Optimal Redistribution via Income Taxation and Market Design

Paweł Doligalski ⓡ Piotr Dworczak
Mohammad Akbarpour ⓡ Scott Duke Kominers*

December 1, 2025

Abstract

Policymakers often distort goods markets to effect redistribution—for example, via price controls, differential taxation, or in-kind transfers. We investigate the optimality of such policies alongside the (optimally-designed) income tax. In our framework, agents differ in both their ability to generate income and their consumption preferences, and a planner maximizes a social welfare function subject to incentive and resource constraints. We uncover a generalization of the Atkinson-Stiglitz theorem by showing that goods markets should be undistorted if the heterogeneous consumption tastes (i) do not affect the marginal utility of disposable income, (ii) do not enter into the social welfare weights and (iii) are statistically independent of ability. We also show, however, that market interventions play a role in the optimal resolution of the equity-efficiency trade-off if any of the three assumptions is relaxed. In a special case of our model with linear utilities, binary ability, and continuous willingness to pay for a single good, we characterize the globally optimal mechanism and show that it may feature means-tested consumption subsidies, in-kind transfers, and differential commodity taxation.

*The orders of the first and second as well as third and fourth co-authors' names are randomized, as indicated by the ⓡ symbol. Dworczak and Doligalski acknowledge the support received under the ERC Starting grant IMD-101040122. Kominers acknowledges the support of the Ng Fund and the Mathematics in Economics Research Fund of the Harvard Center of Mathematical Sciences and Applications, the Washington Center for Equitable Growth, and the Digital Data Design (D^3) Institute at Harvard. Kominers is a Research Partner at a16z crypto (for general a16z disclosures, see https://www.a16z.com/disclosures/). Notwithstanding, the ideas and opinions expressed herein are those of the authors, rather than of the EU, the ERC, or a16z and its affiliates. We would like to thank Lassi Ahlvik, Andrea Attar, Pierre Boyer, Gabriel Carroll, Jonathan Heathcote, Tai-Wei Hu, Mingshi Kang, Abdoulaye Ndiaye, Alexandru Nichifor, Karl Schulz, Joanna Tyrowicz, Nicolas Werquin, Hakki Yazici, and several seminar audiences for helpful feedback. *Doligalski:* University of Bristol. *Dworczak:* Northwestern University; and Group for Research in Applied Economics. *Akbarpour:* Graduate School of Business, Stanford University. *Kominers:* Entrepreneurial Management Unit, Harvard Business School; Department of Economics, Harvard University; and a16z crypto.

1 Introduction

Policymakers often distort market allocations as a way of addressing inequality. In developed countries, for example, it is common for local housing authorities to provide affordable housing, impose rent-control policies, or subsidize construction of rental housing for low-income tenants. Food assistance is also prevalent—for instance, in the form of food stamps in the United States. And recently, many European countries imposed caps on electricity prices to shield vulnerable households from sharp increases in energy bills. In developing countries, likewise, in-kind provision of food items and subsidized energy consumption have long been an important part of the safety net.

Market-level redistributive policies defy conventional economic wisdom—rooted in the welfare theorems—that market interventions compromise efficiency and thus should be avoided (absent market failures). A recent literature on *inequality-aware market design* has pointed out, however, that policies such as price controls and rationing may be justified on welfare grounds if policymakers lack the information or instruments needed to effect redistribution through targeted lump-sum transfers.[1] The key intuition is that agents' behavior in the market may reveal information about their welfare weights: in settings where agents' redistribution-worthiness is not directly observed, willingness to pay for a good may be correlated with welfare-relevant characteristics (such as income level or wealth). By modifying market-clearing rules to induce appropriate self-selection, the planner can trade off efficiency in the market with equity—distorting the allocation in order to effect redistribution to agents with higher levels of need.

However, arguments in favor of market interventions remain incomplete without considering the role of broader policy instruments that address inequality. In particular, income taxation is often thought of as the primary—and ideal—tool for effecting redistribution in the presence of incentive constraints (see, e.g., Kaplow (2011) and the references therein). Thus, we ask: Can redistribution through markets be justified if the policymaker also controls income taxation? And if yes, *how do market interventions interact with the income tax to strike the balance between equity and efficiency*?

The public finance literature provided the answer in a core benchmark case: By the Atkinson-Stiglitz theorem, if agents only differ in their ability to generate income (and preferences satisfy a weak-separability assumption), then income taxes alone are sufficient to maximize social welfare for any set of welfare weights (Atkinson and Stiglitz (1976)); i.e., goods

[1] See, for example, the work of Condorelli (2013), Dworczak ⓡ al. (2021), and Akbarpour ⓡ al. (2024)—building on classical insights such as those of Weitzman (1977), Spence (1977), Nichols and Zeckhauser (1982), Guesnerie and Roberts (1984) and Blackorby and Donaldson (1988).

market interventions are redundant (and potentially harmful). Under this perspective, redistribution through markets can be defended only if the policymaker lacks the ability to adjust income taxes.

However, in this paper, we show that market interventions can be a valuable redistributive tool more broadly. Specifically, we examine the desirability of interventions in goods markets in the presence of *multidimensional* heterogeneity, where individuals differ both in their productive ability and in their consumption tastes. In line with the public finance literature studying the robustness of the Atkinson-Stiglitz theorem, we find that once heterogeneity in tastes is introduced, it may be beneficial to supplement the income tax with market interventions. In contrast to most of that literature, by employing a mechanism design framework under a few simplifying assumptions, we are able to characterize the optimal *combination* of income taxation and redistributive market design.

Our analysis has two parts. First, as a baseline, we prove an extension of the Atkinson-Stiglitz theorem to our setting with multidimensional heterogeneity under three strong assumptions (on top of the separability assumption on preferences); specifically, we show that goods market interventions are redundant when:

A1. Consumption tastes do not affect the marginal utility of disposable income (when goods are priced at marginal costs);

A2. Welfare weights depend only on agents' ability levels, and not on their consumption tastes;

A3. Consumption tastes and ability are statistically independent.

Assumption A1 means that the utility an agent obtains from a given disposable income does not vary with their preference type, assuming goods are priced at marginal costs. Therefore—conditional on the same ability—two agents with different consumption tastes must value an additional dollar of income in the same way, and in particular make the same labor supply decisions. Together with Assumption A2, this implies that the planner has no desire to redistribute between agents with different tastes; the planner is concerned only with redistribution across ability levels. By Assumption A3, meanwhile, willingness to pay for goods is uninformative about ability. Thus, under Assumptions A1–A3, distorting choices in the goods market neither serves a valuable redistributive role on its own, nor screens agents' ability levels—and hence it should be avoided.

In the second part of our analysis, we show that the assumptions used to derive the Atkinson-Stiglitz result in our setting are "tight" in the sense that if we relax any one of them, the conclusion of the theorem fails. We characterize the optimal mix of income tax and market intervention in a simplified specification that permits a tractable analysis of globally optimal mechanisms in the presence of multidimensional heterogeneity (explicitly allowing for the possibility of "ironing," or "bunching"). The simplified framework features an intensive-margin choice of two commodities: a good x, which is singled out for potential redistributive market intervention, and a numeraire c, which can be thought of as an aggregate consumption good. There is a continuum of taste types, but ability is a binary characteristic, with low-ability agents unable to generate any income, and the planner attaching a weakly higher average welfare weight to low-ability agents. The simplicity in modeling the labor market is what allows us to consider rich heterogeneity in consumption tastes and develop detailed market-design implications, which have received far less attention in the literature than the optimal income tax.

The simplified specification makes income taxation stylized since only two income levels (corresponding to two ability levels) are generated in the second-best solution. While limiting from a practical perspective, this feature of our model makes the role of goods market interventions more transparent. In particular, the optimal income tax leads to efficient labor market outcomes even when it is strongly redistributive. Consequently, interventions in goods markets are not aimed at reducing distortions in the labor market; rather, they expand the set of feasible redistributive allocations.

We first relax Assumption A1, according to which tastes do not affect the marginal utility of disposable income. When agents choose between good x and the numeraire, Assumption A1 holds provided that utilities are *linear* in the numeraire, allowing agents to freely adjust consumption until the marginal utility of good x is equated with the constant marginal utility of the numeraire. We violate Assumption A1 by supposing that consumption of the numeraire cannot fall below some subsistence level $\underline{c}$. Absent any market intervention, agents with low ability (and, thus, low income) and high taste for good x must limit their consumption of x to maintain subsistence. As a result, their marginal utility of income is higher than that of other agents.

The optimal mechanism in this setup looks as follows: First, the income tax redistributes from rich (high-ability) to poor (low-ability) as much as possible subject to maintaining the high-ability agents' incentive to work. Second, good x is subsidized at low consumption levels and taxed at high consumption levels. The subsidy at low consumption levels creates a (first-order) welfare gain by relaxing low-ability, high-taste agents' subsistence constraint—allowing them to consume more of the good. This positive effect dominates the (second-order) negative distortion of making some agents over-consume good x. Meanwhile, because high levels of consumption of good x can only be reached by agents with labor income, the planner can tax purchases of x on the margin to raise more revenue from high-taste, high-ability agents. That additional revenue is then used to partially finance the subsidy for low consumption of good x. In particular, if the planner cares only about low-ability agents, the marginal after-tax price for high-enough quantities should be set to the price that would be chosen by a revenue-maximizing monopolist. The resulting nonlinear pricing scheme can be implemented via a combination of an in-kind transfer and a commodity tax on transactions in the private market.

The structure of the optimal policy—in particular, the optimality of nonlinear pricing in the goods market—is qualitatively shared by any specification of agents' utilities that violates Assumption A1. Imposing a subsistence constraint allows us to obtain a closed-form solution in the baseline framework. In an extension, we assume instead that agents' utility from consuming the numeraire is a strictly concave function (which also violates Assumption A1). Then, agents with higher taste for good x will consume less numeraire and thus have higher marginal utility of income. By subsidizing low levels of consumption of good x, the planner redistributes the numeraire towards high-taste, low-ability agents who have high marginal utility of income. The planner may also tax high levels of consumption of good x, which allows her to further redistribute from high-ability to low-ability agents. Thus, qualitative features and intuitions behind nonlinear pricing in the goods markets are preserved in the more realistic framework with curvature in the utility function.

Summarizing—barring the restrictive linear case in which Assumption A1 holds—taste heterogeneity generates differences in marginal utility between agents with the same ability level. These differences create an endogenous motive to redistribute across agents depending on their tastes. In such cases, income taxes alone are never sufficient to achieve optimality because distinguishing between agents with different tastes requires conditioning on agents' consumption choices—creating a scope for market interventions.

Meanwhile, relaxing either Assumption A2 or Assumption A3 creates a correlation between welfare weights and tastes and, thus, leads to a direct motive to redistribute across taste types. These two cases are conceptually similar and we discuss them together. As a general principle, income redistribution is complemented by a simple market intervention: Purchases of the good are *taxed* when agents consuming it have lower welfare weights on average (e.g., when there is positive correlation between taste and ability) and *subsidized* otherwise (e.g., when there is negative correlation between taste and ability). However, there is also a sense in which income taxation and market design become substitutes. Since income reveals information about ability, it serves as a useful instrument for (third-degree) price discrimination in the goods market—precisely when either Assumption A2 or A3 fails. Specifically, under our assumptions, the planner would like to offer a lower price to low-ability agents. However, the degree to which this is feasible is limited by incentive constraints. When the income tax is maximally redistributive subject to respecting the incentive constraints of high-ability agents, prices in the market cannot depend on income. This outcome is optimal when the planner has strong vertical redistributive preferences, measured by the gap between the average welfare weights of low- and high-ability agents. However, when the planner's vertical redistributive preferences are weak, it becomes optimal to reduce the income tax paid by high-ability agents, and use the slack in the incentive constraints to offer income-dependent prices in the goods market. Thus, in this case, the role of the income tax is limited to allow for richer instruments in the design of the market—namely, a means-tested subsidy for the good.

Our analysis suggests that there may be a number of circumstances under which it is optimal to combine income tax policies with redistributive market design. The key intuition for why this happens aligns closely with the motivation for redistributive market design described above: *in many goods markets, it is possible to use agents' purchasing behavior to infer welfare-relevant information that an income tax policy has no way of conditioning on directly.* Thus, distortions can be justified in markets where consumption choices are informative for redistribution, e.g., because they identify agents with high marginal utility of income (related to A1), directly correlate with welfare weights (related to A2), or induce correlation with welfare weights through a statistical link to ability (related to A3). This lends qualified theoretical support for redistributive market interventions, including some policies used in practice, while at the same time informing the optimal interaction of such interventions with the income tax.

The classic Atkinson-Stiglitz framework masks this potential role for redistributive market design by implicitly assuming that all of the welfare-relevant information revealed through market behavior is redundant (because it is already revealed by agents' labor supply choices). To illustrate the difference, it is instructive to contrast the reasoning just presented with the logic of the original Atkinson-Stiglitz theorem in the context of a specific example, for instance healthcare. In both frameworks, richer agents tend to consume more healthcare. However, because agents in the model of Atkinson and Stiglitz (1976) only differ in their earnings ability, high consumption of healthcare is not a signal of need but merely a manifestation of high income. It is thus most efficient to redistribute income directly by taxing it—any intervention in the healthcare market is an imperfect substitute for the optimal income tax policy. Yet in practice, demand for healthcare stems from a combination of ability to pay (i.e., income) and need (i.e., taste). And indeed, in our framework, heterogeneity in preferences allows for need to shape demand for healthcare. Then, subsidies for low consumption (or low quality) of healthcare services can improve redistribution relative to income tax alone because they endogenously target the low-ability agents who have a particularly high marginal utility of income—i.e., the agents who are consuming low amounts of healthcare relative to need.

The remainder of the paper is organized as follows: The next subsection describes the related literature. We introduce our framework in Section 2. In Section 3, we extend the Atkinson-Stiglitz theorem to a setting with multidimensional heterogeneity under assumptions A1–A3. In Section 4, we relax each of the three assumptions in turn, and characterize the optimal interaction between the income tax and goods market interventions in the specialized framework. We discuss the general structure of the optimal mechanism and our solution method in Section 5. We present brief concluding remarks in Section 6.

1.1 Related literature

Our work bridges a recent literature on inequality-aware market design with the classical public-finance literature on optimal taxation of commodities and income. We discuss these two literatures and our respective contributions next.

Research on inequality-aware market design (e.g., Condorelli (2013), Dworczak ⓡ al. (2021), Kang and Zheng (2022), Akbarpour ⓡ al. (2024), Kang (2023), Kang and Watt (2024b)) focused on the problem of optimal allocation of a single good in the presence of socioeconomic inequality observed in the market. The underlying assumption of that approach has been that the designer does not control other redistributive tools such as income taxation—for example, because the designer is a local authority and taxes are set at the national level, or because political economy frictions block the government from adjusting taxes to the optimal redistributive target. We show that redistribution through markets may be optimal *even when the designer does control income taxation*. At the same time, the interaction between the two tools is non-trivial. For example, the market intervention may take the form of a means-tested subsidy for consumption or an income-independent in-kind transfer combined with a tax on top-up consumption; income taxes may sometimes be lowered in order to incentivize labor provision by high-ability agents when low-ability agents face lower prices in the goods market.

A closely related strand of the literature—including the work of Kang (2024), Pai and Strack (2024), Bierbrauer (2024), and Ahlvik et al. (2024)—studied the problem of optimal regulation of consumption goods generating an externality (e.g., pollution) under taste heterogeneity and redistributive social preferences. Some of these papers also considered the interaction of the regulation with income taxation. Pai and Strack (2024) showed that their main results are unaffected when income taxes are present but are set exogenously. Bierbrauer (2024) derived the welfare impact of small deviations from the uniform tax on externalities in a setting with a nonlinear income tax and general equilibrium effects. Ahlvik et al. (2024) allowed for joint design of an income tax and a consumption tax. These works showed that Pigovian taxation—a classical solution to the externality problem—must be appropriately modified to account for redistributive concerns, and that the income tax alone is not sufficient to address those concerns in the presence of taste heterogeneity.

A rich public finance literature has studied the optimal design of income and consumption taxes under heterogeneity in abilities and preferences. Assuming homogeneous tastes over goods, Atkinson and Stiglitz (1976) proved that consumption taxes are redundant if the government can use a nonlinear income tax and agents' utilities are weakly-separable between labor supply and consumption. Mirrlees (1976, 1986) allowed for differences in tastes by assuming that tastes and abilities are perfectly correlated (and the heterogeneity is, thus, one-dimensional) and described the optimal consumption taxes; he also concluded that the multidimensional case, while realistic, is largely intractable. Several papers utilized the assumption of one-dimensional heterogeneity to investigate policies such as taxes on capital (Golosov et al. (2013); Scheuer and Slemrod (2020); Gerritsen et al. (2025); Schulz (2021); Hellwig and Werquin (2024)) or on housing (Cremer and Gahvari (1998)). Within the work that considers multidimensional settings, Cremer et al. (2001, 2003), Blomquist and Christiansen (2008), Diamond and Spinnewijn (2011) and Gauthier and Henriet (2018) characterized a nonlinear income tax and linear consumption (including capital and inheritance) taxes. Moser and Olea de Souza e Silva (2019) considered agents who are heterogeneous in abilities and in the strength of present-bias, and studied the optimal (paternalistic) savings policies. Finally, Gaubert et al. (2025) analyzed income taxation in a spatial equilibrium model and showed that heterogeneity in locational preferences can lead to optimality of place-based redistribution (akin to differential commodity taxation).[2]

We contribute to this literature by providing precise conditions under which a nonlinear income tax is sufficient (and interventions in the goods markets are redundant) in the multidimensional setting. These conditions are restrictive, suggesting that market interventions are (at least theoretically) useful in resolving the equity-efficiency trade-off under a broad set of circumstances. Furthermore, by leveraging recent developments in mechanism design, we provide a multidimensional framework where the optimal interaction between the income tax and market design is rich but can be characterized in closed form.

There are also related studies that examine desirability of goods market interventions us-

[2]Other related studies include Kaplow (2008) on how preference heterogeneity affects classical results from public finance, Kleven et al. (2009), Golosov and Krasikov (2025), Spiritus et al. (2025) and Bierbrauer et al. (2024) on tax treatment of couples, and Gomes et al. (2017) and Boerma et al. (2022) on sector- or task-specific taxation in a model with multidimensional skills.

ing a sufficient-statistics approach. Saez (2002) considered a setting in which individuals differ in abilities and tastes, and derived conditions under which introducing a small tax on one of the goods cannot improve social welfare if the policymaker can use a nonlinear income tax. Several papers extended this result by deriving the formulas for optimal income and commodity (or capital) taxes and estimating the relevant sufficient statistics from the data (Piketty and Saez (2013); Saez and Stantcheva (2018); Allcott et al. (2019); Ferey et al. (2024)). Our results are consistent with those in this literature but provide a complementary perspective by fleshing out the role of model primitives—rather than endogenous sufficient statistics. Working directly with agents' utility functions and exogenous welfare weights uncovers natural mechanisms through which taste heterogeneity generates differences in marginal utilities (which would be subsumed in marginal social welfare weights in the sufficient statistics approach). These differences create scope for market interventions to improve redistribution. In particular, we show in Section 4.3 that the Atkinson-Stiglitz theorem *always* fails under taste heterogeneity when agents' marginal utilities decline with their level of consumption (even if Assumptions A2 and A3 hold). Moreover, we derive the globally optimal policy with respect to arbitrary mechanisms that could include rationing, quotas or public provision of goods. By contrast, conditions based on sufficient statistics are necessarily local and informative about the effects of small tax reforms only.

Restricting attention to *linear* taxes (on income and goods), Deaton (1979) proved the sufficiency of an income tax when, in addition to the original assumptions of the Atkinson-Stiglitz theorem (i.e., preference homogeneity and weak separability), demand for goods is linear in disposable income. Deaton and Stern (1986) extended this result to a setting with (limited) taste heterogeneity by allowing the linear Engel curves to have taste-dependent intercepts, under the assumption that tastes are uncorrelated with social marginal utility of income. This result on redundancy of *linear* consumption distortions under taste heterogeneity parallels our baseline result on the redundancy of arbitrary *nonlinear* distortions. Our proof strategy, however, is very different, and it allows us to accommodate a richer class of preferences, including those with nonlinear Engel curves.

Finally, our work makes a technical contribution to the multidimensional screening literature; we propose a novel solution technique for a class of models in which one dimension of the type space is continuous and the other binary. Models with a similar structure (although in different economic contexts) have been studied by Fiat et al. (2016) and Li (2021). The key idea in our approach is to represent the incentive compatibility constraint across the binary (ability) type as an outside option constraint, as in the work of Jullien (2000), using the fact that high-ability agents in our framework can mimic low-ability agents but not vice versa. We then adopt a recent procedure developed by Dworczak and Muir (2024) to solve for the optimal allocation rule for high-ability agents, when the low-ability agents' allocation is fixed. Finally, we use the linearity of the problem to argue that the full solution features allocation rules that are step functions with a limited number of steps, explicitly accounting for the possibility of "ironing." We discuss our methodology and the relationship to existing approaches in Section 5 which contains our main technical result.

2 General Framework

Our framework features agents who are heterogeneous in ability and taste and make optimal labor supply and consumption choices, as well as a planner who aims to maximize social welfare subject to incentive-compatibility and resource constraints.

There is a unit mass of agents who differ in both their *consumption taste* $t \in \Theta_t$ and their *earning ability* $a \in \Theta_a$. Types $\theta = (t, a) \in [\Theta_t \times \Theta_a] \equiv \Theta \subseteq \mathbb{R}^2$ are jointly distributed according to $F(\theta)$. Agents have preferences over a vector of *goods* $x \in \mathbb{R}^L_+$, a *numeraire* consumption good $c \in \mathbb{R}$, and *earnings* $z \in \mathbb{R}_+$, as given by the utility function

$$U((c, x, z), (t, a)) = u(c) + v(x, t) - w(z, a), \tag{1}$$

which is upper semi-continuous in (c, x, z) and measurable in (t, a).[3] The separation of consumption into a vector of goods x and a one-dimensional numeraire c is convenient for studying different cases of the model. Conceptually, one can think of x as a set of goods that are singled out for potential intervention, and of c as aggregating the consumption of all remaining goods into a single composite commodity. We assume that $u(c)$ is strictly increasing. While we do not (at this point) assume specific functional forms for the different components of utility (we have not even imposed any single-crossing conditions), we assumed that utility is additively separable between the numeraire c, goods x, and earnings z.

The planner chooses an *allocation rule* $Y = (c, x, z) : \Theta \to [\mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}^L_+]$ to maximize the expected utility of agents weighted with welfare weights $\lambda(\theta) \geq 0$. The average welfare weight is normalized to 1. That is, the social objective is

$$\int \lambda(\theta)\, U(Y(\theta), \theta)\, dF(\theta). \tag{2}$$

[3] As is common in the public finance literature, we model labor choices by letting individuals select the level of earnings z directly, with ability determining the cost of generating it; this is equivalent to a model in which agents choose labor supply and ability determines the earnings generated per unit of labor.

The planner faces the resource constraint

$$\int [z(\theta) - c(\theta) - k \cdot x(\theta)] \, dF(\theta) \geq G, \tag{3}$$

where $k \in \mathbb{R}^L_{++}$ is the marginal cost of producing goods x in terms of the numeraire and G represents the minimum revenue requirement (which could be negative, representing exogenous revenue sources).

The planner does not observe individual agents' types. Thus, the mechanism must satisfy the standard incentive compatibility (IC) constraints

$$U(Y(\theta), \theta) \geq U(Y(\theta'), \theta), \quad \forall \theta, \theta' \in \Theta. \tag{4}$$

The constraints (4) prevent agents from misreporting their taste, ability, or both. A mechanism that satisfies the resource constraint (3) and the IC constraints (4) is called *feasible*.

To define two concepts crucial for the analysis, consider a consumption problem of an agent with taste type t and disposable income m who faces prices equal to marginal costs:

$$V(m, t) := \max_{y \in \mathbb{R}^L_+} \{u(m - k \cdot y) + v(y, t)\}; \tag{5}$$

we assume that this problem has a solution. Consider some allocation rule $Y = (c, x, z)$ and let $m(\theta) := c(\theta) + k \cdot x(\theta)$ be the implied resource cost of consumption of type θ (in terms of numeraire). Consumption choices c and x are efficient if it is not possible to increase utility for a positive measure of agents by adjusting their consumption while keeping the resource cost fixed. Formally, we say that allocation rule Y induces a *(Pareto) efficient allocation of goods* if, for almost all θ, $x(\theta)$ is a solution to problem (5) with disposable income $m = m(\theta)$. An allocation of goods that is not Pareto efficient is called *distorted*.

Finally, we define the *marginal utility of disposable income* as $\partial V(m, t)/\partial m$ (whenever the derivative exists). Note that the marginal utility of disposable income is tied to the curvature of the function u: when $u(c)$ is differentiable, the envelope formula implies

$$\frac{\partial V(m, t)}{\partial m} = u'(m - k \cdot x^\star(m, t)),$$

where $x^\star(m, t)$ is a consumption bundle solving problem (5). In particular, the marginal utility of disposable income is constant in consumption taste t when the utility function u is *linear*.

3 Baseline: When Interventions in Goods Markets are Redundant

In this section, we establish a baseline result that organizes the rest of the analysis. We introduce three assumptions that jointly imply that the planner can achieve an optimal allocation with an income tax alone, without intervening in the goods markets.

Assumption A1. *Consumption tastes do not affect the marginal utility of disposable income:* $u(c) = c$, $\forall c \in \mathbb{R}$.[4]

Assumption A2. *Welfare weights depend only on ability:* $\lambda(t, a) \equiv \bar{\lambda}(a)$.

Assumption A3. *Ability and tastes are statistically independent:* $F(t, a) \equiv F_t(t)F_a(a)$, *where* F_t *and* F_a *denote the marginal distributions.*

For the following result, we assume that an optimal mechanism exists. Moreover, we assume that there exists a Lagrange multiplier $\alpha > 0$ on the resource constraint (3) (interpreted as the marginal value of public funds) such that any optimal mechanism maximizes the Lagrangian (i.e., strong duality holds). These regularity assumptions allow us to provide a proof that highlights the economic intuitions; in Online Appendix A, we prove a version of the result that does not require such assumptions.

Theorem 1. *Suppose Assumptions A1–A3 hold. Then, any optimal mechanism induces an efficient allocation of goods, and can be implemented with a competitive goods market and an income tax.*

3.1 Intuition

Theorem 1 can be understood as a generalization of the Atkinson-Stiglitz result to settings with multidimensional heterogeneity. For intuition, note that distortions in goods market always lead to a deadweight loss, so they can be justified only if the planner can achieve better redistribution. However, as we explain below, Assumptions A1–A3 imply that *(i)* the planner does not want to redistribute across taste types, and *(ii)* distortions in goods markets do not aid redistribution across ability types.

Because the standard Atkinson-Stiglitz setting does not feature preference heterogeneity, it vacuously satisfies property *(i)*. In our setting, Assumptions A1 and A2 are needed to guarantee that the planner does not have a direct motive to redistribute across taste types.

[4]A careful reader will note that linearity of u is sufficient but not necessary to eliminate the taste-driven variation in marginal utility of income. Nevertheless, the quasi-linear case simplifies the exposition. The substantive result relies directly on the marginal utility of income being constant in taste, as we demonstrate in Online Appendix A.

The role of Assumption A1 is to ensure that the marginal utility of disposable income (evaluated at efficient consumption choices) does not depend on the agent's type; taste t can be associated with a high need for a particular good but never a high need for more resources overall (a high need for a particular good can always be accommodated by reducing the consumption of other goods).[5] By contrast, if $u(c)$ were strictly concave, agents with a particularly high taste for certain goods could also have higher marginal utility of income, creating a motive for the planner to redistribute to them. Property *(ii)*, meanwhile, lies at the heart of the Atkinson-Stiglitz logic and is guaranteed by weak separability in the one-dimensional setting. Under preference heterogeneity, an additional assumption is needed because distorting goods choices could potentially reduce labor market distortions or minimize information rents from privately observed ability; Assumption A3 rules out that possibility by ensuring that tastes are uninformative about ability.

While Theorem 1 shows that we can recover the Atkinson-Stiglitz result in our multidimensional setting, the intuition just described already suggests what we prove in the sequel: all three of the assumptions are essentially necessary for the Atkinson-Stiglitz conclusion to hold. We thus interpret Theorem 1 as in effect revealing three channels through which intervention in goods markets may become optimal *even when nonlinear income tax instruments are available*.

3.2 Sketch of the proof of Theorem 1

We prove Theorem 1 in two steps (see Appendix A.1 for details). In the first step, we consider a relaxed problem in which the planner is able to directly observe the taste type t. Then, the problem can be solved for each t separately, ignoring agents' incentives to misreport their tastes. For each t, the relaxed problem becomes a one-dimensional optimal taxation problem in which agents have identical consumption preferences. Thus, the original Atkinson-Stiglitz theorem implies that the allocation of goods should be undistorted and redistribution should be conducted via (potentially nonlinear) income taxes alone.

In the second step, we show that under our three assumptions, the income tax schedule that solves the relaxed problem in fact does not depend on the taste type t, and hence is feasible (and therefore optimal) in the original problem in which the planner does not observe tastes.[6] This step highlights the key role played by Assumptions A1–A3. The social welfare

[5]Interestingly, Eden and Freitas (2024) found a condition closely related to Assumption A1 when studying utilitarian social welfare functions that satisfy "income anonymity"—a property that each person's disposable income matters equally for social welfare.

[6]Our approach is inspired by the proof strategy of Haghpanah and Hartline (2020) and Yang (2022) who—in the context of a revenue-maximizing monopolist—first derive an optimal mechanism assuming that the

function has the same shape for every t: The planner neither conditions welfare weights on t directly (Assumption A2), nor learns about the distribution of ability a by observing t (Assumption A3). Moreover, taste t does not affect the marginal utility of disposable income (Assumption A1). As a result, the income tax that solves the relaxed subproblem does not condition on t even when t is freely observable. Finally, because the allocation of goods is efficient in the relaxed problem, it can be implemented in an incentive-compatible way by pricing goods at marginal cost and letting agents make unrestricted consumption choices, given their disposable incomes implemented by the optimal income tax.

3.3 Discussion

Theorem 1 can be extended along several dimensions (see Online Appendix A for details). First, we show that neither the existence of the numeraire good nor quasi-linear preferences are essential for the result. In line with the intuition discussed in Section 3.1, it suffices to directly assume that the marginal utility of disposable income does not depend on taste type t. Second, the additive separability of agents' preferences over earnings and consumption can be weakened to the standard weak-separability assumption.[7] Third, we can relax the technical assumptions on the existence of the optimal mechanism and on strong duality that we used to prove Theorem 1. Formally, we show that any mechanism that distorts the goods markets can be improved upon by a mechanism that only uses income taxation to redistribute, and that strong duality is not needed if the income tax is allowed to be stochastic; we also provide conditions under which strong duality holds and the improving mechanism is deterministic.

It is instructive to compare Theorem 1 to the canonical multidimensional screening framework used to study, e.g., nonlinear pricing by a multiproduct monopolist. Rochet and Choné (1998) showed that in such problems it is typically optimal to distort *all* dimensions of the allocation by bunching, i.e., assigning identical bundles of goods to different types. Full separation of types is suboptimal due to the tension between the participation constraints and the second-order incentive constraints. In contrast, we found that the allocation of goods x should never be distorted and, thus, bunching of different taste types can be easily ruled out. This stark difference in conclusions is due to the absence of participation constraints in our framework; indeed, Theorem 1 would fail if such constraints were included and were binding. In particular, if the planner in our framework wanted to

designer observes an auxiliary signal, and then show that the optimal mechanism does not need to condition on that signal.

[7] Relaxing the weak-separability of preferences would introduce a separate motive to tax goods, depending on their complementarity to leisure—see Corlett and Hague (1953) and Christiansen (1984).

maximize revenue, then a participation constraint would be needed to make the problem well-defined, and the optimal mechanism would likely distort all decisions.[8]

Finally, we note that the logic behind Theorem 1 can be applied more broadly to derive conditions under which only a subset of available instruments is needed to maximize the planner's objective. For illustration, one could ask whether it is ever optimal to focus *solely* on redistributive market mechanisms (as in the framework of Dworczak ⓡ al. (2021)) and set incomes taxes to zero. By flipping the roles of taste t and ability a, as well as goods x and earnings z in the proof of Theorem 1, we can deduce the following corollary.

Corollary 1. *Suppose that Assumptions A1 and A3 hold and—instead of Assumption A2—welfare weights depend only on tastes: $\lambda(t,a) \equiv \tilde{\lambda}(t)$. Then, any optimal mechanism induces an efficient choice of earnings: $z(t,a) \in \arg\max_{z'}\{z' - w(z',a)\}, \forall (t,a) \in \Theta$, and can be implemented with a (possibly distortionary) mechanism allocating goods x and no income taxes.*[9]

4 Main Analysis: Interaction of Income Taxation and Market Design

In this and the next section, we characterize the optimal design of goods markets and income taxes when the assumptions of Theorem 1 fail. To that end, we introduce additional simplifying assumptions that allow us to find the optimal mechanism in closed form. Section 4 focuses on economic insights and discusses the role of redistributive market design when each of the three assumptions of Theorem 1 is relaxed individually. The underlying proof technique and its broader relation to multidimensional screening are explained in Section 5. In particular, the proofs of all results in Section 4 are based on Proposition 2 in Section 5, which characterizes the structure of optimal mechanisms when Assumption A1–A3 are relaxed all at once.

4.1 Simplified framework

Relative to our baseline framework from Section 2, we make the following additional assumptions. An agent with type $\theta = (t,a)$ has a utility function

$$u(c) + tx - \frac{z}{a},$$

[8]A Rawlsian planner would choose a mechanism similar to the revenue-maximizing planner (except for the choice of the lump-sum payment) but Rawlsian preferences additionally violate Assumption A2 because who the worst-off agent is depends on both the ability type *and* the taste type.

[9]Note that the role of quasi-linearity in Assumption A1 for this case is to ensure that the marginal utility of the numeraire does not vary with *ability* a at efficient labor supply decisions.

with $c \in \mathbb{R}$, $x \in [0,1]$, and $z \in [0,\bar{z}]$. Unless stated otherwise, we take

$$u(c) = \begin{cases} c & c \geq \underline{c}, \\ -\infty & c < \underline{c}. \end{cases}$$

The interpretation is that c represents numeraire consumption whose level cannot fall below a "subsistence threshold" denoted $\underline{c}$. The introduction of a subsistence constraint for numeraire consumption provides a simple way of breaking the assumption of quasi-linear utility (Assumption A1) that we used in Theorem 1. While this way of relaxing quasi-linearity is stylized, it allows us to derive tight predictions about the optimal mechanism; as we discuss later, the qualitative conclusions continue to hold whenever $u(c)$ is a smooth and strictly concave function.

The variable x represents the level of consumption of a good or commodity; x can thus represent quantity, quality, or probability of allocation (in case the good is indivisible). For concreteness, we refer to x as "quality" throughout. Note that we normalize the maximal quality of x to be 1 (which is convenient if we want to interpret x as probability). All individuals have the same preferences over goods that comprise the numeraire c, but they differ (through the taste parameter t) in their marginal rates of substitution between c and x. Our specification also assumes a linear disutility function for generating earnings z, and imposes a finite bound $\bar{z}$ on earnings.

For tractability, we assume that the ability type is binary: $a \in \{l, h\}$. We call agents with $a = h$ the high-ability types, and agents with $a = l$ the low-ability types. We let μ_a denote the mass of agents with ability type a. We assume that it is efficient for high-ability agents to work, $h > 1$, and that low-ability agents are effectively unable to work, $l = 0$.[10]

Taste types t are distributed according to a cumulative distribution function (cdf) F_a with strictly positive, absolutely continuous densities f_a, for $a \in \{l,h\}$, supported on the same interval $[0, \bar{t}]$. The assumption that the lowest type t is 0 is convenient because it implies that—without loss of generality—the lowest taste type will only consume the numeraire, which allows us to interpret $c_a(0)$ as the lump-sum transfer to group a.[11]

[10]We adopt the convention that $0/0 = 0$, so that low-ability agents receive utility 0 from not working. Our analysis only requires that low-ability agents are sufficiently unproductive relative to high-ability agents.

[11]The assumption that there exist agents who have no value for the good has economic implications (see Akbarpour ⓡ al. (2024) for a detailed analysis of the case when all values are bounded away from zero); in our setting, the main implication is that offering a lower price of the good to high-ability agents (compared to low-ability agents) cannot be used to incentivize *all* high-ability agents to work.

Letting $\lambda_a(t) \equiv \lambda(t, a)$, the objective function of the planner is to maximize

$$\sum_{a \in \{l,h\}} \mu_a \int_0^{\bar{t}} \lambda_a(t) \left(c_a(t) + t x_a(t) - \frac{z_a(t)}{a} \right) dF_a(t) \tag{6}$$

over $c_a(t) \geq \underline{c}$, $z_a(t) \in [0, \bar{z}]$, and $x_a(t) \in [0, 1]$, subject to the incentive-compatibility constraint

$$c_a(t) + t x_a(t) - \frac{z_a(t)}{a} \geq c_{a'}(t) + t x_{a'}(t') - \frac{z_{a'}(t')}{a}, \quad \forall t, t' \in [0, \bar{t}], \forall a, a' \in \{l, h\}, \tag{7}$$

and the resource constraint

$$\sum_{a \in \{l,h\}} \mu_a \int_0^{\bar{t}} \left(z_a(t) - c_a(t) - k x_a(t) \right) dF_a(t) \geq G. \tag{8}$$

We assume that the average welfare weight $\bar{\lambda}_l$ on low-ability agents is weakly larger than the average welfare weight $\bar{\lambda}_h$ on high-ability agents. To avoid trivial cases, we also assume that $k \in (0, \bar{t})$, and that G is low enough that there exist feasible allocations at which all agents' utilities are finite. Under these assumptions, restricting attention to allocations satisfying the "subsistence constraint," $c_a(t) \geq \underline{c}$ for all t, is without loss of generality.

4.2 Preliminary observations

In our simplified framework, Pareto efficiency requires that all high-ability agents choose the maximal earnings $\bar{z}$, and that low-ability agents do not work. We first show that redistribution via income tax does not conflict with labor market efficiency in our setting.

Lemma 1. *In any optimal mechanism, labor supply is efficient: $z_h(t) = \bar{z}$, $z_l(t) = 0$, $\forall t \in [0, \bar{t}]$.*

Lemma 1 implies that our simple model can be considered a "best-case scenario" for income taxation: Labor markets remain efficient even under strong redistributive motives.

In contrast—as we will show—the optimal mechanism might distort the goods markets; it is thus instructive to formally define the properties of Pareto efficient allocations of goods (see Online Appendix B for formal supporting results). By linearity of utility functions, agents with taste type t below the marginal cost k should not consume good x in any efficient allocation; agents with taste type t above the marginal cost k should consume the maximal quality of good x subject to the subsistence constraint. Thus, interior consumption $x \in (0, 1)$ of the good is consistent with Pareto efficiency only if $t = k$ (due to the agent's indifference) or $t > k$ and $c = \underline{c}$ (due to the binding subsistence constraint).

A simple test of Pareto efficiency of the goods allocation is to check whether the per-unit price of good x equals its marginal cost. For an incentive-compatible direct mechanism $(z_a(t), x_a(t), c_a(t))$, we define the *per-unit price* for a good with (strictly positive) quality $q \in \text{Im}(x_a)$ faced by ability type a as

$$p_a(q) := \frac{c_a(0) - c_a(x_a^{-1}(q))}{q}. \tag{9}$$

Intuitively, the numerator is equal to the total payment made by any type t consuming quality q (i.e., $q = x_a(t)$), compared to a type $t = 0$ who does not consume the good at all. Dividing by q turns the total payment into the per-unit price. Pareto efficiency requires that all agents face a per-unit price for the good equal to its marginal cost k: $p_a(q) = k$ for any $a \in \{l, h\}$ and any strictly positive $q \in \text{Im}(x_a)$.

4.3 Optimality of market distortions without Assumption A1

We first consider the case in which the utility for numeraire consumption is no longer linear—allowing for the marginal utility of disposable income to vary with taste type—while maintaining Assumptions A2 and A3. We start by investigating the properties of optimal mechanisms when the subsistence constraint binds for the low-ability agents. Then, we show that the main results are robust to introducing strictly concave utility in the numeraire. Whenever Assumption A3 holds, we use "F" (without a subscript) to denote the (common across the two ability types) cdf of the distribution of taste types.

Theorem 2. *Suppose that Assumptions A2 and A3 hold. Furthermore, assume that $\frac{1-F(t)}{f(t)}$ is non-increasing, $\frac{F(t)}{f(t)}$ is non-decreasing, and*

$$\mu_h \bar{z}\left(1 - \frac{1}{h}\right) - G < k + \underline{c} \quad \text{and} \quad \frac{\bar{z}}{h} \geq \bar{t}. \tag{10}$$

Then, there exists an optimal mechanism in which every agent:

1. *chooses how much to work, with each unit of earnings taxed at the rate $1 - 1/h$;*
2. *receives a lump-sum transfer equal to $\underline{c} + p_l q_l$; and*
3. *can purchase any quality $q \in [0, 1]$ of the good at a per-unit price $p(q)$, where*

$$p(q) = \begin{cases} p_l & q \leq q_l, \\ p_l \frac{q_l}{q} + p_h\left(1 - \frac{q_l}{q}\right) & q > q_l, \end{cases}$$

for some quality q_l and $p_h > k > p_l$. That is, the marginal price is below marginal cost for qualities lower than q_l and above marginal cost for qualities above q_l. Moreover, if $\lambda_h = 0$, then p_h is the price that maximizes the net revenue of the planner.

Under the optimal mechanism, agents face a nonlinear price schedule in the goods market, making the allocation of goods inefficient. The mechanism offers low-quality goods at an average price strictly below marginal cost. Low-ability agents of sufficiently high taste type choose to consume the subsidized quality q_l (which is the maximal quality they can afford without labor income). At the same time, high-ability agents—who work and thus have higher disposable income—can "top up" their consumption of the good; the marginal price charged for qualities above q_l is strictly above marginal cost, and thus extracts revenue from wealthier agents. In fact, when the welfare weight on high-ability agents is 0, the top-up marginal price p_h is chosen to maximize the net revenue of the planner. In other words, p_h is the price that would have been chosen by a revenue-maximizing monopolist producing the good at marginal cost k.

Because the lump-sum payment is equal to subsistence consumption plus $p_l q_l$—the total price of the subsidized quality q_l—the optimal mechanism can be decentralized by letting agents choose between receiving the good with quality q_l in-kind, or opting for a higher cash transfer, together with a flat commodity tax on additional purchases of the good in the market.[12] If, instead, quality represents the probability of allocation of a homogeneous good, then an alternative interpretation is that the low-quality option is implemented via price controls (setting the price below the market-clearing level) and rationing (interior probability of receiving the good).

Note that the consumption subsidy in the optimal mechanism is not means-tested: It is available also to high-ability agents. This is because income taxes are set at a level that extracts all surplus from working (i.e., the net wage is just enough to cover the cost of labor supply). If high-ability agents were excluded from the subsidy, given the tax regime, some of them would choose not to work—which Lemma 1 shows is not optimal. At the same time, though, the subsidy is phased out by charging agents a higher marginal price for topping up. This combination of prices achieves the screening effect that underlies the main results of Dworczak ⓡ al. (2021): high-ability agents with strong enough preferences for the good choose to top up; because low-ability agents cannot afford to do so, market sorting in effect identifies the higher-ability agents with particularly strong preferences for

[12]In particular, in our model, implementing the optimal mechanism from Theorem 2 does not require monitoring the consumption levels of individual agents. However, agents must be able to top-up the publicly provided option with private consumption—this is an appropriate assumption if quality is interpreted as quantity but may fail in general; see Kang and Watt (2024a) and Kang and Watt (2024b).

the good, and extracts surplus from them to redistribute via the lump-sum transfer.

The following intuition helps explain why it is optimal to lower the price of the low-quality good below marginal cost. Suppose that the good is initially offered at marginal cost k. Low-ability agents with $t > k$ spend all their disposable income (equal to the lump-sum payment) on the good with quality $q < 1$ which puts their consumption c at subsistence level. Crucially, the marginal utility of disposable income for these agents is higher than of anybody else, since they would (prefer to) spend an additional dollar of income on higher quality of x (yielding utility gain $t/k > 1$) rather than on the numeraire (yielding utility gain 1). Now, perturb the price to a slightly lower level $k/(1+\epsilon)$, where $\epsilon > 0$. This perturbation has a negative effect due to allocative inefficiency (some agents with $t < k$ consume the good); however, this effect is of second-order in ϵ (it is an order-ϵ distortion for an order-ϵ mass of agents). The perturbation also has a positive effect, which is that all low-ability agents with types $t > k$ (whose marginal utility of income exceeds the marginal value of public funds) now consume a quality higher by $q\epsilon$; this is a first-order effect because even inframarginal taste types enjoy an increase in utility at the order of ϵ. Thus, for small ϵ, the positive effect dominates the negative effect, and it is optimal to lower the price below marginal cost.

For additional illustration, consider an example in which the good is treatment for a medical condition. The taste type captures whether (and to what extent) treatment is needed. When the cost of treatment is high enough, low-income agents are constrained by subsistence and can afford at most a low-quality treatment, while some of them have a very high marginal utility from treatment quality. Thus, even if the designer has no inherent preference for redistribution to agents who are sick (i.e., welfare weights do not depend on the taste type) and the likelihood of getting sick is unrelated to ability (taste type and ability types are independent), she still wants to redistribute towards agents who consume low-quality treatment because of the associated higher marginal utility of disposable income; the planner achieves this by subsidizing the price of treatment below its marginal cost.[13]

Summarizing, we find that *the mechanism subsidizes the consumption of the low-quality good because choosing low quality is a signal of high marginal value of income*. Low-ability agents

[13]The same story could be told with a luxury good (perhaps a yacht) replacing treatment for a disease, in which case the appeal of the intuition would be lost. One way to distinguish between these two cases (essential healthcare versus yachts) is to consider a social welfare function in which agents with higher *absolute* level of utility receive a lower social welfare weight. An agent consuming yachts could have high marginal utility from them but since their overall utility is presumably already high, their contribution to social welfare would be small, turning off the effect described here. In contrast, an agent in need of medical treatment who chooses low-quality healthcare is likely to have both high marginal utility and low overall utility, resulting in high redistribution-worthiness. This distinction arises naturally when the planner maximizes a strictly concave transformation of agents' utilities; in our current framework with exogenous welfare weights, we could replicate that logic by making welfare weights a function of tastes (i.e., by relaxing Assumption A2).

who consume the good must be at subsistence level, which means that the marginal utility they would derive from any extra income is high. It is thus optimal to give more resources to them—and if high-ability agents have enough demand that they choose to top-up their consumption, taxing that consumption helps finance the subsidy.

Finally, we comment on the regularity assumptions in Theorem 2: First, the monotonicity of hazard rates is a mathematically restrictive assumption but its only role in the proof is to rule out ironing and ensure that first-order conditions are sufficient.[14] With ironing, as we explain in Section 5, the optimal mechanism may need to offer additional options to high-ability agents, which complicates the exposition without adding new economic insight. The condition $\mu_h \bar{z}(1 - 1/h) - G < k + \underline{c}$, by contrast, is economically important: it states that the economy does not have enough resources to put all agents strictly above the subsistence level. If that condition fails, then the subsistence constraint is moot, and optimality of efficient provision of the good (pricing at marginal cost) follows from Theorem 1 because all agents have identical marginal utility of disposable income (effectively, restoring Assumption A1). Finally, the assumption $\bar{z}/h \geq \bar{t}$ ensures that the subsistence constraint is slack for high-ability agents—it states that high-ability agents are sufficiently productive relative to the strongest possible taste for buying the good. This assumption is not crucial; if the subsistence constraint is binding also for some high types, then the solution is virtually the same, with the only difference that high-ability agents with strong enough taste will consume some quality level q_h strictly lower than 1 (see Proposition 2 for the form of the optimal mechanism without any regularity assumptions).

Robustness to curvature in utility functions. A potential concern with Theorem 2 is that its conclusion could be driven by the stylized form of the subsistence constraint (that we used to invalidate A1). An alternative way of relaxing quasi-linearity of preferences is to consider a utility from numeraire that is *nonlinear*. Suppose that preferences take the general form in equation (1) with the functions u, v, and $-w$ assumed smooth and strictly concave. While a full characterization of the optimal mechanism is no longer attainable in that case, we find that the main features of the optimal mechanism from Theorem 2 are maintained. Here, we summarize and discuss the findings; the formal results are relegated to Online Appendix C.

First, under the optimal mechanism, consumption of good x is distorted for at least some types. Second, if the consumption of the low-ability agents is distorted, then it must be

[14] The condition is stronger than usual because the analog of virtual surplus in our analysis is endogenous to the Lagrange multiplier on the subsistence constraint—see the proof in Appendix A for details.

distorted upwards, meaning that it is optimal to subsidize good x for them.[15] Third, if the planner does not value the utility of high-ability types, then it is optimal to charge a revenue-maximizing price (on the margin) for qualities of x above what the low-ability agents consume. Thus, the optimal mechanism under curvature in the utility function preserves the main features of the optimum under the subsistence constraints.

To build intuition, notice that agents with high taste type choose relatively higher quality of x and lower consumption of numeraire c. That results in higher marginal utility of numeraire $u'(c)$ and, hence, higher marginal utility of disposable income. The concavity of $u(c)$ thus creates a positive correlation between purchases of x and the marginal utility of income. The optimal mechanism exploits this correlation by subsidizing purchases of good x to support individuals with high marginal utility. Note that preference heterogeneity is crucial for the force just described to be present. Agents with different tastes endogenously differ in their marginal value for disposable income, which creates a motive for redistribution that cannot be addressed through the income tax alone.

Summary of the impact of relaxing Assumption A1. We have shown that taste heterogeneity creates an endogenous motive to redistribute between agents with same abilities but different tastes. Whenever agents' preferences depart from the restrictive quasi-linear case—either due to subsistence constraints or concavity in the utility function—differences in tastes generate differences in marginal utilities. Thus, a utilitarian planner benefits from redistributing across tastes within ability level. An income tax is ill-suited to conduct such redistribution and, hence, the optimal mechanism involves market distortions (in the form of nonlinear pricing). Importantly, this redistributive motive due to taste heterogeneity has nothing to do with potential dependence of the exogenous welfare weights on taste (which was in fact ruled out in this section by Assumption A2).

4.4 Optimality of distortions without Assumptions A2 or A3

Next, we assume that utilities are linear in the numeraire consumption (taking $\underline{c} = -\infty$ in our specification), restoring Assumption A1, and examine what happens when we relax Assumptions A2 and/or A3 of Theorem 1. Because both relaxations have a similar effect, we exposit them together.

Even though we assumed that the designer cares more about low-ability agents than high-ability agents on average, we cannot conclude that she would prefer to charge a lower price for the good to low-ability agents. For example, it may be optimal to post a lower price for

[15]To prove this property, we assume that ironing is not required in the optimal mechanism.

high-ability agents if their distribution of taste types is lower (in an appropriate sense) than that for the low-ability agents. To limit the number of cases to consider, we rule out this (perhaps less economically interesting) case by assuming

$$(1 - \Lambda_h(t))\gamma_h(t) \geq (1 - \Lambda_l(t))\gamma_l(t), \quad \forall t \in [0, \bar{t}], \tag{11}$$

where $\Lambda_a(t)$ is the average welfare weight on agents with ability a and taste type above t, and $\gamma_a(t) = (1 - F_a(t))/f_a(t)$ is the inverse hazard rate. To interpret condition (11), note that if the welfare weights on low- and high-ability agents are the same, then the assumption states that the distribution of high-ability agents' tastes is higher—in the hazard-rate order—than that of low-ability agents. Analogously, if the taste distribution in the two different groups is the same, then the assumption states that the welfare weights on low-ability agents are weakly higher than on high-ability agents, conditional on the taste type exceeding any threshold.

Theorem 3. *Suppose that Assumption A1 and inequality (11) hold, and that*

$$(t - k - (1 - \Lambda_h(t))\gamma_h(t))\, f_h(t) \tag{12}$$

is non-decreasing whenever it is negative. Then, there exists an optimal mechanism that takes one of two forms:

1. *Labor income is taxed at a rate $1 - 1/h$ and all agents can purchase the good at a constant per-unit price p satisfying*
$$p = k + (1 - \Lambda(p))\gamma(p),$$
where $\Lambda(p)$ is the average welfare weight on all agents with taste type t above p, and γ is the inverse hazard rate of the unconditional distribution of taste; or

2. *Labor income is taxed at a rate strictly lower than $1 - 1/h$; agents with no labor income can purchase the good at a constant per-unit price p_l, while agents with positive labor income can purchase the good at a constant per-unit price p_h, where*
$$k + (1 - \Lambda_h(p_h))\gamma_h(p_h) \geq p_h > p_l \geq k + (1 - \Lambda_l(p_l))\gamma_l(p_l).$$

Without the subsistence constraint, the optimal mechanism in our framework becomes simple. In particular, after ironing is ruled out by the regularity conditions, it is always optimal to sell the good to each group at a single price.[16] If income taxes are set to the max-

[16]Like with our anti-ironing condition from Theorem 2, the assumption that (12) is non-decreasing when-

imal level (subject to maintaining the high-ability agents' incentive to work), then prices are in fact the same for both groups; but when income taxes leave high-ability agents with a strictly positive surplus from working, the price for the good faced by agents with labor income (i.e., high-ability agents) is higher than the price faced by agents with no labor income (i.e., low-ability agents).

In case 1, efficiency in the goods market is generically suboptimal when either Assumption A2 or A3 fails. Indeed, efficiency in the goods market obtains only if the average welfare weight on agents buying the good, $\Lambda(p)$, is equal to the unconditional average of 1—and when welfare weights depend directly on the taste type (i.e., when Assumption A2 is violated), $\Lambda_a(t)$ (and hence its average over a) will typically diverge from 1. Moreover, even if welfare weights do not depend on the taste type directly, $\Lambda(p)$ might deviate from 1 if taste types are correlated with ability types (i.e., when Assumption A3 is violated).

For intuition, suppose that $\bar{\lambda}_h = 0$, so that the designer only cares about low-ability agents, and that all low-ability agents have the same welfare weight (Assumption A2 holds). It is then straightforward to show that

$$\Lambda(p) < 1 \quad \iff \quad \frac{1 - F_l(p)}{\mu_l(1 - F_l(p)) + \mu_h(1 - F_h(p))} < 1 \quad \iff \quad F_l(p) > F_h(p).$$

That is, the good is taxed if low-ability agents have a lower distribution of the taste type (in the sense of first-order stochastic dominance), and is subsidized if low-ability agents have a higher distribution of the taste type. In both cases, the designer uses the market for the good to transfer more resources from high-ability agents to low-ability agents, relying on the statistical dependence between taste and ability. Note the role played by taste heterogeneity: Although the designer cannot take away more resources from all high-ability agents directly (their income is already taxed as much as is possible while still satisfying the incentive constraints), she can transfer more resources away from high-ability agents with high taste by taxing the good (when high-ability agents have higher taste on average) or from high-ability agents with low taste by lowering the lump-sum transfer and subsidizing the good instead (in the case that low-ability agents have higher taste on average).

In case 2, the allocation of the good is never efficient because high-ability agents face a strictly higher price than low-ability agents (hence, it cannot be that both prices are equal

ever it is negative is mathematically restrictive but not essential for our results. The simplicity of the mechanism is of course driven by the linearity of our model, which in particular implies that agents with the same taste choose the same consumption bundle *regardless* of their income. We certainly do not claim that offering a single quality at a single price is a policy recommendation emerging from our analysis. Rather, we point out that even when other forces (e.g., income effects) are turned off, violation of either Assumption A2 or A3 leads to an optimal adjustment of prices away from marginal costs.

to marginal cost). The mechanism can be implemented as a means-tested subsidy for the good (in the sense that the subsidy is only available to agents who have no earnings). For this mechanism to be incentive-compatible, earnings cannot be taxed maximally, which is why the tax rate is strictly below $1 - 1/h$.

To provide further intuition, it is instructive to consider the benchmark case in which ability is observed (but revenue is still valued at the average welfare weight of 1). Then, it is optimal to offer quality 1 to agents with ability a at price $p_a^\star$ that (at an interior solution) must satisfy[17]

$$p_a^\star = k + (1 - \Lambda_a(p_a^\star))\gamma_a(p_a^\star), \quad a \in \{l, h\}. \tag{13}$$

Under Assumption (11), $p_h^\star > p_l^\star$. However, to implement $p_h > p_l$ when ability is not observed, the designer must provide enough rents to high-ability types to maintain incentive compatibility—and the larger the gap $p_h - p_l$, the larger the rent that high-ability agents must receive in the labor market. This creates a trade-off and implies that the optimal prices p_h and p_l will be closer together than the benchmark prices $p_h^\star$ and $p_l^\star$ given by (13).

It remains to discuss the determinants of which of the two candidate optimal mechanisms is used. The result below shows that both mechanisms are sometimes optimal, and the choice between them depends on whether the designer has a strong motive to redistribute across ability types.

Proposition 1. *Parametrize the welfare weights by $\lambda_a(t) \equiv \bar{\lambda}_a \cdot \omega_a(t)$, where $\omega_a(t)$ is a function with mean 1 with respect to distribution F_a, and $\bar{\lambda}_h = (1 - \mu_l \bar{\lambda}_l)/\mu_h$ (to keep the overall average welfare weight normalized to 1). Suppose that the assumptions of Theorem 3 hold, and that $(t - k - (1 - \Lambda_l(t))\gamma_l(t))\, f_l(t)$ is non-decreasing whenever it is positive, for any $\bar{\lambda}_l \geq 1$. Then, there exists a cutoff value $\bar{\lambda}_l^0 \geq 1$ such that mechanism 1 from Theorem 3 is optimal if $\bar{\lambda}_l > \bar{\lambda}_l^0$ and mechanism 2 is optimal if $\bar{\lambda}_l < \bar{\lambda}_l^0$ (moreover, $\bar{\lambda}_l^0 > 1$ if inequality (11) is strict for all interior t).*

For intuition, recall that the planner would in general benefit from using income-dependent prices in the market: Income reveals information about ability, and formula (13) shows that conditioning prices on ability is useful whenever social welfare weights depend on taste differently for low- and high-ability agents (Assumption A2 fails) or ability is correlated with taste (Assumption A3 fails). This observation is in fact a straightforward extension of the classical idea of third-degree price discrimination, which has been studied in detail in the context of inequality-aware market design by Akbarpour ⓡ al. (2024). The key difference is that income in our setting is not exogenous, leading to a trade-off: More income redistribution through the income tax tightens the incentive-compatibility constraint

[17] This follows, for example, from the analysis of Akbarpour ⓡ al. (2024) by replacing their fixed-supply assumption with a constant marginal cost.

between high- and low-ability agents, restricting the use of income as an observable characteristic used for third-degree price discrimination in the goods market.

The trade-off explains the comparative statics in Proposition 1. The weight $\bar{\lambda}_l$ determines the strength of the planner's vertical redistributive preferences, and hence the benefits from redistributive income taxation. When $\bar{\lambda}_l$ is high, the planner is primarily concerned with redistribution from high- to low-ability agents. In such cases, the income tax should be maximally redistributive and, as a result, income cannot be used for price discrimination in the goods market—mechanism 1 is used. In contrast, when $\bar{\lambda}_l$ is close to 1 (and hence to $\bar{\lambda}_h$), the social cost of limiting the redistribution through the income tax is small. Then, the planner finds it optimal to lower income taxes (introducing slack in the incentive constraint) in order to implement income-dependent goods prices (up to the point at which the incentive constraint binds again)—mechanism 2 is used.

One takeaway is that when Assumptions A2 and A3 are violated, income taxation and redistribution through markets become to some extent *substitutes*. This is because their degree is jointly limited by a common incentive constraint: Sometimes the income tax is scaled down to allow for the use of an additional instrument in the market design, and sometimes market design avoids using this instrument to allow for more redistribution via the income tax.

5 General Structure of the Optimal Mechanism

In this section, we explore the general structure of optimal mechanisms in the simplified setting introduced in Section 4. In particular, we uncover the reasons for the relatively simple mechanism form identified in Theorems 2 and 3, and sketch the proofs of these results. The main goal here is to explain the key technical ideas behind our construction that could be useful in studying similar multidimensional screening problems.

We center our discussion in this section around the following result that predicts the form of the optimal mechanism in the absence of any regularity conditions:

Proposition 2. *In the framework of Section 4, there always exists an optimal mechanism in which:*

1. *Low-ability agents consume one quality of the good:* $x_l(t) \in \{0, q_l\}$, *where* $q_l \in (0, 1]$;
2. *High-ability agents consume at most three distinct qualities of the good:* $x_h(t) \in \{0, q_i, q_l, q_h\}$, *where* $0 \leq q_i \leq q_l \leq q_h$;
3. *If the subsistence constraint does not bind, then the highest quality consumed is* 1*: If* $c_l(\bar{t}) > \underline{c}$, *then* $q_l = q_h = 1$, *and if* $c_h(\bar{t}) > \underline{c}$, *then* $q_h = 1$;

4. *High-ability agents either (i) receive a post-tax wage equal to the cost of labor provision but face a weakly better average price for the good consumed by low-ability agents, or (ii) receive a post-tax wage strictly higher than the cost of labor provision but face a higher average price for the good consumed by low-ability agents: Whenever $q_l \in Im(x_h)$, either*
   (a) $c_h(0) - \frac{\bar{z}}{h} = c_l(0)$ *and* $p_h(q_l) \leq p_l(q_l)$, *or*
   (b) $c_h(0) - \frac{\bar{z}}{h} > c_l(0)$ *and* $p_h(q_l) > p_l(q_l)$.

Note that Proposition 2 describes the optimal *direct* mechanism. Given the linearity of our model, discrete quality choices can be decentralized with a piece-wise linear—and potentially income-dependent—price schedule similar to the one described in Theorem 2.

In an optimal mechanism, low-ability agents face a simple choice (recall that they do not work, by Lemma 1). They receive a lump-sum transfer, equal to $c_l(0)$, and decide whether to buy the good with quality q_l at a per-unit price of p_l, or spend their entire disposable income on the numeraire. Whether $q_l = 1$ or $q_l < 1$ depends on whether the subsistence constraint binds: $q_l < 1$ can only be optimal if high-taste low-ability agents consume at the subsistence level.

High-ability agents face a potentially more complicated choice: They choose from up to three distinct quality levels. First, intermediate-taste high-ability agents consume the same quality q_l as low-ability agents as a result of the binding downward IC constraint (in ability). Second, when the subsistence constraint binds for low-ability agents so that $q_l < 1$, a higher quality $q_h \in (q_l, 1]$ is offered to the highest-taste high-ability agents who can afford it because of their labor income. If high-ability agents are sufficiently productive, the subsistence constraint is slack for them, and q_h is equal to the maximal quality 1 (as in Theorem 2). Finally, the lowest quality level q_i is needed if the optimal solution requires ironing. Intuitively, if the planner's objective function is non-monotone in the taste type, it may be preferable to satisfy the downward IC constraint (in ability) by offering a low-quality good at a low price to high-ability agents (this possibility is ruled out by the regularity conditions imposed in Theorems 2 and 3).

Finally, the price for quality q_l may be different for the two ability types in the optimal mechanism. The trade-off was already explained in the context of Theorem 3. In case *(i)*, the planner maximizes the income tax paid by high-ability agents subject to all of them working (which is optimal by Lemma 1) and is hence constrained to offer them a weakly lower price for the good. In case *(ii)*, the income tax is reduced and a strictly better price for quality q_l is offered to low-ability agents.

5.1 Sketch of argument

The proof of Proposition 2—which is then specialized to prove Theorems 2 and 3—is relatively involved but can be decomposed into several steps, most of which use familiar ideas. First, we reduce the problem to maximizing the welfare function over allocation rules $x_l(t)$ and $x_h(t)$ that are monotone in t, using the envelope formula to express consumption of the numeraire in terms of the allocation rules and ability-specific lump-sum transfers. We use the resource constraint to pin down the lump-sum transfer to the low-ability agents.

Second, we argue that the subsistence constraints can only bind for the highest–taste type within each ability level. This is intuitive, as agents with the same ability share the same disposable income, by Lemma 1. This observation allows us to incorporate all subsistence constraints into the objective function via a pair of Lagrange multipliers, after we parametrize the highest quality level consumed by each ability type. In the final stage of the construction, we optimize over these highest quality levels; intuitively, the highest quality level is 1 (the maximal consumption of good x) if the subsistence constraints for the given ability type turn out to be slack, but it is interior otherwise.

Third, we argue that the incentive constraint preventing low-ability agents from pretending to have high ability is slack (as a consequence of Lemma 1). The opposite constraint generally binds but—conditional on misreporting ability—high-ability agents find it optimal to report their taste truthfully (due to separability of preferences and the fact that all low-ability agents have equal earnings). This key step reduces the incentive constraints in our multidimensional model to two standard one-dimensional constraints (within each ability level) and an outside-option–like constraint for the high-ability agents: High-ability agents must receive a minimal utility level pinned down by the allocation to low-ability agents with the corresponding taste type. Fixing the allocation rule for low-ability agents, our problem thus becomes a standard one-dimensional screening problem with a type-dependent outside option (as in Jullien (2000)).

Fourth, we fix the allocation of low-ability agents and solve for the optimal allocation rule for high-ability agents. We rely on an ironing technique—recently introduced by Dworczak and Muir (2024)—that extends the analysis of Myerson (1981) to problems with type-dependent outside options. The key take-away is that the allocation rule for high-ability agents is *linear* in their outside option, i.e., in the allocation rule for low-ability agents.

Fifth, we maximize the Lagrangian over the allocation rule for low-ability agents, accounting for how that choice affects the optimal allocation rule for high-ability agents. As a consequence of the previous step, this problem is *linear*, with no constraints. By a standard argument, the optimal allocation rule for low-ability agents is therefore a posted price for a single quality level (pinned down by the subsistence constraint).

Sixth, we show that the allocation rule for high-ability agents derives its simple structure from the cutoff allocation rule for low-ability agents. Up to two additional quality (and price) levels may be needed for high-ability agents: a lower quality may be introduced by the ironing procedure, while a higher quality level may be required due to a more permissive subsistence constraint—consistent with the informal discussion of Proposition 2.

Finally, part 4 of Proposition 2 follows from the optimal choice of a lump-sum payment to high-ability agents. Intuitively, the planner faces a trade-off: She can satisfy the endogenous outside-option constraint for high-ability agents (created by the allocation to low-ability agents) by either *(i)* increasing the allocation to high-ability agents, or *(ii)* giving those agents a higher monetary payment (implemented in an incentive-compatible way as a reduction in income taxes). The generalized ironing procedure determines how to optimally use options *(i)* and *(ii)*, leading to the two cases in part 4 of Proposition 2.

5.2 Literature notes

As the proof sketch makes clear, the tractability of our model relies crucially on the simplifying assumption of a binary ability type (with an incentive constraint that can only bind in one direction). A mathematically similar structure arises in the so-called "FedEx problem" in which agents differ in their (continuous) value for receiving a package, as well as a discrete (possibly binary) deadline by which they need to receive it. Relying on duality techniques, Fiat et al. (2016) derived the structure of the revenue-maximizing mechanism in such an environment; Saxena et al. (2018) showed that the number of prices required for full optimality grows exponentially with the number of "deadline" types, suggesting that the optimal mechanism in our framework would become increasingly complex to describe with more ability types.

Ahlvik et al. (2024) solved a multidimensional screening problem (assuming no bunching in the solution) in a model with a binary (and deterministic) choice in the goods market. We instead assume that productivity is binary but allow for continuous types and choices in the goods market. This makes the two papers complementary: we obtain a richer design of the goods market (e.g., with several levels of consumption, rationing etc.), while Ahlvik et al. (2024) obtain richer predictions about the optimal income tax schedule. Additionally, the techniques used to solve the respective problems are different (Ahlvik et al. (2024) rely on first-order conditions using a perturbation approach and optimal-control methods).

Our way of modeling the subsistence constraint (as a means of relaxing Assumption A1) connects our framework to models with budget-constrained agents. Most closely related is the work of Che et al. (2013) and Li (2021) who studied optimal allocation of resources when agents differ in values for the good and budgets (with Li (2021) additionally allowing for costly verification of the agents' types). The main similarity is the emergence of "rationing" (which we interpreted as interior levels of quality) resulting from the need to satisfy the subsistence/ budget constraint. The main difference is that these papers study *efficient* allocations, without the redistributive concerns that are the core focus of our work.

Another benchmark for our analysis is that of Dworczak ⓡ al. (2021) who solved an analogous market-level redistribution problem without integrating income taxation. The optimality of offering a single quality to low-ability agents (and at most three quality levels to high-ability agents) is a consequence of the assumption of constant marginal cost in our framework; if, instead, we had assumed a fixed supply of goods as in Dworczak ⓡ al. (2021), an additional quality level might be needed in the optimal mechanism.[18] (Avoiding this additional complexity in the optimal mechanism is why we decided to work with a fixed–marginal cost model, which is also closer to the original work of Atkinson and Stiglitz (1976).) With a fixed supply and linear utilities as in the setting of Dworczak ⓡ al. (2021), rationing (i.e., an interior quality level) is always inefficient. In our model, rationing may be efficient for agents whose numeraire consumption is at subsistence—and instead, inefficiency is manifested by the price diverging from marginal cost.

6 Concluding Remarks

We investigated the problem of joint design of income taxation and goods markets using a mechanism design framework, and showed that goods market interventions play an important role in balancing equity and efficiency under redistributive preferences. Market interventions are generally useful when consumption decisions reveal welfare-relevant information, due to consumption tastes generating heterogeneity in marginal utilities in disposable income, correlating with ability, or directly entering the social welfare weights. In each of these cases, restricting attention to income taxes means giving up on additional information revealed by market behavior, and hence not achieving the second best.

Our work provides qualified theoretical support for several types of observed market interventions. First, we showed that nonlinear pricing of goods and services (with subsidies or in-kind transfers of low qualities and taxes on high qualities) can complement optimal nonlinear income taxation. This conclusion holds in particular when individuals differ in their

[18]Mathematically, the Dworczak ⓡ al. (2021) way of modeling scarcity is closely related to ours because a fixed supply constraint results in a Lagrange multiplier that enters the problem in the same way as a constant marginal cost. However, the need to satisfy the supply constraint may result in the optimal mechanism being a convex combination of different maximizers of the Lagrangian (see, e.g., Doval and Skreta (2024)).

marginal utilities of disposable income (e.g., because of different needs) and purchases of certain qualities of goods act as a signal of high marginal utility. This may be arguably the case for essential goods and services that constitute a large share of some households' budgets, like housing or healthcare. And indeed, many countries have programs guaranteeing access to low-quality housing and healthcare to all citizens, while often imposing sales taxes on transactions in the private market.

Second, differential commodity taxation may be useful—along the lines suggested by Diamond (1975)—even in the presence of an optimally designed income tax. That happens when goods purchases provide additional information about welfare weights, on top of what can be inferred from the level of earnings. For example, a redistributive government that already imposes a strongly progressive income tax can reap further equity gains with a tax on goods and services catering to high-ability individuals.

Finally, when the redistributive social preferences across the income distribution are not very strong, it may be optimal to use means-tested programs, potentially in combination with sales taxes on private transactions—effectively using income for price discrimination in certain goods markets. For example, it may be optimal to offer programs such as food stamps in the United States, where low income is a primary eligibility criterion. At the same time, however, our framework suggests that income taxes should become less progressive in response to the introduction of such programs.

That said, we would like to stress that our framework is designed to understand the equity-efficiency trade-off at an abstract level rather than to directly inform policy. Any market intervention is associated with costs and drawbacks that are not captured by our model. Moreover, optimal policies depend on the details of population distributions and social preferences. Thus, further theoretical and empirical research is needed to understand the structure of market-level redistributive policies in realistic settings. The main point of our work is to highlight—by fully characterizing second-best mechanisms in a simple model—that effective redistribution may need to combine traditional public finance policies and equitable design of markets for goods and services. We expect that many important insights remain to be discovered at the intersection of these two research fields.

References

AHLVIK, L., M. LISKI, AND M. MÄKIMATTILA (2024): "Pigouvian Income Taxation," CEPR Discussion Paper.

AKBARPOUR, M. ⓡ P. DWORCZAK ⓡ S. D. KOMINERS (2024): "Redistributive Allocation Mechanisms," *Journal of Political Economy*, 132, 1831–1875.

ALLCOTT, H., B. B. LOCKWOOD, AND D. TAUBINSKY (2019): "Regressive sin taxes, with an application to the optimal soda tax," *Quarterly Journal of Economics*, 134, 1557–1626.

ATKINSON, A. B. AND J. E. STIGLITZ (1976): "The design of tax structure: Direct versus indirect taxation," *Journal of Public Economics*, 6, 55–75.

BIERBRAUER, F. (2024): "Is a uniform price on Carbon desirable? A public finance perspective," CESifo Working Paper.

BIERBRAUER, F., P. BOYER, A. PEICHL, AND D. WEISHAAR (2024): "Pareto-Improvements, Welfare Trade-Offs and the Taxation of Couples," .

BLACKORBY, C. AND D. DONALDSON (1988): "Cash versus kind, self-selection, and efficient transfers," *American Economic Review*, 691–700.

BLOMQUIST, S. AND V. CHRISTIANSEN (2008): "Taxation and heterogeneous preferences," *FinanzArchiv/Public Finance Analysis*, 218–244.

BOERMA, J., A. TSYVINSKI, AND A. P. ZIMIN (2022): "Bunching and taxing multidimensional skills," National Bureau of Economic Research Working Paper.

BONNANS, J. F. AND A. SHAPIRO (2000): *Perturbation Analysis of Optimization Problems*, Springer.

CHE, Y.-K., I. GALE, AND J. KIM (2013): "Assigning resources to budget-constrained agents," *Review of Economic Studies*, 80, 73–107.

CHRISTIANSEN, V. (1984): "Which commodity taxes should supplement the income tax?" *Journal of Public Economics*, 24, 195–220.

CONDORELLI, D. (2013): "Market and non-market mechanisms for the optimal allocation of scarce resources," *Games and Economic Behavior*, 82, 582–591.

CORLETT, W. J. AND D. C. HAGUE (1953): "Complementarity and the excess burden of taxation," *The Review of Economic Studies*, 21, 21–30.

CREMER, H. AND F. GAHVARI (1998): "On optimal taxation of housing," *Journal of Urban Economics*, 43, 315–335.

CREMER, H., P. PESTIEAU, AND J.-C. ROCHET (2001): "Direct versus indirect taxation: the design of the tax structure revisited," *International Economic Review*, 42, 781–800.

——— (2003): "Capital income taxation when inherited wealth is not observable," *Journal of Public Economics*, 87, 2475–2490.

DEATON, A. (1979): "Optimally uniform commodity taxes," *Economics Letters*, 2, 357–361.

DEATON, A. AND J. MUELLBAUER (1980): "An almost ideal demand system," *American Economic Review*, 70, 312–326.

DEATON, A. AND N. STERN (1986): "Optimally uniform commodity taxes, taste differences and lump-sum grants," *Economics Letters*, 20, 263–266.

DIAMOND, P. (1975): "A many-person Ramsey tax rule," *Journal of Public Economics*, 4, 335–342.

DIAMOND, P. AND J. SPINNEWIJN (2011): "Capital income taxes with heterogeneous discount rates," *American Economic Journal: Economic Policy*, 3, 52–76.

DOLIGALSKI, P., P. DWORCZAK, J. KRYSTA, AND F. TOKARSKI (2025): "Incentive separability," *Journal of Political Economy Microeconomics*, 3, 539–567.

DOVAL, L. AND V. SKRETA (2024): "Constrained Information Design," *Mathematics of Operations Research*, 49, 78–106.

DWORCZAK, P. AND E. MUIR (2024): "A mechanism-design approach to property rights," Working Paper.

DWORCZAK, P. ⓡ S. D. KOMINERS ⓡ M. AKBARPOUR (2021): "Redistribution through Markets," *Econometrica*, 89, 1665–1698.

EDEN, M. AND L. M. FREITAS (2024): "Income Anonymity," Working Paper.

FEREY, A., B. B. LOCKWOOD, AND D. TAUBINSKY (2024): "Sufficient statistics for nonlinear tax systems with general across-income heterogeneity," *American Economic Review*, 114, 3206–3249.

FIAT, A., K. GOLDNER, A. R. KARLIN, AND E. KOUTSOUPIAS (2016): "The FedEx Problem," in *Proceedings of the 2016 ACM Conference on Economics and Computation*, Association for Computing Machinery, 21–22.

GAUBERT, C., P. KLINE, D. VERGARA, AND D. YAGAN (2025): "Place-Based Redistribution," *American Economic Review*, 115, 3415–50.

GAUTHIER, S. AND F. HENRIET (2018): "Commodity taxes and taste heterogeneity," *European Economic Review*, 101, 284–296.

GERRITSEN, A., B. JACOBS, K. SPIRITUS, AND A. V. RUSU (2025): "Optimal taxation of capital income with heterogeneous rates of return," *The Economic Journal*, 135, 180–211.

GOLOSOV, M. AND I. KRASIKOV (2025): "The optimal taxation of couples," *The Quarterly Journal of Economics*, qjaf021.

GOLOSOV, M., M. TROSHKIN, A. TSYVINSKI, AND M. WEINZIERL (2013): "Preference heterogeneity and optimal capital income taxation," *Journal of Public Economics*, 97, 160–175.

GOMES, R., J.-M. LOZACHMEUR, AND A. PAVAN (2017): "Differential Taxation and Occupational Choice," *The Review of Economic Studies*, 85, 511–557.

GUESNERIE, R. AND K. ROBERTS (1984): "Effective policy tools and quantity controls," *Econometrica: Journal of the Econometric Society*, 59–86.

HAGHPANAH, N. AND J. HARTLINE (2020): "When Is Pure Bundling Optimal?" *Review of Economic Studies*, 88, 1127–1156.

HELLWIG, C. AND N. WERQUIN (2024): "Using Consumption Data to Derive Optimal Income and Capital Tax Rates," Working Paper.

JULLIEN, B. (2000): "Participation constraints in adverse selection models," *Journal of Economic Theory*, 93, 1–47.

KANG, M. AND C. Z. ZHENG (2022): "Optimal Design for Redistributions among Endogenous Buyers and Sellers," *Economic Theory*, 75, 1141–1180.

KANG, Z. Y. (2023): "The Public Option and Optimal Redistribution," Working Paper.

——— (2024): "Optimal Indirect Regulation of Externalities," Working Paper.

KANG, Z. Y. AND M. WATT (2024a): "Optimal In-Kind Redistribution," Working Paper.

——— (2024b): "Optimal Redistribution Through Subsidies," Working Paper.

KAPLOW, L. (2008): "Optimal Policy with Heterogeneous Preferences," *The BE Journal of Economic Analysis & Policy*, 8, 1–30.

——— (2011): *The Theory of Taxation and Public Economics*, Princeton University Press.

KLEVEN, H. J., C. T. KREINER, AND E. SAEZ (2009): "The optimal income taxation of couples," *Econometrica*, 77, 537–560.

LI, Y. (2021): "Mechanism design with financially constrained agents and costly verification," *Theoretical Economics*, 16, 1139–1194.

MIRRLEES, J. A. (1976): "Optimal tax theory: A synthesis," *Journal of Public Economics*, 6, 327–358.

——— (1986): "The theory of optimal taxation," *Handbook of Mathematical Economics*, 3, 1197–1249.

MOSER, C. AND P. OLEA DE SOUZA E SILVA (2019): "Optimal paternalistic savings policies," Columbia Business School Research Paper.

MYERSON, R. B. (1981): "Optimal auction design," *Mathematics of Operations Research*, 6, 58–73.

NICHOLS, A. L. AND R. J. ZECKHAUSER (1982): "Targeting transfers through restrictions on recipients," *American Economic Review*, 72, 372–377.

PAI, M. AND P. STRACK (2024): "Taxing Externalities Without Hurting the Poor," Working Paper.

PIKETTY, T. AND E. SAEZ (2013): "A theory of optimal inheritance taxation," *Econometrica*, 81, 1851–1886.

ROCHET, J.-C. AND P. CHONÉ (1998): "Ironing, sweeping, and multidimensional screening," *Econometrica*, 66, 783–826.

SAEZ, E. (2002): "The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes," *Journal of Public Economics*, 83, 217–230.

SAEZ, E. AND S. STANTCHEVA (2018): "A simpler theory of optimal capital taxation," *Journal of Public Economics*, 162, 120–142.

SAXENA, R. R., A. SCHVARTZMAN, AND S. M. WEINBERG (2018): "The menu complexity of 'one-and-a-half-dimensional' mechanism design," in *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, 2026–2035.

SCHEUER, F. AND J. SLEMROD (2020): "Taxing our wealth," National Bureau of Economic Research Working Paper.

SCHULZ, K. (2021): "Redistribution of return inequality," Working Paper.

SPENCE, M. (1977): "Nonlinear prices and welfare," *Journal of Public Economics*, 8, 1–18.

SPIRITUS, K., E. LEHMANN, S. RENES, AND F. T. ZOUTMAN (2025): "Optimal taxation with multiple incomes and types," *Theoretical Economics*, 20, 543–581.

WEITZMAN, M. L. (1977): "Is the price system or rationing more effective in getting a commodity to those who need it most?" *Bell Journal of Economics*, 8, 517–524.

YANG, F. (2022): "Costly Multidimensional Screening," Working Paper.

A Proofs

A.1 Proof of Theorem 1

Let $(c^\star, x^\star, z^\star)$ be an optimal mechanism for the planner; then, $(c^\star, x^\star, z^\star)$ maximizes the Lagrangian, for some Lagrange multiplier $\alpha > 0$ on the resource constraint,[19]

$$\int \lambda(\theta)\, U(c(\theta), x(\theta), z(\theta)), \theta)\, dF(\theta) + \alpha \int [z(\theta) - c(\theta) - k \cdot x(\theta)]\, dF(\theta) \tag{14}$$

over all incentive-compatible mechanisms:

$$U(c(\theta), x(\theta), z(\theta)), \theta) \geq U(c(\theta'), x(\theta'), z(\theta')), \theta), \quad \forall \theta, \theta' \in \Theta. \tag{15}$$

Suppose that $(c^\star, x^\star, z^\star)$ has a distorted allocation of goods for a positive measure of types. We will find an incentive-compatible mechanism that strictly increases the value of the Lagrangian, which will contradict the optimality of $(c^\star, x^\star, z^\star)$.

Let $L^\star$ be the value of the Lagrangian (14) achieved by $(c^\star, x^\star, z^\star)$. We will construct a mechanism that improves this value in two steps. First, we consider a relaxed problem in which some incentive-compatibility constraints are dropped, and argue that a solution to the relaxed problem achieves a value of (14) strictly above $L^\star$. Second, we construct an implementation of that relaxed solution that satisfies all the incentive-compatibility constraints.

Consider a relaxed problem in which the planner can observe the taste type, that is, incentive-compatibility constraints only apply to the ability type:

$$U((c(t,a), x(t,a), z(t,a)), t, a) \geq U((c(t,a'), x(t,a'), z(t,a')), t, a), \quad \forall t \in \Theta_t, a, a' \in \Theta_a. \tag{16}$$

[19] Under Assumption A1, the Lagrange multiplier α must be 1 if strong duality holds.

Intuitively, when the planer can observe the taste type t, the standard Atkinson-Stiglitz theorem applies conditional on every t: Distortions in goods markets are redundant.[20] Formally, we rely on the results from Doligalski et al. (2025) that extend the logic of the Atkinson-Stiglitz theorem to general mechanism-design problems in which some decisions are "incentive-separable." Under the relaxed incentive constraint (16), consumption decisions are incentive-separable; thus, by Lemma 1 and Lemma 2 in Doligalski et al. (2025), any mechanism can be weakly improved upon by a mechanism that *(i)* achieves the same utilities for all agents, *(ii)* achieves higher revenue $\int [z(\theta) - c(\theta) - k \cdot x(\theta)]\, dF(\theta)$ (strictly, if the original mechanism distorted the goods markets), and *(iii)* can be implemented by letting agents purchase goods at their marginal costs subject to type-dependent budgets $m(\theta)$.[21] It is thus without loss of generality to restrict attention to mechanisms with these properties when solving the relaxed problem. Let

$$v^\star(t) = \max_x \{v(x,t) - k \cdot x\}$$

denote the indirect utility from the efficient allocation of goods x to an agent with taste type t (relying on Assumption A1 to simplify the definition). Define disposable income as $m = c + k \cdot x$. Using Assumptions A1–A3, the relaxed problem on the subset of mechanisms satisfying property *(iii)* takes the form:

$$\sup_{m,z} \iint \bar{\lambda}(a)\,[m(t,a) + v^\star(t) - w(z(t,a),a)] dF_t(t) dF_a(a) + \alpha \iint [z(t,a) - m(t,a)]\, dF_t(t) dF_a(a) \tag{17}$$

subject to

$$m(t,a) - w(z(t,a),a) \geq m(t,a') - w(z(t,a'),a), \quad \forall t \in \Theta_t, a, a' \in \Theta_a. \tag{18}$$

Moreover, we know that the value of this relaxed problem strictly exceeds $L^\star$, by properties *(i)* and *(ii)*, since $(c^\star, x^\star, z^\star)$ distorts goods choices. Note that the objective function (17) can be maximized point-wise in t and that the term $v^\star(t)$ does not affect the solution. As a result, we can find a feasible allocation $(\tilde{m}, \tilde{z})$ that does not depend on t (none of the remaining terms in the objective or constraints depend on t) and achieves a value $L > L^\star$.[22]

[20] Note that we assumed additive separability of utility functions which is stronger than weak separability required for the Atkinson-Stiglitz theorem.

[21] Doligalski et al. (2025) make an additional assumption that the utility each agent receives from the initial mechanism is higher than the utility from consuming nothing—this is needed for consumer duality to hold; we do not need that assumption because of Assumption A1 which is sufficient for consumer duality.

[22] We could take $(\tilde{m}, \tilde{z})$ to be an optimal mechanism for the relaxed problem if not for the fact that we have not guaranteed existence of solutions to the relaxed problem; for our purposes, it suffices to take a feasible mechanism that achieves a higher value of the objective than the original mechanism $(c^\star, x^\star, z^\star)$.

This finishes the first step.

Under the mechanism $(\tilde{m},\tilde{z})$ that we constructed, agents report their ability a, and are recommended to choose earnings $\tilde{z}(a)$ that lead to a disposable income $\tilde{m}(a)$ (by the taxation principle, this step can be achieved by imposing a nonlinear income tax on z); then, they spend their disposable income $\tilde{m}$ optimally on consumption of c and x priced at marginal costs. Denote the resulting allocation rule by (c,x,z). In the second step, we prove that the direct mechanism (c,x,z) satisfies all incentive constraints (15); since by construction (c,x,z) achieves a higher value of the Lagrangian than $(c^\star,x^\star,z^\star)$, this will finish the proof.[23]

Using Assumption A1, we have, for all θ, $\theta' \in \Theta$,

$$\begin{aligned} U(c(\theta),x(\theta),z(\theta)),\theta) &= c(t,a)+v(x(t,a),t)-w(z(t,a),a)=\tilde{m}(a)-w(\tilde{z}(a),a)+v^\star(t) \\ &\geq \tilde{m}(a')-w(\tilde{z}(a'),a)+v(x(t',a'),t)-k\cdot x(t',a') \\ &= c(a',t')+v(x(t',a'),t)-w(\tilde{z}(a'),a)=U(c(\theta'),x(\theta'),z(\theta')),\theta), \end{aligned}$$

where the key inequality follows from two observations. First, $\tilde{m}(a)-w(\tilde{z}(a),a)\geq \tilde{m}(a')-w(\tilde{z}(a'),a)$ by inequality (18); second, $v^\star(t)\geq v(x,t)-k\cdot x$ for any x by definition of $v^\star(t)$.

A.2 Proof of Lemma 1

It is without loss of generality to assume that $z_l(t)=0$ (given the welfare objective function and the fact that $l=0$). Thus, we only have to solve for the earnings choice of high-ability agents. We will prove that it is optimal to choose $z_h(t)=\bar{z}$ for all t. Suppose that it is not the case. Then, we can adjust all high types' allocations so that their $z_h(t)$ increases to $\bar{z}$ and their $c_h(t)$ increases just enough to make their overall utility unchanged. This adjustment does not affect the objective function and relaxes the resource constraint (as well as the IC constraints of low-ability types). Since the relaxation of the resource constraint is strict, and increasing the lump-sum payment increases social welfare, it is always strictly preferred to set $z_h(t)\equiv\bar{z}$.

A.3 Proof of Proposition 2

Consider the incentive constraints (7). Since all high-ability agents work (by Lemma 1), low-ability agents cannot mimic the high-ability agents. By standard arguments, their incentive constraint can be represented as a monotonicity constraint on the allocation $x_l(t)$

[23] Note that the proof already establishes that any optimal mechanism can be decentralized via a (possibly nonlinear) income tax and a competitive goods market in which all goods are priced at marginal costs.

and an integral condition pinning down consumption $c_l(t)$ (up to a lump-sum payment). Letting $T_l := c_l(0)$ denote the lump-sum payment to low-ability agents, we have:

$$x_l(t) \text{ is non-decreasing, } c_l(t) + t x_l(t) = T_l + \int_0^t x_l(\tau) d\tau.$$

The constraint $c_l(t) \geq \underline{c}$ can only bind for the highest taste type $t = \bar{t}$, since the above representation implies that $c_l(t)$ is non-increasing in t. Therefore, the constraint $c_l(t) \geq \underline{c}$ for all $t \in [0, \bar{t}]$ is equivalent to requiring

$$T_l + \int_0^{\bar{t}} x_l(t) dt - \bar{t} x_l(\bar{t}) \geq \underline{c}.$$

The incentive constraint for high-ability agents preventing them from misreporting their taste type alone leads to a similar representation (let $T_h := c_h(0)$):

$$x_h(t) \text{ is non-decreasing, } c_h(t) + t x_h(t) = T_h + \int_0^t x_h(\tau) d\tau,\ T_h + \int_0^{\bar{t}} x_h(\tau) d\tau - \bar{t} x_h(\bar{t}) \geq \underline{c}.$$

However, we must additionally satisfy the incentive constraint that high-ability agents do not want to mimic one of the low-ability types, which can now be represented as

$$T_h + \int_0^t x_h(\tau) d\tau - \frac{\bar{z}}{h} \geq T_l + \int_0^{t'} x_l(\tau) d\tau + (t - t') x_l(t'), \quad \forall t, t' \in [0, \bar{t}].$$

Note that

$$\int_0^{t'} x_l(\tau) d\tau + (t - t') x_l(t') \leq \int_0^t x_l(\tau) d\tau$$

by the monotonicity of $x_l(t)$ in t, and hence—conditional on misreporting the ability type—it is optimal to report the taste type truthfully. Thus, the constraint simplifies to

$$T_h + \int_0^t x_h(\tau) d\tau - \frac{\bar{z}}{h} \geq T_l + \int_0^t x_l(\tau) d\tau, \forall t. \tag{19}$$

Next, using the above formulas for $c_a(t)$, and after a few standard transformations (integration by part), we can rewrite the objective function (up to a term that is constant in the remaining choice variables) as

$$\sum_{a \in \{l, h\}} \mu_a \left(\bar{\lambda}_a T_a + \mu_a \int \Lambda_a(t) \gamma_a(t) x_a(t) dF_a(t) \right),$$

where $\gamma_a(t) = (1 - F_a(t))/f_a(t)$ is the inverse hazard rate and $\Lambda_a(t) = \mathbb{E}\left[\lambda_a(\tilde{t}) | \tilde{t} \geq t, a\right]$ is the average welfare weight on types above t, conditional on ability a. Similarly, the resource constraint (8) can be rewritten as

$$\mu_h \bar{z} \geq G + \sum_{a \in \{l,h\}} \mu_a \left(T_a + \int_0^{\bar{t}} (k - J_a(t))\, x_a(t) dF_a(t) \right),$$

where $J_a(t) = t - \gamma_a(t)$ is the virtual surplus function, conditional on ability a.

We reparameterize the problem by denoting $T = T_l$ and $\Delta T = T_h - \frac{\bar{z}}{h} - T$. That is, T is the lump-sum payment to all agents, and ΔT is the additional monetary payment that high-ability agents receive on top of the lump-sum transfer and the compensation for disutility of labor. By the incentive constraint (19), $\Delta T \geq 0$. Intuitively, when $\Delta T = 0$, the post-tax wage received by high-ability agents who work is just enough to offset the disutility from labor that they incur (corresponding to case *(a)* in point 4 of Proposition 2); when $\Delta T > 0$, high-ability agents enjoy a strictly positive surplus from working (corresponding to case *(b)* in point 4 of Proposition 2).

We summarize the progress made so far by restating the full problem as

$$\max_{x_h(t), x_l(t), T, \Delta T \geq 0} T + \mu_h \bar{\lambda}_h \Delta T + \sum_{a \in \{l,h\}} \mu_a \int_0^{\bar{t}} \Lambda_a(t) \gamma_a(t) x_a(t) dF_a(t),$$

subject to

$$x_l(t) \text{ is non-decreasing, } T \geq \underline{c} + \bar{t} x_l(\bar{t}) - \int_0^{\bar{t}} x_l(\tau) d\tau,$$

$$x_h(t) \text{ is non-decreasing, } T + \frac{\bar{z}}{h} + \Delta T \geq \underline{c} + \bar{t} x_h(\bar{t}) - \int_0^{\bar{t}} x_h(\tau) d\tau,$$

$$\Delta T + \int_0^t x_h(\tau) d\tau \geq \int_0^t x_l(\tau) d\tau, \forall t,$$

$$\mu_h \bar{z} \left(1 - \frac{1}{h}\right) \geq G + T + \mu_h \Delta T + \sum_{a \in \{l,h\}} \mu_a \left(\int_0^{\bar{t}} (k - J_a(t))\, x_a(t) dF_a(t) \right).$$

Let us parameterize the problem by imposing an additional constraint $x_l(\bar{t}) \leq \bar{x}_l$ and $x_h(\bar{t}) \leq \bar{x}_h$, and optimizing separately over $\bar{x}_l$ and $\bar{x}_h$. Intuitively, the need to bound the allocation rule from above by a number less than 1 may come from the subsistence constraint. Note that, as long as constraint (19) and the subsistence constraint hold for the low-ability agents, we have

$$T + \frac{\bar{z}}{h} + \Delta T \geq \underline{c} + \bar{t} x_h(\bar{t}) - \int_0^{\bar{t}} x_h(\tau) d\tau + \frac{\bar{z}}{h} + \bar{t}(x_l(\bar{t}) - x_h(\bar{t})).$$

It follows that we can increase $x_h(\bar{t})$ to be at least at the level of $x_l(\bar{t})$ while preserving the subsistence constraint for the high-ability agents; thus, it is without loss of generality to assume that $\bar{x}_h \geq \bar{x}_l$. This argument also establishes that the subsistence constraint is slack for high-ability agents if $\bar{z}/h \geq \bar{t}$.

We solve the problem by introducing two Lagrange multipliers, $\eta_l \geq 0$ and $\eta_h \geq 0$, on the subsistence constraints for the low- and high-ability agents, respectively. The resource constraint must hold with equality at the optimal mechanism, which allows us to substitute T in the objective function. The Lagrangian—fixing $\bar{x}_l$ and $\bar{x}_h$—is then maximized over non-decreasing $x_l(t)$ and $x_h(t)$, as well as $\Delta T \geq 0$,

$$\max_{x_l(t)\leq \bar{x}_l,\, x_h(t)\leq \bar{x}_h,\, \Delta T\geq 0} \sum_{a\in\{l,h\}} \mu_a \int_0^{\bar{t}} \left[\Lambda_a(t)\gamma_a(t) + \frac{\eta_a}{\mu_a f_a(t)} + (1+\eta_l+\eta_h)(J_a(t)-k)\right] x_a(t)dF_a(t)$$

$$+ \left(\mu_h \bar{\lambda}_h + \eta_h - (1+\eta_l+\eta_h)\mu_h\right)\Delta T,$$

subject to a single constraint

$$\Delta T + \int_0^t x_h(\tau)d\tau \geq \int_0^t x_l(\tau)d\tau, \forall t.$$

First, we will derive the optimal $x_h(t)$ and ΔT holding fixed $x_l(t)$. Let

$$\phi_h(t) := \left(\Lambda_h(t)\gamma_h(t) + \frac{\eta_h}{\mu_h f_h(t)} + (1+\eta_l+\eta_h)(J_h(t)-k)\right)\mu_h f_h(t),$$

$$\psi := \mu_h \bar{\lambda}_h + \eta_h - (1+\eta_l+\eta_h)\mu_h,$$

so that this auxiliary problem can be written succinctly as

$$\max_{x_h(t)\leq \bar{x}_h,\, \Delta T\geq 0} \int_0^{\bar{t}} \phi_h(t)x_h(t)dt + \psi\Delta T$$

subject to

$$\Delta T + \int_0^t x_h(\tau)d\tau \geq \int_0^t x_l(\tau)d\tau, \forall t.$$

Note that we must have $\psi \leq 0$ as otherwise the problem would not have a solution (and the problem does admit a solution by standard arguments).

The above problem is a linear mechanism design problem with a type-dependent outside option constraint pinned down by the allocation rule for the low-ability agents. Such a problem can be solved using existing techniques.

Lemma 2 (Dworczak and Muir (2024)). *Define*

$$\Phi_h(t) = \int_t^{\bar{t}} \phi_h(\tau)d\tau \text{ and } \bar{\Phi}_h(t) = \mathrm{co}(\Phi_h)(t),\ \bar{\phi}_h(t) = -\bar{\Phi}_h'(t),$$

where $\mathrm{co}(\cdot)$ *is the concave closure of a function. Let* t_0 *be defined as the smallest solution to* $\bar{\phi}_h(t_0) = \psi$ *(*$t_0 = 0$ *if* $\bar{\phi}_h(t) > \psi$ *for all* t*), and let* t_1 *be defined as the largest solution to* $\bar{\phi}_h(t_1) = 0$ *(*$t_1 = \bar{t}$ *if* $\bar{\phi}_h(t) < 0$ *for all* t*). (Note that* $t_0 \leq t_1$ *because* $\bar{\phi}_h(t)$ *is non-decreasing.) Then,*

$$\max_{x_h(t)\leq \bar{x}_h,\, \Delta T \geq 0} \left\{ \int_0^{\bar{t}} \phi_h(t) x_h(t) dt + \psi \Delta T \right\} = \int_{t_0}^{t_1} \bar{\phi}_h(t) x_l(t) dt + \bar{x}_h \int_{t_1}^{\bar{t}} \bar{\phi}_h(t) dt + \psi \int_0^{t_0} x_l(t) dt.$$

Moreover, the optimal solution is given by

$$\Delta T^\star = \int_0^{t_0} x_l(t) dt,$$

$$x_h^\star(t) = \begin{cases} 0 & t \leq t_0, \\ x_l(t) & t \in [a,b] \text{ for every maximal } [a,b] \text{ such that } \Phi_h \equiv \bar{\Phi}_h \text{ on } [a,b], \\ \frac{\int_a^b x_l(\tau) d\tau}{b-a} & t \in (a,b) \text{ for every maximal } (a,b) \text{ such that } \Phi_h < \bar{\Phi}_h \text{ on } (a,b), \\ \bar{x}_h & t \geq t_1. \end{cases}$$

Explaining Lemma 2 is beyond the scope of this paper.[24] The important take-aways for our purposes are that the problem of choosing the optimal $x_h(t)$ for a fixed $x_l(t)$ admits a closed-form solution characterized by two cutoffs, t_0 and t_1, and (possibly) a number of ironing intervals. Ignoring the possibility of ironing (formally, ironing is not needed if $\phi_h(t)$ is monotone), the intuition for the cutoffs t_0 and $t_1$is as follows. The planner chooses an allocation rule for high-ability agents to maximize welfare subject to delivering a certain minimal level of utility to high-ability agents, where the lower bound on utility comes from the possibility of mimicking a low-ability type. It is better to give a cash transfer $\Delta T^\star$ to types $t \leq t_0$ than to let these types consume the allocation for the low-ability agents. Note that $\Delta T^\star > 0$ only if low-ability agents of taste type below t_0 consume the good: $x_l(t) > 0$ for some $t < t_0$. These considerations (after endogenizing $x_l(t)$) ultimately determine whether or not high-ability agents receive strictly positive surplus from working. For types $t \in [t_0, t_1]$, it is optimal to satisfy the constraint by letting them consume what the low-ability agents with analogous taste types consume (again, this is further complicated if ironing is needed). Finally, types $t \geq t_1$ should consume the maximal amount $\bar{x}_h$ regardless of the outside option (here, we rely on the fact that $\bar{x}_h \geq \bar{x}_l \geq x_l(t)$ for all t).

Note that the definition of t_0 and t_1 does not depend on $x_l(t)$. The closed-form expression for the maximized objective function thus allows us to maximize over $x_l(t)$ in the next step.

[24] The reader is referred to Dworczak and Muir (2024) for a discussion.

Let

$$\phi_l(t) := \left[\Lambda_l(t)\gamma_l(t) + \frac{\eta_l}{\mu_l f_l(t)} + (1+\eta_l+\eta_h)(J_l(t)-k)\right]\mu_l f_l(t).$$

Then, the problem of maximizing over $x_l(t)$ (assuming that $x_h(t)$ and ΔT are chosen optimally for any $x_l(t)$, as described by Lemma 2), becomes (fixing $\bar{x}_l$)

$$\max_{x_l(r)\leq \bar{x}_l} \int_0^{\bar{t}} \phi_l(t)x_l(t)dt + \int_{t_0}^{t_1} \bar{\phi}_h(t)x_l(t)dt + \bar{x}_h \int_{t_1}^{\bar{t}} \bar{\phi}_h(t)dt + \psi \int_0^{t_0} x_l(t)dt.$$

This problem is linear in $x_l(t)$ with no additional constraints (other than monotonicity of $x_l(t)$), so there exists an optimal solution that takes the form $x_l(t) = \bar{x}_l \mathbf{1}_{\{t\geq t_l\}}$ for some t_l (formally, these are the extreme points of the set of non-decreasing functions on $[0, \bar{t}]$ bounded below by 0 and above by $\bar{x}_l$.)

Finally, consider maximizing the Lagrangian over $\bar{x}_h$ and $\bar{x}_l$ (at the optimal solution, without loss of generality, $x_l(\bar{t}) = \bar{x}_l$ and $x_h(\bar{t}) = \bar{x}_h$):

$$\max_{\bar{x}_l,\bar{x}_h} \left\{ \bar{x}_l \int_{t_l}^{\bar{t}} \phi_l(t)dt + \bar{x}_l \int_{t_0}^{t_1} \bar{\phi}_h(t)dt + \bar{x}_h \int_{t_1}^{\bar{t}} \bar{\phi}_h(t)dt + \psi \bar{x}_l(t_0 - t_l)_+ - \eta_l \bar{x}_l - \eta_h \bar{x}_h \right\}. \tag{20}$$

The problem is linear. There are two possibilities. First, the subsistence constraint could be slack for both ability types, in which case $\eta_a = 0$ and $\bar{x}_a = 1$, for $a \in \{l, h\}$. Moreover, if the subsistence constraint is slack for low-ability agents, then it is also slack for high-ability agents. Second, the subsistence constraint could bind (for low-ability types, or both types). In that case, η_a is set so that the coefficient on $\bar{x}_a$ in the Lagrangian (20) is zero; this allows us to choose $\bar{x}_a$ to satisfy the subsistence constraint with equality (by assumption, we restricted attention to cases in which we can satisfy the subsistence constraint when agents do not consume the good, so there is some intermediate level of consumption that satisfies the constraint with equality). In either case, we conclude that the solution described above is a solution to the original problem for *some* choice of η_l and η_h.

We are now ready to finish the proof of Proposition 2. Part 1 follows from the fact that the optimal $x_l(t)$ is a cutoff allocation rule (we set $q_l = \bar{x}_l$). Part 2 follows from the following observation: Since the optimal $x_l(t)$ is a cutoff allocation rule, the optimal allocation rule $x_h(t)$—as predicted by Lemma 2—can take on at most one value, which we call q_i, other than 0, $q_l = \bar{x}_l$, and $q_h := \bar{x}_h$. Specifically, $q_i \in (0, q_l)$ if and only if $t_l \in (a, b) \subseteq (t_0, t_1)$ for some maximal interval (a, b) such that $\Phi_h < \bar{\Phi}_h$ on (a, b): then, $q_i = q_l(b - t_l)/(b - a)$. In this sense, q_i is a result of ironing that is required when the objective function $\phi_h(t)$ is not monotone (so that Φ_h lies below its concave closure $\bar{\Phi}_h$ on some interval). Part 3 follows from the analysis of Lagrange multipliers η_l and η_h above. Finally, to prove part 4, let us separately analyze the form of the optimal mechanism when *(a)* $t_l \geq t_0$, *(b)* when $t_l < t_0$.

In case *(a)*, by Lemma 2, $\Delta T^\star = 0$. This means that $T_h = T_l + \bar{z}/h$ or $c_h(0) - \bar{z}/h = c_l(0)$. By incentive compatibility, $p_l(q_l) = t_l$ since type t_l is the cutoff type consuming quality q_l. To determine the average prices paid by high-ability type, we consider three subcases:

(i) If $t_1 \leq t_l$, then $x_h(t) = \bar{x}_h(t)\mathbf{1}_{\{t \geq t_1\}}$, so $\mathrm{Im}(x_h) = \{0, \bar{x}_h\}$ and $p_h(q_h) = t_1 \leq t_l = p_l(q_l)$. If $q_l \in \mathrm{Im}(x_h)$, then it follows that $q_l = q_h$ and hence $p_h(q_l) \leq p_l(q_l)$.

(ii) If $t_1 > t_l$ and no ironing is required (q_i is not offered), then it follows that $x_h(t) = x_l(t) = \mathbf{1}_{\{t \geq t_l\}}$ for all $t \in (t_0, t_1)$, and hence $p_l(q_l) = p_h(q_l) = t_l$.

(iii) If $t_1 > t_l$ but ironing is required, then we have $x_h(t) = q_i$ for $t \in [a, b)$, $x_h(t) = q_l$ for $t \in [b, t_1)$, and $x_h(t) = \bar{x}_h$ for $t \geq t_1$, for some $t_0 \leq a \leq t_l \leq b \leq t_1$. In this case, quality q_i must be offered at the average price a, and type b must be indifferent between buying quality q_i at a per-unit price of a, or buying q_l at a per-unit price of $p_h(q_l)$:

$$(b - a)q_i = (b - p_h(q_l))q_l \iff t_l = p_h(q_l).$$

We conclude that in all subcases when q_l is offered to high-ability agents ($q_l \in \mathrm{Im}(x_h)$), we have $p_h(q_l) \leq p_l(q_l)$.

In case *(b)*, $\Delta T^\star = \bar{x}_l(t_0 - t_l) > 0$, by Lemma 2. A further consequence of the lemma is that—since x_l is constant in $[t_0, t_1]$—x_h must be equal to $\bar{x}_l$ on $[t_0, t_1]$, and hence $\mathrm{Im}(x_h) \subseteq \{0, q_l, q_h\}$. Since $t_l < t_0$, high-ability agents must face a higher per-unit price for consuming q_l.

Cases *(a)* and *(b)* above thus correspond to the analogous cases in Proposition 2, which finishes its proof.

A.4 Proof of Theorem 2

The proof relies on Proposition 2 and its proof found in Appendix A.3.

We will construct a solution in which the subsistence constraint is slack for high-ability agents and binds for low-ability agents (the condition on the aggregate resources in the statement of the theorem ensures that we will be able to verify that property). Using the notation from the proof of Proposition 2, we set $\eta_l = \eta$ and $\eta_h = 0$.

First, we prove a technical lemma showing that under our regularity conditions, ironing is not required in the optimal allocation rule for high-ability agents.

Lemma 3. *Under the assumptions of Theorem 2, the solution described by Lemma 2 does not involve ironing ($\phi_h = \bar{\phi}_h$ in the relevant range).*

Proof. Under the current assumptions, we have

$$\phi_h(t) \equiv \left((t-k)(1+\eta)\mu_h + \psi\frac{1-F(t)}{f(t)}\right)f(t),$$

$$\psi = -\mu_h(1+\eta-\lambda_h),$$

where λ_h is the (constant) welfare weight on high-ability agents. We will first show that $\phi_h(t)$ is non-decreasing over $[t_0, t_1]$, so that ironing is not required in the optimal mechanism, where t_0 and t_1 are defined by

$$(t_1-k)(1+\eta)\mu_h + \psi\frac{1-F(t_1)}{f(t_1)} = 0,$$

$$(t_0-k)(1+\eta)\mu_h - \psi\frac{F(t_0)}{f(t_0)} = 0.$$

(We will later verify that these definitions coincide with the definition in Lemma 2.) Note that our regularity assumptions imply that t_0 and t_1 are uniquely defined (recall that $\psi \le 0$).

We need to show that, for $t \in (t_0, t_1)$,

$$((1+\eta)\mu_h + \psi\gamma'(t))\,f(t) + ((t-k)(1+\eta)\mu_h + \psi\gamma(t))\,f'(t) > 0.$$

We know that, in the relevant range,

$$\psi\frac{F(t)}{f(t)} < (t-k)(1+\eta)\mu_h < -\psi\frac{1-F(t)}{f(t)}.$$

When $f'(t) \ge 0$, we have

$$((1+\eta)\mu_h + \psi\gamma'(t))\,f(t) + ((t-k)(1+\eta)\mu_h + \psi\gamma(t))\,f'(t) > \left(\psi\frac{F(t)}{f(t)} + \psi\frac{1-F(t)}{f(t)}\right)f'(t) \ge 0.$$

When $f'(t) < 0$, we have

$$((1+\eta)\mu_h + \psi\gamma'(t))\,f(t) + ((t-k)(1+\eta)\mu_h + \psi\gamma(t))\,f'(t) > \left(-\psi\frac{1-F(t)}{f(t)} + \psi\frac{1-F(t)}{f(t)}\right)f'(t) = 0.$$

This shows that $\phi_h(t)$ is non-decreasing over $[t_0, t_1]$.

Next, notice that $\phi_h(t)$ crosses zero once from below, and hence $\phi_h(t) \ge 0$ for all $t \ge t_1$.

Similarly, we want to show that $\phi_h(t) \le \psi$ for all $t \le t_0$. For $t \le t_0$,we have

$$\left((t-k)(1+\eta)\mu_h + \psi\frac{1-F(t)}{f(t)}\right)f(t) = \underbrace{\left((t-k)(1+\eta)\mu_h - \psi\frac{F(t)}{f(t)}\right)}_{\le 0} f(t) + \psi \le \psi.$$

We have thus shown that $\phi_h(t) = \bar{\phi}_h(t)$ over $[t_0, t_1]$, and moreover that t_0 and t_1 defined above coincide with those defined in Lemma 2. It follows that no ironing is needed: For a fixed x_l, the optimal allocation rule for high-ability agents is given by

$$x_h^\star(t) = \begin{cases} 0 & t < t_0, \\ x_l(t) & t \in [t_0, t_1), \\ \bar{x}_h & t \ge t_1. \end{cases}$$

□

Combining Lemma 3 with Proposition 2, we conclude that the optimal solution is parameterized by: t_l, t_0, t_1, $\bar{x}_l$ (which we keep fixed for now), and $\bar{x}_h$ (which we conjecture will be equal to 1). Note that as long as $t_0 < t_l$, the value of t_0 does not affect the mechanism (since $x_l(t) = 0$ for $t \le t_l$). Thus, it is without loss of generality to assume that $t_0 \ge t_l$. Under that assumption, we have $\Delta T^\star = \bar{x}_l(t_0 - t_l)$. The resulting Lagrangian—which is maximized over t_l, t_0, and t_1—takes the form:

$$\begin{aligned} &-(1+\eta)\left[\mu_h\bar{x}_l\int_{t_0}^{t_1}(k-J(t))\,dF(t) + \mu_h\int_{t_1}^{\bar{t}}(k-J(t))\,dF(t) + \mu_l\bar{x}_l\int_{t_l}^{\bar{t}}(k-J(t))\,dF(t)\right] - \eta\bar{x}_l t_l \\ &-(1-\lambda_h+\eta)\mu_h\bar{x}_l(t_0-t_l) + \mu_l\lambda_l\bar{x}_l\int_{t_l}^{\bar{t}}\gamma(t)dF(t) + \mu_h\lambda_h\bar{x}_l\int_{t_0}^{t_1}\gamma(t)dF(t) + \mu_h\lambda_h\int_{t_1}^{\bar{t}}\gamma(t)dF(t). \end{aligned} \tag{21}$$

We will argue that $t_0 = t_l$ in the optimal mechanism. Since we know that $t_0 \ge t_l$, towards a contradiction, suppose that $t_0 > t_l$; then, the first-order conditions for optimal t_0 and t_l must hold, which would require (after some transformations, and in particular substituting $\lambda_l = (1-\mu_h\lambda_h)/\mu_l$):

$$(1+\eta)(t_0-k)f(t_0) + (1+\eta-\lambda_h)F(t_0) = 0,$$

$$(1+\eta)(t_l-k)f(t_l) + (1+\eta-\lambda_l)F(t_l) = 0.$$

The first condition states that $\phi_h(t_0) = \psi$, and since t_0 is the smallest solution to this equation, we know that $\phi_h(t) < \psi$ for all $t < t_0$, and thus in particular,

$$(1+\eta)(t_l - k)f(t_l) + (1+\eta-\lambda_h)F(t_l) < 0.$$

But this clearly contradicts the second condition (since $\lambda_h \leq \lambda_l$).

Thus, we have proven that $t_0 = t_l$. In particular, $\Delta T^\star = 0$, so we are in case *(a)* in part 4 of Proposition 2. The solution is characterized (up to pinning down $\bar{x}_l$ and confirming that $\bar{x}_h = 1$) by the first-order conditions for optimal t_0 and t_1:

$$[(1+\eta)(k - J(t_0)) - h(t_0)]\, f(t_0) = \eta,$$

$$(1+\eta)(k - J(t_1)) = \lambda_h h(t_1).$$

We can rewrite the FOCs as

$$t_0 = k - \frac{\eta}{1+\eta}\frac{F(t_0)}{f(t_0)},$$

$$t_1 = k + \left(1 - \frac{\lambda_h}{1+\eta}\right)\frac{1-F(t_1)}{f(t_1)}.$$

Note that, as long as $\eta > 0$, we have $t_0 < k < t_1$. In the indirect implementation, the per-unit price $p_l = t_l = t_0$ for quality $\bar{x}_l = q_l$ is thus below marginal cost. The total price $p_h(1)$ for the good with quality 1 must make type t_1 indifferent:

$$t_1 - p_h(1) = q_l(t_1 - p_l) \implies p_h(1) = p_l q_l + t_1(1 - q_l).$$

This verifies point *3* of Theorem 2 if we define $p_h = t_1$ (note that p_h is the revenue-maximizing price if $\lambda_h = 0$). Note that the piece-wise linear price schedule from Theorem 2 implements the same quality choices as the mechanism described here (because of linearity of agents' utilities and the fact that the subsistence constraint for low-ability agents becomes binding precisely when they consume a good with quality q_l). Point *1* of Theorem 2 follows from the fact that $\Delta T^\star = 0$ (the income tax makes high-ability agents indifferent between working or not). Finally, the binding subsistence constraint for low-ability agents implies that $c_l(0) = \underline{c} + p_l q_l$, verifying point *2*.

It remains to verify that *(i)* the subsistence constraint binds for low-ability agents (so that $\eta > 0$) and *(ii)* the subsistence constraint is slack for high-ability agents (which will verify our conjecture that $\eta_h = 0$ and $\bar{x}_h = 1$).

The resource constraint states that

$$\mu_h \bar{z}\left(1-\frac{1}{h}\right) = G + T - \mu_h(1-\bar{x}_l)(t_1-k)(1-F(t_1)) - \bar{x}_l(t_0-k)(1-F(t_0)).$$

Towards a contradiction, suppose that $\eta = 0$. Then, it is optimal to set $\bar{x}_l = 1$, $t_0 = t_1 = k$, and the resource constraint becomes

$$\mu_h \bar{z}\left(1-\frac{1}{h}\right) \geq G + T.$$

Since low-ability agents do not work but can afford to buy one unit of the good at price k, it must be that $T \geq \underline{c} + k$. Thus, we must have

$$\mu_h \bar{z}\left(1-\frac{1}{h}\right) \geq G + \underline{c} + k,$$

which is ruled out by the condition assumed in Theorem 2.

Finally, we make sure that in the solution we have constructed consumption of the high ability agents exceeds the subsistence level. We know that $T = \underline{c} + t_0\bar{x}_l$. Thus, it suffices to show that

$$\underline{c} + t_0\bar{x}_l + \frac{\bar{z}}{h} \geq \underline{c} + t_0\bar{x}_l + t_1(1-\bar{x}_l) \iff \frac{\bar{z}}{h} \geq t_1(1-\bar{x}_l).$$

A sufficient condition is that $\bar{z}/h \geq \bar{t}$, which is what we assumed.

A.5 Proof of Theorem 3 and Proposition 1

The proof relies on Proposition 2 and its proof found in Appendix A.3.

Since we assumed that

$$\phi_h(t) \equiv (t - k - (1-\Lambda_h(t))\gamma_h(t))\, f_h(t)$$

is non-decreasing whenever it is negative, it follows from Lemma 2 that no ironing is required to describe the optimal $x_h(t)$. Moreover, combining this observation with Proposition 2, we conclude that the allocation rule $x_l(t)$ takes the form $\mathbf{1}_{\{t \geq t_l\}}$, from which it follows that $x_h(t) = \mathbf{1}_{\{t \geq t_h\}}$, for some t_h. It remains to characterize t_h and t_l.

The optimization problem—based on the derivation in the proof of Proposition 2—becomes

$$\max_{t_h, t_l} \sum_{a \in \{l,h\}} \mu_a \int_{t_a}^{\bar{t}} (J_a(t) - k + \Lambda_a(t)\gamma_a(t))\, dF_a(t) - (1-\bar{\lambda}_h)\mu_h(t_h - t_l)_+.$$

The FOCs for an interior solution (in particular, when $t_h \neq t_l$) are

$$\text{FOC } t_h : -(1-\bar{\lambda}_h)\mathbf{1}_{\{t_h \geq t_l\}} - (t_h - k - (1-\Lambda_h(t_h))\gamma_h(t_h))\, f_h(t_h) = 0,$$

$$\text{FOC } t_l : \frac{\mu_h}{\mu_l}(1-\bar{\lambda}_h)\mathbf{1}_{\{t_h \geq t_l\}} - (t_l - k - (1-\Lambda_l(t_l))\gamma_l(t_l))\, f_l(t_l) = 0.$$

We argue that it is without loss of generality to assume that the optimal t_h and t_l satisfy $t_h \geq t_l$. By assumption, $t - k - (1-\Lambda_h(t))\gamma_h(t)$ can cross zero from below at most once. Conditional on $t_h < t_l$, the point (or interval) at which $t - k - (1-\Lambda_h(t))\gamma_h(t)$ crosses zero from below defines the optimal t_h (in case the crossing is an interval, we can without loss of generality take t_h to be the right end point of the interval, since every point in the interval is optimal). But then assumption (11) guarantees that

$$t - k - (1-\Lambda_l(t))\gamma_l(t) \geq t - k - (1-\Lambda_h(t))\gamma_h(t),$$

which implies that any t_l satisfying the FOC under the hypothesis $t_h < t_l$ must in fact be smaller than t_h.

We will consider the two cases, *(i)* $t_h = t_l$ and *(ii)* $t_h > t_l$, separately.

In case *(i)*, we immediately obtain that $\Delta T^\star = 0$ (which means that high-ability agents get no utility surplus from working or, equivalently, that income is taxed at the rate $1 - 1/h$ per unit of earnings) and that all agents face the same price p in the market. This price p must be equal to $t_h = t_l$. Since the same mechanism is offered to low- and high-ability agents, we can use the unconditional distribution F of taste types. Let us also denote by $\Lambda(p)$ the unconditional (over ability types) expectation of the welfare weight on agents with taste type above p. The FOC for that price p is

$$p - k - (1-\Lambda(p))\gamma(p) = 0,$$

which gives us the formula from point 1 in Theorem 3.

In case *(ii)*, we conclude that $\Delta T^\star > 0$, so that high-ability agents receive a strictly positive surplus from working. In this case, the two FOCs must hold, and thus (using the fact that $\bar{\lambda}_h \leq 1$)

$$t_h - k - (1-\Lambda_h(t_h))\gamma_h(t_h) \leq 0,$$

$$t_l - k - (1-\Lambda_l(t_l))\gamma_l(t_l) \geq 0.$$

This gives us the string of inequalities on the prices from point 2 in Theorem 3.

Finally, we prove Proposition 1. Under the additional assumptions we made, the first-order conditions described above are necessary and sufficient (under the convention that the equality becomes an inequality at the boundaries 0 or $\bar{t}$). Therefore, mechanism 2 is optimal for welfare parameters $(\bar{\lambda}_h, \bar{\lambda}_l)$ (that must satisfy $\mu_h \bar{\lambda}_h + \mu_l \bar{\lambda}_l = 1$) if there exists a solution to the system of equations

$$\mathcal{H}_h(\bar{\lambda}_h, t_h) \equiv \bar{\lambda}_h - 1 - (t_h - k - (1 - \Lambda_h(t_h))\gamma_h(t_h))\, f_h(t_h) = 0,$$
$$\mathcal{H}_l(\bar{\lambda}_l, t_l) \equiv \bar{\lambda}_l - 1 - (t_l - k - (1 - \Lambda_l(t_l))\gamma_l(t_l))\, f_l(t_l) = 0,$$

that satisfies $t_h > t_l$. Our goal is to show that if a solution exists for some $\bar{\lambda}_l$, then it must also exist for all lower $\bar{\lambda}_l$.

First, note that $\mathcal{H}_h(\bar{\lambda}_h, t_h)$ is non-increasing in t_h in the relevant range (since we assumed that its second term is non-decreasing whenever it is negative). Furthermore, $\mathcal{H}_h(\bar{\lambda}_h, t_h)$ is non-decreasing in $\bar{\lambda}_h$:

$$\frac{\partial \mathcal{H}_h(\bar{\lambda}_h, t_h)}{\partial \bar{\lambda}_h} = 1 - \int_{t_h}^{\bar{t}} \omega_h(t) dF_h(t) = \int_0^{t_h} \omega_h(t) dF_h(t) \geq 0.$$

Thus, when $\bar{\lambda}_l$ goes down, $\bar{\lambda}_h$ goes up (since $\mu_h \bar{\lambda}_h + \mu_l \bar{\lambda}_l = 1$), and the first-order condition can be satisfied either by keeping t_h constant or by increasing it, potentially all the way to $\bar{t}$ (at which point the first-order condition may hold as inequality).

Using analogous reasoning as before, we argue that the first-order condition $\mathcal{H}_l(\bar{\lambda}_l, t_l) = 0$ can be satisfied with lower $\bar{\lambda}_l$ by either keeping t_l constant or decreasing it. First, $\mathcal{H}_l(\bar{\lambda}_l, t_l)$ is non-increasing in t_l in the relevant range, which follows from the assumption made in the proposition. Second, $\mathcal{H}_l(\bar{\lambda}_l, t_l)$ is non-decreasing in $\bar{\lambda}_l$:

$$\frac{\partial \mathcal{H}_l(\bar{\lambda}_l, t_l)}{\partial \bar{\lambda}_l} = 1 - \int_{t_l}^{\bar{t}} \omega_l(t) dF_l(t) = \int_0^{t_l} \omega_l(t) dF_l(t) \geq 0.$$

Thus, when $\bar{\lambda}_l$ goes down, the first-order condition can be satisfied either by keeping t_l constant or by decreasing it, potentially all the way to 0. Therefore, we can find a solution to the system of first-order conditions that respects $t_h > t_l$.

We conclude that there exists a cutoff $\bar{\lambda}_l^0 \in [1, \infty]$ such that mechanism 2 is optimal if $\bar{\lambda}_l < \bar{\lambda}_l^0$ and mechanism 1 is optimal if $\bar{\lambda}_l > \bar{\lambda}_l^0$. The case $\bar{\lambda}_l^0 = \infty$ can be ruled out: When $\bar{\lambda}_l$ becomes sufficiently large, the second condition cannot hold and thus mechanism 1 must be optimal. On the other hand, $\bar{\lambda}_l^0 = 1$ can be ruled out when inequality (11) is strict for all interior t. As $\bar{\lambda}_l$ converges to 1, $\bar{\lambda}_h$ must also converge to 1, and inequality (11) then implies that the solution to the system of first-order conditions must satisfy $t_h > t_l$.

Online Appendix

In Online Appendix A we prove a version of Theorem 1 under weaker assumptions than in the main body. Online Appendix B describes the properties of Pareto efficient allocations in the simplified framework from Section 4. Finally, in Online Appendix C we characterize the optimal goods market distortions under a strictly concave utility from numeraire.

A Theorem 1 under weaker assumptions

In this appendix, we prove a version of Theorem 1 while weakening a number of assumptions made in the main body. Specifically, we do not identify any of the goods as a numeraire; we relax some restrictions on agents' utility functions (including quasi-linearity); we prove the existence of an improving mechanism with efficient allocation of goods; and our argument does not rely on the validity of the Lagrangian representation of the planner's problem (strong duality). The role of the extension is to show that the simplifying assumptions made in the main text do not play a major economic role in our overall analysis; the generalized setup supports the argument that knife-edge conditions appear to be needed to extend the Atkinson-Stiglitz conclusion to multidimensional settings.

A.1 Main result

An agent consumes a vector of L goods $x \in \mathcal{X} = \Pi_{i=1}^{L}[\underline{x}_i, \infty)$, where $\underline{x}_i \in \mathbb{R} \cup \{-\infty\}$ for all $i \in \{1, ..., L\}$, and earns $z \in \mathbb{R}_+$. The utility function is $U(x, z, t, a)$ which is continuous in (x, z) and locally non-satiated in x.

The following assumption relaxes the additive separability between consumption and earnings that we imposed in the main body of the paper. This "weak separability" assumption is required by the Atkinson-Stiglitz theorem even in one-dimensional case.

Assumption A0. *The ability type does not affect the marginal rate of substitution between goods: $U(x, z, t, a) \equiv \mathcal{U}(v(x, z, t), z, a)$, for some functions $\mathcal{U}$ and v, with $\mathcal{U}$ strictly increasing in the first argument.*

Denote the indirect utility function of type (t, a) with disposable income m, earnings z, and facing goods' prices equal to the marginal costs by

$$V(m, z, t, a) := \max_{x \in \mathcal{X}} \{U(x, z, t, a) : x \cdot k \leq m\}. \tag{22}$$

We assume that the solution to this problem always exists.

In the framework from Section 2 and under Assumption A1, we have $V(m,z,t,a) = m - w(z,a) + \max_x\{v(x,t) - k \cdot x\}$. The following assumption relaxes this property by only requiring that the taste type does not influence the indirect utility from disposable income and earnings (when goods are consumed efficiently):

Assumption A1′. *The indirect utility function is additively separable in allocation (m,z) and tastes t: $V(m,z,t,a) \equiv \mathcal{W}(m,z,a) + \mathcal{W}_0(t,a)$, for some functions $\mathcal{W}$ and $\mathcal{W}_0$.*

We note that—much like Assumption A1—Assumption A1′ plays a dual role. On the one hand, it implies that the taste type does not influence agents' *choices* over combinations of earnings z and disposable income m. This is an ordinal property that ensures that distorting consumption of goods does not allow the planner to relax incentive constraints in a useful way. On the other hand, combined with Assumption A2, Assumption A1′ implies that the planner does not have a motive to redistribute across taste types. This is a cardinal property of the utility functions. In particular, fixing agents' ordinal preferences, it may be possible to *select* a cardinal representation of these preferences that satisfies Assumption A1′; however, different cardinal representations generally correspond to different social preferences under welfare maximization (if welfare weights are fixed). Thus, Assumption A1′ is also a restriction on social preferences.[25] We return to this point later when we exploit various examples of preferences satisfying Assumption A1′.

In the setup without quasi-linear preferences we need to impose an additional restriction on the initial mechanism. Define a mechanism (x,z) to be *admissible* if it is feasible and satisfies $U(x(t,a),z(t,a),t,a) \geq U(\underline{x},z(t,a),t,a)$ for all $(t,a) \in \Theta$. Recall that $\underline{x}$ is the smallest possible consumption vector (for instance, it is common to have $\underline{x} = \mathbf{0}$). Admissibility is, hence, a very weak requirement that rules out mechanisms in which the planner spends resources on "bads" to make some agents worse off relative to consuming the smallest possible consumption bundle $\underline{x}$.

We can now state the generalization of Theorem 1.

Theorem 1′. *Suppose that Assumptions A0, A1′, A2, and A3 hold. Then, for any admissible mechanism, there exists a feasible mechanism that (weakly) improves the planner's objective, induces an efficient allocation of goods, and can be implemented with a competitive goods market and a (potentially stochastic) income tax.*

Proof. Denote the original mechanism by (x_0,z_0). Suppose that the planner can observe taste types t. (Later we will show that this relaxation of the problem leads to an improving

[25] Under Assumption A1, quasi-linearity of preferences corresponds to an ordinal assumption on agents' preferences, while the normalization of all agents' marginal utility of disposable income to 1 corresponds to an assumption on social preferences.

mechanism that can still be implemented even when taste is not observed.) Thus, the planner faces incentive constraints only in ability:

$$\mathcal{U}(v(x(t,a),z(t,a),t),z(t,a),a) \geq \mathcal{U}(v(x(t,a'),z(t,a'),t),z(t,a'),a), \quad \forall t \in \Theta_t, a, a' \in \Theta_a. \tag{23}$$

Under these incentive constraints, goods x are incentive-separable, as defined by Doligalski et al. (2025). By their Lemma 1 and Lemma 2, the planner can provide all agents with the same utility while saving resources and keeping earnings unchanged by allocating disposable income $m(t,a)$ and allowing agents to purchase goods at their marginal costs (i.e., in a competitive goods market). Let $m_1(t,a)$ be the disposable income required to achieve the original utility level:

$$V(m_1(t,a),z_0(t,a),t,a) = \mathcal{U}(v(x_0(t,a),z_0(t,a),t),z_0(t,a),a), \quad \forall (t,a) \in \Theta. \tag{24}$$

Thus, allocating disposable income $m_1(t,a)$ and earnings $z_0(t,a)$ in conjunction with a competitive goods market feasibly improves (at least in a weak sense) the planner's objective relative to the original mechanism (x_0,z_0). However, the allocation potentially depends on the taste type t and may be infeasible without the planner knowing the agents' preferences.

We will now construct a mechanism that delivers the same social objective as (m_1,z_0) without conditioning on agents' tastes. The main technical challenge in this step—compared to the proof of Theorem 1 from the main text—is that without strong duality we cannot rule out the possibility that the taste type in the relaxed problem is used to implement a *de facto* stochastic income tax (which is known to be sometimes better than a deterministic income tax). The trick is to replicate the randomness provided by conditioning on the taste type with an explicit random variable with the same distribution.

Given the random variable (t,a) describing the distribution of agents' types, define a random variable (τ,t,a), where τ has the same marginal distribution as t but is independent of (t,a). With slight abuse of notation, consider a mechanism that conditions allocations on (τ,a) rather than (t,a): For an agent with "extended" type (τ,t,a), the planner allocates disposable income $\tilde{m}_1(\tau,t,a) = m_1(\tau,a)$ and labor earnings $\tilde{z}_0(\tau,t,a) = z_0(\tau,a)$. Clearly, $(\tilde{m}_1, \tilde{z}_0)$ can be implemented even if the planner does not know agents' true taste parameters t.

The value of the social objective under $(\tilde{m}_1, \tilde{z}_0)$ is

$$\iint \lambda(t,a)V(m_1(\tau,a),z_0(\tau,a),t,a)\, dF_t(\tau)dF(t,a). \tag{25}$$

Applying Assumptions A1′, A2, and A3 allows us to rewrite it as

$$\iint \bar{\lambda}(a)\mathcal{W}(m_1(\tau,a),z_0(\tau,a),a)\,dF_t(\tau)dF_a(a) + \iint \bar{\lambda}(a)\mathcal{W}_0(t,a)dF_t(t)dF_a(a). \tag{26}$$

Combining the two integrals and applying Assumptions A1′, A2 and A3 in reverse yields the value of the social objective:

$$\int \lambda(\tau,a)V(m_1(\tau,a),z_0(\tau,a),\tau,a)\,dF(\tau,a), \tag{27}$$

showing that $(\tilde{m}_1, \tilde{z}_0)$ achieves the same social welfare as the original mechanism (m_1, z_0).

Next, we show that $(\tilde{m}_1, \tilde{z}_0)$ is feasible. The mechanism $(\tilde{m}_1, \tilde{z}_0)$ is incentive-compatible if, for all $t,\tau \in \Theta_t$ and $a,a' \in \Theta_a$:

$$V(m_1(\tau,a),z_0(\tau,a),t,a) \geq V(m_1(\tau,a'),z_0(\tau,a'),t,a). \tag{28}$$

By Assumption A1′, this is equivalent to requiring that

$$\mathcal{W}(m_1(\tau,a),z_0(\tau,a),a) \geq \mathcal{W}(m_1(\tau,a'),z_0(\tau,a'),a). \tag{29}$$

Since (m_1,z_0) is a feasible mechanism in the relaxed problem, we must have

$$V(m_1(\tau,a),z_0(\tau,a),\tau,a) \geq V(m_1(\tau,a'),z_0(\tau,a'),\tau,a), \tag{30}$$

implying that—under Assumption A1′—the constraint (28) also holds.

Finally, the resource constraint clearly holds, since $(\tilde{m}_1, \tilde{z}_0)$ induces the same distribution of disposable income and earnings as (m_1,z_0).

The decentralization with an income tax follows from the taxation principle. Note that the income tax will depend on the choice of earnings and, potentially, on the auxiliary type τ (which means that the income tax could be stochastic). □

A.2 When is the improving income tax deterministic?

Theorem 1′ allows for an income tax that is *stochastic*, meaning that two otherwise identical agents may be facing a different income tax schedule. This is mostly a technical complication: A stochastic income tax in the improving mechanism is not needed if strong duality holds so that the resource constraint can be incorporated into the objective function with a Lagrange multiplier—in that case, we can apply the proof of Theorem 1′ to the Lagrangian and "purify" the optimal mechanism. (We note that strong duality is implicitly assumed by almost the entire literature on optimal taxation.) Formal conditions under which strong duality holds are necessarily somewhat restrictive in the current context, as the planner's problem must be convex for standard constraint qualifications to apply. For completeness, we provide sufficient conditions for an existence of the improving mechanism with a *deterministic* income tax.

Proposition 3. *Strengthen Assumption A1′ to: $V(m, z, t, a) \equiv w_1(m) + w_2(z)w_3(a) + w_4(t, a)$, where $w_1(m)$ and $-w_2(z)$ are concave and $w_2(z)$ is strictly monotone. Then, the decentralizing income tax in Theorem 1′ can be made deterministic.*

The essential assumption in Proposition 3 is that ability affects preferences over earnings *multiplicatively*. For example, it is satisfied by the common isoelastic disutility function:

$$w_2(z)w_3(a) = -\frac{1}{1+\sigma}\left(\frac{z}{a}\right)^{1+\sigma}, \quad \sigma \geq 0. \tag{31}$$

Proof of Proposition 3. Following the first part of Theorem 1′, we know that the planner can weakly improve the objective by allocating disposable income $m_1(t, a)$ and earnings $z_0(t, a)$, and allowing the agents to purchase goods at the competitive market. This mechanism may require the planner to know the taste types. In the second part of Theorem 1′ we removed this requirement by introducing a randomly drawn auxiliary type τ that, for the purpose of allocating disposable income and earnings, replaced the true taste type. Now we will follow an alternative strategy and show that, under stronger assumptions, the planner can remove the dependence of the mechanism on preference type by "averaging" (in the proper sense) the allocation for each ability level—all the while maintaining the value of the social objective and feasibility.

Reformulate the planner's problem such that the choice variables are the utility from disposable income $\omega_1(t, a) = w_1(m(t, a))$ and the utility from earnings $\omega_2(t, a) = w_2(z(t, a))$ for each type $(t, a) \in \Theta$. The social objective becomes

$$\int \bar{\lambda}(a)\left(\omega_1(t, a) + \omega_2(t, a)w_3(a) + w_4(t, a)\right) dF(t, a), \tag{32}$$

the incentive constraints are

$$\omega_1(t, a) + \omega_2(t, a)w_3(a) \geq \omega_1(t, a') + \omega_2(t, a')w(a), \quad \forall t \in \Theta_t, a, a' \in \Theta_a, \tag{33}$$

and the resource constraint is

$$\int w_1^{-1}\left(\omega_1(t, a)\right) - w_2^{-1}\left(\omega_2(t, a)\right) dF(t, a) + G \leq 0. \tag{34}$$

Note that the objective and the incentive constraints are linear in choice variables, while the left-hand side of the resource constraint is convex in the choice variables. Thus, the constraint set of the planner's problem is convex.

Construct a new mechanism by averaging the utility implied by the mechanism (m_1, z_0) across the taste types:

$$\tilde{\omega}_1(t,a) = \int w_1(m_1(t,a))\, dF_t(t), \quad \tilde{\omega}_2(t,a) = \int w_2(z_0(t,a))\, dF_t(t), \quad \forall (t,a) \in \Theta. \tag{35}$$

Since the mechanism (m_1, z_0) is feasible and the constraint set (expressed in terms of utilities) is convex, the new mechanism belongs to the constraint set and is, thus, feasible. Since the objective is linear in the utilities, the new mechanism delivers the same value of the social objective as (m_1, z_0). Finally, the new mechanism does not condition on t and, hence, is feasible also when the planner does not know taste types. □

A.3 Preferences satisfying Assumption A1′

While Assumption A1′ has the virtue of capturing the required economic properties at a higher level of generality than Assumption A1, it is more abstract and difficult to interpret.[26] In this subsection, we explore the consequences of this assumption and consider a few examples.

Recall that Assumption A1′ has implications both for agents' ordinal preferences and for their cardinal utilities (which matter for the social welfare function). To flesh out the ordinal consequences of Assumption A1′, we will assume that it holds for all price vectors in a (small) neighborhood of the vector of marginal costs k.[27] Assume that V is twice continuously differentiable and denote the undistorted choice of goods by $x^*(m,z,t,a)$. Then it follows from Roy's identity that, for any good $i \in \{1, ..., L\}$,

$$\frac{\partial x_i^*(m,z,t,a)}{\partial m} = \alpha_i(m,z,a) + \beta_i(m,z,a) x_i^*(m,z,t,a), \tag{36}$$

for some functions α_i and β_i that *do not* depend on the taste type t.[28] Thus, the slope of

[26] Eden and Freitas (2024) find that a condition analogous to Assumption A1′ is necessary for the utilitarian social welfare function to treat each person's disposable income equally—a property they call "income anonymity."

[27] Economically, this seems to be without loss of generality because it ensures that a small perturbation of marginal costs does not alter the conclusion regarding the desirability of distorting goods markets. Formally, this strengthening of the assumption is needed to use calculus to characterize optimal demand functions.

[28] By Roy's identity:

$$x_i^*(m,z,t,a) = -\frac{\partial V}{\partial k_i}\left(\frac{\partial V}{\partial m}\right)^{-1} = -\left(\frac{\partial \mathcal{W}(m,z,a)}{\partial k_i} + \frac{\partial \mathcal{W}_0(t,a)}{\partial k_i}\right)\left(\frac{\partial \mathcal{W}(m,z,a)}{\partial m}\right)^{-1}.$$

the Engel curve for any good i, conditional on the consumption level of that good x_i^*, is independent of the preference type t. As a result, the Engel curves of any two taste types—keeping z and a constant—either never cross, or perfectly overlay.

Although this property is already quite restrictive, it is satisfied by commonly used families of preferences such as quasi-linear preferences, homothetic (more generally, Stone-Geary) preferences, as well as "Almost Ideal Demand Systems" of Deaton and Muellbauer (1980). Next, we look at these examples one by one, and argue that a particular cardinal representation must be used to make Assumption A1′ hold. In this sense, despite being more permissive than Assumption A1, Assumption A1′ remains a knife-edge condition.

Example 1 (Quasi-linear preferences)**.**

$$U(x,z,t,a) = x_1 + v(x_2,...,x_L,t) - w(z,a), \tag{37}$$

where $x \in \mathbb{R} \times \mathbb{R}_+^{L-1}$. The efficient choice of good $i \in \{2,...,L\}$ depends on the taste type but not on the disposable income m; denote that choice by $x_i^*(t)$. Then, the indirect utility function is

$$V(m,z,t,a) = \frac{m - \sum_{i=2}^{L} k_i x_i^*(t)}{k_1} + v(x_2^*(t),...,x_L^*(t),t) - w(z,a). \tag{38}$$

This indirect utility function satisfies Assumption A1′. Note, however, that it was important that we normalized the marginal utility of good x_1 to 1; for example, the utility function

$$U(x,z,t,a) = v_1(t)\left(x_1 + v(x_2,...,x_L,t) - w(z,a)\right),$$

represents the same ordinal preferences but violates Assumption A1′. This is because $v_1(t)$ effectively acts as a taste-dependent welfare weight which could break the conclusion of Theorem 1′.[29]

Example 2 (Homothetic preferences[30])**.**

$$U(x,z,t,a) = \sum_{i=1}^{L} \gamma_i(t)\log(x_i) - w(z,a), \tag{39}$$

where $\gamma_i(t) \geq 0$, and $x \in \mathbb{R}_+^L$. Here, the parameter $\gamma_i(t)$ is proportional to the optimal expenditure

Differentiate it with respect to m and substitute in the above expression for x_i^* to obtain (36).

[29]This example shows that we could slightly weaken the assumptions of Theorem 1′ (at the cost of complicating the notation) by imposing Assumption A1′ on the product of the welfare weight and the indirect utility function and relaxing Assumption A2.

[30]This example easily generalizes to the class of Stone-Geary preferences given by $U(x,z,t,a) = \sum_{i=1}^{L} \gamma_i(t)\log(x_i - \underline{x}_i) - w(z,a)$, where $\underline{x}_i$ is the minimal consumption level of good i.

share of good i. Then, the indirect utility function is

$$V(m,z,t,a) = \sum_{i=1}^{L} \gamma_i(t) \cdot \log(m) + \sum_{i=1}^{L} \gamma_i(t) \log\left(\frac{\gamma_i(t)}{\sum_{j=1}^{L} \gamma_j(t) \cdot k_i}\right) - w(z,a), \tag{40}$$

and Assumption A1′ holds if and only if $\sum_{i=1}^{L} \gamma_i(t)$ is constant in t (i.e., agents can differ in proportions of income they spend on different goods, but they cannot differ in total utility of income).

Example 3 (Almost Ideal Demand System)**.** *Rather than specifying the utility function, Deaton and Muellbauer (1980) directly posit the following indirect utility from goods:*

$$V(m,z,t,a) = \frac{\log(m) - \log(P(t))}{\Pi_{i=1}^{L} k_i^{\beta_i(t)}} - w(z,a)$$

where the price index $P(t)$ is defined as

$$\log(P(t)) = \alpha_0(t) + \sum_{i=1}^{L} \alpha_i(t) k_i + \frac{1}{2} \sum_{i=1}^{L} \sum_{j=1}^{L} \gamma_{ij}(t) \log(k_i) \log(k_j),$$

and parameters satisfy, for all $t \in \Theta_t$:

$$\sum_{i=1}^{L} \alpha_i(t) = 1, \quad \sum_{i=1}^{L} \beta_i(t) = 0, \quad \sum_{i=1}^{L} \gamma_{ij}(t) = 0, \ \forall j, \quad \gamma_{ij}(t) = \gamma_{ji}(t), \ \forall i,j.$$

Note that this indirect utility function satisfies Assumption A1′ if parameters $\{\beta_i(t)\}_{i=1}^{L}$ are constant in the taste type t. Then, the implied demand follows a log-linear equation in disposable income:

$$x_i^*(m,t) = \frac{m}{k_i}\left(\alpha_i(t) + \sum_{j=1}^{L} \gamma_{ij}(t) \log(k_j)\right) + \frac{\beta_i m}{k_i} \log\left(\frac{m}{P(t)}\right). \tag{41}$$

The two previous examples featured preferences with linear Engel curves. This example demonstrates that Assumption A1′ is more permissive and allows for some specifications with Engel curves that are nonlinear.

B Pareto efficiency in the simplified framework

In this appendix, we formally derive properties of Pareto efficient allocations in the simplified framework of Section 4. (An allocation is Pareto efficient if there does not exist another allocation preserving the resource constraint that makes a positive measure of agents strictly better off without making anyone worse off.) The first lemma pins down the key condition for an efficient allocation under subsistence constraints.

Lemma 4. *The allocation $(z_a(t), c_a(t), x_a(t))$ is efficient if and only if the resource constraint (8) holds with equality and, for almost all t,*

$$z_h(t) = \bar{z} \textit{ and } z_l(t) = 0,$$

$$x_a(t) = \begin{cases} 1 & t \geq k,\, c_a(t) > \underline{c} \\ \in [0, 1] & t \geq k,\, c_a(t) = \underline{c} \\ 0 & \textit{otherwise} \end{cases}. \tag{42}$$

The properties of an efficient allocation in our setting are straightforward given the linear-utility model. For agents above the subsistence level, the taste type t is equal to their willingness to pay (WTP) for the good, and efficiency requires that these agents consume the good if and only if their WTP is above marginal cost. However, for agents whose numeraire consumption is at the subsistence level $\underline{c}$, WTP is not uniquely defined. Intuitively, agents at the subsistence constraint have a rate of substitution t for buying slightly less of the good, and a rate of substitution 0 for buying slightly more of the good (since this would shift their numeraire consumption below subsistence). Consequently, for an agent at the subsistence constraint who has taste type t, any level of consumption of the good consistent with WTP being between 0 and t is Pareto efficient.

Next, we formalize the intuitive result that a Pareto efficient allocation of goods requires pricing them at their marginal costs.

Definition 1. *For an incentive-compatible mechanism $(z_a(t), x_a(t), c_a(t))$, we define the per-unit price for a good with (strictly positive) quality $q \in Im(x_a)$ faced by ability type a as*

$$p_a(q) := \frac{c_a(0) - c_a(x_a^{-1}(q))}{q}. \tag{43}$$

Lemma 5. *If an incentive-compatible mechanism $(z_a(t), x_a(t), c_a(t))$ is efficient, then for any $a \in \{l, h\}$ and any strictly positive $q \in Im(x_a)$, we must have $p_a(q) = k$.*

B.1 Proofs of Lemma 4 and Lemma 5

Proof of Lemma 4. It is immediate that labor supply must satisfy $z_h(t) = \bar{z}$ and $z_l(t) = 0$ in any efficient allocation. Since we restricted attention to allocations in which agents' utilities are finite, we can assume that $c_a(t) \geq \underline{c}$ for all a and t. It is also clear that the resource constraint (8) must be binding in any efficient allocation.

We will first prove that condition (42) is necessary. Suppose condition (42) fails for a positive mass of agents such that $c_a(t) > \underline{c}$ and $t \geq k$. Find $\epsilon > 0$ such that a strictly positive mass of agents have $c_a(t) \geq \underline{c} + \epsilon k$ and $x_a(t) \leq 1 - \epsilon$; then, decrease their $c_a(t)$ by ϵk, and increase their $x_a(t)$ by ϵ. This leaves the resource constraint unaffected and raises the utility of these agents (almost all of them strictly), which contradicts Pareto efficiency. Condition (42) does not restrict $x_a(t)$ for agents with $c_a(t) = \underline{c}$. Finally, suppose that condition (42) fails for a positive mass of agents such that $c_a(t) = \underline{c}$ and $t < k$. Then, an analogous argument as for the first case shows that the utility of a positive mass of such agents can be improved.

We will now prove that condition (42) is sufficient. Fix an allocation satisfying condition (42) and suppose that there is a Pareto improvement. Notice that there cannot be a Pareto improvement for agents for whom $c_a(t) > \underline{c}$ unless these agents consume more resources in total (understood as a decrease in the left-hand side of constraint (8)). Hence, if there is a Pareto improvement, there is also a Pareto improvement in which only agents with $c_a(t) = \underline{c}$ are affected. Similarly, the utility of agents with $c_a(t) = \underline{c}$ and $t < k$ can only be increased by giving them more resources, so we can find a Pareto improvement among agents with $c_a(t) = \underline{c}$ and $t \geq k$. Fix such a Pareto improvement, and denote the set of affected agent types by A. Let $\Delta c_a(t) \geq 0$, $\Delta x_a(t)$ denote the change in their allocation of c and x in the Pareto improvement. It must be that $\Delta c_a(t) + t\Delta x_a(t) \geq 0$ for all $(t, a) \in A$, with a strict inequality for a positive mass of agents within A. To preserve the resource constraint, it must be that, $\mathbb{E}[\Delta c_a(t) + k\Delta x_a(t)|A] \leq 0$, where the expectation is taken over (a, t) conditional on A. We have

$$0 \geq \mathbb{E}[\Delta c_a(t) + k\Delta x_a(t)|A] \geq \mathbb{E}[\underbrace{\Delta c_a(t) + t\min\{\Delta x_a(t), 0\}}_{\geq 0}|A] > 0,$$

where the last inequality is strict because $\Delta c_a(t) + t\Delta x_a(t) > 0$ for a positive mass of agents in the set A. Contradiction.

Proof of Lemma 5. First, suppose that $c_a(\bar{t}) > \underline{c}$ (which implies that, in an incentive-compatible mechanism, $c_a(t) > \underline{c}$ for all t). Then, Pareto efficiency requires that $x_a(t) = \mathbf{1}_{\{t \geq k\}}$. Fixing a, incentive compatibility implies that $c_a(t)$ jumps downward at $t = k$ by k, and is constant otherwise. In particular, only quality $q = 1$ is offered. Plugging this into the definition of the per-unit price, we obtain that $p_a(1) = k$, as required.

Next, let us assume that $c_a(\bar{t}) = \underline{c}$. By incentive-compatibility, there must exist a type $t^\star$ such that, for $t \in [t^\star, \bar{t}]$, $c_a(t) = \underline{c}$, while for types $t < t^\star$, $c_a(t) > \underline{c}$. For types $t < t^\star$, Pareto efficiency requires that $x_a(t) = \mathbf{1}_{\{t \geq k\}}$. For types $t \geq t^\star$, incentive compatibility requires that $x_a(t) = x_a(\bar{t})$, while Pareto efficiency requires that $t^\star \geq k$. However, if $t^\star > k$, then the resulting $x_a(t)$ would not be monotone on $[0, \bar{t}]$, which contradicts incentive-compatibility.

We conclude that $x_a(t) = x_a(\bar{t})\mathbf{1}_{\{t \geq k\}}$. The rest of the proof is analogous to the previous case.

C Results under curvature in the utility function

In this appendix, we examine the robustness of findings from the model with subsistence constraints (Theorem 2) to a utility function that is smooth and strictly concave in the numeraire. We also allow for a strictly concave utility from good x and a strictly convex disutility from working. Specifically, assume the utility of type (t, a) is given by

$$u(c) + v(x,t) - (\mathbb{1}_{a=h} w(z) + \mathbb{1}_{a=l} \bar{w} z) \tag{44}$$

that is twice continuously differentiable in all arguments and where: $u(c)$ is strictly increasing, strictly concave, and either $c \in \mathbb{R}$ or $c \geq 0$ and $\lim_{c \to 0} u'(c) = \infty$; $v(x,t)$ is concave in $x \in \mathbb{R}_+$ and satisfies the single-crossing property: $v_{xt}(x,t) > 0$ for all $t \in [0, \bar{t}]$ and $x \geq 0$; $w(z)$ is strictly increasing and strictly convex in $z \in \mathbb{R}_+$ and $\bar{w}$ is high enough that low-ability types neither work nor prefer to mimic high-ability types in the optimum. The rest of the model is the same as in Section 4.3.

C.1 Preliminary results

No earnings distortion. Define the efficient choice of earnings of high-ability agents given numeraire c as $z^*(c) := w'^{-1}(u'(c))$. Suppose there exists type (t, h) with distorted earnings: $z_h(t) \neq z^*(c_h(t))$. Perturb $z_h(t)$ towards $z^*(c_h(t))$ and adjust $c_h(t)$ to keep the utility of this type constant. The perturbation improves the planner's objective: It relaxes the resource constraint and preserves all incentive constraints, since high-ability types can be mimicked only by other high-ability agents, who are indifferent to this alteration. Thus, earnings of high-ability types are undistorted at the optimum.

Given this result, it will be convenient to define the utility from numeraire net of disutility from working as

$$\tilde{u}(c) := u(c) - w(z^*(c)).$$

Summarizing incentive constraints. Take some $t, t' \in \Theta_t$ and assume that the IC of type (t, h) mimicking (t, l) and of (t, l) mimicking (t', l) are satisfied. Then

$$u(c_h(t)) + v(x_h(t), t) - w(z_h(t)) \geq u(c_l(t)) + v(x_l(t), t) \geq u(c_l(t')) + v(x_l(t'), t). \tag{45}$$

Comparing the left-hand and the right-hand sides, we see that type (t, h) has no incentives to mimic (t', l). Thus, provided that other ICs are satisfied, the ICs corresponding to joint deviations in ability and taste are redundant.

Denote the utility level of type (t, a) by $U_a(t) = u(c_a(t)) + v(x_a(t), t) - \mathbb{1}_{a=h} w(z_h(t))$. The downward ICs in ability can be written as:

$$U_h(t) \geq U_l(t), \forall t \in \Theta_t. \tag{46}$$

Regarding the ICs in taste dimension, given the single-crossing assumption, it is standard to summarize them as

$$U_a(t) = U_a(0) + \int_0^t v_t(x_a(t), t)\, dt, \quad \forall t \in \Theta_t, a \in \{h, l\}, \tag{47}$$

combined with a requirement that $x_l(\cdot)$ and $x_h(\cdot)$ are non-decreasing. Note that $U_a'(t) = v_t(x_a(t), t)$, whenever it exists.

Note that ICs (in taste) imply that $c_l(t)$ must be non-increasing, and strictly decreasing whenever $x_l(t)$ is strictly increasing. The same is true for high-ability agents. To see that, suppose that $x_h(t') \geq x_h(t)$ and $c_h(t') > c_h(t)$ for some $t' > t$. Since earnings are undistorted, $z_h(t') \leq z_h(t)$. Thus, type (t, h) strictly gains from mimicking (t', h)—a contradiction.

Reformulating the resource constraint. Let $u_a(t)$ represent the utility from numeraire net of the cost of working of type (t, a). The resource constraint can be written as a function of $\{u_a(\cdot), x_a(\cdot)\}_{a \in \{h,l\}}$:

$$\int \Big(\mu_h(z^*(\tilde{u}^{-1}(u_h(t))) - \tilde{u}^{-1}(u_h(t)) - kx_h(t)) - \mu_l(u^{-1}(u_l(t)) + kx_l(t))\Big)\, dF(t) \geq G. \tag{48}$$

Furthermore, $u_a(t)$ is pinned down by $U_a(0)$ and $x_a(\cdot)$:

$$u_a(t) := U_a(t) - v_t(x_a(t), t) = U_a(0) + \int_0^t v_t(x_a(\tau), \tau)\, d\tau - v_t(x_a(t), t). \tag{49}$$

Thus, we effectively expressed the resource constraint as a function of $\{U_a(0), x_a(\cdot)\}_{a \in \{h,l\}}$.

Reformulating the objective. Incorporate the downward incentive constraints in ability (46) into the objective function by forming a Lagrangian:

$$\mathcal{L} = \lambda_h \mu_h \int U_h(t)\, dF(t) + \lambda_l \mu_l \int U_l(t)\, dF(t) + \int (U_h(t) - U_l(t))\, d\Gamma(t), \tag{50}$$

where $\Gamma(t)$ stands for the value of marginally relaxing the downward incentive constraints (in ability) for all types in the interval $[0,t]$—see Jullien (2000) for an analogous formulation in the model with type-dependent outside options. We assume that $\Gamma(t)$ corresponding to the optimal mechanism exists.[31] Note that $\Gamma(t)$ is non-negative and non-decreasing, equal to zero for $t<0$ and constant for $t\geq\bar{t}$. The multiplier $\Gamma(t)$ can be discontinuous. For instance, $\Gamma(\bar{t})=\Gamma(0)>0$ means that the IC in ability binds only for the lowest taste type $t=0$, while $\Gamma(\bar{t})>0$ and $\Gamma(t)=0, \forall t<\bar{t}$, means that this constraint binds only for the highest taste type $t=\bar{t}$. An intermediate case, with $\Gamma(t)$ increasing over the interval of types, is also possible.

Integrate the objective by parts, starting with high-ability agents:

$$\begin{aligned}\int \lambda_h\mu_h U_h(t)\,dF(t)+\int U_h(t)\,d\Gamma(t) &= (\lambda_h\mu_h+\Gamma(\bar{t}))U_h(\bar{t})-\int(\lambda_h\mu_h F(t)+\Gamma(t))U_h'(t)\,dt\\ &= (\lambda_h\mu_h+\Gamma(\bar{t}))U_h(0)+\int(\lambda_h\mu_h(1-F(t))+\Gamma(\bar{t})-\Gamma(t))v_t(x_h(t),t)\,dt,\end{aligned}$$

and similarly for the low-ability agents:

$$\begin{aligned}\int \lambda_l\mu_l U_l(t)\,dF(t)-\int U_l(t)\,d\Gamma(t) &= (\lambda_l\mu_l-\Gamma(\bar{t}))U_l(\bar{t})-\int(\lambda_l\mu_l F(t)-\Gamma(t))U_l'(t)\,dt\\ &= (\lambda_l\mu_l-\Gamma(\bar{t}))U_l(0)+\int(\lambda_l\mu_l(1-F(t))+\Gamma(t)-\Gamma(\bar{t}))v_t(x_l(t),t)\,dt,\end{aligned}$$

where we used $U_a'(t)=v_t(x_a(t),t)$, implied by the local ICs in taste.

Planner's problem. We can write the planner's problem as

$$\begin{aligned}\max_{\{U_a(0),x_a(\cdot)\}_{a\in\{h,l\}}} &(\lambda_h\mu_h+\Gamma(\bar{t}))U_h(0)+\int(\lambda_h\mu_h(1-F(t))+\Gamma(\bar{t})-\Gamma(t))v_t(x_h(t),t)\,dt\\ &+(\lambda_l\mu_l-\Gamma(\bar{t}))U_l(0)+\int(\lambda_l\mu_l(1-F(t))+\Gamma(t)-\Gamma(\bar{t}))v_t(x_l(t),t)\,dt \qquad (51)\end{aligned}$$

subject to the resource constraint (48) and the monotonicity constraints that require $x_h(\cdot)$ and $x_l(\cdot)$ to be non-decreasing. We define the relaxed problem as the planner's problem with the monotonicity constraints dropped.

[31] Sufficient conditions for the existence of $\Gamma(t)$ are the convexity of the planner's problem and the generalized Slater condition (based on Theorem 3.4 in Bonnans and Shapiro (2000)). Since our problem does not include individual rationality constraints, the generalized Slater condition is always satisfied. The convexity of the planner's problem can be proven when the taste type affects the utility from good x multiplicatively—see Proposition 3 in Online Appendix A for an analogous result.

FOCs of the relaxed problem. It will be convenient to define $g_a(t) := 1/u'(c_a(t))$. Note that $g_a(t)$ is a strictly increasing transformation of $c_a(t)$. Thus, $g_a(t)$ is non-increasing in t, and strictly decreasing when $x_a(t)$ is strictly increasing.

$g_a(t)$ represents a resource benefit of marginally lowering the utility from numeraire (net of labor cost) of an agent with type (t, a). For the low-ability agents, this can be verified by differentiating the resource constraint (48) with respect to $u_l(t)$. For the high-ability types, the resource impact of perturbing $u_h(t)$ is given by:

$$\frac{d[z^*(\tilde{u}^{-1}(u_h(t))) - \tilde{u}^{-1}(u_h(t))]}{du_h(t)} = \frac{dz^*(c(t))}{dc(t)} \frac{1}{\tilde{u}'(c(t))} - \frac{1}{\tilde{u}'(c(t))}. \tag{52}$$

Furthermore, $\tilde{u}'(c(t)) = u'(c(t)) - w'(z^*(c(t))\frac{dz^*(c)}{dc}$. If earnings are on the boundary and $\frac{dz^*(c)}{dc} = 0$, then it follows that the resource impact is $g_h(t)$. Otherwise, given that earnings are undistorted, we have $u'(c) = w'(z^*(c))$, which implies $\frac{dz^*(c)}{dc} = \frac{u''(c)}{w''(z^*(c))}$. Plugging these in, we obtain

$$\frac{d[z^*(\tilde{u}^{-1}(u_h(t))) - \tilde{u}^{-1}(u_h(t))]}{du_h(t)} = -\frac{1}{u'(c_h(t))} = -g_h(t). \tag{53}$$

Intuitively, since earnings are undistorted, the planner is indifferent between adjusting $c_h(t)$ or $z_h(t)$ to achieve a given change of $u_h(t)$.

The first-order conditions of the relaxed problem with respect to $U_h(0)$ and $U_l(0)$ are:

$$\lambda_h \mu_h + \Gamma(\bar{t}) - \alpha \mu_h \int g_h(t) dF(t) = 0, \tag{54}$$

$$\lambda_l \mu_l - \Gamma(\bar{t}) - \alpha \mu_l \int g_l(t) dF(t) = 0. \tag{55}$$

Summed, they pin down the multiplier on the resource constraint α:

$$\frac{1}{\alpha} = \mathbb{E}_{a,t}\left[g_a(t)\right] =: \bar{g}. \tag{56}$$

We can also rewrite these first-order conditions as

$$\Gamma(\bar{t}) = \mu_h \int \left(\frac{g_h(t)}{\bar{g}} - \lambda_h\right) dF(t) = -\mu_l \int \left(\frac{g_l(t)}{\bar{g}} - \lambda_l\right) dF(t). \tag{57}$$

The first-order conditions with respect to $x_a(t), a \in \{h, l\}$, accounting for the potential corner solution at $x_a(t) = 0$, require

$$(\lambda_h \mu_h (1 - F(t)) + \Gamma(\bar{t}) - \Gamma(t)) v_{tx}(x_h(t), t) \\ + \alpha \left[\left(\frac{v_x(x_h(t), t)}{u'(c_h(t))} - k \right) \mu_h f(t) - v_{tx}(x_h(t), t) \mu_h \int_t^{\bar{t}} g_h(t') dF(t') \right] \le 0, \quad (58)$$

and

$$(\lambda_l \mu_l (1 - F(t)) + \Gamma(t) - \Gamma(\bar{t})) v_{tx}(x_l(t), t) \\ + \alpha \left[\left(\frac{v_x(x_l(t), t)}{u'(c_l(t))} - k \right) \mu_l f(t) - v_{tx}(x_l(t), t) \mu_l \int_t^{\bar{t}} g_l(t') dF(t') \right] \le 0. \quad (59)$$

Define the good x wedge as $\tau_a(t) := \frac{v_x(x_a(t),t)}{u'(c_a(t))} - k$. A positive (respectively, negative) value of the wedge implies that allocation x is distorted downwards (resp., distorted upwards, provided that $x_a(t) > 0$). We can express the FOCs as

$$\tau_h(t) \frac{\mu_h f(t)}{v_{tx}(x_h(t), t)} \frac{1}{\bar{g}} \le \mu_h \int_t^{\bar{t}} \left(\frac{g_h(t')}{\bar{g}} - \lambda_h \right) dF(t') - \Gamma(\bar{t}) + \Gamma(t), \quad (60)$$

$$\tau_l(t) \frac{\mu_l f(t)}{v_{tx}(x_l(t), t)} \frac{1}{\bar{g}} \le \mu_l \int_t^{\bar{t}} \left(\frac{g_l(t')}{\bar{g}} - \lambda_l \right) dF(t') + \Gamma(\bar{t}) - \Gamma(t). \quad (61)$$

Sum them up and multiply by $\bar{g}$ to get

$$\tau_h(t) \frac{\mu_h f(t)}{v_{tx}(x_h(t), t)} + \tau_l(t) \frac{\mu_l f(t)}{v_{tx}(x_l(t), t)} \le (1 - F(t)) \mathbb{E}_{a,t'} \left[g_a(t') - \bar{g} \mid t' \ge t \right]. \quad (62)$$

Since $g_a(t)$ is non-increasing with taste, the right-hand side is (weakly) negative. Thus, either $\tau_h(t)$ or $\tau_l(t)$ must be (weakly) negative for any $t \in \Theta_t$.

C.2 Optimal goods distortions

The following proposition characterizes the optimal goods market distortions with curvature in the utility function. To rule out an uninteresting case, we assume that in the optimum a positive measure of agents receives $x > 0$. We discuss this proposition and provide intuition in the main body of the paper (Section 4.3).

Proposition 4. *Suppose that Assumptions A2 and A3 hold, and that agents' preferences are given by formula (44). The optimal mechanism has the following properties:*

1. *Distortions to good x are optimal: There can be no interval of taste types $[i_1, i_2] \subseteq [0, \bar{t}]$ where, for all $t \in [i_1, i_2]$, $\max\{x_h(t), x_l(t)\} > 0$ and both $x_h(t)$ and $x_l(t)$ are undistorted.*

2. *Assume that the optimum does not require ironing.*[32] *The optimal allocation of good x of the low-ability types is either distorted upwards or undistorted.*

3. *Assume $\lambda_h = 0$ and that at the optimum $x_h(t) > x_l(t)$ for all $t \geq t_0$. The optimal allocation of good x of the high-ability types with taste $t \in (t_0, \bar{t}]$ coincides with the solution to the one-dimensional monopolistic screening problem (with the reservation value given by the utility of type (t_0, h) and the lower bound on feasible allocations of x given by $x_h(t_0)$).*

Proof. **Part 1.** Consider an optimal allocation rule. Suppose there exist $i_1, i_2 \in \Theta_t$, $i_2 > i_1$, such that for all $t \in [i_1, i_2]$ both $x_h(t)$ and $x_l(t)$ are undistorted and at least one of them is strictly positive.

We will start by showing that when $x_a(t)$ is undistorted and strictly positive over the taste interval $[i_1, i_2]$ then $x_a(t)$ is strictly increasing and $c_a(t)$ strictly decreasing in t over this interval, for all $a \in \{h, l\}$. This is useful since it means that the monotonicity constraints are slack for any $t \in (i_1, i_2)$ and the FOC from the relaxed problem must hold at the optimum. Suppose that $x_a(t_1) = x_a(t_2) = \bar{x} > 0$ for some $i_1 \leq t_1 < t_2 \leq i_2$. Since markets are not distorted

$$ku'(c_a(t_2)) - ku'(c_a(t_1)) = v_x(\bar{x}, t_2) - v_x(\bar{x}, t_1) = \int_{t_1}^{t_2} v_{xt}(\bar{x}, t)dt > 0, \tag{63}$$

which means that $c_a(t_2) < c_a(t_1)$. However, then type (t_2, a) would mimic (t_1, a), which is a contradiction. Furthermore, if $x_a(t_2) > x_a(t_1)$ then $c_a(t_2) < c_a(t_1)$, since otherwise type (t_1, a) would mimic (t_2, a).

Suppose that low-ability types with taste above threshold t_l, where $t_l < i_2$, consume good x in a positive quality. Define $\tilde{t} := \max\{i_1, t_l\}$. Then the high-ability types with taste from $[\tilde{t}, i_2]$ must also consume x in a positive quality, which follows from their allocation of x being undistorted. Since both $x_h(t)$ and $x_l(t)$ are strictly increasing over $[\tilde{t}, i_2]$, the monotonicity constraints are slack for all $t \in (\tilde{t}, i_2)$ and the FOCs from the relaxed problem inform us of the welfare impact of a small perturbation within this open interval. Consider (62). Since $g_a(t)$ is a monotone transformation of $c_a(t)$, it is strictly increasing over $[\tilde{t}, i_2]$, and the right-hand side of (62) is strictly negative for any $t \in (\tilde{t}, i_2)$. On the other hand, the left-hand side is equal to zero, since good x is undistorted. The FOC is violated and the planner can improve the allocation by perturbing both $x_l(t)$ and $x_h(t)$ upward for all $t \in (\tilde{t}, i_2)$.

Next suppose that within the taste interval $[i_1, i_2]$ none of the low-ability types and all of the high-ability types consume a positive quality of x. Recall that $U_a(t) = v_t(x_a(t), t)$, for all $a \in \{h, l\}$, where the right-hand side is increasing in $x_a(t)$. Since $U_h(i_1) \geq U_l(i_1)$ and

[32] That is, we can drop the monotonicity constraints on $x_l(t)$ and $x_h(t)$ without affecting the solution.

$x_h(t) > 0 = x_l(t)$ for all $t \geq i_1$, it follows that $U_h'(t) > U_l'(t)$ and $U_h(t) > U_l(t)$ for all $t > i_1$. Thus, $\Gamma(t) = \Gamma(i_1)$ for all $t \in (i_1, i_2)$. The FOC for $x_h(t)$ becomes

$$\tau_h(t)\frac{\mu_h f(t)}{v_{tx}(x_h(t),t)}\frac{1}{\bar{g}} = \mu_h \int_t^1 \left(\frac{g_h(t')}{\bar{g}} - \lambda_h\right) dF(t') - \Gamma(\bar{t}) + \Gamma(i_1). \tag{64}$$

The derivative of the right-hand side with respect to t is proportional to $\lambda_h - \frac{g_h(t)}{\bar{g}}$. Given that $g_h(t)$ is strictly decreasing over the interval (i_1, i_2), the right-hand side can be either strictly decreasing, or strictly increasing, or first strictly decreasing and then strictly increasing. Either way, there are at most two values of t for which the right-hand side is zero. Thus, for almost all $t \in (i_1, i_2)$ the first-order condition is violated and the planner can improve the allocation by distorting $x_h(t)$.

Part 2. Consider the relaxed problem. Define

$$\phi(t) := \mu_l \int_0^t \left(\lambda_l - \frac{g_l(t)}{\bar{g}}\right) dF(t) - \Gamma(t). \tag{65}$$

Combining FOCs with respect to $x_l(t)$ and $U_l(0)$ yields

$$\phi(t) \geq \tau_l(t)\frac{\mu_l f(t)}{v_{tx}(x_l(t),t)}\frac{1}{\bar{g}}. \tag{66}$$

Let's characterize the behavior of $\phi(t)$. From the FOC with respect to $U_l(0)$ we know that

$$\mu_l \int_0^{\bar{t}} \left(\lambda_l - \frac{g_l(t)}{\bar{g}}\right) dF(t) = \Gamma(\bar{t}) \geq 0. \tag{67}$$

Since $g_l(t)$ is non-increasing, there exists a threshold $\tilde{t} \geq 0$ such that $\lambda_l \geq \frac{g_l(t)}{\bar{g}}$ for all $t \geq \tilde{t}$. It follows that the first term of $\phi(t)$ is continuous in t, equal to 0 at $t = 0$, and (weakly) decreasing until $\tilde{t}$, at which point it becomes (weakly) increasing, eventually reaching $\Gamma(\bar{t})$. The second term, $-\Gamma(t)$, is right-continuous (which follows from the definition of $\Gamma(t)$) and (weakly) decreasing in t, eventually reaching $-\Gamma(\bar{t})$. Thus, $\phi(0) \leq 0$ and $\phi(\bar{t}) = 0$. In addition, $\phi(t)$ can be strictly positive only when $t > \tilde{t}$, i.e., in the region where the first term is increasing.

Now, suppose there exists $t_0 \in \Theta_t$ for which $\tau_l(t_0) > 0$, which would contradict the proposition. That requires $\phi(t_0) > 0$. Since $\phi(\cdot)$ is right-continuous and it eventually reaches the value $\phi(\bar{t}) = 0$, there must exist $t_1 > t_0$ such that $\phi(t) > 0$ for all $t \in (t_0, t_1)$ and $\phi(t_1^-) > \phi(t)$ for all $t > t_1$. Since $\phi(\cdot)$ is strictly decreasing at t_1, $\Gamma(\cdot)$ is strictly increasing at this point, implying $U_h(t_1) = U_l(t_1)$.

Given that $U_h(t_0) \geq U_l(t_0)$, it follows that

$$U_h(t_1) - U_h(t_0) \leq U_l(t_1) - U_l(t_0), \tag{68}$$

which can be restated as

$$\int_{t_0}^{t_1} U_h'(s)ds \leq \int_{t_0}^{t_1} U_l'(s)ds. \tag{69}$$

Thus, there must exists $t' \in (t_0, t_1)$ such that $U_h'(t') < U_l'(t')$ or, equivalently, $x_h(t') \leq x_l(t')$.

If $x_l(t') > 0$, then the FOC with respect to $x_l(t')$ holds as an equality and $\tau_l(t') > 0$. Since $x_h(t') \leq x_l(t')$ and $c_h(t') > c_l(t')$—which must hold, as otherwise type (t, h) would mimic (t, l)—it follows that $\tau_h(t') > 0$. That contradicts (62), which requires that either $\tau_h(t')$ or $\tau_l(t')$ is non-positive.

If $x_l(t') = x_h(t') = 0$, then, by monotonicity, $x_l(t_0) = x_h(t_0) = 0$. Similarly as in the previous case, $\tau_l(t_0) > 0$ and $x_h(t_0) \leq x_l(t_0)$ implies that $\tau_h(t_0) > 0$, which contradicts (62).

Part 3. We will consider the planner's subproblem of choosing the allocation of high-ability types with taste $t > t_0$ taking as given the rest of the allocation rule. We will show that it can be written as a one-dimensional monopolistic screening problem.

By assumption, the welfare weight is 0 for the high-ability types. Furthermore, the marginal value of public funds is always positive. Thus, the planner's objective with respect to the high-ability types is to maximize revenue.

We will show that the downward ICs in ability are slack for $t > t_0$. Note that the ICs in taste require that $U_a'(t) = v_t(x_a(t), t)$, with the right-hand side strictly increasing in $x_a(t)$. Thus, given that the optimal allocation involves $U_h(t_0) \geq U_l(t_0)$ and $x_h(t) > x_l(t)$ for all $t \geq t_0$, it follows that $U_h'(t) > U_l'(t)$ and $U_h(t) > U_l(t)$ for all $t > t_0$.

Define $p_h(t) := z^*(c_h(t)) - c_h(t)$ as a transfer from type (t, h) to the planner. Let $P(c) := z^*(c) - c$, which is strictly decreasing, and define the disutility from transfer as $d(p) := -\tilde{u}(P^{-1}(p))$. It follows that the utility of type (t, h) from allocation (c, x, z) where $z = z^*(c)$ can be described as $v(x, t) - d(p)$ where $p = z - c$.

Now, we can write the planner's subproblem over $\{x_h(t), p_h(t)\}_{t \in (t_0, \bar{t}]}$, taking the allocation of remaining high-ability types as given, as:

$$\max_{\{x_h(t), p_h(t)\}_{t \in (t_0, \bar{t}]}} \int_{t_0}^{\bar{t}} (p_h(t) - kx_h(t)) \, dF(t) \tag{70}$$

subject to incentive-compatibility constraints

$$v(x_h(t), t) - d(p_h(t)) \geq v(x_h(t'), t) - d(p_h(t')), \quad \forall t, t' \in [0, \bar{t}]. \tag{71}$$

The single-crossing condition implies that the local incentive constraints are sufficient and we only need to keep track of incentives to deviate to within the set $[t_0, \bar{t}]$. We can summarize these incentive constraints as

$$U_h(t) = U_h(t_0) + \int_{t_0}^{t} v_t(x_h(t), t), \quad \forall t \in (t_0, \bar{t}], \tag{72}$$

together with the requirement that $x_h(t)$ is non-decreasing over $[t_0, \bar{t}]$. Note that $U_h(t_0)$ and $x_h(t_0)$, which are taken as given, play the roles of the reservation value and the lower bound on the feasible allocation of x, respectively. Thus, we can rewrite the problem as

$$\max_{\{x_h(t), p_h(t)\}_{t \in (t_0, \bar{t}]}} \int_{t_0}^{\bar{t}} (p_h(t) - k x_h(t)) \, dF(t) \tag{73}$$

subject to (72), $x_h(\cdot)$ being non-decreasing and $x_h(t) \geq x_h(t_0), \forall t \in (t_0, \bar{t}]$, which concludes the proof. □