Screening with endogenous preferences*

Lizi Chen
MIT

Casey Rothschild
Wellesley College

July 2015

Abstract

A general framework is developed for studying screening in many-agent discrete type environments wherein each agent's preferences depend endogenously on the allocations received by the other agents. Applications include optimal income taxation, performance contracting with across-worker externalities, and insurance with endogenous risks. The solution to the principal's problem is analyzed by decomposing it into an inner problem with fixed preferences and an outer problem which determines preferences. Because the outer problem is typically discontinuous at points where the preferences of two or more types endogenously coincide, the principal frequently finds it optimal to select allocations which involve two or more types with endogenously coinciding preferences, even though such allocations are ex-ante highly unusual. Assuming that types are strictly ordered by their single-crossing preferences is, therefore, not innocuous in endogenous preference environments.

*The authors thank Eduardo Azevedo, Dominik Sachs, Florian Scheuer, Glen Weyl and participants in the Cologne Taxation Theory Conference for helpful comments on earlier drafts. Any remaining errors are our own. Casey thanks the Radcliffe Institute for Advanced Study at Harvard University for generous support.

1 Introduction

Principal-agent contracting models of screening developed by Rothschild and Stiglitz (1976), Mirrlees (1971), Mussa and Rosen (1978), and others have been applied to a wide variety of economic settings with informational asymmetries. Examples include insurance contracting, the design of redistributive income taxes, and compensation via performance contracts. This paper develops a technique for extending principal-agent screening models to allow for across-agent interactions—specifically, to allow for the possibility that each agent's preferences over contracts may depend on the contracts received by other agents.[1]

We discuss three main applications. First, different income earners may be employed in complementary (or otherwise interacting) activities, as in Stiglitz (1982), so that the earnings of workers in one activity can indirectly affect the wages of workers in another sector. Second, firms typically hire and offer performance contracts to multiple workers who interact in teams—and effort by one individual within that team may affect the measured performance of the other individuals on that team. Third, an insured's exposure to risk may depend on other insureds' coverage. For example, one individual's coverage level may affect her accident avoidance effort, which in turn may affect another individual's accident risk and hence preferences over coverage levels. Similarly, exposure to financial risk from a not-at-fault accident with another driver may depend on the coverage level of that other driver.

We incorporate agents with interdependent preferences into screening problems by allowing each agent's preferences to depend, endogenously, on other agents' contracts through a (possibly multi-dimensional) function X of *all* agents' contracts. We make standard single crossing assumptions about preferences that facilitate the application of standard screening techniques at any given value x of the function X. We then solve the principal's problem by decomposing it into two problems: an inner problem for a fixed x—but which requires the allocations to be consistent with $X = x$—and an outer problem for determining the optimal value of x (and hence individuals' endogenous preferences). We combine first order necessary conditions for these two problems to provide a general characterization of

[1]Our paper can alternatively be thought of as extending Segal (1999)—who, like us, studies a principal-agent problem with multiple agents and across-agent externalities—to screening contexts with an uninformed principal.

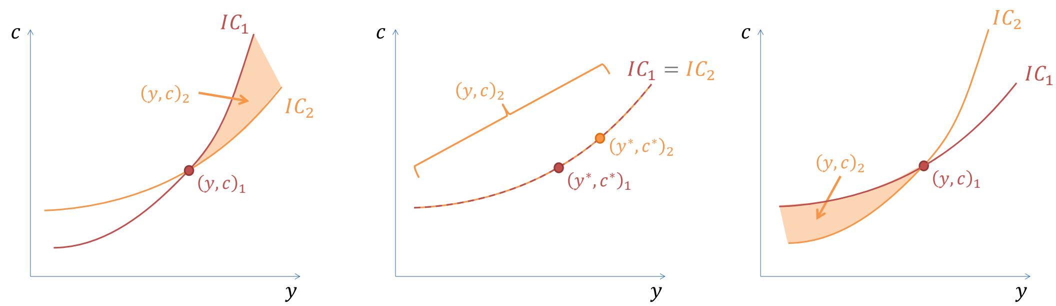
the optimal distortions of individual allocations away from first best.

Two distinct types of optima are possible: "separating" optima, and (partial-) "pooling" optima. A separating optimum occurs at a value of x for which the types are strictly ordered by their single crossing preferences. A pooling optimum occur at a value of x for which the types are only weakly ordered because two or more types have endogenously coincident preferences.

A central result of our paper is to establish that pooling optima are surprisingly common: despite the fact that the set of x-values which involve preference-pooling is generically sparse, it will frequently be *optimal* to choose such an x. For example, consider an optimal tax context, where an individual's market wage w_i is a sufficient statistic for her preferences over income-consumption bundles (y, c). Our results indicate that when individuals' wages depend endogenously on the earnings (or efforts) of other workers, it will frequently be optimal for a Paretian social planner to choose an income tax code that endogenously leads two or more individuals to have identical wages.

The intuition for this "pooling is common" result lies in a failure of lower-hemicontinuity (in x) of the set of incentive compatible allocations at values of x with pooled types. Figure 1 illustrates the basic intuition by depicting the topology of incentive compatibility between two types ("1" and "2") in a standard optimal tax framework where contracts can be described by the pair (y, c) of pre-tax income y and after-tax income c and where individuals' preferences over contracts are given by $u(c, y/w_i)$, where $w \in \{w_1, w_2\}$. The figure depicts three distinct cases. In case (a), $w_2 > w_1$, so type 2 has the higher wage (and hence the shallower indifference curve under standard assumptions). In case (b) $w_2 = w_1$. In case (c), and $w_2 < w_1$. In each case and for any fixed type-1 allocation $(y, c)_1$, incentive compatibility requires that type 2's allocation $(y, c)_2$ lies on or below the type-1 indifference IC_1 through $(y, c)_1$ and on or above the type-2 indifference curve IC_2 through $(y, c)_1$. When $w_2 > w_1$, $(y, c)_2$ must therefore lie in the shaded wedge in Panel (a) above and to the right of $(y, c)_1$. When $w_2 < w_1$, $(y, c)_2$ must lie in the shaded wedge in Panel (c) below and to the left of $(y, c)_1$. When $w_1 = w_2$, IC_1 and IC_2 coincide, and incentive compatibility requires that $(y, c)_2$ lies on these coincident indifference curves.

Now suppose that the wages of the two types vary smoothly with some scalar parameter x, and that, for some x^*, $w_1(x^*) = w_2(x^*)$, $w_2(x) > w_1(x)$ for $x < x^*$,



(a) $w_2 > w_1$ (b) $w_2 = w_1$ (c) $w_2 < w_1$

Figure 1: Discontinuous incentive constraints as a force for a pooling optima

and $w_2(x) < w_1(x)$ for $x > x^*$. As $x \to x^*$ from below, the incentive compatible "wedge" in Panel (a) of Figure 1 collapses to the right-hand half of the coincident curve $IC_1 = IC_2$ in Panel (b). Similarly, as $x \to x^*$ from above, the incentive compatible wedge in Panel (c) collapses to the left-hand half of the same curve.

Finally, suppose that—for some as-yet-unspecified objective function for the principal—there is a unique optimal allocation $((y^*, c^*)_1, (y^*, c^*)_2)$ *given* x^*, and further that this unique allocation has $(y^*, c^*)_2 \gg (y^*, c^*)_1$, as depicted in Panel (b) of Figure 1. Per the topology of Panel (c), there is *no* incentive compatible allocation near $((y^*, c^*)_1, (y^*, c^*)_2)$ for any $x = x^* + \varepsilon$ for any $\varepsilon > 0$, however small. On the other hand, per Panel (a)'s topology, we typically *can* expect to find feasible allocations close to $((y^*, c^*)_1, (y^*, c^*)_2)$ for $x = x^* - \varepsilon$ for sufficiently small ε. This suggests that the value $T(x)$ of the principal's optimized-given-x objective will be left-continuous and will exhibit a downward jump discontinuity to the right of x^*. Figure 2 plots an objective function for a simple optimal-tax model we develop in greater detail in Section 5.3 below. It exhibits exactly this sort of downward jump discontinuity.

The computational example plotted in Figure 2 is a simple three-type extension of Stiglitz's (1982) model of optimal redistributive taxation with two varieties of labor which are complements in a constant-returns-to-scale production function. As in Stiglitz, there is a low-skill type (type 0) and a high-skill type (type 2) whose labor supplies are complementary to each other. In addition, there is a third type (type 1) who has high-skill and whose effective labor effort is a perfect substitute for type 0's labor effort (and thus complements type 2's). Labor markets are competitive, so each worker is paid a wage equal to her marginal product. Constant-

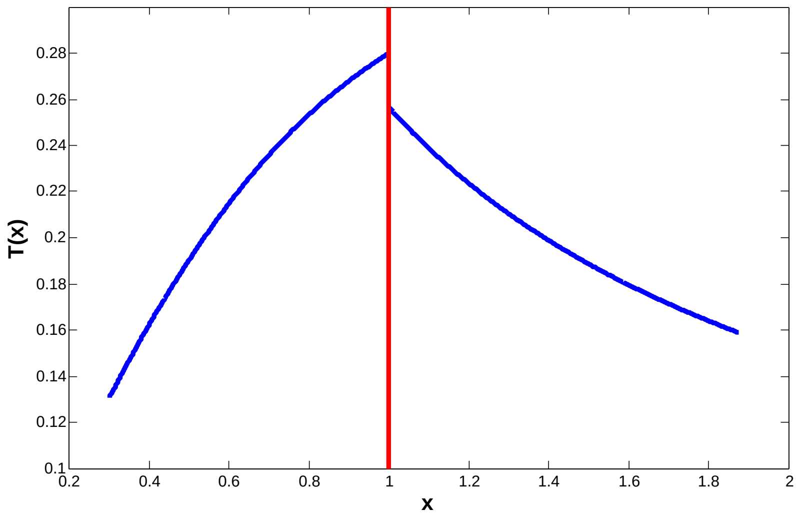


Figure 2: An example with a jump-pooling optimum

returns-to-scale therefore implies that the ratio x of effective effort of type 2 to the combined effective efforts of types 0 and 1 is a sufficient statistic for each types' wage.

As described below, the parameters of this computational example are set up so that, at $x = x^* = 1$, the wages of types 1 and 2 coincide. Exactly as in Panel (b) of Figure 1, the Rawlsian optimum at $x^* = 1$ for this parameterized model involves $(y^*, c^*)_2 \gg (y^*, c^*)_1$. Intuitively, this is because type 2's labor is a complement to the worst-off type 0's labor, while type 1's labor is a substitute for it. As such, distorting type 2's labor supply *up* and distorting type 1's labor supply *down* both indirectly redistribute to type 0 by raising her wage. Since $x < x^*$ implies $w_2 > w_1$ and $x > x^*$ implies $w_2 < w_1$, the argument following Figure 1 applies: the optimum at x^* can be approximated for $x < x^*$ but not for $x > x^*$, and the value function exhibits the downward jump-discontinuity at $x^* = 1$ that is readily apparent in Figure 2. Also apparent is that the global optimum occurs at $x = x^* = 1$, where two types have tied wages, and that the optimality of this pooled-wage optimum is robust to small changes in the underlying model. The optimality of tied wages is not a knife-edge result.

Proposition 2 below formalizes and generalizes the preceding intuition: jump-discontinuities in the value function occur at tied-wage values x^* of the parameter

x precisely when it is optimal to separate the tied types at x^*.[2] When such separation is optimal, these jump-discontinuities will frequently yield robust local optima, and global optima are thus likely to involve tied wages—even if tied wages might appear rare, ex-ante.

A number of other papers, including Weymark (1987), Brett and Weymark (2008, 2011), and Simula (2010), analyze the comparative statics of optimal taxation with respect to the distribution of wages in a discrete type framework. These comparative statics are qualitatively similar to our analysis of the outer problem, since both analyses consider how the optimal tax varies as the wage distribution changes. Our problem differs fundamentally from theirs precisely because our outer-problem wage variation is an endogenous choice for the principal. Our "wage pooling is common" result arises because it is frequently optimal for the principal to *choose* a pooling allocation. Absent this active choice, abstracting from wage pooling by assuming fully order wages—as the earlier papers do—is reasonable. Moreover, and relatedly, even if wages of two distinct are tied in the exogenously-varying wage settings of these earlier papers, then it will typically be optimal to bunch the tied types at a common allocation rather than separating them as in Panel (b) of Figure 1.[3] Indeed, in the three-type optimal tax example underlying Figure 2, it is precisely the endogeneity of preferences—the fact the the contracts given to the two distinct types with tied wages types differentially affect the preferences of the third type—that incents the principal to assign different contracts to the two tied-wage types. In exogenous-wage environments, there is no such incentive, so the tied types can typically be treated as a single (more common) type.[4]

We view our "pooling is common" result as having at least two important im-

[2]Guesnerie and Seade (1982) show that related discontinuities quite generally arise when the principal *indirectly* optimizes over the contract menu rather than directly optimizing over incentive compatible allocations. The discontinuities they identify arise because individual types will "jump" to distinct, far-away contracts in response to local contracting changes, leading to a corresponding jump in the principal's objective. In the present context we instead consider direct incentive-compatible allocations, but a similar effect arises at points where wages are endogenously tied: a small change in x discontinuously changes the set of incentive compatible allocations, which *forces* a jump in the directly-assigned contracts and hence in the principal's objective.

[3]Viz Blackorby et al (2007, Lemma 4) and Weymark (1986, Section 5), for the basic argument.

[4]For comparative statics over x in exogenous settings, it is potentially important to note, however, that the failure of lower hemicontinuity in the incentive constraints can cause a kink-discontinuity in the value function over x. See Section 5.4 for a related discussion.

plications for theoretical and applied work on optimal taxation and elsewhere. First, it calls for caution in assuming strictly-ordered wages in the discrete-type optimal tax environments with endogenous wages that are increasingly important in applied work (viz Ales et al. (2015)). Strictly-ordered wages can be justified ex-ante by sufficiently strong assumptions about wages—for example, Ales et al. make (2015) an absolute-advantage assumption which exogenously fixes the wage order of types while still allowing changes in the absolute and relative magnitude of individual types' wages—but not by appeals to genericity. Second, it is important for rationalizing the connection between discrete type models and their continuum limits. For example, terms related to the overlapping wages of distinct types feature prominently in Rothschild and Scheuer's (2013, 2014a, 2014b) work on optimal income taxation with multidimensional skill heterogeneity. Wage overlap is generic in their continuum-of-types model; absent our result, it might seem highly unusual, and hence puzzling, with discrete types.

This paper proceeds as follows. Section 2 describes our basic analytical framework. Section 3 illustrates how the framework can be applied in various economic applications, namely: optimal income tax settings with endogenous wages, optimal contracting in a production-in-teams multitask setting, and insurance markets with ex-ante moral-hazard and accident prevention costs which are endogenous to the effort prevention activities of other agents. Section 4 contains our main analysis. We define the inner and outer problems, characterize the first order conditions for both, and then combine them to provide a characterization of the optimal distortions for the different types for separating optima where there are no types with identical preferences. We then discuss pooling optima where one or more types endogenously have the same preferences. Proposition 2 presents our central analytical result about jump-discontinuities in the value function at pooled wages. Section 5 provides several explicitly worked examples, including ones which feature separating optima and ones which feature pooling optima. We then relate our work to Rothschild and Scheuer (2014b) in Section 6 before concluding.

2 Setup

Consider an economy with N types of agent $i = 1, ..., N$, each with probability mass f_i. Type i agents have preferences

$$u(c, y, w_i(\vec{x})). \tag{1}$$

The notation in (1) is tailored to a Mirrleesian (1971) income tax setting, where c is "consumption," y is "income," and w_i is a "wage", which depends endogenously on some $\vec{x}$. The value of $\vec{x}$ is a function $X(\vec{Y}) : \mathbb{R}^N \to \mathbb{R}^K$ with $K \leq N$, which maps the N-vector $\vec{Y}$ of total incomes (with components $f_i \mathbb{E}[y_i|i] \equiv f_i \bar{y}_i$) of the various types of agents into a parameter vector which is sufficient for determining $\{w_i\}_{i=1}^N$. Within the income-tax setting and more generally, w_i can be interpreted as any measure that affects the marginal rate of substitution between c and y (e.g., a disutility of labor parameter). We discuss additional applications in Section 3.

We assume that $u(c, y, w)$ has $u_c > 0$, $u_y < 0$, $u_w > 0$, and satisfies the single crossing property:

$$\frac{\partial}{\partial w}\left(\frac{\partial u/\partial y}{\partial u/\partial c}\right) < 0. \tag{2}$$

We also assume that $w_i(\vec{x})$ is continuously differentiable in $\vec{x}$.

The principal's goal is to maximize surplus $g(\vec{Y}, \vec{C})$, where $\bar{c}_i \equiv \mathbb{E}[c_i|i]$ is the average consumption of i-types, g is a "production function", assumed increasing in each component of $\vec{Y}$ and decreasing in each component of $\vec{C}$, and $\vec{C}$ is the vector with components $f_i \mathbb{E}[c_i|i] \equiv f_i \bar{c}_i$).[5] The principal must provide a minimum utility $\bar{u}_i\left(X(\vec{Y})\right)$ to type-i agents. (Note that we allow $\bar{u}_i$ to depend endogenously on $X(\vec{Y})$.) The principal can observe only the consumptions c and incomes y of each agent, but not that agent's type i. We adopt the standard mechanism design approach of treating the principal's problem as one of directly assigning incentive compatible allocations (y, c) to all agents.

We start with a formal definition of allocation.

Definition 1. *Let $\mathcal{N} = \{0, 1, ..., N\}$ be the set of all types. A type-profile is a correspon-*

[5] The case with $g(\vec{Y}, \vec{C}) = \sum_i (\bar{y}_i - \bar{c}_i) f_i$ is of particular interest, and we focus on it in much of our analysis.

dence $p: I \to 2^{\mathbb{R}^2}$ *mapping each type into a measurable set of* (y, c) *pairs. An* allocation *is a collection* $\Pi_{i\in\mathcal{N}} p(i)$ *of type profiles* $p(i)$ *and probability measure* M_i *on* $p(i)$ *for each type* i*. An allocation is* simple *if* $p(i)$ *is a singleton for all* i*.*

Incentive compatibility of an allocation with type profiles $\Pi_{i\in\mathcal{N}} p(i)$ then requires that:

$$u\left(c, y, w_i(X(\vec{Y}))\right) \geq u\left(\hat{c}, \hat{y}, w_i(X(\vec{Y}))\right) \;\; \forall i,\; (y, c) \in p(i),\; (\hat{y}, \hat{c}) \in \cup_{i\in\mathcal{N}} p(i). \tag{3}$$

The principal must offer each individual some minimum utility level $\bar{u}_i$:

$$u\left(c, y, w_i(X(\vec{Y}))\right) \geq \bar{u}_i\left(X(\vec{Y})\right) \quad \forall (y, c) \in p(i). \tag{4}$$

The following lemma allows us to restrict attention to simple allocations, which will do from this point forward.

Lemma 1. *If* u *is quasiconcave then, for any allocation* $\Pi_{i\in\mathcal{N}} p(i)$ *and* $\Pi_{i\in\mathcal{N}} M_i$ *satisfying (3) and (4), there exists a simple allocation* (y_i', c_i'), $i \in \mathcal{N}$ *satisfying (3), (4) and* $g(\vec{Y}', \vec{C}') \geq g\left(\vec{Y}, \vec{C}\right)$*, where* $\vec{Y}$ *is the vector with components* $f_i \mathbb{E}_{M_i}[\tilde{y}|i]$ *and* $\vec{Y}'$ *is the vector with components* $f_i y_i'$*.*

Proof. Let $\tilde{u}_i$ denote the utility of all i types in the initial allocation, which, by (3), is well-defined. Take $y_i' \equiv \mathbb{E}_{M_i}[\tilde{y}|i]$. Take c_i' such that $u(c_i', y_i', w(X(\vec{Y}))) = \tilde{u}_i(X(\vec{Y}))$. By construction, $\vec{Y}' = \vec{Y}$, so the new allocation leaves wages and reservation utilities unchanged. So (4) remains satisfied. By quasiconcavity of u, $\tilde{c}_i' \leq \mathbb{E}_{M_i}[\tilde{c}|i]$, so the principal's objective is at least weakly improved. Finally, incentive compatibility of the new allocation follows from single crossing and the fact that $\tilde{y}_i' \in [\inf\{\tilde{y}|\exists\tilde{c} \text{ s.t. } (\tilde{y}, \tilde{c}) \in p(i)\}, \sup\{\tilde{y}|\exists\tilde{c} \text{ s.t. } (\tilde{y}, \tilde{c}) \in p(i)\}]$. □

We can thus write the principal's problem as:

$$\max_{(\vec{y}, \vec{c})} g(\vec{Y}, \vec{C}) \tag{5}$$

$$\text{subject to} \qquad u\left(c_i, y_i, w_i\left(X(\vec{Y})\right)\right) \geq \bar{u}_i\left(X(\vec{Y})\right) \quad \forall i \tag{6}$$

$$\text{and} \qquad u\left(c_i, y_i, w_i\left(X(\vec{Y})\right)\right) \geq u\left(c_j, y_j, w_i\left(X(\vec{Y})\right)\right) \quad \forall i, j \tag{7}$$

Before characterizing the solution to this problem we discuss some applications that fit in this general framework.

3 Applications

This general framework can be applied to a broad range of economic problems. We discuss three here: an optimal tax application, a multitask-in-teams application, and an insurance market application.

3.1 Optimal Taxes with Multidimensional Skill Heterogeneity

Stiglitz (1982) considers a simple optimal tax model with two types whose labor efforts are complementary. We first construct a simple example—using our notation—which is consistent with Stiglitz's framework; we return to this example in section 5.1.

Suppose there are two equal-mass types $i = 1, 2$, with wages given by $w_1(x) = \frac{1}{3}\left[\frac{1}{3} + \frac{2}{3}x^{\frac{1}{2}}\right]$ and $w_2(x) = \frac{2}{3}\left[\frac{1}{3}x^{-\frac{1}{2}} + \frac{2}{3}\right]$, with $x = X(\vec{y}) = \frac{1}{4}\left(\frac{y_2}{y_1}\right)^2$. Both types have utility function $u(c, y, w) = c - \frac{1}{2}\left(\frac{y}{w}\right)^2$ and a minimum utility of $\bar{u}_i = 0$. The principal's objective is to maximize $\sum_{i=1,2} \frac{1}{2}(y_i - c_i)$.

It is straightforward to show (see Section 5.1 for details) that the optimum will occur at an x for which $w_1(x) < w_2(x)$. This (and $\bar{u}_1 = \bar{u}_2 = 0$) implies that the only binding incentive constraint is the one constraining type 2 not to imitate type 1. It also implies that only type 1's individual rationality constraint is binding. By strong duality, the principal's problem is equivalent to that of maximizing type 1's utility subject to downward incentive compatibility (7) and a resource constraint $\sum_{i=1,2} \frac{1}{2}(y_i - c_i) \geq \bar{R}$ for some $\bar{R}$. The principal's problem is thus equivalent to an optimal tax problem for a Rawlsian social planner. In fact, we show in section 5.1, that this optimal tax problem is a special case of Stiglitz (1982) model with production function $Y = \left[\frac{1}{3}E_1^{\frac{1}{2}} + \frac{2}{3}E_2^{\frac{1}{2}}\right]^2$, where $E_i = y_i/w_i$ measures agents' effort.

Rothschild and Scheuer (2013) extend Stiglitz (1982) to a general two-sector economy with a continuum of types i with preferences $u(c, y/w_i)$ who endogenously choose to work in one of the two sectors. As in Stiglitz aggregated sectoral efforts E_θ and E_φ are complements in a constant returns to scale production technology $Y(E) = Y(E_\theta, E_\varphi)$, and workers are assumed to earn a wage equal to their marginal product. Each individual is characterized by a skill vector $(\theta, \varphi)_i$ drawn from a continuous distribution, and each chooses to work in the sector which provides the highest wage; individuals thus achieves the wage $w_i(\rho) \equiv \max\{\theta \partial Y/\partial E_\theta, \varphi \partial Y/\partial E_\varphi\}$ where $\rho = E_\theta/E_\varphi$. The discrete analog of their model is readily subsumed into our framework—the only complication being that one needs to translate between writing wages in terms of incomes y versus efforts E. See appendix C for details.

Rothschild and Scheuer (2014b) generalize this model to have an arbitrary number of sectors, arbitrary output functions $Y(E^1, ..., E^K) = Y(\vec{E})$, and arbitrary return functions $r^k(\vec{E})$[6]. This generalized model can be similarly adapted to our framework, as we discuss in Section 6.

3.2 Multitask in Teams

In this section, we adapt to a team setting the framework of Baker (2002), who studies an optimal incentive contracting problem for an agent who can pursue multiple different types of work but can only be compensated based on an imperfect uni-dimensional performance measure.[7] We show that this adaptation is a special case of our framework.

Suppose there is a continuum of N types of worker $i = 1, \cdots, N$. Each can simultaneously pursue K distinct tasks $k = 1, \cdots, K$. Conditional on some parameter vector $\vec{\beta}$, with elements $\beta_k > 0$, total firm output depends linearly on the total effort devoted to each task, via $Y = \sum_k \beta_k \sum f_i e_k^i$. The firm cares about total output net of total compensation $Y - \sum f_i c_i$. The firm wishes to incent effort through performance contracts, but cannot directly observe individual effort or types. The firm can only observe the uni-dimensional performance measure $y_i = \sum \beta_k e_k^i = \vec{\beta} \cdot \vec{e}^i$. The firm's problem is to choose a non-linear compensation scheme $c(y)$, or, equivalently, a menu of (y, c) pairs.

Workers' preferences depend separably on their compensation utility $v(c)$ and on a type-specific cost of effort function $\frac{1}{\alpha} m_i(e_1, ..., e_K; \vec{x})$, where m_i is assumed to be homogenous of degree α (henceforth HODα) and strictly quasiconvex in $\vec{e}$ and may

[6]With the property that $\sum_k r^k(\vec{E}) E^k = Y(\vec{E})$.

[7]See also Hölmstrom and Milgrom (1991).

also depend on some parameter vector $\vec{x}$. Since $\vec{e}$ is unobservable, an individual of type i who wishes to achieve a given level of her total performance measure y_i will chooses the $\vec{e}$ that minimizes $m_i(\vec{e};\vec{x})$ subject to $\vec{\beta}\cdot\vec{e}=y$. Since m_i is HODα, the function

$$m^*(y;\vec{x},\vec{\beta}) \equiv \min_{\vec{e}} m_i(\vec{e};\vec{x}) \quad \text{subject to } \vec{\beta}\cdot\vec{e}=y$$

is proportional to y^α. Hence, $w_i(\vec{x},\vec{\beta}) \equiv \left(y^\alpha / m^*(y;\vec{x},\vec{\beta})\right)^{\frac{1}{\alpha}}$, is well defined. We can thus write a worker's utility as

$$u(c,y,w_i(\vec{x},\vec{\beta})) = v(c) - \frac{1}{\alpha}\left(y/w_i(\vec{x},\vec{\beta})\right)^\alpha. \tag{8}$$

The firm's goal is to maximize $\sum_i f_i(y_i - c_i))$ subject to incentive compatibility and some participation constraints of the form $u(c_i,y_i,w_i(\vec{x},\vec{\beta})) \geq \bar{u}_i$.

For a fixed $\vec{x}$ and an exogenous $\vec{\beta}$, this is a standard screening problem. We can endogenize preferences by allowing $\vec{\beta}$ and m_i to depend endogenously, via some aggregator $\vec{x} = \check{X}(\vec{e})$, on the efforts of others. The direct dependence of m_i on the efforts of others captures the possibility that the efforts of others may affect, positively or negatively, the ease with which a given individual can accomplish any given task. The indirect effect, via $\vec{\beta}$, may capture production complementarities. For example, if no one is putting in effort e_2 at making a given product user-friendly, then engineering efforts e_1 may be wasted. In formally: β_1 may be increasing in aggregated efforts e_2.[8]

For any given $\vec{y}$, fixing any given $\vec{x}$ pins down $\beta(\vec{x})$ and hence the optimal effort vectors $\vec{e}_i^*(y_i,\beta(\vec{x}),\vec{x})$ of all types. These effort vectors in turn determine $\check{X}(\vec{e}_1,\cdots,\vec{e}_N)$. We can then define $X(\vec{y})$ as the fixed point of this mapping $\vec{x} \to \{\vec{e}_i\}_{i=1}^N \to \check{X}$. Insofar as this fixed point mapping is well defined (and we give an explicit example in section 5.4 where it is), this problem is a special case of our general framework.

[8] More generally, endogenizing β allows for quite general non-linear production function.

3.3 Insurance Provision with Endogenous Avoidance Cost

For our third application, consider a simple monopoly insurance market with *ex-ante* moral-hazard in the spirit of Bond and Crocker (1991). Each individual $k = 1, \cdots, N$ faces a potential loss of size D out of an initial wealth W. The probability p of experiencing the loss depends on the individual's costly self-protection effort. Without loss of generality, parameterize effort (negatively) by p.

Insurance contracts offer indemnities in the loss state in exchange for state-independent premiums. We can describe such contracts in terms of the consumptions c_L and c_{NL} induced in the loss and no loss states. Individual preferences over such contracts are given by the indirect utility function:

$$\hat{U}(c_L, c_{NL}, \theta) = \max_{p \in [0,1]} pu(c_L) + (1-p)u(c_{NL}) - h(p, \theta), \tag{9}$$

where u is a standard utility function with $u'(x) > 0$ and $u''(x) < 0$ and $h(p, \theta)$ a type-specific cost-of-prevention function with $\partial h / \partial p < 0$ and $\partial^2 h / \partial p^2 > 0$, and $\partial^2 h / \partial \theta \partial p < 0$.

A contract (c_L, c_{NL}) can equivalently be uniquely described by $\Delta = u(c_{NL}) - u(c_L)$ and $V_{NL} = u(c_{NL})$. The unique solution $p^*(\Delta, \theta)$ to (9) is independent of V_{NL} in this formulation. It is nonincreasing in Δ, since a bigger utility gap incents greater effort and hence lower risk, and it is nondecreasing in θ—and strictly decreasing in Δ and strictly increasing in θ except at the $p^* = 0$ and $p^* = 1$ corners.

In (Δ, V_{NL}) space, preferences are $U(\Delta, V_{NL}, \theta) = V_{NL} - p^*(\Delta, \theta)\Delta - h(p^*(\Delta, \theta), \theta)$. U is increasing in V_{NL} and decreasing in Δ (since $p^*(\Delta) + \Delta \frac{\partial p^*(\Delta, \theta)}{\partial \Delta} > 0$). By the envelope theorem, the marginal rate of substitution between V_{NL} and Δ is given by $p^*(\Delta, \theta)$ at any point. Hence, single crossing is satisfied (except for types who are at the same corner).

Finally, the participation constraints are:

$$U(\Delta, V_{NL}, \theta) \geq \bar{u}_k \equiv \hat{U}(W - D, W, \theta_k), \tag{10}$$

and the risk-neutral monopolist-principal's expected profit from the sale of a contract menu (Δ^k, V_{NL}^k) to a set of types θ_k, $k = 1, \cdots, N$ with probability masses f_k

is:

$$\sum_{k=1}^{N} f_k \left(W - \left[p^*(\Delta^k, \theta_k) \left(D + u^{-1}(V_{NL}^k - \Delta^k) \right) \right] - \left[\left(1 - p^*(\Delta^k, \theta_k) \right) u^{-1}(V_{NL}^k) \right] \right). \tag{11}$$

All of the preceding is independent of whether the θ_k are endogenous or exogenous. We might reasonably expect, however, that greater accident prevention by *others* reduces the cost of achieving any given risk level through one's own costly effort. A simple way to model this is to take θ_k to be some function $\theta_k(\vec{p}^*)$, where $\vec{p}^*$ is the vector of risks (hence efforts) of the N types. Since, for any given θ, the optimal p^* is a monotonic function of $\vec{\Delta}$, we can (typically) equivalently write instead $\theta_k(\vec{\Delta})$—i.e., types' preferences depend endogenously on the vector of allocations $\vec{\Delta}$ to the individuals in the economy. See Section 5.2 for an explicitly worked example.

This model is again a special case of our general model, with $\theta \leftrightarrow w$, $\Delta \leftrightarrow y$, and $V_{NL} \leftrightarrow c$.

4 Analysis

The only non-standard element of the principal's problem is the endogeneity of the preference parameters w_i. *Conditional* on the value of $X(\vec{Y}) = \vec{x}$, the w_i are determined, and the problem is essentially standard. This motivates a decomposition into an "outer" problem for the optimal $\vec{x}$ and an "inner" problem which takes x as fixed and imposes the additional consistency constraint $X(\vec{Y}) = \vec{x}$. For simplicity, we focus in the main text on the simpler case with exogenous $\bar{u}_i$ and a linear objective function $\sum f_i(y_i - c_i)$. Appendix D has a general derivation; the results are entirely analogous.

The inner problem in this simpler case is:

$$T(\vec{x}) \equiv \max_{(\vec{y},\vec{c})} \sum_i f_i(y_i - c_i) \quad \text{subject to} \tag{12}$$

$$u\left(c_i, y_i, w_i\left(\vec{x}\right)\right) \geq \bar{u}_i \quad \forall i \tag{13}$$

$$u\left(c_i, y_i, w_i\left(\vec{x}\right)\right) \geq u\left(c_j, y_j, w_i\left(\vec{x}\right)\right) \quad \forall i, j, \text{ and} \tag{14}$$

$$X(\vec{Y}) = \vec{x}, \tag{15}$$

where (15) is the consistency constraint: since $\vec{x}$ is fixed for the inner problem, the principal must choose an allocation $(\vec{y}, \vec{c})$ consistent with this $\vec{x}$.

The outer problem is then to choose $\vec{x}$ to maximize $T(\vec{x})$. We define $\Gamma(\vec{x})$ as the set of all $(\vec{y}, \vec{c})$ satisfying (13), (14), and (15). It will be useful to define $Z(\vec{x})$ as the set of all allocations consistent with (15):

$$Z(\vec{x}) \equiv \{(\vec{y}, \vec{c}) | \overrightarrow{X}(\vec{Y}) = \vec{x}\}. \tag{16}$$

4.1 Solving the inner problem

Since fixing $\vec{x}$ fixes $w_i(\vec{x})$ for all i, we temporarily suppress the argument $\vec{x}$ for notational convenience in this subsection.

4.1.1 Preliminary results

Under the following mild regularity condition, which we assume to hold hereafter, existence of a solution to the inner problem is straightforward.

Assumption 1. *The function u is strictly quasiconcave, and for any $\bar{u}_i$ and any finite w, there exists a (finite) (y^*, c^*) such that $MRS(y^*, c^*) \equiv -\frac{u_y(c^*,y^*)}{u_c(c^*,y^*)} = 1$.*

An immediate implication of this assumption is that $y - c \to -\infty$ along any indifference curve as $y \to \pm\infty$. This ensures that we can restrict attention to a compact subset of the set $\Gamma(\vec{x})$ of feasible allocations given $\vec{x}$ and yields the following lemma.

Lemma 2. *For any $\vec{x}$, a solution to the inner problem exists.*

Proof. See Appendix A.1. □

The following lemma collects some simple implications of incentive compatibility and single-crossing for the solution to the inner problem

Lemma 3. *In any solution to the inner problem, the following properties hold*

1. *For any two types i and j: $c_i \geq c_j \Leftrightarrow y_i \geq y_j$;*

2. *For any two types i and j: $w_i > w_j \Rightarrow c_i \geq c_j$;*

3. *For any three types i, j and k with $w_i \geq w_j \geq w_k$: $u(c_i, y_i, w_i) \geq u(c_j, y_j, w_i)$ and $u(c_j, y_j, w_j) \geq u(c_k, y_k, w_j)$ together imply $u(c_i, y_i, w_i) \geq u(c_k, y_k, w_i)$;*

4. *For any three types i, j and k with $w_i \geq w_j \geq w_k$: $u(c_k, y_k, w_k) \geq u(c_j, y_j, w_k)$ and $u(c_j, y_j, w_j) \geq u(c_i, y_i, w_j)$ together imply $u(c_k, y_k, w_k) \geq u(c_i, y_i, w_k)$.*

Proof. See Appendix A.2. □

Lemma 3 establishes the standard result that incentive compatible allocations must be monotonically increasing in w, and that local incentive constraints are sufficient. Importantly, note that when $w_i = w_j$, the lemma permits $c_i > c_j$, $c_j > c_i$, or $c_i = c_j$.

4.1.2 The inner problem: first order conditions

By Lemma 3, we can restrict incentive constraints to the set of local incentive constraints. Without loss of generality, order the types so that wages are non-decreasing in i.

Define $\tau_i \equiv 1 - MRS_i$. This is interpretable as the implicit "tax wedge" at i's allocation—i.e., as the distortion away from the first best, at which $MRS_i = 1$.

Denote by ξ_i, $\eta_{i,j}$ ($j = i \pm 1$), and $\overrightarrow{\mu}$ the respective multipliers on the individual rationality constraints (13), the incentive constraints (14), and the consistency constraints (15). As above, we write $\vec{Y} = (f_1 y_1, ..., f_K y_N)$. Finally, denote by $\widehat{MRS}_i^j$ the marginal rate of substitution $\equiv -u_y(c_i, y_i, w_j)/u_c(c_i, y_i, w_j)$ for a j type given i's allocation.

The following lemma uses the first order conditions from the Lagrangian formulation of the inner problem to characterize the optimal tax wedges $\vec{\tau}$ for any given x. It is most easily expressed in terms of $\vec{f}\tau$ which is the vector with components $f_i \tau_i$.

Lemma 4. *For any given x, the optimal tax wedge satisfies:*

$$\vec{f}\tau = -(\partial X)\,\vec{\mu} + \vec{D}, \tag{17}$$

where $\vec{D}$ is the N-vector with elements

$$D_i = \sum_{j \in \{i-1, i+1\} \cap \mathcal{N}} \eta_{j,i} u_c(c_i, y_i, w_j) \left(MRS_i - \widehat{MRS}_i^j \right) \tag{18}$$

and ∂X is the $N \times K$ matrix with $(i,k)^{\text{th}}$ element $\frac{\partial x_k}{\partial y_i}$.

Proof. See appendix A.3. □

Equation (17) reveals two sources of distortions: the term $\vec{D}$ and the term $-(\partial X)\vec{\mu}$. The former is an entirely standard distortion due to incentive effects. The latter arises from the endogeneity of preferences, as captured by consistency constraints.

We will next use the outer problem to characterize $(\partial X)\vec{\mu}$ and show that it can be interpreted as a (modified) Pigouvian corrective term.

4.2 Characterizing the outer problem

Towards solving the outer problem of finding the optimal $\vec{x}$—i.e., the $\vec{x}$ that maximizes $T(\vec{x})$ from the inner problem—it is useful to define $X^* = \{\vec{x} | \exists i \neq j \text{ with } w_i(\vec{x}) = w_j(\vec{x})\}$, i.e., the set of all $\vec{x}$ with "tied" types.

We make three technical assumptions. The first ensures that tied wages are unusual from an ex-ante perspective:

Assumption 2. $\nabla \left(\frac{w_i(\vec{x}^*)}{w_j(\vec{x}^*)} \right) \neq 0$ *for all $i \neq j$ and all $\vec{x}^* \in X^*$ with $w_i(\vec{x}^*) = w_j(\vec{x}^*)$.*

Under Assumption 2, the wage ratio $w_i(\vec{x}^*)/w_j(\vec{x}^*)$ is not locally identical to unity. In particular, for any $\vec{x} \in X^*$, there is some smooth curve passing through $\vec{x}$ along which the tied wages are pulled apart.

The partitions $X^*$and $\mathbb{R}^K \setminus X^*$ naturally divide optima into two classes: separating optima, which occur at some $x \in \mathbb{R}^K \setminus X^*$ and are characterized by N distinct wages; and pooling optima, which occur at some $\vec{x} \in X^*$ and are characterized by fewer than N distinct wages.

Our second technical assumption is sufficient for establishing differentiability of $T(\vec{x})$ at points $\vec{x} \notin X^*$.

Assumption 3. *At any $\vec{x} \notin X^*$ and $(\vec{y}, \vec{c})$ satisfying the KKT conditions, the corresponding vector of Lagrange multipliers $(\vec{\zeta}, \vec{\mu}, \vec{\eta})$ is unique.*

It is well known that the uniqueness of KKT multipliers is equivalent to the Strict Mangasarian-Fromowitz Constraint Qualification (SMFCQ) condition (see for example Kyparisis (1985)). By Theorem 7 in Morand et al. (2015), SMFCQ together with appropriate smoothness assumptions on u and w and compactness of (a restriction of) $\Gamma(\vec{x})$, guarantees that $T(\vec{x})$ is directionally differentiable, and the Envelope Theorem applies. We exploit this in computing the directional derivative of $T(\vec{x})$ in the following section.

The third technical assumption will be useful in ensuring that the constraint set $\Gamma(\vec{x})$ is reasonably well-behaved at $\vec{x} \in X^*$. The following definition is necessary:

Definition 2 (Perverseness)**.** *The point $\vec{x}^*$ is said to be* non-perverse *if there is a neighborhood V of $\vec{x}^*$ such that for each $\vec{x} \in V$ there is a solution $(\vec{y}, \vec{c})$ to the inner problem (12)-(15) and a sequence $\vec{y}^n \to \vec{y}$ such that (i) $(\vec{y}^n, \vec{c})$ satisfies (15) for all n and (ii) $y_i^n \neq y_j^n$ for all n and all $i \neq j$. Otherwise, $\vec{x}$ is said to be* perverse.

Assumption 4. *All $\vec{x} \in X^*$ are non-perverse.*

Note that $\Gamma(\vec{x})$ is quite generally upper-hemicontinuous in $\vec{x} \notin X^*$. Lower hemicontinuity can fail for three reasons. First, it will typically fail at $\vec{x} \in X^*$, since, as we discuss in the introduction and in Proposition 2 below, the set $R(\vec{x})$ of allocations satisfying (13) and (14) is discontinuous at such points. Second, it can fail at $\vec{x} \notin X^*$ where $Z(\vec{x})$ is discontinuous. Finally, it can potentially fail at perverse $\vec{x} \notin X^*$, even when $Z(\vec{x})$ is continuous. At non-perverse $\vec{x} \notin X^*$ where $Z(\vec{x})$ is continuous, $\Gamma(\vec{x})$ is continuous, and, by Berge's Maximum Theorem, $T(\vec{x})$ is continuous.[9]

[9] To see why assumption 4 is useful, note that at non-perverse $\vec{x}$, the optimum can always be approximated by non-bunching allocations which give distinct types different allocations; moreover, since (15) does not depend on $\vec{c}$, the optimum can be approximated by non-bunching allocations in the interior of $R(\vec{x})$. Hence, $\Gamma(\vec{x}) = Cl\left[Int\left(R(\vec{x})\right) \cap Z(\vec{x})\right]$ at non-perverse $\vec{x}$. Theorem 11.21 in Border (1989) and the fact that closure preserves lower hemicontinuity then implies $\Gamma(\vec{x})$ is lower hemicontinous.

4.2.1 Decomposing $\nabla T(x)$ in the outer problem

We now compute $\nabla T(x)$ using the Envelope Theorem and the Lagrangian formulation. Specifically, the proof of the following lemma differentiates the inner-problem Lagrangian with respect to $\vec{x}$ holding fixed c_i and $z_i \equiv u(c_i, y_i, w_i(\vec{x}))$ for all i. Suppose a small change in x leads to a change dw_i in the i type's wage. Then y_i will change by $\left(\frac{\partial y_i}{\partial w_i}\right)_{z_i, c_i} dw_i = -\frac{u_w(c_i,y_i,w_i(\vec{x}))}{u_y(c_i,y_i,w_i(\vec{x}))} dw_i$. We denote $-\frac{u_w(c_i,y_i,w_i(\vec{x}))}{u_y(c_i,y_i,w_i(\vec{x}))}$ by $\left(-\frac{u_w}{u_y}\right)^i$. It is also useful to define an analogous term for the j type imitating i: $\widehat{\left(-\frac{u_w^i}{u_y^i}\right)}^j \equiv -\frac{u_w(c_i,y_i,w_j(\vec{x}))}{u_y(c_i,y_i,w_j(\vec{x}))}$.

Finally, define $\vec{\tau}^p$ as the N-vector with elements

$$\tau_j^p \equiv -\sum_i f_i \left(-\frac{u_w}{u_y}\right)^i \sum_k \frac{\partial X^k}{\partial Y_j} \frac{\partial w_i}{\partial X^k}. \tag{19}$$

These elements are interpretable as Pigouvian corrective taxes on agents of the various types j. To see this, note that the term $\sum_k \frac{\partial X^k}{\partial Y_j} \frac{\partial w_i}{\partial X^k}$ measures the amount by which the wage of i types would change in response to a small increase in Y_j, and $\left(-\frac{u_w}{u_y}\right)^i$ measures the amount by which y_i would change in response to a wage change, holding constant i's utility and consumption. Hence τ_j^p is a measure of how much the total output is indirectly affected by wage changes induced by a marginal increase in earnings by a (non-atomic) type-i individual. We define $\vec{f\tau}^p$ as the vector with components $f_i \tau_i^p$ (by analogy with $\vec{f\tau}$ defined in section 4.1.2).

Lemma 5. *When $T(\vec{x})$ is differentiable, its derivative satisfies:*

$$(\partial X)\, \nabla T(\vec{x}) = (\partial X)\, P\vec{\mu} - \vec{f\tau}^p + (\partial X)\, I, \tag{20}$$

where

$$I = \sum_{i=1}^{N} \sum_{j \in \{i-1, i+1\} \cap \{1, \cdots, N\}} \eta_{j,i} u_y(c_i, y_i, w_j) \left[\widehat{\left(\frac{u_w^i}{u_y^i}\right)}_j \nabla w_j(\vec{x}) - \left(\frac{u_w^i}{u_y^i}\right) \nabla w_i(\vec{x}) \right], \tag{21}$$

and where $P = \bar{C}\,(\partial X) - \mathbb{I}_K$, with $\bar{C}$ is the $K \times N$ matrix with elements $\bar{C}_{ji} = \left(-\frac{u_w}{u_y}\right)^i \left(\frac{\partial w_i}{\partial x_j}\right)$

and $\mathbb{I}_K$ is the $K \times K$ identity matrix,

Proof. See Appendix B.1. □

4.3 Combining the inner and outer problems

We will now combine the first order conditions from lemmas 4 and 5 for the inner and outer problem to characterize the optimal distortions at separating optima. To that end, let $Q = \mathbb{I}_N - (\partial X)\bar{C}$ be an $N \times N$ matrix. We assume Q is invertible.[10]

Proposition 1. *If $T(\vec{x})$ is differentiable at a separating optimum, then*

$$\vec{f\tau} = D + Q^{-1}\vec{f\tau}^p - Q^{-1}(\partial X)\, I \tag{22}$$

Proof. Follows directly from lemmas 4 and 5 and the necessary condition $\nabla T = 0$. □

Equation (22) can be understood as decomposing the optimal distortion for each type into three components. The first, D, is the standard distortion that arises from incentive compatibility.

The other two components are a result of the endogeneity of preferences. One is a Pigouvian correction $\vec{f\tau}^p$. It is modified by the term Q^{-1}, which can be understood as a general equilibrium correction. Recall that $\vec{\tau}^p$ is interpretable as the (negative of) the externality caused by a small increase in $Y_i = f_i y_i$, computed holding each type's utility and consumption constant. Specifically, an increase in Y_i changes $X(\vec{Y})$, and therefore the preferences of all individuals of all types j. When preferences change, holding utility and consumption constant requires changes in y_j. These changes in turn change $X(\vec{Y})$, which change preferences, which change $\vec{y}$, and so forth. The term Q translates between the partial equilibrium effects of the change in y_i and the general equilibrium effect after all these "rounds" of changes have taken place.

The final component—which is modified by Q^{-1} for the same reason—is a "Stiglitz" effect. It arises because a change in y_i differentially affects the preferences of different types; accordingly, it indirectly relaxes or tightens incentive constraints. Equivalently, one can think of the Stiglitz term $(\partial X)I$ as a term capturing

[10] If Q is singular, then the fixed point mapping induced by the consistency constraints, holding c_i and z_i constant as we change $\vec{x}$, is not well-defined.

the *indirect* redistribution which occurs as a result of the change in preferences induced by a change in y_i. (This is a particularly transparent interpretation when the preference parameter w_i is a wage.)

4.4 Pooling optima

The preceding analysis applies only to separating optima, i.e., when it is optimal to choose an $\vec{x} \notin X^*$. Away from X^*, the value function $T(\vec{x})$ is typically well-behaved, since the incentive and individual rationality constraints are continuous in $\vec{x}$ at such points. Per Figure 1, the incentive constraints are (in general) not lower hemicontinuous at $\vec{x} \in X^*$. We first show that this can imply jump discontinuities in $T(\vec{x})$ for $\vec{x} \in X^*$, and we characterize precise conditions under which such jump discontinuities will occur. The following definitions will be useful in this characterization.

Definition 3. *An $\vec{x} \in \vec{X}^*$ at which $w_i(\vec{x}) = w_j(\vec{x})$ for some $i, j \neq i$ is said to be* unusual *if there exist two distinct solutions $(\vec{y}, \vec{c})$ and $(\vec{y}', \vec{c}')$ such that $c_i > c_j$ and $c'_i < c'_j$. Otherwise, $\vec{x}$ is said to be* usual.

Definition 4. *An $\vec{x} \in \vec{X}^*$ at which $w_i(\vec{x}) = w_j(\vec{x})$ for some $i, j \neq i$ is said to be* degenerate *if there exists a solution $(\vec{y}, \vec{c})$ such that $c_i = c_j$. Otherwise, $\vec{x}$ is said to be* nondegenerate.

Figure 1 and the associated discussion in the introduction, can be used to provide intuition for these definitions in the one-dimensional $\vec{x}$ case.[11] Graphically, an $x^* \in X^*$ would be unusual if *in addition* to the optimum depicted in Panel (b) with $(y^*, c^*)_2$ lying above and to the right of $(y^*, c^*)_1$, there were a second optimum with $(y^{**}, c^{**})_2$ lying below and to the left of $(y^{**}, c^{**})_1$. An $x^* \in X^*$ would be degenerate if the optimum at x^* involved bunching, i.e., had $(y^*, c^*)_2$ coinciding with $(y^*, c^*)_1$. In either case, an optimum allocation at x^* can be approached as $x \to x^*$ both from above and below, so we expect the value function $T(x)$ to be continuous at x^*. At usual, nondegenerate x^*, the limit allocation as $x \to x^*$ either from the right or the left will approach a strictly sub-optimal allocation, and the value function $T(x)$ will have a jump discontinuity at x^*.

[11] Appendix E discusses usualness and provides simple primal conditions which imply it.

The following proposition formalizes this intuition for a general-dimensional X. For simplicity, the proposition is stated only for those $\vec{x}^*$ with exactly two tied types.[12]

Proposition 2. *If $\vec{x}^* \in X^*$ is usual and nondegenerate, then $T(\vec{x})$ has an one-sided jump discontinuity at $\vec{x}^*$. Specifically, $T(\vec{x})$ is continuous at $\vec{x}^*$ viewed as a point in half space $\{\vec{x} \in \mathbb{R}^N | w_i(\vec{x}) \geq w_j(\vec{x})\}$ or $\{\vec{x} \in \mathbb{R}^N | w_i(\vec{x}) \leq w_j(\vec{x})\}$, but not both.*

Conversely, if $\vec{x}^* \in X^*$ is unusual or degenerate, then $T(\vec{x})$ is continuous at $\vec{x}^*$.

Proof. See appendix B.2 □

Even when $\vec{x}^* \in X^*$ is unusual or degenerate, and, per Proposition 2, $T(\vec{x}^*)$ is continuous, T may still be poorly behaved. In particular, T can fail t be differentiable at such points because the linear independence constraint qualification condition fails (as two incentive constraints coincide). So, even if T is continuous, it can still have kink-discontinuities at $\vec{x}^* \in X^*$.

These jump- and kink-discontinuities in $T(\vec{x})$ can be interpreted as "forces" pushing towards pooling optima. Indeed, in the one-dimensional case, $x^* \in X^*$ will be a robust local maximum of T whenever T jumps up at x^* and $T'(x) < 0$ to the right of x^*. Symmetrically, $x^* \in X^*$ will be a robust local maximum of T whenever T jumps down at x^* and $T'(x) > 0$ to the left of x^*. Finally, $x^* \in X^*$ will be a robust local maximum if T is continuous at x^* and $T'(x) < 0$ to the right of x^* and $T'(x) > 0$ to the left.

Suppose, for example, that $x^* \in X^*$ is a local maximum of $T(x)$ because T jumps down at x^* and $T'(x) > 0$ to the left of x^*. Let $\lim_{x \uparrow x^*} T'(x) \equiv C$. Then instead of the optimization condition for the optimal tax $\vec{\tau}$ from equation (22), we will have

$$\vec{f}\tau = D + Q^{-1}\vec{f}\vec{\tau}^p - Q^{-1}(\partial X) I + Q^{-1}C. \quad (23)$$

In other words, we can think of the jump-discontinuity at the optimal pooling allocation as "adding" an extra distortion. Similarly, we can think of kink discontinuities as adding an extra distortion R.

[12]This is completely innocuous if X is one-dimensional (since it holds generically in this case), but it is a substantive restriction with a multidimensional X. With considerable notational clutter, the proposition is readily generalized to the many-tied-type case, when the same logic pushes for additional discontinuities and additional pooling.

We now provide a number of examples which illustrate both separating and pooling optima and the various components of the distortions.

5 Illustrative Examples

In this section, we return to the three applications discussed in Section 3. The first two examples—optimal taxes a la Stiglitz (1982) and insurance provision—feature separating optima. We then present two additional examples with pooling optima: a multitask-in-teams application and a second optimal tax example in the spirit of Rothschild and Scheuer (2014b).

5.1 Relation to Stiglitz (1982)

As discussed in Section 3.1, Stiglitz (1982) is essentially a two-type special case of our framework. A key difference is that Stiglitz ignores the possibility of tied wages by tacitly assuming that one type is more productive than the other (which is problematic in light of endogenous productivity). In this section, we show formally that the model discussed in Section 3.1 is indeed a special case of Stiglitz (1982). We then solve for the optimum and show that wages can *in principle* be tied, but, in practice they will not be.

To recap: There are two equal-mass types $i = 1, 2$, with wages given by $w_1(x) = \frac{1}{3}\left[\frac{1}{3} + \frac{2}{3}x^{\frac{1}{2}}\right]$ and $w_2(x) = \frac{2}{3}\left[\frac{1}{3}x^{-\frac{1}{2}} + \frac{2}{3}\right]$, and $x = X(\vec{y}) = \frac{1}{4}\left(\frac{y_2}{y_1}\right)^2$. Both types are assumed to have $u(c, y, w) = c - \frac{1}{2}\left(\frac{y}{w}\right)^2$ and a minimum utility of $\bar{u}_i = 0$. The principal's objective function is $\sum_{i=1}^{2} \frac{1}{2}(y_i - c_i)$.

Note that $w_1/w_2 = \frac{1}{2}\sqrt{x}$, and define $E_i = \frac{1}{2}\frac{y_i}{w_i}$. The consistency constraint (15), namely $y_2/y_1 = 2\sqrt{x}$ implies $y_2/y_1 = 4w_1/w_2$ and

$$x = \frac{1}{4}\left(\frac{y_2}{y_1}\right)^2 = \left(\frac{y_2/w_2}{y_1/w_1}\right) \equiv \frac{E_2}{E_1}.$$

The total income earned in the economy is:

$$Y = \frac{1}{2}(y_1 + y_2)$$
$$= w_1 E_1 + w_2 E_2$$
$$= \frac{1}{3}\left[\frac{1}{3} + \frac{2}{3}\sqrt{\frac{E_2}{E_1}}\right] E_1 + \frac{1}{3}\left[\frac{1}{3}\sqrt{\frac{E_1}{E_2}} + \frac{2}{3}\right] E_2 = \left[\frac{1}{3}E_1^{\frac{1}{2}} + \frac{2}{3}E_2^{\frac{1}{2}}\right]^2, \qquad (24)$$

whereby $w_i = \frac{\partial Y}{\partial E_i}$. In other words, the problem can equivalently be formulated as a Stiglitz economy with production function (24).

Wages are tied at $\hat{x} = 4$. For $x > \hat{x}$ (resp. $<$) $w_1(x) > w_2(x)$ (resp. $<$), and incentive compatibility requires $y_1 \geq y_2$, (resp. $\leq$). The consistency constraint, re-written as

$$y_2 = 2x^{\frac{1}{2}} y_1, \qquad (25)$$

implies $y_2 \geq 4y_1 > y_1$ for any $x \geq \hat{x} = 4$. Any value $x > \hat{x}$ is therefore infeasible for the principal, since (25) would require $y_2 > y_1$, which is incompatible with incentive compatibility. Moreover, since, per (25), $y_2 > y_1$ at $\hat{x}$, Proposition 2 implies that $T(x)$ is left-continuous at $\hat{x}$.

For $x < \hat{x}$, standard arguments show that only the downward incentive compatibility and type-1 minimum utility constraints bind. The principal's inner problem *sans* the upward incentive constraint is a concave optimization problem when re-written in terms of the utilities v_i and incomes y_i of the two types. The unique solution has

$$y_1(x) = \frac{1 + 2\sqrt{x}}{\frac{1}{w_1(x)^2} + \frac{4x}{w_2(x)^2}}, \qquad y_2(x) = \frac{4x + 2\sqrt{x}}{\frac{1}{w_1(x)^2} + \frac{4x}{w_2(x)^2}}$$

and Lagrange multipliers $\eta_{21} = \frac{1}{2}$ and $\psi_1 = 1$ respectively associated with the binding incentive constraint and the binding minimum utility constraint in the inner problem. Assumption 3 is therefore satisfied, so $T'(x)$ is well defined. Straightforward calculations using the Lagrangian formulation of the inner problem, the envelope theorem (to compute $T'(x)$ holding c_i and y_i/w_i constant), and the first

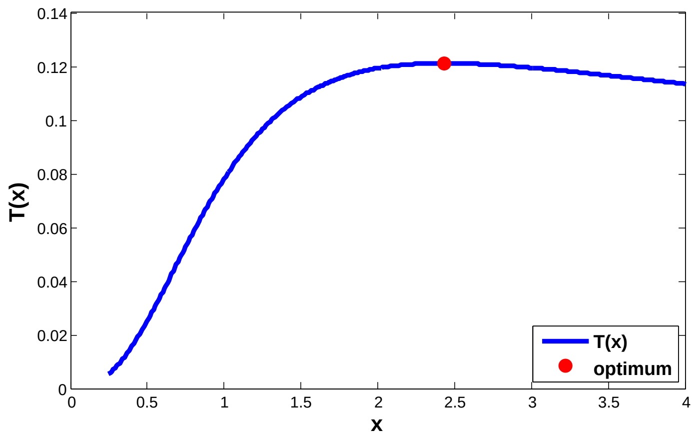


Figure 3: Value function for a Stliglitz (1982)-like separating optimum example

order conditions for the inner problem imply

$$T'(x) = \frac{y_1^2}{w_1^3}\frac{dw_1}{dx} + \frac{1}{2}\frac{y_2^2}{w_2^3}\frac{dw_2}{dx} - \frac{1}{4}\left(\frac{y_2}{w_2^2} - 1\right)\frac{y_2}{x}. \tag{26}$$

Hence,

$$\begin{aligned}\lim_{x \uparrow \hat{x} = 4} T'(x) &= \frac{y_1}{w_1^3}\frac{dw_1}{dx}\left(1 + \frac{1}{2}\left(\frac{y_2}{y_1}\right)^2\left(\frac{dw_2/dx}{dw_1/dx}\right)\right) + \frac{y_2}{4x}\left(1 - \frac{y_2}{w_2^2}\right) \\ &< \frac{y_1}{w_1^3}\frac{dw_1}{dx}\left(1 + \frac{1}{2}(4x)\left(-x^{-1}\right)\right) < 0.\end{aligned}$$

The optimum thus occurs for some $x < \hat{x}$ (with $x \geq \frac{1}{4}$ for consistency). It is safe, per Stiglitz (1982) to assume that $w_1 < w_2$ when computing the optimum. Indeed, $T(x)$ is smooth throughout the domain of consistent x's. It is depicted in Figure 3.

5.2 Insurance with Endogenous Avoidance Costs

Consider next the insurance example from section 3.3, with the following additional assumptions:

- The risk avoidance disutility function is given by $h(p, \theta) = \theta/p$.
- The parameter θ depends endogenously on some x via $\theta_k = \theta_k(x) = \alpha_k x$.

- The value of x is endogenously given by a type-weighted average of the risk avoidance efforts: $x = \sum_i f_i p^*(\Delta_i, \theta_i)$.

Individual optimization for any given $V_{NL} \equiv u(c_{NL})$ and $\Delta \equiv u(c_{NL}) - u(C_L)$ implies $p^*(\Delta, \theta) = (\theta/\Delta)^{\frac{1}{2}}$ and $U(V_{NL}, \Delta, \theta) = \max_p \{V_{NL} - p\Delta - h(p, \theta)\} = V_{NL} - 2\sqrt{\theta\Delta}$. Hence, $x = \sum_i f_i(\theta_i/\Delta_i)^{\frac{1}{2}} = \sqrt{x}\sum_i f_i(\alpha_i/\Delta_i)^{\frac{1}{2}}$, and we can equivalently write

$$x = X(\vec{\Delta}) = \left(\sum_i f_i \left(\alpha_i/\Delta_i\right)^{\frac{1}{2}}\right)^2.$$

In other words, this example is nested within our general framework, with θ_i taking the role of w_i, and (V_{NL}, Δ) playing the role of (c, y).[13]

Note that $\frac{\theta_i}{\theta_j} = \frac{\alpha_i}{\alpha_j}$, so wage overlap of distinct types is impossible here: preferences are endogenous, but the ordering of preferences coincides with the ordering of the exogenous α parameters.[14]

5.3 An Optimal Tax Model with a Jump-Pooling Optimum

The preceding examples involved separating optima. We now further develop the simple three-type tax model with a pooling optimum discussed in the introduction.

Suppose the economy consists of three equal measure of workers with $u(c, y, w) = c - (y/w)^3$. Let $x = X(\vec{y}) = \left(\frac{y_2}{y_1+y_0}\right)^{1/\alpha}$ for some $\alpha < 0$. Let wages be given by $w_0(x) = \frac{1}{2}\left[\frac{1}{2}\left(1 + x^{\alpha}\right)\right]^{\frac{1-\alpha}{\alpha}}$, $w_1(x) = \left[\frac{1}{2}\left(1 + x^{\alpha}\right)\right]^{\frac{1-\alpha}{\alpha}}$, and $w_2(x) = \left[\frac{1}{2}\left(1 + x^{-\alpha}\right)\right]^{\frac{1-\alpha}{\alpha}}$. Work in the dual formulation with a social planner with a Rawlsian social welfare function and a balanced budget requirement. Since utility is quasilinear, it is straightforward to show that this is equivalent to a primal formulation with a linear objective and some common reservation utility $\bar{u}$ for all three types.

Figure 2 in the introduction plots the value function $T(x)$ for this problem for $\alpha = -0.1$. The figure clearly indicates that the optimum occurs at a jump-

[13]Strictly speaking, this nesting only works if we *exogenously* assume that all individuals of the same type receive the same contract rather than rely on Lemma 1. This is because the average within a type of (c, y) is not a sufficient statistic for the firms' profits.

[14]Ales et al. (2015) discuss a generalized discrete type Stiglitz (1982) model with the same property: wages are endogenous, but because of an absolute advantage assumption, the ordering of types is exogenously fixed and wage overlap is impossible. In an appendix, they also treat a continuum-of-types case with wage overlap.

discontinuity at the point $x^* = 1$, where $w_2(x^*) = w_1(x^*)$.[15] The mechanical intuition for this jump-discontinuity is straightforward, and mirrors the intuition from Figure 1 in the introduction. To wit: at $x = 1$, $y_2 = y_1 + y_0$. It turns out to be optimal to have $y_0 > 0$; hence, $y_2 > y_1$ at $x = 1$. Since w_2 is decreasing and w_1 increasing in x, $y_2 \leq y_1$ for any $x > 1$. The optimum at $x^* = 1$ thus cannot be approximated by any feasible sequence of allocations as $x \to 1$ from above—so the value function necessarily jumps down at $x^* = 1$.

The underlying intuition is also reasonably straightforward. It is easiest to see in an equivalent Stiglitz (1982)-like formulation of the same problem, namely: a problem with the production function $Y(E_\theta, E_\phi) = [\frac{1}{2}E_\phi^\alpha + \frac{1}{2}E_\theta^\alpha]^{1/\alpha}$, where $E_\theta = \sum_i \theta_i e_i^\theta$ and $E_\phi = \sum_i \phi_i e_i^\phi$ are the aggregate *efforts* in two sectors, and where there are three equal measure worker types with skills $(\theta, \phi)_0 = (0, 1)$, $(\theta, \phi)_1 = (0, 2)$, $(\theta, \phi)_2 = (2, 0)$. We omit the straightforward calculations, akin to those in section 5.1, establishing equivalence.

In this alternative formulation, it is clear from the production function that type 2's effort increases the wages of the other two types while reducing the wages of other type 2s—and vice-versa for types 0 and 1. Intuitively, this provides an extra incentive to "up-distort" 2s effort and earnings and "down-distort" the effort and earnings of type 0 and type 1, in order to indirectly redistribute towards the worst-off type 0s. At x^* (and to the left of it), this incentive manifests in $y_2 > y_1$. To the right of x^*, $y_2 \leq y_1$ by incentive compatibility, and the objective $T(x)$ drops discretely.

Towards unpacking this intuition and relating it to the formalism in the preceding section, Figure 4 plots the optimal tax distortions on types 1 and 2 from the inner problem, as a function of x, just to the left of the optimum x^*. The dashed lines show the total tax distortion on the two types. Per the preceding intuition, type 1 faces a positive tax—so her earnings are distorted down—while type 2 faces

[15] The figure omits social welfare for $x > \bar{x}$, where $\bar{x}$ is the x at which $w_2 = w_0$. It can be verified, per the following argument, that such solutions yield strictly lower social welfare. For $x > \bar{x}$, $y_1 \geq y_0 \geq y_2$ and $c_1 \geq c_0 \geq c_2$ by incentive compatibility. Hence, $c_2 < Y/3$, where Y is total output. The consistency constraint for $\alpha = -0.1$ and $y_2 > 0$ then requires $x = ((y_1 + y_0)/y_2)^{10} > 2^{10}$ (hence the omission from Figure 2). Type 2's utility, $c_2 - e_2^3$, is no greater than $Y/3 - e_2^3 = 3Ke_2 - e_2^3$, where $K(x) = \frac{2}{27}(.5 + .5x^{-\alpha})^{1/\alpha}$ as is computed from an equivalent Stiglitz (1982)-like formulation of the same problem (see following text). Using the fact that $K'(x) < 0$ and $x > 2^{10}$, we can bound type 2's utility for $x > \bar{x}$ by $\max_{e_2} K(x)e_2 - e_2^3 = 2K(x)^{1.5} \leq 2K(2^{10})^{1.5} = .0026$, which is clearly below the optimum in Figure 2.

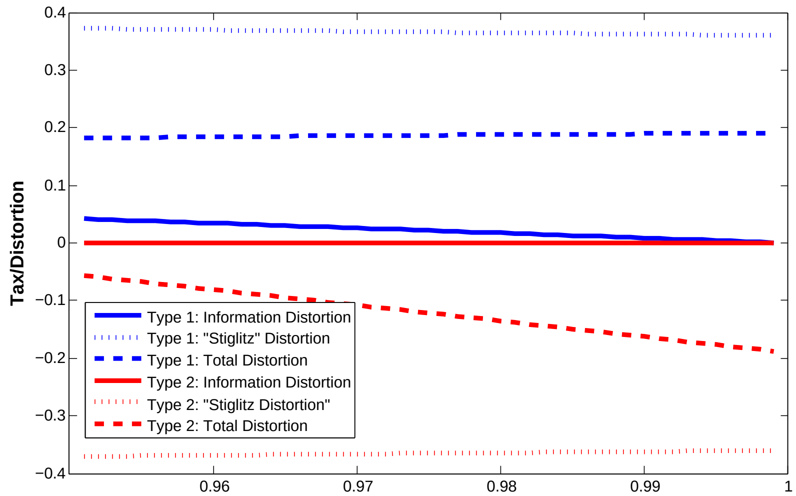


Figure 4: Tax distortion components near $x^* = 1$

a negative tax and has up-distorted earnings.

The figure also decomposes this total distortion into the components per equations (22) and (23) (divided through, componentwise, by f_i). Since individuals are are paid their marginal product of effort in this example, the Pigouvian component $\vec{\tau}^p$ is identically zero, and it is omitted from the plot. Recall that the component D, graphed as the solid lines in Figure 4, is a standard information distortion. It corresponds to the distortion that would appear in a model without endogenous preferences. The information distortion for type 2 is identically 0. Since type 2 is the highest wage type over the plotted range, this corresponds to the standard "no distortion at the top" result. Type 1, who is the the intermediate wage type, faces a positive information distortion, as is to be expected with a progressive social welfare function. This distortion converges to 0 as $x \to 1$ and type 1's wage converges to type 2's.

The "Stiglitz" distortion corresponding to the I term in equation (22) is plotted as the dashed lines. It is large in magnitude for both types, but with opposite signs. Intuitively, distorting type 2 towards higher effort compresses the wage distribution and beneficially redistributes towards the worst-off type 0. Similarly, distorting type 1 towards lower effort beneficially compresses the wage distribution.

It is clear from the figure that the *total* distortion for both types is smaller in magnitude than the sum of the three component distortions from equation (22).

This reflects the fact that $T'(x) > 0$ (*viz* Figure 2). Per equation (23) and the surrounding discussion, the $T' > 0$ can be interpreted as an additional component contributing to the total tax distortion. We infer that this additional distortion partially but not completely offsets the sum of the other distortions in this model. This implies that, at the optimum x^*, the effective marginal tax rates faced by the workers are blunted relative to what one would expect based on information and Stiglitz (and Pigouvian) distortions alone. This is consistent with Rothschild and Scheuer (2013), who show that the wage overlap effect partially (but not completely) offsets the Stiglitz effect in their continuum of types setting.

5.4 A Multitask Example with a Kink-Pooling Optimum

Next, we illustrate another type of pooling optimum—one that occurs with a value function $T(x)$ which is continuous but kinked at the optimal pooling-wage x^*. We first consider the model without explicitly connecting it with a particular application. We then show that it can be re-interpreted as a Multitask-in-Teams model, per Section 3.2.

Suppose that $x \in \mathbb{R}$, and there are three equal-mass types with $w_1(x) = \frac{19}{20} + \frac{1}{10}\frac{1}{1+e^{4-x}}$, $w_2(x) = 2 - w_1(x)$, $w_3(x) = 2$, that $x = y_3$, that $u(y,c,w) = c - \frac{1}{2}\left(\frac{y}{w}\right)^2$, and that $\bar{u}_i = 0$.

The unique x^* at which wages are tied is $x = 4$, where $w_1(4) = w_2(4) = 1 < w_3$. Figure 5 plots the value function $T(x)$ for the inner problem. It indicates that $T(x)$ is continuous and is maximized at x^*. The figure also plots the left and right subtangents at x^*, both of which are non-zero: $T(x)$ is kinked, with a strictly positive derivative to the left and a strictly negative derivative to the right.

The intuition for the kink in this example is simple. First, the solution to the inner problem at x^* has $y_1^* = y_2^* = 8/11 < w_1^2 = 1$, and $y_3^* = 4$. That is, the solution is *degenerate*; the lower wage types are *bunched*. This bunching is to be expected: neither of the types affects x, so by standard arguments there is no reason to pull their allocations apart. Using an envelope argument, we can compute the effects on $T(x)$ of a small change in x while holding type 1 and 2's allocation fixed, and adjusting $y_3 = x$ and c_3 to maintain incentive compatibility. Since type 3 is undistorted at x^*, the adjustment of type 3's allocation has only second order revenue effects. The only first order welfare effects of this change come from the fact that a local ad-

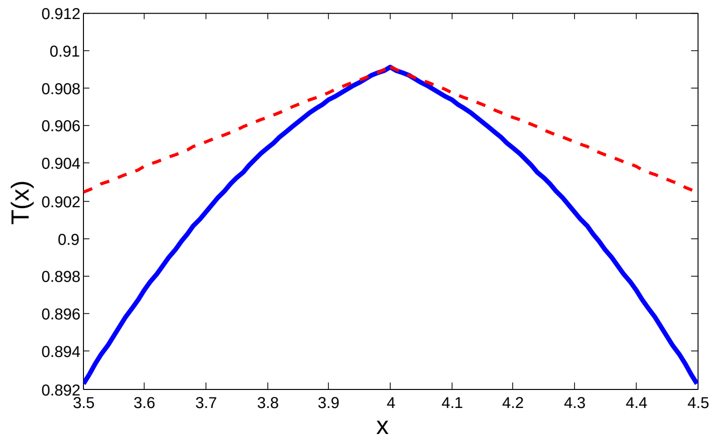


Figure 5: A continuous, kinked value function and its left and right subtangents

justment to x will lower *either* type 1 or type 2's wage to $w_i(x^*) - dw_i$, and hence will violate the participation constraints by $dw_i\left[(y_i^*)^2/w_i(x^*)^3\right] > 0$. In particular, this means that the left-hand derivative is $T'_{left}(x^*) = \frac{dw_1}{dx}\left[(8/11)^2\right] \approx .013 > 0$ and $T'_{right}(x^*) = \frac{dw_2}{dx}\left[(8/11)^2\right] \approx -.013 < 0$. Pulling apart their wages—in either direction—is bad, precisely because it is the *lower* of the two wages which is welfare relevant.

5.4.1 A Multitask Interpretation

Suppose that $\vec{e} = (e_1, e_2, e_3)$—so that there are three potential tasks. Suppose further that the three types of workers respectively have the linear-homogenous effort disutility functions

$$m_1(\vec{e}) = \left((k(e_1)^2 + (e_2)^2)^{\frac{1}{2}} + Me_3\right)^2 \tag{27}$$

$$m_2(\vec{e}) = \left(((e_1)^2 + k(e_2)^2)^{\frac{1}{2}} + Me_3\right)^2 \tag{28}$$

$$m_3(\vec{e}) = (Me_1 + Me_2 + e_3)^2, \tag{29}$$

Now consider the solutions to each worker's problem of minimizing m subject to earning a given income $y = \sum \beta_i e_i$ for some fixed $\vec{\beta} = (\beta_1, \beta_2, \beta_3)$. For sufficiently large M, type 1 and 2 workers will have $e_1 > 0$, $e_2 > 0$, and $e_3 = 0$, while type 3 workers will have $e_1 = e_2 = 0$ and $e_3 > 0$. Moreover, letting $m_i^*(y)$ denote the minimal m_i given y, we have

$$w_1(\vec{\beta}) \equiv \frac{y}{m_1^*(y)} = \left(\beta_1^2 + \frac{1}{k}\beta_2^2\right)^{\frac{1}{2}} \tag{30}$$

$$w_2(\vec{\beta}) \equiv \frac{y}{m_2^*(y)} = \left(\frac{1}{k}\beta_1^2 + \beta_2^2\right)^{\frac{1}{2}} \tag{31}$$

$$w_3(\vec{\beta}) \equiv \frac{y}{m_3^*(y)} = \beta_3. \tag{32}$$

Now endogenize $\vec{\beta}$ as follows. Define $g(x) \equiv \frac{19}{20} + \frac{1}{10}\frac{1}{1+e^{4-x}}$ and let:

$$\beta_1(x) = \left(\frac{k^2 g(x)^2 - k(2-g(x))^2}{k^2-1}\right)^{\frac{1}{2}} \tag{33}$$

$$\beta_2(x) = \left(\frac{k^2(2-g(x))^2 - kg(x)^2}{k^2-1}\right)^{\frac{1}{2}} \tag{34}$$

$$\beta_3(x) = 2. \tag{35}$$

For $k = (22/19)^2$, it is easily verified that $\beta_i > 0$ for all x and that $w_1(\vec{\beta}(x)) = g(x)$, $w_2(\vec{\beta}(x)) = 2 - g(x)$, and $w_3(x) = 2$.[16] This multitask problem thus collapses to the preceding example.

6 Relation to Rothschild and Scheuer (2014b)

In this section, we relate the effects identified in section 4 to the effects identified in Rothschild and Scheuer's (2014b, henceforth RS) model of optimal taxation with general, continuously distributed, multidimensional types.

In their framework, preferences are given by $u(c, m(\vec{e}))$, where m is a homogenous of degree one strictly quasiconvex effort aggregator and $\vec{e}$ is a K-vector of activity-specific efforts with k^{th} element e^k. Total effort in each activity (in a discrete-type version of their model) is given by $E_k = \sum_i f_i e_i^k$ ($\vec{E}$ in vector notation). Types are described by a K-vector θ_i of activity-specific abilities. The pre-tax income earned in activity k for type i is given by $y_i^k = \theta_i^k r_k(\vec{E}) e_i^k$, and total income y_i is

[16]Moreover, since the β_i are bounded, we can choose M large enough to ensure that types 1 and 2 will only pursue activities 1 and 2 while worker 3 will only pursue activity 3 for all x.

$\sum_k y_i^k$. The social planner can observe total income y_i (and after tax income c_i) but not abilities, efforts, or the components y_i^k.

RS show that this problem can be re-expressed as an optimal tax problem with endogenous wages $w_i(\vec{E})$ and consistency constraints

$$\frac{1}{r^k(\vec{E})} \sum_i f_i q_i^k(\vec{E}) y_i = E^k, \quad k = 1, \cdots, K \tag{36}$$

for some functions $q_i^k(\vec{E})$, which are interpreted as activity k income shares for type i.

The essential difference between their framework and the present one is that the wages and consistency constraints depend on the (unobserved) vector of aggregated efforts E rather than the observed incomes y_i. Under suitable invertibility assumptions, we can connect the two by using the consistency constraints to implicitly define $E = X(\vec{Y})$, so that E plays the role of x in the present paper.

Specifically, write (36), stacked into a K-vector, as $\hat{X}(\vec{E}, \vec{Y}) = \vec{E}$, and define $X(\vec{Y})$ implicitly as the solution to $\hat{X}(X(\vec{Y}), \vec{Y}) = X(\vec{Y})$. By the Implicit Function Theorem, $\partial \hat{X} = (\partial X) J$, where J is the (assumed invertible) $K \times K$ matrix with elements $\delta_{mk} - \frac{\partial \hat{X}^m}{\partial E^j}$, and where $\partial \hat{X}$ is the $N \times K$ matrix with elements $\frac{\partial \hat{X}^k}{\partial y_i}$, and, as above, ∂X is the $N \times K$ matrix with elements $\frac{\partial X^k}{\partial y_i}$ (and where $\delta_{mk} = 1$ is the Dirac delta).

There are two ways of formulating the consistency constraints: as $\hat{X}(\vec{E}, \vec{Y}) = \vec{E}$ as in RS, or $X(\vec{Y}) = E$, as in the present paper's notation. Let $\vec{\zeta}$ be the associated Lagrange multipliers in the former formulation and, as in Lemmas 4 and 5 above, let $\vec{\mu}$ be the multiplier in the latter formulation. These multipliers are mechanically related via $\vec{\mu} = J\vec{\zeta}$. From Lemma 4, we can thus use the first order conditions for the inner problem to write the optimal correction for a separating optimum

$$\vec{f\tau} = -\left(\partial \hat{X}\right) \vec{\zeta} + \vec{D}. \tag{37}$$

Note that $(\partial \hat{X})_{ik} = f_i q_i^k / r^k$, so the the k^{th} element of $-\left(\partial \hat{X}\right) \vec{\zeta}$ is $f_i \sum_k q_i^k (\zeta_k / r^k)$. Hence, exactly as in RS, Proposition 1, the optimal correction on income level y_i is the standard (incentive constraint-based) correction D modified by an sectoral-income-share weighted average of the modified multipliers ζ_k / r^k.

The logic of Lemma 5 can similarly be used to write the derivative $\nabla T(\vec{E})$ from the outer problem in simpler terms that accord closely with RS's Lemma 7.

Lemma 6. *If ∇T is differentiable and each type has a distinct wage, then:*

$$\nabla T(\vec{E}) = (S - \mathbb{I}_K)\vec{\xi} - \tilde{\tau}^p + I, \tag{38}$$

where S is the $K \times K$ matrix with elements $S_{mk} \sum_i \frac{\theta_i^k}{w_i(E)} \frac{d}{dE^m} \left(m(\vec{e}_i / e_i^k)\right)^{-1}$, and where $\tilde{\tau}^p$ is the K vector with elements $\tilde{\tau}_k^p = -\sum f_i \left(-\frac{u_w^i}{u_y^i}\right) \frac{dw_i}{dE^k}$.

Proof. See Appendix B.3. □

Note that $\tilde{\tau}_k^p$ is just the (negative of the) change in output induced by a unit change in E_k, so it can be interpreted as a Pigouvian externality on *effort* in sector k. (As opposed to the Pigouvian correction on income from Lemma 5.) The term S is interpretable as a sectoral shift effect: it measures the extent to which individuals shift their sectoral effort allocations in response to the changes in relative sectoral returns induced by a change in $\vec{E}$.

Up to minor notational differences, this matches the outer problem results in RS, Lemma 7, in the special case where RS's "overlapping wage" terms C and R are zero. The absence of C and R-like terms in Lemma 6 is to be expected: it applies away from X^*, where there is no wage overlap.

When the optimum occurs at a point in X^* where wages *do* overlap, the discussion at the end of section 4.4 and examples in sections 5.3 and 5.4 indicate that C and R-like terms can "reappear" respectively when T has a jump or a kink discontinuity. To elaborate, recall that the extra terms arise in RS when wages overlap because a change in $\vec{E}$ will differentially effect the wages of the originally pooled individuals. The originally pooled individuals will be "pulled apart." The term R arises if there are different welfare weights assigned to the originally pooled individuals, so that the change in $\vec{E}$ effectively redistributes in welfare-relevant ways. The term C arises because the originally pooled individuals move along the tax schedule in different directions, and thus see their aggregated efforts $m(\vec{e})$ pulled apart. Since these individuals exert their effort in different activities, this effectively redistributes effort across activities. In other words, C is non zero precisely insofar as a change in $\vec{E}$ will indirectly redistribute effort across activities and thereby affect binding consistency constraints.

To see the connection to pooling optima in the present framework, consider first the kink discontinuity example in section 5.4. Viewed on the domain $x \in [0, x^*]$, the model is, by strong duality, equivalent to an optimal tax model with all social welfare weight placed on type 1 individuals. On this domain, lowering x from x^* is "bad" for the social planner precisely because it is pulling two tied types apart, and pulling them apart in the "wrong" direction: it effectively redistributes from the type 1s (who have all the welfare weight) to the type 2s (who have none of it). So (on the left) $\nabla T > 0$. Symmetrically, above x^*, the model is equivalent to one with all welfare weight on type 2, and, again, moving away from x^* in the upward direction is "bad" because it involves redistributing from type 2 to type 1. In other words, the extra wedge away from $\nabla T = 0$ arises because the two types have different welfare weights and are pulled apart (in an undesirable direction) by a change in x—just as in the R term in RS.

As discussed above, jump discontinuities—and hence jump-pooling optima as in section 5.3—arise because of an incentive to separate the incomes (and hence $m(\vec{e})$) of the two tied-wage types. This desire to separate arises precisely because the two types have differential effects on a binding consistency constraints (15)—again, exactly as in the intuition for the term C in RS.

7 Discussion and Conclusion

We view this paper as making two important contributions.

First, we provide a general technology for studying a wide range of important economic problems where screening is employed and where preferences among the agents being screened are intertwined through contracting. This technology is general enough to be widely applicable and simple enough to be used in practical applications.

Second, we provide insight into the important differences between discrete and continuous-type models when preferences are endogenous—differences that are less important with exogenous preferences. Discrete-type models are frequently easier to work with, especially when one can abstract from pooling of types with different characteristics. Our analysis illustrates that, when preferences are endogenous to contracts, such an abstraction is not typically justified. Indeed, there are two distinct types of discontinuity that occur when types endogenously have the same preferences, and both discontinuities can push a principal to optimally select such pooling points. As such, endogenous pooling is likely to be common.

References

ALES, L., M. KURNAZ, AND C. SLEET (2015): "Technical Change, Wage Inequality and Taxes," *American Economic Review*, To appear.

BAKER, G. (2002): "Distortion and Risk in Optimal Incentive Contracts," *Journal of Human Resources*, 37, 728–751.

BLACKORBY, C., C. BRETT, AND A. CEBREIRO (2007): "Nonlinear Taxes for Spatially Mobile Workers," *International Journal of Economic Theory*, 3, 57–74.

BOND, E. W., AND K. J. CROCKER (1991): "Smoking, skydiving, and knitting: The endogenous categorization of risks in insurance markets with asymmetric information," *Journal of Political Economy*, pp. 177–200.

BORDER, K.C. (1989): *Fixed Point Theorems with Applications to Economics and Game Theory*. Cambridge University Press, Cambridge.

BRETT, C., AND J. WEYMARK (2008): "The impact of changing skill levels on optimal nonlinear income taxes," *Journal of Public Economics*, 92(7), 1765–1771.

——— (2011): "How Optimal Nonlinear Income Taxes Change when the Distribution of the Population Changes," *Journal of Public Economics*, 95.

GUESNERIE, R., AND J. SEADE (1982): "Non-linear Pricing in a Finite Economy," *Journal of Public Economics*, 17, 146–157.

HOLMSTRÖM, B., AND P. MILGROM (1991): "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, and Organization*, 7, 24–52.

KYPARISIS, J. (1985): "On Uniqueness of Kuhn-Tucker Multipliers in Nonlinear Programming," *Mathematical Programming*, 32, 242–246.

MIRRLEES, J. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38, 175–208.

MORAND, O., K. REFFETT, AND S. TARAFDAR (2015): "A Nonsmooth Approach to Envelope Theorems," *SSRN Discussion Paper 2609292*.

MUSSA, M., AND S. ROSEN (1978): "Monopoly and product quality," *Journal of Economic theory*, 18(2), 301–317.

ROTHSCHILD, C., AND F. SCHEUER (2013): "Redistributive Taxation in the Roy Model," *Quarterly Journal of Economics*, 128, 623–668.

——— (2014a): "Optimal Taxation with Rent-Seeking," *Mimeo, Stanford University and Wellesley College*.

——— (2014b): "A Theory of Income Taxation under Multidimensional Skill Heterogeneity," *Mimeo, Wellesley College and Stanford University*.

ROTHSCHILD, M., AND J. STIGLITZ (1976): "Equilibrium in Competitive Insurance Markets: An Essay in the Economics of Incomplete Information," *Quarterly Journal of Economics*, 90, 629–649.

SEGAL, I. (1999): "Contracting with externalities," *Quarterly Journal of Economics*, pp. 337–388.

SIMULA, L. (2010): "Optimal nonlinear income tax and nonlinear pricing: optimality conditions and comparative static properties," *Social Choice and Welfare*, 35(2), 199–220.

STIGLITZ, J. (1982): "Self-Selection and Pareto Efficient Taxation," *Journal of Public Economics*, 17, 213–240.

WERNING, I. (2007): "Pareto-Efficient Income Taxation," *Mimeo, MIT*.

WEYMARK, J. (1986): "A reduced-form optimal nonlinear income tax problem," *Journal of public Economics*, 30(2), 199–217.

——— (1987): "Comparative Static Properties of Optimal Nonlinear Income Taxes," *Econometrica*, 55, 1165–1185.

A Proofs for the inner problem

A.1 Proof of Lemma 2

Note that $\Gamma(\vec{x})$ is closed since it's the intersection of closed sets. It suffices to show that any feasible allocations outside of some bounded subset of $\Gamma(\vec{x})$ yields a lower value of $\sum(y_i - c_i)f_i$ than some allocation within that bounded subset.

To that end, define

$$t_i \equiv \max_{(y_i,c_i)} y_i - c_i \quad \text{s.t.} \quad u(c_i, y_i, w) \geq \bar{u}_i, \quad \text{and}$$

$$t_{min} \equiv \max_{(y,c)} y - c \quad \text{s.t.} \quad u(c, y, w_i) \geq \bar{u}_i \; \forall i,$$

which, by Assumption 1, have some well-defined, finite solutions $(\hat{y}_i, \hat{c}_i)$ and (y_{min}, c_{min}), respectively. (The latter because the lower envelope of the set of (y, c) satisfying $u(c, y, w_i) \geq \bar{u}_i \; \forall i$ is the convex upper envelope of the set of (y, c) satisfying $u(c, y, w_i) \geq \bar{u}_i$.) Since the allocation that assigns $(y_i, c_i) = (y_{min}, c_{min}) \; \forall i$ is trivially incentive compatible, t_{min} is a lower bound on the value $T(\vec{x})$ of the inner problem. Similarly, t_i is an *upper* bound on the resources that can be extracted from i types. Together, these implies that any allocation for which $(y_i - c_i)f_i < t_{min} - \sum_{j \neq i}(y_j - c_j) \equiv t^i_{min}$ for any i is sub-optimal.

Let $\hat{c}(y)$ solve $u(\hat{c}(y), y, w_i) = \bar{u}_i$ (for any y for which it is well defined). By assumption 1 there are two values y_i^{max} and $y_i^{min} \leq y_i^{max}$ satisfying $y_i^{max} - \hat{c}(y_i^{max}) = y_i^{min} - \hat{c}(y_i^{min}) = t^i_{min}$. Then we know that feasible allocations with $y_i > y_i^{max}$ or $c_i < \hat{c}\left(y_i^{min}\right)$ have $\sum(y_i - c_i)f_i < t_{min}$ These, together with 13, respectively imply that allocations with $c_i > \hat{c}\left(y_i^{max}\right)$ or $y_i < y_i^{min}$ also have $\sum(y_i - c_i)f_i < t_{min}$, which is sub-optimal. Hence, we can restrict attention to the bounded subset $\Pi_i \left[y_i^{min}, y_i^{max}\right] \times [\hat{c}(y_i^{min}), \hat{c}(y_i^{max})]$ of $\Gamma(\vec{x})$.

A.2 Proof of Lemma 3

1. Assume WLOG for some types i and j, we have $c_i \geq c_j$ and $y_i < y_j$. This cannot be incentive compatible since $u(c_j, y_j, w_j) \leq u(c_i, y_j, w_j) < u(c_i, y_i, w_j)$.

2. Let $S = (R^2, \preceq)$ be the partially ordered set on R^2 endowed with the lexicographic order $\preceq$, with $(c, y) \preceq (c', y')$ if and only if either $c < c'$ or $c = c'$ and $y < y'$. Let T be the partially ordered set on R endowed with the usual order. By 1, any incentive compatible allocation $\{(c_k, y_k)_{k=1,2,...,N}\}$ is a chain on S. Let $f(s, w) : S \times R \to R$ be defined as $f(s, w) = u(c, y, w)$

where $s = (c, y)$. By Theorem 3 in Milgrom and Shannon (1994), $f(s, w)$ has the single crossing property in $(s; w)$ if and only if $u(c, y, w)$ satisfies the Spence-Mirrlees condition. Since f is (trivially) quasi-supermodular in s and satisfies the single-crossing property in $(s; w)$, the Monotone Selection Theorem (Milgrom and Shannon (1994)) implies $argmax_{x \in S} f(x, w)$ is monotone nondecreasing in w.

3. We have $u(c_i, y_i, w_i) \geq u(c_j, y_j, w_i) \geq u(c_j, y_j, w_j)$ and $-u(c_k, y_k, w_i) \geq -u(c_k, y_k, w_k) \geq -u(c_k, y_k, w_j)$, where the first inequality follows from incentive compatibility and the second follows from $u_w > 0$. Summing up gives

$$u(c_i, y_i, w_i) - u(c_k, y_k, w_i) \geq u(c_j, y_j, w_j) - u(c_k, y_k, w_j) \geq 0$$

where the last step follows from type j's incentive constraint. Claim 4 can be proven analogously.

A.3 Proof of Lemma 4

The Lagrangian of the inner problem is

$$\begin{aligned} L(\vec{c}, \vec{y}; x) = \sum_i (y_i - c_i) f_i + \sum_i \xi_i (u(c_i, y_i, w_i(x)) - \bar{u}_i) + \vec{\mu}^T (X(\vec{Y}) - x) \\ + \sum_{i>0} \eta_{i,i-1} [u(c_i, y_i, w_i(x)) - u(c_{i-1}, y_{i-1}, w_i(x))] \\ + \sum_{i<N} \eta_{i,i+1} [u(c_i, y_i, w_i(x)) - u(c_{i+1}, y_{i+1}, w_i(x))] \end{aligned} \tag{39}$$

The first order conditions with respect to c_i and y_i are, respectively,

$$\begin{aligned} f_i = \xi_i u_c(c_i, y_i, w_i(x)) + \eta_{i,i-1} u_c(c_i, y_i, w_i(x)) + \eta_{i,i+1} u_c(c_i, y_i, w_i(x)) \\ - \eta_{i-1,i} u_c(c_i, y_i, w_{i-1}(x)) - \eta_{i+1,i} u_c(c_i, y_i, w_{i+1}(x)), \end{aligned} \tag{40}$$

and

$$\begin{aligned} -f_i - \vec{\mu}^T \overrightarrow{X_{y_i}}(\vec{Y}) = \xi_i u_y(c_i, y_i, w_i(x)) + \eta_{i,i-1} u_y(c_i, y_i, w_i(x)) + \eta_{i,i+1} u_y(c_i, y_i, w_i(x)) \\ - \eta_{i-1,i} u_y(c_i, y_i, w_{i-1}(x)) - \eta_{i+1,i} u_y(c_i, y_i, w_{i+1}(x)) \end{aligned} \tag{41}$$

Multiplying equation (40) by $MRS_i = -u_y(c_i, y_i, w_i)/u_c(c_i, y_i, w_i)$, adding to equation (41), and re-arranging yields

$$\begin{aligned} f_i (1 - MRS_i) = - \vec{\mu}^T \overrightarrow{X_{y_i}}(\vec{Y}) + \eta_{i-1,i} u_c(c_i, y_i, w_{i-1}) \left(MRS_i - \widehat{MRS}_i^{i-1} \right) \\ + \eta_{i+1,i} u_c(c_i, y_i, w_{i+1}) \left(MRS_i - \widehat{MRS}_i^{i+1} \right), \end{aligned} \tag{42}$$

which is equivalent to (17)

B Proofs for the Outer and Overall Problems

B.1 Proof of Lemma 5

Differentiating the Lagrangian with respect to x_k, holding constant c_i and $z_i \equiv u(c_i, y_i, w_i(\vec{x}))$ for all i yields

$$\frac{\partial T(\vec{x})}{\partial x_k} = \sum_j \mu_j \sum_i \frac{\partial X_j}{\partial y_i} \frac{\partial y_i}{\partial w_i} \frac{\partial w_i}{\partial x_k} - \mu_k - \sum_i f_i \frac{u_w^i}{u_y^i} \frac{\partial w_i}{\partial x_k} + \sum_i \sum_{\substack{j \in \\ \{i-1, i+1\} \\ \cap \{1, \ldots, N\}}} \eta_{j,i} u_y(c_i, y_i, w_j) \left[\widehat{\left(\frac{u_w^i}{u_y^i} \right)_j} \frac{\partial w_j}{\partial x_k} - \left(\frac{u_w^i}{u_y^i} \right) \frac{\partial w_i}{\partial x_k} \right] \tag{43}$$

The $(k, j)^{\text{th}}$ element of $\bar{C}(\partial X)$ is $\sum_i \left(-\frac{u_w}{u_y} \right)^i \left(\frac{\partial w_i}{\partial x_k} \right) \frac{\partial x_j}{\partial y_i}$, so the k^{th} element of $\bar{C}(\partial X) \vec{\mu}$ is $\sum_j \sum_i \left(-\frac{u_w}{u_y} \right)^i \left(\frac{\partial w_i}{\partial x_k} \right) \frac{\partial x_j}{\partial y_i} \mu_j$, i.e., the first term in (43).

Stacking the K equations in (43) thus yields

$$\nabla T(\vec{x}) = \bar{C}\left(\partial X\right) \vec{\mu} - \vec{\mu} + \sum f_i \left(-\frac{u_w^i}{u_y^i} \right) \nabla w_i(\vec{x}) + I \tag{44}$$

The lemma follows from the fact that $\vec{f} \vec{\tau}^p = -\left(\partial X\right) \sum_i f_i \left(-\frac{u_w^i}{u_y^i} \right) \nabla w_i(\vec{x})$, as is readily verified by an explicit computation.

B.2 Proof of Proposition 2

Consider an $\vec{x}^* \in X^*$ at which $w_i(\vec{x}^*) = w_j(\vec{x}^*)$. We claim that $\Gamma(\vec{x})$ is not lower hemicontinuous at $\vec{x}^*$ if $\vec{x}^*$ is usual and nondegenerate. Specifically, we show that lower hemicontinuity fails at $\vec{x}^*$ along the ε-parameterized line $\vec{x}(\varepsilon) = \vec{x}^* + \varepsilon \nabla \left(\frac{w_i(\vec{x})}{w_j(\vec{x})} \right)$.

To see this, suppose (without loss of generality given Assumption 2) that $\nabla \left(\frac{w_i(\vec{x})}{w_j(\vec{x})} \right) > 0$, i.e., that $w_i(\vec{x}(\varepsilon)) - w_j(\vec{x}(\varepsilon))$ is increasing in ε at $\vec{x}^*$.

Define $\Gamma_1(\vec{x}) = \{(\vec{y}, \vec{c}) \in \Gamma(\vec{x}) | \, c_i \geq c_j\}$ and $\Gamma_2(\vec{x}) = \{(\vec{y}, \vec{c}) \in \Gamma(\vec{x}) | \, c_i \leq c_j\}$, and let $T_1(\vec{x})$ and $T_2(\vec{x})$ be the value function associated with replacing the constraint set $\Gamma(\vec{x})$ respectively by $\Gamma_1(\vec{x})$ or $\Gamma_2(\vec{x})$.

By Lemma 3, we have $\Gamma_1(\vec{x}(\varepsilon)) = \Gamma(\vec{x}(\varepsilon))$ for $\varepsilon > 0$ and $\Gamma_2(\vec{x}(\varepsilon)) = \Gamma(\vec{x}(\varepsilon))$ for $\varepsilon < 0$. Under assumption 4, $\Gamma_1(\vec{x}(\varepsilon))$ and $\Gamma_2(\vec{x}(\varepsilon))$ are respectively right- and left-continuous in ε near $\varepsilon = 0$. Hence, Berge's Maximum Theorem implies $T_1(\vec{x}(\varepsilon))$ and $T_2(\vec{x}(\varepsilon))$ are respectively right- and left-continuous for ε near $\varepsilon = 0$.

If $c_i \geq c_j$ at some solution to the inner problem at $\vec{x}^*$ and $\vec{x}$, $T_2(\vec{x}^*) \leq T_1(\vec{x}^*) = T(\vec{x}^*)$. Similarly, $T_1(\vec{x}^*) \leq T_2(\vec{x}^*) = T(\vec{x}^*)$ if $c_j \geq c_i$ at some solution to the inner problem at $\vec{x}^*$. More generally, we have:

$$T(\vec{x}(\varepsilon)) = \begin{cases} T_2\left(\vec{x}(\varepsilon)\right) & \text{if } \varepsilon < 0 \\ \max\left\{T_1\left(\vec{x}(\varepsilon)\right), T_2\left(\vec{x}(\varepsilon)\right)\right\} & \text{if } \varepsilon = 0 \\ T_1\left(\vec{x}(\varepsilon)\right) & \text{if } \varepsilon > 0 \end{cases} \tag{45}$$

If x^* is usual and non-degenerate, then *either* all solutions at x^* have $c_i > c_j$, in which case $T_1(\vec{x}^*) > T_2(\vec{x}^*)$, or else all solutions have $c_j > c_i$, in which case $T_2(\vec{x}^*) > T_1(\vec{x}^*)$. That is: $T(\cdot)$ has a jump discontinuity.

For the converse, note simply that $T_1(\vec{x}^*) = T_2(\vec{x}^*)$ if x^* is unusual or degenerate.

Finally, note that the same logic applies along any curve through $\vec{x}^*$ that crosses the set $\{\vec{x} \in \mathbb{R}^N | w_i(\vec{x}) = w_j(\vec{x})\}$ at $\vec{x}^*$, completing the half-space portion of the proof.

B.3 Proofs of Lemma 6

Following the logic in the proof of Lemma 5 and using the relation between ∂X and $\partial \hat{X}$ and $\vec{\xi}$ and $\vec{\mu}$, we have

$$\nabla T(\vec{E}) = \left(\bar{C}\partial \hat{X} + \frac{\partial \hat{X}}{\partial \vec{E}} - \mathbb{I}_K\right)\vec{\xi} - \tilde{\tau}^p + I. \tag{46}$$

Write $\frac{q_i^k(\vec{E})}{r^k(\vec{E})} = \frac{\theta_i^k}{w_i(E)}\left(m(\vec{e}_i/e_i^k)\right)^{-1}$. Using $m(\vec{e}/e_i^k)^{-1} = e_i^k/(m(\vec{e}))$ and $y = wm(\vec{e})$:

$$\frac{\partial \hat{X}^k}{\partial E^m} = \sum_i f_i y_i \frac{\partial}{\partial E^m}\left(\frac{q_i^k(\vec{E})}{r^k(\vec{E})}\right) = -\sum_i \frac{f_i \theta_i^k e_i^k}{w_i(\vec{E})}\frac{\partial w_i(\vec{E})}{\partial E^m} + S^k. \tag{47}$$

Using the functional form $u(c, y/w)$, we have $-\frac{u_w^i}{u_y^i} = \frac{y}{w}$. We also have $\partial \hat{X}^k/\partial y_i = f_i \theta_i^k e_i^k / y_i$. So the $(m,k)^{\text{th}}$ element of $\bar{C}\partial\hat{X}$ is $\sum_i \left(\frac{y_i}{w_i(E)}\right)\frac{dw_i(E)}{dE^m}\frac{f_i\theta_i^k e_i^k}{y_i}$. Hence, $\left(\bar{C}\partial\hat{X} + \frac{\partial \hat{X}}{\partial \vec{E}} - \mathbb{I}_K\right) = S - \mathbb{I}_K$, from which the lemma follows.

C On Subsuming Rothschild-Scheuer (2013)

In Rothschild and Scheuer (2013) extend Stiglitz (1982) to a general two-sector economy with a continuum of types i with preferences $u(c, y/w_i)$ who endogenously choose to work in one of the two sectors. As in Stiglitz, sectoral efforts E_θ and E_φ are complements in a constant returns to scale production technology $Y(E) = Y(E_\theta, E_\varphi)$. Each individual is characterized by a skill vector $(\theta, \varphi)_i$ drawn from a continuous distribution and chooses to work in the sector which provides the highest wage and thus achieves the wage $w_i(\rho) \equiv \max\{\theta\partial Y/\partial E_\theta, \varphi \partial Y/\partial E_\varphi\}$ where $\rho = E_\theta/E_\varphi$. Utility is given by $v(c, y/w)$.

The discrete analog of their model is readily subsumed into our framework—the key trick being, as above, to translate between writing wages in terms of incomes y instead of efforts E. To accomplish this, let $\mathcal{N}^k(\rho)$ denote the set of individuals who strictly prefer sector k at a given ρ and suppose, for expositional simplicity, that the output function Y takes the constant elasticity of substitution (CES) form $Y = (E_\theta^\alpha + E_\varphi^\alpha)^{1/\alpha}$ for some $\alpha \in (0,1)$. In this case, the ratio of total incomes earned in the two sectors is given by ρ^α, which increases with ρ, while the relative wage w_1/w_2 for a type with skills (θ, φ) is given by $\frac{\theta}{\varphi}\rho^{\alpha-1}$, which decreases with ρ.

Let $\mathcal{N}^k(\rho)$ denote the set of types i who strictly prefer to work in sector k given ρ. Define a correspondence $\Lambda(\rho; \vec{y})$ via

$$\Lambda(\rho; \vec{y}) = \left[\frac{\sum_{i \in \mathcal{N}^\theta(\rho)} y_i}{\sum_{i \notin \mathcal{N}^\theta(\rho)} y_i}, \frac{\sum_{i \notin \mathcal{N}^\varphi(\rho)} y_i}{\sum_{i \in \mathcal{N}^\varphi(\rho)} y_i}\right].$$

At values of ρ for which no type is indifferent between the two sectors, $\Lambda(\rho; \vec{y})$ is single valued and corresponds to the ratio of incomes earned in the two sectors given $\vec{y}$ and optimal sectoral choice by all individuals. At values of ρ with types who are indifferent between the two sectors,

$\Lambda(\rho;\vec{y})$ gives the interval of sectoral income ratios incomes which are consistent with $\vec{y}$ and optimal sectoral choice. The lower limit corresponds to the case where all indifferent individuals work in the φ sector, and the upper limit corresponds to the case where they all work in the θ sector. $\Lambda(\cdot;\vec{y})$ is decreasing and upper hemicontinuous for all ρ.

So, on the one hand ρ mechanically implies the income ratio ρ^{α}, which is increasing in ρ. On the other hand, for any given $\vec{y}$, the income ratio $\Lambda(\rho;\vec{y})$ is decreasing in ρ. For any $\vec{y}$, there is therefore a unique solving $X(\vec{y})^{\alpha} = \Lambda(X(\vec{y});\vec{y})$. Implementing an allocation with incomes $\vec{y}$ requires $\rho = X(\vec{y})$, and hence wages $w_i(X(\vec{y}))$. In other words, with an appropriate choice of $X(\vec{y})$ we have converted the problem from one with wages that depend on sectoral efforts E to one in which wages depend on the income allocation $\vec{y}$. Up to unimportant differences—notably that Rothschild and Scheuer consider the Pareto problem of maximizing a weighted average of types' welfare subject to a resource constraint rather than the dual problem of maximizing revenue subject to minimum utility constraints—the discrete version is thus a special case of our framework.

D Derivations with General Functional Forms

We generalize the objective function to $g(\vec{Y},\vec{C})$, and allow outside utility options to be a function of $\vec{x}$. The inner problem is then:

$$T(x) \equiv \max_{(\vec{y},\vec{c})} g(\vec{Y},\vec{C}) \tag{48}$$

subject to

$$u\left(c_i, y_i, w_i\left(x\right)\right) \geq \bar{u}_i(x) \quad \forall i \tag{49}$$

$$u\left(c_i, y_i, w_i\left(x\right)\right) \geq u\left(c_j, y_j, w_i\left(x\right)\right) \quad \forall i,j, \text{ and} \tag{50}$$

$$X(\vec{Y}) = x, \tag{51}$$

The outer problem is then to choose x to maximize $T(x)$.

The inner problem is almost identical to our baseline model. Denote by ξ_i, $\eta_{i,j}$ ($j = i \pm 1$), and $\overrightarrow{\mu}$ the respective multipliers on the minimal utility constraints, the incentive constraints, and the consistency constraint. Write $\vec{Y} = (f_i y_i, ..., f_K y_N)$. Define $\partial G = [g_Y(\vec{Y},\vec{C})\, g_C(\vec{Y},\vec{C})]$. Define ∂X_Y as the $N \times K$ matrix with element $\frac{\partial x_k}{\partial y_i}$. Define ∂X_C as the $N \times K$ matrix with element $\frac{\partial x_k}{\partial c_i}$.

The Lagrangian of the inner problem is

$$\begin{aligned} L(\vec{c},\vec{y};x) = g(\vec{Y},\vec{C}) &+ \sum_i \xi_i(u(c_i,y_i,w_i(x)) - \bar{u}_i(x)) + \vec{\mu}^T(X(\vec{Y}) - x) && (52)\\ &+ \sum_{i>0} \eta_{i,i-1}[u(c_i,y_i,w_i(x)) - u(c_{i-1},y_{i-1},w_i(x))] && (53)\\ &+ \sum_{i<N} \eta_{i,i+1}[u(c_i,y_i,w_i(x)) - u(c_{i+1},y_{i+1},w_i(x))] && (54)\end{aligned}$$

The first order conditions with respect to c_i and y_i are, respectively,

$$\begin{aligned} -f_i g_{C_i}(\vec{Y},\vec{C}) = {}& \xi_i u_c(c_i,y_i,w_i(x)) + \eta_{i,i-1} u_c(c_i,y_i,w_i(x)) + \eta_{i,i+1} u_c(c_i,y_i,w_i(x)) \\ & - \eta_{i-1,i} u_c(c_i,y_i,w_{i-1}(x)) - \eta_{i+1,i} u_c(c_i,y_i,w_{i+1}(x)), \end{aligned} \tag{55}$$

and

$$-f_i g_{Y_i}(\vec{Y}) - \vec{\mu}^T \overrightarrow{X_{y_i}}(\vec{Y}) = \xi_i u_y(c_i, y_i, w_i(x)) + \eta_{i,i-1} u_y(c_i, y_i, w_i(x)) + \eta_{i,i+1} u_y(c_i, y_i, w_i(x)) - \eta_{i-1,i} u_y(c_i, y_i, w_{i-1}(x)) - \eta_{i+1,i} u_y(c_i, y_i, w_{i+1}(x)) \tag{56}$$

Multiplying equation (40) by $MRS_i = -u_y(c_i, y_i, w_i)/u_c(c_i, y_i, w_i)$, adding to equation (41), and re-arranging yields

$$f_i g_{C_i}(\vec{Y}, \vec{C}) \left(g_{Y_i}(\vec{Y}) / g_{C_i}(\vec{Y}, \vec{C}) - MRS_i \right) = - \vec{\mu}^T \overrightarrow{X_{y_i}}(\vec{Y}) + \eta_{i-1,i} u_c(c_i, y_i, w_{i-1}) \left(MRS_i - \widehat{MRS}_i^{i-1} \right) + \eta_{i+1,i} u_c(c_i, y_i, w_{i+1}) \left(MRS_i - \widehat{MRS}_i^{i+1} \right), \tag{57}$$

where $\widehat{MRS}_i^j \equiv -u_y(c_i, y_i, w_j)/u_c(c_i, y_i, w_j)$ is the j type's marginal rate of substitution at the i type's allocation.

Observe that $g_{Y_i}(\vec{Y}, \vec{C})/g_{C_i}(\vec{Y}, \vec{C})$ can be interpreted as the "marginal rate of substitution" between $f_i y_i$and $f_i c_i$ in the social planner's objective $g(\vec{Y}, \vec{C})$.

Define $\tau_i \equiv g_{Y_i}(\vec{Y}, \vec{C})/g_{C_i}(\vec{Y}, \vec{C}) - MRS_i$. Write $\check{\tau}_i = g_{C_i}(\vec{Y}, \vec{C})\tau_i$. We can write equation (42) as

$$f_i \check{\tau}_i = -\vec{\mu}^T \overrightarrow{X_{y_i}}(\vec{Y}) + \sum_{j \in \{i-1, i+1\} \cap \mathcal{N}} \eta_{j,i} u_c(c_i, y_i, w_j) \left(MRS_i - \widehat{MRS}_i^j \right). \tag{58}$$

Define ∂X as the $N \times K$ matrix with element $\frac{\partial x_k}{\partial y_i}$.

Written compactly, the formula for the optimal tax "wedge" from the inner problem becomes

$$f \check{\tau} = -(\partial X)\vec{\mu} + D, \tag{59}$$

where $D = \sum_{j \in \{i-1, i+1\} \cap \mathcal{N}} \eta_{j,i} u_c(c_i, y_i, w_j) \left(MRS_i - \widehat{MRS}_i^j \right)$.

For the outer problem, we have

$$\nabla T(\vec{x}) = \left(\sum_{i=1}^{N} \overrightarrow{X_{y_i}}(\vec{Y}) \left[- \left(\frac{u_w^i}{u_y^i} \right)_j \nabla w_j(\vec{x}) \right] - \mathbb{I}_K \right) \vec{\mu} + \nabla \bar{u}_i(\vec{x}) \cdot \xi + \sum f_i \overrightarrow{g_{Y_i}}(\vec{Y}, \vec{C}) \left[- \left(\frac{u_w^i}{u_y^i} \right) \nabla w_i(\vec{x}) \right] + I \tag{60}$$

where $\mathbb{I}_K$ is the $K \times K$ identity matrix and

$$I = \sum_{i=1}^{N} \sum_{j \in \{i-1, i+1\} \cap \{1, \cdots, N\}} \eta_{j,i} u_y(c_i, y_i, w_j) \left[\widehat{\left(\frac{u_w^i}{u_y^i} \right)}_j \nabla w_j(\vec{x}) - \left(\frac{u_w^i}{u_y^i} \right) \nabla w_i(\vec{x}) \right], \tag{61}$$

with $\widehat{\left(\frac{u_w^i}{u_y^i} \right)}_j \equiv \frac{u_w(c_i, y_i, w_j(\vec{x}))}{u_y(c_i, y_i, w_j(\vec{x}))}$, collects the incentive-term effects.

Denote by $\bar{C}$ the $K \times N$ matrix with elements $\bar{C}_{ji} = \left(-\frac{u_w}{u_y} \right)^i \left(\frac{\partial w_i}{\partial x_j} \right)$. The term $P \equiv \bar{C}(\partial X) - \mathbb{I}_K$ translates this change to effect on the consistency constraints. We can then re-write the first order condition for an optimal $x \notin X^*$, namely $\nabla T(\vec{x}) = 0$ using expression 60 as:

$$P\vec{\mu} + \nabla \bar{u}_i(\vec{x}) \cdot \xi + \sum f_i \overrightarrow{g_{Y_i}}(\vec{Y}, \vec{C}) \left[- \left(\frac{u_w^i}{u_y^i} \right) \nabla w_i(\vec{x}) \right] + I = 0. \tag{62}$$

where the first term collects effects on the consistency constraint, the second term collects effects on the minimal utility constraints, and the last term collects the effect on the incentive constraint.

Define

$$\tau_j^p \equiv - \sum_i f_i g_{Y_i}(\vec{Y}, \vec{C}) \left(- \frac{u_w}{u_y} \right)^i \sum_k \frac{\partial X^k}{\partial y_j} \frac{\partial w_i}{\partial X^k}, \tag{63}$$

which is interpretable as a Pigouvian corrective tax on j.

In compact notation,

$$\vec{\tau}^p = -(\partial X)(fg)_Y(\vec{Y}, \vec{C}) \left(- \frac{u_w^i}{u_y^i} \right) \nabla w_i(\vec{x})$$

Define

$$B_j^p \equiv \sum_i \xi_i \sum_k \frac{\partial X^k}{\partial y_j} \frac{\partial \bar{u}_i}{\partial X^k}.$$

Note that the term $\sum_k \frac{\partial X^k}{\partial y_j} \frac{\partial \bar{u}_i}{\partial X^k}$ measures the amount by which the reservation utility of i types would change in response to a small increase in y_j, and ξ_i translates this change into change in the objective (holding constant i's utility and consumption). The optimality condition

$$P\vec{\mu} + \nabla \bar{u}_i(\vec{x}) \cdot \xi + \sum f_i \overrightarrow{g_{Y_i}}(\vec{Y}, \vec{C}) \left[\left(\frac{u_w^i}{u_y^i} \right) \nabla w_i(\vec{x}) \right] + I = 0. \tag{64}$$

given in Equation 62 is equivalent to (through left-multiplication of ∂X)

$$Q \partial X \vec{\mu} - \vec{B}^p + \vec{\tau}^p - \partial X I = 0$$

Substituting in the optimality condition from the inner problem $\check{f\tau} = -(\partial X)\vec{\mu} + D$, we have

$$Q(D - \check{f\tau}) - \vec{B}^p + \vec{\tau}^p - \partial X I = 0,$$

from which we obtain

$$\check{f\tau} = D + Q^{-1} \vec{\tau}^p - Q^{-1} \vec{B}^p - Q^{-1} (\partial X) I.$$

In the above equation, $Q^{-1}\vec{B}^p$ captures the endogeneity of reservation utility and $Q^{-1}\vec{\tau}^p$ measures the Pigouvian correction. Note that the $\vec{\tau}^p$ term incorporates the generalized objective function.

E Primal Conditions for "usualness"

There is good reason to think that, in fact, $\vec{x}^* \in \vec{X}^*$ will "usually" be usual. The informal intuition is most straightforward with a single dimensional $\vec{x}$. In this case, suppose (WLOG) that $w_i(x^*) = w_{i+1}(x^*)$ and w_i / w_{i+1} is increasing in x at x^*. An increase in x will induce a change in the wage distribution. Ignore, for the moment, the consistency constraint. Then we can ask whether the change in the wage distribution is beneficial, harmful, or neutral to the principal. If it is neutral, then we are justified in ignoring the consistency constraint, and x^* will be degenerate, since the principal will want to pool i and $i+1$ types (as the only reason to separate them is because they have differential effects on X). On the other hand, if, for example, an increase in x is beneficial, then, intuitively, the principal has an extra incentive to increase y_i if and only if $\partial X/\partial y_i > \partial X/\partial y_{i+1}$, in other words, if i's earnings increase X by more than $i+1$'s. Thus, the endogenous wages provides a "force" for separating rather than pooling the two types.

We now formalize this intuition by construct a concrete (and fairly general) example in which all $\vec{x}^* \in \vec{X}^*$ are indeed *usual*. In particular, suppose

Assumption 5.

1. *$u(c, y, w) = v(c) - w^{-1}h(y)$ with v concave and y convex.*
2. *$\vec{X}$ (henceforth X) is single dimensional).*
3. *The constraint $X(\vec{y}) = x$ can be written as an affine function of y_i for any* given *x. I.e., $\sum \alpha_i(x)y_i + K(x) \geq 0$ with $\alpha_i \geq 0$.*

This formulation of utility is equivalent to Werning (2007), where we interpret w^{-1} as a disutility of earnings parameter rather than a wage.[17]

As in Werning (2007), instead of optimizing over (y, c) vectors, we can equivalently optimize over (V, H) vectors, with the obvious mapping $V = v(c)$ and $H = h(y)$. The objective is strictly concave in (V, H), since $c = v^{-1}(V)$ in concave in V and $y = h^{-1}(H)$ is convex in h. The incentive and minimum utility constraints are *linear* in (V, H).

Unlike in Werning (2007) , the problem as a whole will not generally be convex in (V, H), however, since the equality constraint $\sum \alpha_i(x)y_i + k(x) \geq 0$ is linear in y, and hence convex in H unless h is linear.

If, h is linear—so preferences are quasilinear in y, then all constraints are linear, the (minimization of $\sum(c_i - y_i)$) problem is convex, and the solution to the inner problem is unique, and all $x^* \in X^*$ are trivially usual.[18]

Proposition 3 below establishes that all $x^* \in X^*$ are usual even with convex, non-linear h. The following lemma are useful for proving that Proposition.

Lemma 7. *Suppose Assumption 5 holds, and take any solution $(\vec{y}^*, \vec{c}^*)$ to the inner problem at some $x^* \in X^*$, for which $w_k(x^*) = w_{k+1}(x^*) \equiv w^*$. Then $p_k = 1$, $p_{k+1} = 0$, $(y_j, c_j) = (y_j^*, c_j^*)$, $j = k, k+1$ solves the problem [RP] of*

$$\max_{\{(p_j, y_j, c_j)\}_{j=k,k+1}} \quad (y_{k+1} - c_{k+1})(f_k(1 - p_k) + f_{k+1}(1 - p_{k+1})) + (y_k - c_k)(f_k p_k + f_{k+1} p_{k+1}) + \sum_{j \neq k, k+1} (y_j - c_j) f_j \tag{65}$$

subject to:

$$\sum_{i \in \{k,k+1\}} \alpha_i(x^*) f_i (p_i y_k + (1 - p_i) y_{k+1}) + \sum_{i \notin \{k,k+1\}} \alpha_i(x^*) f_i y_i^* + K(x^*) = 0; \tag{66}$$

the relevant incentive constraints $u(c_k, y_k, w^*) = u(c_{k+1}, y_{k+1}, w^*)$, $u(c_k, y_k, w^*) \geq u(c_{k-1}, y_{k-1}, w^*)$, $u(c_k, y_k, w^*) \geq u(c_{k+2}, y_{k+2}, w^*)$, $u(c_{k-1}, y_{k-1}, w_{k-1}(x^*)) \geq u(c_{k+1}, y_{k+1}, w_{k-1}(x^*))$, $u(c_{k-1}, y_{k-1}, w_{k-1}(x^*)) \geq$

[17]With $h(y) = y^k$, $k > 0$, $w^{1/k}$ can be interpreted as wage.

[18]Such quasilinear-in-effort models are used regularly in the optimal tax literature. See for example Brett et al. (2011)

$u(c_k, y_k, w_{k-1}(x^*)), u(c_{k+2}, y_{k+2}, w_{k+2}(x^*)) \geq u(c_{k+1}, y_{k+1}, w_{k+2}(x^*))$, and $u(c_{k+2}, y_{k+2}, w_{k+2}(x^*)) \geq u(c_k, y_k, w_{k+2}(x^*))$ (if $k = 1$ or $k+1 = N$, the constraints involving $k-1$ or $k+2$ are irrelevant); and $p_j \in [0,1]$, $j = k, k+1$.

Proof. If we set $p_k = 1$ and $p_{k+1} = 0$, then problem [RP] is equivalent to the inner problem with fixed (y_i^*, c_i^*) for all $i \neq k, k+1$. By assumption, $(y_j, c_j) = (y_j^*, c_j^*)$, $j = k, k+1$ solves this restricted problem. Adding the variables p_k and p_{k+1} and the constraints $p_j \in [0,1]$, $j = k, k+1$ allows for the possibility that some type k (and $k+1$) individuals receive (y_k, c_k) and others receive (y_{k+1}, c_{k+1}). But, by the logic of Lemma 1, allowing for this possibility does not improve the value of the inner problem. Hence, $p_k = 1$ and $p_{k+1} = 0$ are optimal in the unrestricted problem. □

The Lagrangian for problem [RP] (defined in Lemma 7) can be written as

$$\begin{aligned}\mathcal{L} = (y_k - c_k)(f_k p_k + f_{k+1} p_{k+1}) + (y_{k+1} - c_{k+1})(f_k(1 - p_k) + f_{k+1}(1 - p_{k+1})) \\ + \mu \sum_{i \in \{k,k+1\}} \alpha_i(x^*) f_i (p_i y_k + (1 - p_i) y_{k+1}) + M \\ + \Lambda(c_k, y_k, c_{k+1}, y_{k+1}) - \zeta_k p_k + \zeta_{k+1} p_{k+1} \end{aligned} \tag{67}$$

where μ is the Lagrange multiplier on (66), ζ_j is a composite Lagrange multiplier from the p_j bounding constraints, $\Lambda()$ captures incentive effects, and M is independent of (p_j, y_j, c_j) for $j \in \{k, k+1\}$.

Note that at the optimum with $p_k = 1$, and $p_{k+1} = 0$, $\zeta_k \geq 0$ and $\zeta_{k+1} \geq 0$.

Lemma 8. *$\mu = 0 \Rightarrow y_k^* = y_{k+1}^*$.*

Proof. Suppose by way of contradiction, that $\mu = 0$ and (without loss of generality) $y_{k+1}^* > y_k^*$. There are three possibilities. If $y_{k+1}^* - c_{k+1}^* < y_k^* - c_k^*$, then $\frac{\partial \mathcal{L}}{\partial p_{k+1}} = ((y_k - c_k) - (y_{k+1} - c_{k+1})) f_{k+1} + \zeta_{k+1} > 0$ and thus the allocation is non optimal. Similarly, if $y_{k+1}^* - c_{k+1}^* > y_k^* - c_k^*$, then $\frac{\partial \mathcal{L}}{\partial p_k} < 0$ and again the allocation is optimal.

To rule out the $y_{k+1}^* - c_{k+1}^* = y_k^* - c_k^*$ case, it is useful to define the differential operator $\nabla^* \equiv \frac{\partial}{\partial y_k} + MRS^* \frac{\partial}{\partial c_k}$, where MRS is the marginal rate of substitution $-u_y/u_c$ evaluated at the purported optimum (y_k^*, c_k^*). Intuitively, ∇_k is the derivative with respect to the movement of (y_k, c_k) up and to the right along the k-type's indifference curve. The proof is competed by computing $\nabla_k \mathcal{L} = f_k \nabla_k (y_k - c_k) + \nabla_k \Lambda > 0$, and hence cannot be optimal. This last step follows from (i) $\nabla_k \Lambda \geq 0$ by $y_{k+1}^* > y_k^*$ and single-crossing; and (ii) $\nabla_k (y_k - c_k) > 0$ by the strict convexity of indifference curves and the fact that $y_{k+1}^* - c_{k+1}^* = y_k^* - c_k^*$. □

Lemma 9. *If $\alpha_k > \alpha_{k+1}$ then, at any $x \in X^*$: $y_k > y_{k+1} \Rightarrow \mu > 0$ and $y_k < y_{k+1} \Rightarrow \mu < 0$.*

Proof. At an optimum,

$$\left(f_k \frac{\partial}{\partial p_{k+1}} - f_{k+1} \frac{\partial}{\partial p_k} \right) \mathcal{L} = \mu f_k f_{k+1} (y_k - y_{k+1}) (\alpha_{k+1} - \alpha_k) + \zeta_{k+1} + \zeta_k = 0.$$

Since $\zeta_{k+1} + \zeta_k \geq 0$, this implies $\mu f_k f_{k+1} (y_k - y_{k+1}) (\alpha_k - \alpha_{k+1}) \geq 0$. The lemma follows immediately from this inequality and Lemma 8. □

Corollary 1. *At an optimum for some $x \in X^*$, the multiplier μ' on the consistency constraint in the inner problem (written in the Lagrangian as $\mu \sum_{i=1}^N \alpha_i(x^*) f_i (p_i y_k + (1 - p_i) y_{k+1} + K(x^*) = 0)$ satisfies the following: $y_k > y_{k+1} \Rightarrow \mu' > 0$, and $y_k < y_{k+1} \Rightarrow \mu' < 0$.*

Proposition 3. *All $x \in X^*$ are usual under Assumption 5.*

Proof. Suppose that there are two solutions $(\vec{y}^1, \vec{c}^1)$ and $(\vec{y}^2, \vec{c}^2)$ to the inner problem at some $x^* \in X^*$. Re-write the problem in terms of the variables (V_i, H_i), as above, and let $(\vec{V}^1, \vec{H}^1)$ and $(\vec{V}^2, \vec{H}^2)$ be the two solutions in these variables.

Now consider the line-segment $(\vec{V}(\beta), \vec{H}(\beta)) = \beta(\vec{V}^1, \vec{H}^1) + (1-\beta)(\vec{V}^2, \vec{H}^2)$, $\beta \in [0,1]$ connecting the two optimal allocations. By linearity, $(\vec{V}(\beta), \vec{H}(\beta))$ satisfies the incentive compatibility and minimum utility constraints for all $\beta \in [0,1]$. Since the objective is strictly concave in either c or y (and since both endpoints yield the same objective value), the directional derivative $\nabla_{(\vec{V}^1-\vec{V}^2, \vec{H}^1-\vec{H}^2)}$ of the objective is strictly positive at $(\vec{V}^2, \vec{H}^2)$ and the directional derivative $\nabla_{(\vec{V}^2-\vec{V}^1, \vec{H}^2-\vec{H}^1)}$ of the objective is strictly positive at $(\vec{V}^1, \vec{H}^1)$.

Form Lagrangians $\mathcal{L}^1$ and $\mathcal{L}^2$ for the inner problem at each of the two endpoints with consistency constraint terms

$$\Omega^1 = \mu^1 \left(\sum_{i=1}^{N} \alpha_i(x^*) f_i(p_i y_k^1 + (1-p_i) y_{k+1}^1 + K(x^*)) \right)$$

and

$$\Omega^2 \equiv \mu^2 \left(\sum_{i=1}^{N} \alpha_i(x^*) f_i(p_i y_k^1 + (1-p_i) y_{k+1}^1 + K(x^*)) \right).$$

By the preceding paragraph [19] and the necessary conditions $\nabla_{(\vec{V}^2-\vec{V}^1, \vec{H}^2-\vec{H}^1)} \mathcal{L}^1 = \nabla_{(\vec{V}^1-\vec{V}^2, \vec{H}^1-\vec{H}^2)} \mathcal{L}^2 = 0$, we have $\nabla_{(\vec{V}^2-\vec{V}^1, \vec{H}^1-\vec{H}^2)} \Omega^1 < 0$ and $\nabla_{(\vec{V}^1-\vec{V}^2, \vec{H}^1-\vec{H}^2)} \Omega^2 < 0$.

By Assumption 5.2, $\sum_{i=1}^{N} \alpha_i(x^*) f_i(p_i y_k + (1-p_i) y_{k+1} + K(x^*))$ is an affine function of $\vec{y}$ and hence convex in H. Hence, $\nabla_{(\vec{V}^2-\vec{V}^1, \vec{H}^1-\vec{V}^2)} \Omega^1 < 0$ and $\nabla_{(\vec{V}^1-\vec{V}^2, \vec{H}^1-\vec{V}^2)} \Omega^2 < 0$ respectively imply $\mu^1 < 0$ and $\mu^2 < 0$. The Proposition then follows from Corollary 1. □

[19] Specifically, moving towards allocation 2 from allocation 1 increases the objective and can only ease, not tighten, any binding incentive constraint.