to promote re-distributive goals. A key observation is that the extent to which taxes can raise revenues is inversely related to the price elasticity of conspicuous consumption, whereas the extent to which taxes can be used to achieve a more egalitarian distribution of social status is positively related to the price elasticity of such consumption. Thus, to the extent that conspicuous consumption is highly elastic, it is better to rely on the status channel for redistribution, than to rely on the classical Ramsey (1927) logic in determining optimal tax rates to raise revenue to support the poor (namely, setting taxes at their Laffer rates).

Our results underscore the importance of considering the equity gains associated with a high tax rate beyond its revenue-raising effect. High commodity tax rates can contribute to a fairer distribution of status, even if they generate minimal government revenue. This highlights the importance of looking beyond revenue generation when assessing the impact of commodity tax rates on promoting fairness and equity. The role played by the supplementary status channel is related to the recent strand in the optimal tax literature on the pre-distributive merits of taxation (see, e.g., Stiglitz 2018, Bozio et al. 2020, Bastani et al. 2024).

Our results also bear implications for the design of optimal transfer and anti-poverty programs. These programs should consider not only resource allocation, but also strategies to minimize negative exposure and stigma, adding a new perspective to the concept of the "visible poor" (Blau 1992). A notable example of a policy reform in this direction is the switch to the use of 'debit cards' when implementing the federal SNAP (formerly known as 'Food Stamps') poverty alleviation program in the US, which aimed at minimizing the exposure stigma entailed by potential claimants.

References

Aronsson, T. and Johansson-Stenman, O. (2018). Paternalism against veblen: Optimal taxation and non-respected preferences for social comparisons. *American Economic Journal: Economic Policy*, 10(1):39–76.

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.

Bastani, S., Blumkin, T., and Micheletto, L. (2024). Optimal redistribution and education signaling. *Working Paper, Department of Economics, Uppsala University.*

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

Blau, J. (1992). *The Visible Poor: Homelessness in the United States*. Oxford University Press.

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.

Glazer, A. and Konrad, K. A. (1996). A signaling explanation for charity. *The American Economic Review*, 86(4):1019–1028.

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.

Ramsey, F. P. (1927). A contribution to the theory of taxation. *The Economic Journal*, 37(145):47–61.

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.

Stiglitz, J. E. (2018). Pareto efficient taxation and expenditures: Pre- and re-distribution. *Journal of Public Economics*, 162:101–119. In Honor of Sir Tony Atkinson (1944-2017).

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 Online Appendix: Proofs and Derivations

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 (7). 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 $\alpha_2(t)$.

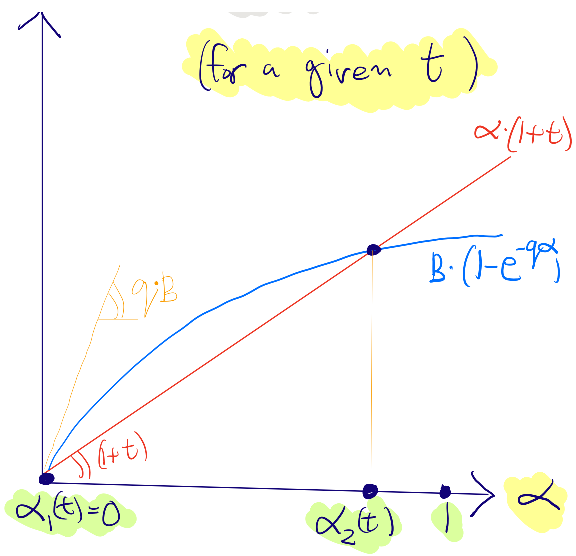

Figure 4: Illustration of the IC constraint (6) .

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:

- $H(t)=0$ 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.

We will now consider two separate cases. Assume first that $H(0)>0$. 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.

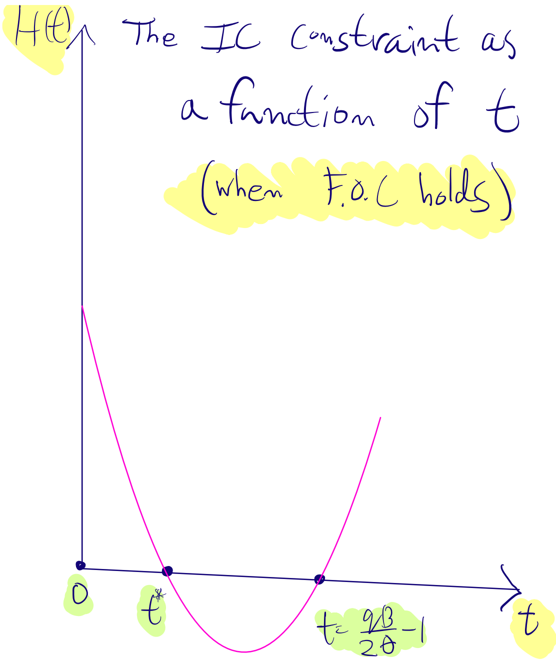

Figure 5: An illustration of $H(t)$ when $\theta < 1/2$.

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:

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

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 $0 \leq t < t^*$.

Consider next the case where $H(0) \leq 0$. It follows, by virtue of the properties of the function $H(t)$ (see figure 5), that the IC constraint is necessarily slack over the entire range, hence for all $t \in (0, \frac{qB}{2\theta} - 1)$, the optimal solution is given by condition (A4).

A.3 Proof of Proposition 2

By virtue of Proposition 1, there exists a threshold, $t^* \geq 0$, such that for $t \in [t^*, \frac{qB}{2\theta} - 1]$, the IC constraint is slack, and $\alpha(t)$ is given by the unconstrained demand function $\alpha(t) = \frac{1}{q} \ln \frac{qB}{2\theta(1+t)}$. Hence, we have that $\alpha'(t) = -[q(1+t)]^{-1}$ and:

$$\frac{1}{|\eta|} = \ln \frac{qB}{2\theta(1+t)}. \tag{A17}$$

Furthermore, when $t^* > 0$, the IC constraint is binding for $t \in [0, t^*)$, and $\alpha(t)$ is dictated by the binding IC constraint, implying that $(1+t)\alpha = (1 - e^{-q\alpha}) B$ and $\alpha'(t) = -\frac{\alpha}{1+t-qBe^{-q\alpha}}$. It hence follows that

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

or equivalently, exploiting (15),

$$\frac{1}{|\eta|} = 1 - 2\frac{\theta - \lambda}{1 - 2\lambda} = \frac{1 - 2\theta}{1 - 2\lambda}. \tag{A18}$$

Recall that t^*, when bounded away from zero, is the threshold value for t that separates the region where the IC constraint is binding ($t < t^*$) from the region where it is slack ($t > t^*$). It hence 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:

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

As we have shown in the Proof of Proposition 1, when t^* is strictly positive, 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 (A18)) $\frac{1}{|\eta|} = 1 - 2\theta$ (i.e., $\lim_{t \to t^{*-}} \frac{1}{|\eta|} = \lim_{\lambda \to 0} \frac{1-2\theta}{1-2\lambda} = 1 - 2\theta$). Now consider (A17) and evaluate at which level for t we get that $\frac{1}{|\eta|} = 1 - 2\theta$. Solving the equation

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

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 $\lim_{\theta \to 1/2} \frac{qB}{2\theta}e^{2\theta-1} = qB$, 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 (A17) is a decreasing function of t, it also follows that $\lim_{t \to t^{*+}} \frac{1}{|\eta|} > 1 - 2\theta$. 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 (14) 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

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

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

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

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 (A18)) that $|\eta|^{-1}$ is a monotonically decreasing function for $t \in (0, t^*)$.

A.4 Proof of Proposition 3

Proof: Throughout the poof we will rely heavily on the first-order condition for the government optimization program. By virtue of Proposition 2, when the IC constraint is slack, , the first order condition (16) can be rewritten as:

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

In contrast, when the IC constraint is binding, by virtue of Proposition 2, the first order condition (16) can be rewritten as:

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

.

We will separate between two cases in the proof depending on whether the IC constraint may bind for some values of t.

Case I: $t^* > 0$

Consider first the case where $t^* > 0$,hence the IC constraint binds for $0 < t < t^*$.

A preliminary result which would prove useful in the characterization is stated as the following lemma:

Lemma 1: The RHS of (A25) is either monotonically decreasing or monotonically increasing in t when, respectively , $\delta > (1+\theta)^{-1}$or $\delta < (1+\theta)^{-1}$.

Moreover, in case $\delta \leq (1+\theta)^{-1}$ the optimal value of t (weakly) exceeds t^*.

Proof: 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 (12) 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 $\alpha\,(1+t) = (1-e^{-q\alpha})\,B$. Notice that a necessary condition for α, as defined by (A29), to be positive is that $1-\delta-\delta\theta < 0$, i.e. $\delta > (1+\theta)^{-1}$. 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 $1-\delta-\delta\theta < 0$. QED

We next turn to establish the solution for the government optimization program.

Lemma 2: The government optimization program has a well defined solution for all δ.

Proof: The solution for the government optimization program is given either by an interior solution, $0 < t < qB/2\theta - 1$,or by a corner solution (which suppresses signaling altogether , i.e. , $\alpha(t) = 0$), $t = qB/2\theta - 1$.

Notice that conditions (A24) and (A25) represent the candidate interior solutions (given by the first order conditions of the government optimization program) associated with the scenarios where the IC constraint is slack or binding ,respectively.

When the (LHS) expressions of (A24) and (A25) are smaller than the corresponding RHS expressions, an increase in t induces an increase in social welfare and vice versa.

It is straightforward to verify that The RHS of (A24) is decreasing in t. By virtue of lemma 1 , the RHS of (A25) , provided that (A25) has a feasible solution ($t < t^*$), is also decreasing in t. As the LHS of (A24) and (A25) are both increasing in t , the interior solutions for (A24) and (A25) , if they exist , define local optima (which may constitute the social optimum).

Notice that there is a also a possibility for a corner solution in which for all $t < t^*$,the RHS of (A25) exceeds its LHS and , for all $t^* \leq t < qB/2\theta - 1$,the RHS of (A24) exceeds its LHS , in which case social welfare is increasing in t over the entire range and the social optimum is given by $t = qB/2\theta - 1$.

Due to the kinked demand for status goods (as the IC constraint may be either binding or slack for different tax rates) , there is a possibility for multiple local optima associated with the range in which the IC constraint binds ($t < t^*$) which is necessarily an interior solution and the range in which it is slack ($t \geq t^*$) which is either an interior or a corner solution.

When multiple local optima emerge (there could be at most two such optima as the RHS of (A24) and (A25) decrease in t whereas their respective LHS increase in t) the one which yields a higher level of social welfare constitutes the social optimum.

The existence of a solution for all δ can be proved by negation.

To see this assume that for some δ no solution exists for the government maximization program.

Then, for this value of δ, (A25) has no feasible solution and hence , as the LHS of (A25) is smaller than its RHS at $t = 0$, it necessarily follows that for $t \to t^{*-}$, the LHS of (A25) is smaller than its RHS.

Due to the discontinuity at $t = t^*$, the RHS of (A24) at $t = t^*$ is larger than the RHS of (A25) at $t \to t^{*-}$ and is hence larger than the LHS of (A24) at $t = t^*$.

As the RHS of (A24) is decreasing in t, whereas its LHS is increasing in t, there are two possibilities to consider: either the LHS of (A24) is larger than its RHS at $t = qB/2\theta - 1$, in which case by continuity a (unique) solution for (A24) exists; or , the LHS of (A24) is (weakly) smaller than its RHS at $t = qB/2\theta - 1$, in which case a corner solution is obtained where $t = qB/2\theta - 1$. We therefore obtain the desired contradiction. QED

We turn next to show that for values of δ sufficiently small the optimal solution is given by a corner solution.

Lemma 3: For $\delta \leq min\left(\frac{2qB}{3qB-2\theta}, (1+\theta)^{-1}\right)$, $t(\delta) = qB/2\theta - 1$ which implies that $\alpha(t) = 0$.

Proof: Suppose that $\delta \leq min\left(\frac{2qB}{3qB-2\theta}, (1+\theta)^{-1}\right)$. By virtue of lemma 1 the optimal solution satisfies $t(\delta) \geq t^*$. We hence focus on the first order condition associated with the case where the IC constraint is slack given by (A24).

Substituting $\delta = \frac{2qB}{3qB-2\theta}$ into the RHS of (A24) yields that $\bar{t}(\delta) = \frac{qB}{2\theta} - 1$, the tax rate which suppresses signaling altogether ($\alpha(t) = 0$). As the RHS of (A24) is decreasing in t, whereas, its LHS is increasing it t, it follows that over the range $t^* \leq t < qB/2\theta - 1$, the RHS of (A24) is larger than its LHS hence social welfare increases in t.

Further notice that the RHS of (A24) is decreasing in δ. It follows that the RHS of (A24) is larger than its LHS for all $\delta < \frac{2qB}{3qB-2\theta}$ and $t^* \leq t < qB/2\theta - 1$. The optimum is given by a corner solution, hence, for all $\delta \leq min\left(\frac{2qB}{3qB-2\theta}, (1+\theta)^{-1}\right)$ as needed. QED

We next show that the set of values of δ for which the social optimum is given by an interior solution in the range $0 \leq t < t^*$, if the set in non-empty, is connected.

Lemma 4: If for some δ' the social optimum is given by the solution for (A25) , then for all $\delta > \delta'$ the social optimum will be given by the solution for (A25) .

Proof: We need to separate between two different cases.

One possibility is that a local optimum associated with δ' in the range $t^* \leq t \leq qB/2\theta - 1$ exists.

This local optimum may be either given by an interior or a corner solution.

Another possibility is that no local optimum in the range exists in which case the LHS of (A24) is larger than its RHS for all $t^* \leq t < qB/2\theta - 1$.

The interior solution for (A25) associated with δ' (which constitutes by presumption the social optimum) is given by $\underline{t}(\delta')$.

Suppose first that a local optimum in the range $t^* \leq t \leq qB/2\theta - 1$, associated with δ', exists.

Denote this local optimum by $\bar{t}(\delta')$,which is given either by solution to (A24) or by $t = qB/2\theta - 1$, depending on whether there is an interior or a corner solution.

By definition of the two local optima it follows that $\bar{t}(\delta') > \underline{t}(\delta') > 0$. As $\alpha'(t) < 0$ for all $0 \leq t < qB/2\theta - 1$, it further follows that $0 \leq \alpha[\bar{t}(\delta')] < \alpha[\underline{t}(\delta']$, where $\alpha[\bar{t}(\delta')] = 0$ in the case of a corner solution and is strictly positive otherwise.

The social welfare is given by:

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

By our presumption $W[\delta', \underline{t}(\delta')] > W[\delta', \bar{t}(\delta')]$.

It hence follows , as $\alpha[\bar{t}(\delta')] < \alpha[\underline{t}(\delta']$, that $\alpha[\underline{t}(\delta')] \cdot \underline{t}(\delta') > \alpha[\bar{t}(\delta')] \cdot \bar{t}(\delta')$.

Differentiation with respect to δ' , employing the envelope theorem (which trivially holds under a corner solution), yields:

$\frac{\partial\left\{W[\delta', \underline{t}(\delta')] - W[\delta', \bar{t}(\delta')]\right\}}{\partial \delta'} = \frac{1}{2} \cdot \theta \cdot \left[\alpha[\underline{t}(\delta')] \cdot \underline{t}(\delta') - \alpha[\bar{t}(\delta')] \cdot \bar{t}(\delta')\right] - \frac{1}{2} \cdot B \cdot \left[e^{-q\alpha[\underline{t}(\delta')]} - e^{-q\alpha[\bar{t}(\delta')]}\right] > 0.$

Thus , for larger values of δ,the local optimum in the range $0 \leq t < t^*$ continues to form the social optimum.

Supoose next that a local optimum in the range $t^* \leq t \leq qB/2\theta - 1$, associated with δ', does not exist.

Thus , the LHS of (A24) is larger than its RHS for all $t^* < t \leq qB/2\theta - 1$. As the RHS of (A24) is strictly decreasing in δ, it follows that for any $\delta > \delta'$,the LHS of (A24) would be larger than its RHS for all $t^* < t \leq qB/2\theta - 1$. The only local optimum is in the range $0 \leq t < t^*$, hence it constitutes the social optimum.QED

The next lemma proves that for values of δ sufficiently large the social optimum is given by an interior solution.

Lemma 5: For $\delta > \frac{2qB}{3qB - 2\theta}$, $0 < t(\delta) < qB/2\theta - 1$.

Proof: As $\delta > \frac{2qB}{3qB - 2\theta}$, the RHS of (A24) is smaller than its LHS at $t = qB/2\theta - 1$. A corner solution is suboptimal hence as one can reduce t slighly below $t = qB/2\theta - 1$ and increase social welfare.

By virtue of lemma 2 the government optimization program has a well-defined solution which is hence necessarily an interior solution. QED

By virtue of lemmas 1-5 one can establish the following lemma:

Lemma 6: If $t^* > 0$, there exist threshold levels, $0.5 < \hat{\delta} < 1$ and $\hat{\delta} \leq \tilde{\delta} \leq 1$, such that the optimal solution for the government problem is given by: (i) a corner solution if $0 \leq \delta \leq \hat{\delta}$; (ii) an interior solution given by (A24) if $\hat{\delta} < \delta \leq \tilde{\delta}$; (iii) an interior solution given by (A25) if

$\tilde{\delta} < \delta \leq 1$.

Proof: Suppose first that for some δ' the social optimum is given by the solution for (A25) then there is a connected set $[\underline{\delta}, 1]$, where $\underline{\delta} > min\left(\frac{2qB}{3qB-2\theta}, (1+\theta)^{-1}\right)$, such that the social optimum is given by a solution to (A25) if-and-only-if $\delta \in [\underline{\delta}, 1]$. If $\underline{\delta} \leq \frac{2qB}{3qB-2\theta}$, then for all $\delta < \underline{\delta}$ the social optimum is given by a corner solution.

To see this notice that for $\delta < \underline{\delta} \leq \frac{2qB}{3qB-2\theta}$, the RHS of (A24) is larger than its LHS at t=qB/2θ-1. As the RHS of (A24) is decreasing it t, whereas the LHS of (A24) is increasing in t for all $t^* \leq t < qB/2\theta - 1$ there is clearly no interior solution in the range $t^* \leq t < qB/2\theta - 1$. As by construction, for $\delta < \underline{\delta}$ there exists no interior solution in the range $0 \leq t < t^*$, it follows that the social optimum for all $\delta < \underline{\delta}$, which exists by lemma 2, is necessarily a corner solution.

If instead $\underline{\delta} > \frac{2qB}{3qB-2\theta}$, then within the range $(\frac{2qB}{3qB-2\theta}, \underline{\delta})$ the social optimum (which cannot be a solution for (A25) by construction, or a corner solution by virtue of lemma 5) is given by a solution for (A24).

In the range $\delta \leq \frac{2qB}{3qB-2\theta}$, the social optimum is given by a corner solution (replicating the arguments used above).

Suppose alternatively that there exists no δ for which the social optimum is given by a solution for (A25) , hence the only possibilities are a solution for (A24) or a corner solution.

This scenario is equivalent to setting $\underline{\delta} = 1$, hence one can replicate the same arguments and establish that the social optimum is given by the solution for (A24) for $\delta > \frac{2qB}{3qB-2\theta}$ and a corner solution otherwise.

This completes the proof. QED

Our next lemma shows that in the case where $t^* > 0$, the set of values of δ for which the social optimum is given by a solution for (A24) is non-empty.

Lemma 7: If $t^* > 0$, then $\tilde{\delta} > \hat{\delta}$.

Proof: To prove the lemma it suffices to show that a solution for (A25) and a corner solution cannot co-exist. Hence by continuity the social optimum will shift from a corner solution to an interior solution given by (A24) in response to a slight increase of δ from $\hat{\delta}$.

Assume by negation that two such local optima co-exist: a corner solution and an interior solution givne by (A25).

By virtue of lemma 1 the existence of a solution for (A25) implies that for $t \rightarrow t^{*-}$,the LHS of (A25) is larger than its RHS.

Formally, noting that for $t \rightarrow t^{*-}$, $\lambda(t) \rightarrow 0$, taking the limit of both sides of (A25) yields the following expression:

(A31) $\frac{t^*}{1+t^*} > \frac{2(1-\delta)}{\delta} + (1-2\theta)$

By virtue of lemma 5 the existence of a corner solution implies that the following condition holds:

(A32) $\delta \leq \frac{2qB}{3qB-2\theta}$

As $0 < t^* < qB/2\theta - 1$ and $t/(1+t)$ is increasing in t, it follows that:

$(A33)\frac{t^*}{1+t^*} < \frac{qB/2\theta-1}{qB/2\theta} = 1 - \frac{2\theta}{qB}$

Employing (A31) hence implies that:
(A34) $1 - \frac{2\theta}{qB} > \frac{2(1-\delta)}{\delta} + (1 - 2\theta)$,
which upon re-arrangement yields:
(A35) $\delta > \frac{qB}{qB\cdot(1+\theta)-\theta}$
Combining (A32) and (A35) then yields:
(A36) $\frac{2qB}{3qB-2\theta} > \frac{qB}{qB\cdot(1+\theta)-\theta}$
which upon re-arrangement yields:
(A37) $\theta > 1/2$
We have therefore obtained the desired contradiction as $\theta < 1/2$, by presumption. QED

This concludes the proof for the case where $t^* > 0$. We turn next to the case where the IC constraint is slack throughout the entire range.

Case II: $t^* = 0$

Consider the case where $t^* = 0$, in which the only possible social optima are given by a solution for (A24) or a corner solution.

The following lemma concludes our proof.

Lemma 8: If $t^* = 0$, the social optimum is given by a corner solution for $\delta \leq \hat{\delta} \equiv \frac{2qB}{3qB-2\theta}$ and by an interior solution given by (A24) otherwise.

Proof: Invoking the same arguments as those used in the proof of lemma 6 in the scenrio where a solution for (A25) is infeasible, the social optimum is given by a corner solution if $\delta \leq \frac{2qB}{3qB-2\theta}$ and by a solution for (A24) otherwise. QED

Finally, we turn to show that the optimal (interior) solution is strictly decreasing in δ.
Lemme 9: $t(\delta)$ is strictly decreasing with respect to δ for all $\delta > \hat{\delta}$.

Proof: When $\delta > \hat{\delta}$, an interior solution exists, which is given either by (A24) or (A25). The LHS of (A24) and the LHS of (A25) are both increasing in t, whereas, the RHS of (A24), and, by virtue of lemma 1, the RHS of (A25), are both decreasing in t. The lemma then follows by noting that both the RHS of (A24) and that of (A25) are decreasing in δ. QED

This completes the proof .QED

B Online Appendix: The status production function and optimal tax differentiation

In our present analysis, for the sake of clarity and simplicity, we have made the assumption that all signaling goods exhibit symmetry in terms of their acquisition costs, visibility, and benefits derived from status. However, it is possible to extend the model and consider certain asymmetries. One approach to achieve this while maintaining tractability is to introduce a status-production technology that demonstrates perfect substitutability.

Consider a scenario where we have two distinct categories of signaling goods denoted as x_{ki}, where i ranges from 1 to n_k and k takes values of 1 or 2. Each category corresponds to a particular group of agents: for instance, category $k = 1$ may pertain to colleagues in a professional setting, while category $k = 2$ could be aimed at friends or relatives. The visibility parameters for each category are given by q_k/n_k, and the unit costs are denoted as θ_k/n_k. It is worth noting that these parameters are specific to their respective categories. Within this framework, it is also possible for the benefits derived from signaling to differ between the two groups, potentially due to inherent variations or disparities in group sizes. Let us denote the benefit associated with group k as B_k.

Under the assumption of separable technology, the status obtained by type-2 agents can be expressed as:

$$Status^2(\alpha_1, \alpha_2) = \sum_{k=1}^{2} \left[1 - \frac{1}{2} \cdot e^{-\alpha_k q_k}\right] \cdot B_k, \tag{B1}$$

In this equation, α_k represents the proportion of type-2 agents' wealth allocated to the purchase of signaling goods in category k. The visibility parameter for each category is denoted by q_k, and B_k represents the corresponding benefit associated with signaling to the specific group.

On the other hand, the status derived by type-1 agents can be described as:

$$Status^1(\alpha_1, \alpha_2) = \sum_{k=1}^{2} \left[\frac{1}{2} \cdot e^{-\alpha_k q_k}\right] \cdot B_k, \tag{B2}$$

In this equation, the term $\frac{1}{2} \cdot e^{-\alpha_k q_k}$ represents the contribution of type-1 agents' signaling goods in category k to their status.

By assuming separability, we imply that there are zero cross-tax elasticities between the two categories. This allows us to extend our analysis by introducing differentiated commodity tax rates, denoted as t_k, for the two categories of consumption goods ($k = 1, 2$).

In extending our analysis, we generalize formula (A24), which characterizes the optimal tax formula in the baseline case with a single category of consumption goods when the inequality constraint (IC) is slack. Assuming that the (IC) constraint is slack, as demonstrated earlier for an intermediate range of δ values, we have the following expression:

$$\frac{t_k}{1 + t_k} + \ln(1 + t_k) = 2\frac{1 - \delta}{\delta} + \ln \frac{q_k B_k}{2\theta_k}. \tag{B3}$$

In this equation, t_k represents the tax rate applied to category k of consumption goods. The left-hand side of (B3) increases with t_k. This implies that a higher tax rate is levied on goods that yield a greater return on signaling.

In the case of separable technology, tax differentiation relies on the variation in the returns on status signaling across different categories of consumption goods. By observing the increasing relationship on the left-hand side of (B3) with respect to t_k, we can conclude that a higher tax rate is imposed on goods exhibiting a higher return on signaling.

It is worth noting that an alternative assumption of full complementarity between the categories of consumption goods could be considered. Under such an assumption, all agents would observe both categories of consumption and form beliefs about the agent's type. The utility of status (B) remains the same as in the single-category case. The full benefit is obtained if at least one signal from each category is observed, otherwise the surplus is divided equally between the two types. This alternative technology would result in 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 $Status^1(\alpha_1, \alpha_2) = [\frac{1}{2} \cdot e^{-\sum_{k=1}^{2} \alpha_k q_k}] \cdot B$. In the case of full complementarity, the tax formula would take into account cross-tax elasticities.