Redistributive Allocation Mechanisms*

Mohammad Akbarpour ® Piotr Dworczak ® Scott Duke Kominers†

June 22, 2023

Abstract

Many scarce public resources are allocated at below-market-clearing prices, and sometimes for free. Such "non-market" mechanisms sacrifice some surplus, yet they can potentially improve equity. We develop a model of mechanism design with redistributive concerns. Agents are characterized by a privately observed willingness to pay for quality, a publicly observed label, and a social welfare weight. A market designer controls allocation and pricing of a set of objects of heterogeneous quality, and maximizes the expectation of a welfare function. The designer does not directly observe individuals' social welfare weights. We derive structural insights about the form of the optimal mechanism, leading to a framework for determining how and when to use non-market mechanisms. The key determinant is the strength of the statistical correlation of the unobserved social welfare weights with the label and willingness to pay.

Keywords: optimal mechanism design, redistribution, inequality, welfare

JEL codes: C78, D47, D61, D63, D82

*The "®" symbol indicates that the authors' names are in certified random order. The authors thank Cassiano Machado Alves, Sebastián Bauer, Yeon-Koo Che, Jeff Ely, Francesco Giovannoni, Ravi Jagadeesan, Andrew Kahn, Simon Loertscher, Ellen Muir, Paula Onuchic, Alessandro Pavan, Natalia Rigol, Ben Roth, Molly Schnell, Stefanie Stantcheva, Tomasz Strzalecki, Bruno Strulovici, Alex Teytelboym, Christine Valente, numerous seminar audiences, the referees, and the editor (Emir Kamenica) for helpful comments. The authors gratefully acknowledge the support of the Washington Center for Equitable Growth and thank the Becker Friedman Institute both for sparking ideas that led to this work and for hosting the authors' collaboration. Additionally, Dworczak gratefully acknowledges funding from the European Union (ERC, IMD-101040122): Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Research Council; Kominers gratefully acknowledges the support of National Science Foundation grant SES-1459912, the Ng Fund and the Mathematics in Economics Research Fund of the Harvard Center of Mathematical Sciences and Applications (CMSA), an Alfred P. Sloan Foundation grant to the CMSA, and the Human Capital and Economic Opportunity (HCEO) Working Group sponsored by the Institute for New Economic Thinking (INET). Xiaoyun Qiu provided excellent research assistance.

†*Akbarpour:* Graduate School of Business, Stanford University; mohamwad@stanford.edu. *Dworczak:* Department of Economics, Northwestern University; and Group for Research in Applied Economics; piotr.dworczak@northwestern.edu. *Kominers:* Entrepreneurial Management Unit, Harvard Business School; Department of Economics and Center of Mathematical Sciences and Applications, Harvard University; and a16z crypto; kominers@fas.harvard.edu.

1 Introduction

Many goods and services—such as certain types of housing, food, and health care, as well as national park permits, road access, and various public services—are allocated at below-market-clearing prices, and sometimes for free. Such "non-market" mechanisms naturally raise concerns among economists because they sacrifice some allocative surplus by failing to allocate resources to those who value them the most. However, policymakers often justify non-market mechanisms on fairness grounds: If resources were allocated using market-clearing prices, they argue, agents with the lowest willingness to pay would be excluded from enjoying their benefits. Because low willingness to pay for many goods and services is likely to be correlated with adverse social and economic circumstances—such as low wealth, health problems, or unemployment—marketplace designers may be especially concerned about the welfare of such agents. But how should we think about the resulting efficiency–equity trade-off?

We study a model in which a market designer allocates a fixed supply of goods with heterogeneous quality. Each agent's utility is linear in the quality of the received good and in monetary transfers—allowing us to parameterize agents' preferences by a single parameter called willingness to pay (or WTP for short). Besides the privately observed willingness to pay for quality, each agent is characterized by a publicly observed label, and an unobserved social welfare weight. The welfare weight is a reduced form representation of designer's redistributive preferences; it measures the social value of giving one unit of money to an agent. Unobservability of welfare weights captures the idea that the designer may not have direct access to information about the agent—such as her detailed financial, social, and economic situation—that determines the welfare weight.

We characterize the optimal incentive-compatible and individually-rational allocation mechanism for a designer who seeks to maximize the expectation of the welfare function, given by the sum of agents' utilities weighted by their social welfare weights. The welfare function places some weight on revenue as well, with the weight interpreted as the marginal value for the designer of spending a dollar on the most valuable social cause (for example, the weight on revenue is equal to the average welfare weight when revenue is returned to agents in the form of a lump-sum transfer).

The market-design approach that we develop is complementary to the classical public finance approach. Our designer decides about the allocation of a single type of good without considering the interaction of this allocation process with macro-level redistribution. The supply of goods and the social welfare weights are thus modeled as exogenous. While these assumptions are limiting in some contexts, they are natural descriptions of many relevant policy problems. For example, public assistance programs (such as allocation of public housing or food stamps) are often run by municipalities or local governments that have limited control over tax policies. Central governments also resort to non-market mechanisms when allocating scarce resources, despite having access to conventional redistributive tools; for example, Covid-19 vaccines were allocated free of charge in most countries, at least partially because setting positive prices would disadvantage poorer populations, an outcome viewed by many as raising moral and fairness issues.[1]

The key tension in our market-design framework is that the designer has redistributive preferences but does not directly observe the social welfare weights. For example, the designer might want to allocate public housing only to people that have both low income and low expected future income. While current income can perhaps be observed relatively precisely (as can be captured by labels in our model), agents are likely to have private information relevant to determining their future financial situation. The designer cannot elicit this information truthfully in any (static) allocation mechanism: for example, if the designer were to promise better terms to those who declare low expectations of future income, it would be beneficial for everyone to make such a claim. In the absence of additional tools, the designer is thus forced to rely exclusively on two types of information: labels—which are publicly observed—and willingness to pay for quality—which can be truthfully elicited by mechanisms with transfers. As a result, the designer uses the *statistical correlation* of labels and willingness to pay with the unobserved welfare weights to "forecast" who is most in need from the perspective of social welfare. Non-market mechanisms—understood as allocating qualities at prices that do not clear the market and thus necessitate rationing—are optimal precisely when the observable and elicitable information reveals inequalities in the underlying welfare weights.

Our first main result shows that, fixing a group of agents with the same label (for example, income below a certain threshold), a non-market allocation is used for agents with the lowest willingness to pay under two conditions: *(1)* the expected welfare weight conditional on the label is strictly higher than the weight on revenue, and *(2)* all agents in the group have a strictly positive willingness to pay for quality. The first condition means that the label identifies a group that the social value of giving one unit of money to a random member of that group is higher than the weight on revenue, but that the designer cannot give a lump-sum payment exclusively to this group.[2] The second condition means that the good is

[1]See for example Schmidt et al. (2020), Pathak et al. (2020), Pathak et al. (2022), and the references therein. In follow-up work, Akbarpour ® al. (2023) study the problem of vaccine allocation by extending the framework of the current paper to a setting where agents have socioeconomic and health externalities.

[2]If a lump-sum payment could be given to all agents with the same label in a frictionless manner, the weight on revenue would be at least equal to the value of giving these agents a lump-sum payment, which

"universally desired," a property that is likely to hold for essential goods such as housing or basic health care. The result then predicts that lowest qualities of universally desired goods are offered free of charge (but subject to rationing and/or random allocation) to agents with the lowest willingness to pay. Agents with higher willingness to pay will often be offered higher qualities at positive prices; however, prices are lower than what they would be in the market allocation, since the free allocation of lowest qualities allows the designer to decrease prices for higher qualities.

We call the first reason for using non-market mechanisms "label-revealed inequality." Effectively, the designer uses information revealed by labels to identify groups of agents with high welfare weights on average; she then subsidizes these groups using a combination of free allocation to low-WTP agents and reduced prices for high-WTP agents. We provide conditions under which the optimal mechanism takes the simple form—often encountered in practice—in which the allocation to the whole group is free. In this case, the assignment of quality is independent of agents' WTP; in the model, it corresponds to a fully random allocation. Assumption *(1)* described in the preceding paragraph is necessary for such fully random allocation to be optimal. In particular, it is *never* optimal to allocate the good for free to all agents when label-contingent lump-sum payments are available to the designer.

Our second main result identifies a distinct reason to use non-market mechanisms. Under the assumption that the weight on revenue is weakly above the average welfare weight in a given group (for example, when the designer *can* give a lump-sum payment to the group), we show that the market allocation is optimal if and only if a certain function—a weighted sum of virtual surplus and welfare-weighted information rents—is non-decreasing. We then argue that this function fails to be non-decreasing when, conditional on the label, there is strong negative correlation between willingness to pay and social welfare weights. We call this effect "WTP-revealed inequality." When lower willingness to pay reveals higher expected welfare weights, the designer chooses to distort the market mechanism and provide some of the goods at reduced prices. Unlike in the case of label-revealed inequality, the designer targets the policy specifically to those agents within the group who select the random-allocation reduced-price option. The policy achieves the redistributive goal because agents with highest willingness to pay are incentivized to select the non-random high-price option.

The key condition of strong negative correlation between willingness to pay and welfare weights is likely to hold in contexts where the variation in willingness to pay stems mainly from the variation in the *ability* to pay—which could depend on individuals' wealth—and

is, by definition, their average welfare weight. Practical reasons for the inability of the designer to condition lump-sum payments on the label could include administrative costs, political constraints, or inefficiencies associated with giving cash to agents suffering from behavioral biases.

the designer preferences depend on individuals' ability to pay. The assumption is thus more plausible when labels are less informative (e.g., the designer cannot observe agents' incomes directly); when the good is relatively expensive (so that only high-income individuals would be able to afford it under the market allocation); and when willingness to pay is not too heavily affected by subjective tastes (e.g., the good is essential).

We have so far focused on describing conditions under which non-market mechanisms are optimal. However, our framework predicts that market allocations are optimal in many (and perhaps most) environments—even when the designer has strong redistributive preferences. Market mechanisms are preferred when the weight on revenue is high, which could be because the designer uses the revenue to give a lump-sum payment to a disadvantaged group of agents, or to subsidize an outside cause that is valuable from a welfare perspective. In particular, when allocating public resources to corporations, it is natural to expect that the weight on revenue far exceeds the average welfare weight. Market mechanisms are also optimal when willingness to pay is not strongly correlated with welfare weights. This could arise for two distinct reasons: the labels could be very informative, so that the designer can infer the welfare weights based solely on observable information; or willingness to pay could be shaped primarily by factors such as private tastes that are orthogonal to social preferences. The latter case helps explain why we would not want to use non-market mechanisms for most affordable, everyday goods and services. We further comment on the market-design implications of our results in Section 6.

1.1 Related work

Non-market allocations in our setting can be interpreted as a form of in-kind transfer—and economists have for a long time been interested in the efficiency and redistributive impacts of such mechanisms. Weitzman (1977), for instance, showed that a free, fully random allocation can be better than competitive pricing when the agents' needs (as reflected in the designer's objective function) are not well expressed by their willingness to pay. Guesnerie and Roberts (1984), meanwhile, gave a general argument that in-kind transfers can be optimal in second-best environments. Nichols and Zeckhauser (1982) were among the first to point out that by increasing the cost of participating in transfer programs, the government can deter the rich from participating, as long as the cost affects the poor less than the rich. Many other papers studied self-targeting mechanisms in different settings; see, for instance, Blackorby and Donaldson (1988), who show that in-kind transfers can be superior to cash transfers because they screen for the right type of individuals, Besley and Coate (1991) who study self-targeting for public options, and Gahvari and Mattos (2007) who analyze conditional cash transfers.[3] Additionally, Cremer and Gahvari (1997) showed that in-kind transfers can be useful even in the presence of optimal non-linear income taxation. Currie and Gahvari (2008) provided an excellent survey of this literature and discussed several other justifications for existence of in-kind transfers.[4]

Our key contribution to the study of in-kind transfers is to employ tools from the theory of mechanism design to explore the optimal structure of such redistribution schemes under rich private information and an arbitrary set of observable labels. We share this market-design perspective on redistribution with the work of Condorelli (2013) and Dworczak ® Kominers ® Akbarpour (2021) (henceforth, DKA). Condorelli (2013) provided conditions for the optimality of market and non-market mechanisms for allocating identical objects to agents in an environment where the designer maximizes agents' values that may be different from their willingness to pay. We extend Condorelli's (2013) model and objective function by allowing the designer to have preferences over revenue, and accommodating cases when lump-sum transfers are not feasible. These features lead to new insights and implications—for instance, when lump-sum transfers are restricted, randomization in the mechanism may be optimal even under conditions that would make rationing suboptimal in the setting of Condorelli (2013). Additionally, our model features heterogeneous qualities of objects (so that our allocations are matchings between types and qualities) and groups of agents with the same observable characteristics (giving rise to the novel across-group allocation problem). Finally, we focus on the economic implications of maximizing a redistributive objective function, while Condorelli worked with a generic objective function. DKA studied a closely related question in the context of buyers and sellers with heterogeneous marginal utilities of money trading a homogeneous good. The current paper takes a more practical approach by focusing on the problem of allocating public resources, and incorporating a range of features that play a key role for real-life policymakers: heterogeneous quality of objects, richer preferences over revenue, additional observable information about the agents, and potential restrictions on the use of lump-sum transfers.

A few recent papers have enriched these frameworks in different ways. For example, Kang and Zheng (2020) characterized the set of constrained Pareto optimal mechanisms for allocating a good and a bad to a finite set of asymmetric agents, with each agent's role as a buyer or a seller determined endogenously by the mechanism. Kang (2020b) allowed for

[3] Self-targeting has also been found to be an effective way of allocating resources in development economics context; see, for example, Alatas et al. (2016). A similar mechanism, albeit in a very different context of constitutional design, was explored by Aidt and Giovannoni (2011).

[4] Note that the economic arguments in favor of in-kind transfers that we focus on in the present work are distinct from and independent of the idea that it may make sense to give in-kind transfers for paternalistic reasons, such as in response to agency problems within a household—one of the motivations described by Currie and Gahvari (2008).

an exogenous private market where agents can also purchase the good (of potentially higher quality). Reuter and Groh (2020) studied a similar problem of allocating a finite number of goods to finitely many agents under redistributive preferences, addressing new challenges in analyzing and implementing the optimal allocation mechanism in the finite case. Fan, Chen, and Tang (2021) analyzed optimal allocations of a divisible good when agents have quadratic preferences over quantity, and differ in bargaining power (modeled as a weight in the designer's objective). Finally, Kang and Zheng (2022) studied a buyer-seller market, where agents are entitled to equal shares of a limited resource, and characterized the optimal mechanism for arbitrary Pareto weights.

In our approach, (implicit) socio-economic inequality motivates attaching non-equal social welfare weights to agents. An alternative approach is to model the effects of differences in wealth via budget constraints. The sizable literature on auction design with budget constraints predicts that the designer may resort to non-market allocations even when she is concerned about maximizing allocative efficiency, as in the work of Che, Gale, and Kim (2013b). Non-market mechanisms play a different role in the two approaches: In our framework, a non-market allocation may be *preferred* to the efficient allocation if it redistributes enough surplus to agents with high welfare weights; in the Che, Gale, and Kim (2013b) setting, non-market mechanisms arise when achieving the efficient allocation is not possible due to budget constraints.

In a broad sense, our paper is connected to the canonical frameworks of public finance and optimal taxation literature (Diamond and Mirrlees, 1971 and Atkinson and Stiglitz, 1976). The key distinction is that in our framework the designer takes the inequality in the market as given, and does not take into account how her mechanism might potentially influence the welfare weights. In a sense, our designer cannot change agents' endowments directly—she can only design the rules of the allocation mechanism. The introduction of observable characteristics to our model is a classical idea in the taxation literature. For example, Akerlof (1978) described how "tagging" could be used in the tax system for redistributive purposes. The interpretation of welfare weights in our model is also closely analogous to how they are used in public finance; specifically, Saez and Stantcheva (2016) introduced generalized marginal welfare weights in the context of optimal tax theory and interpreted them as the value that society puts on providing an additional dollar of consumption to a given individual.

From a technical perspective, our paper is related to a recent body of work that generalizes the Myerson (1981) ironing technique (see also Toikka, 2011). In concurrent research, Muir and Loertscher (2022) relied on similar techniques to solve a problem of a revenue-maximizing seller in the presence of resale; Ashlagi, Monachou, and Nikzad (2020) showed that these methods can be also used in designing the optimal dynamic allocation in a multigood environment by optimizing over how much information is disclosed about different types of objects. Kang (2020a) derived a variant of this approach based on a tool called the constrained maximum principle. Finally, Kleiner, Moldovanu, and Strack (2021) demonstrated that all these procedures can be obtained as a special case of a general property of extreme points that arise in optimization problems involving majorization constraints.

2 The model

Framework. A designer allocates a set of objects of heterogeneous quality to a set of agents who differ in both their observable and unobservable characteristics. There is a unit mass of agents, with each agent characterized by a type vector (i, r, λ). The three dimensions of agents' type vectors have a joint distribution in the population that is known to the designer; formally, we can think of (i, r, λ) as the realization of a random variable on some underlying probability space, where $\mathbb{E}[\cdot]$ will be used to denote the expectation operator.[5] The first ingredient of the type vector, called the *label*, takes one of finitely many values from the set I, and is assumed to be publicly observed. Agents with the same label form a *group*; there are (measure) $\mu_i > 0$ agents in group i. The parameter $r \in \mathbb{R}_+$ is the *willingness to pay* (for quality) which is privately observed by the agent. Conditional on label i, the WTP r has a distribution with cumulative distribution function G_i and continuous density g_i, strictly positive on $[\underline{r}_i, \bar{r}_i]$. Finally, $\lambda \in \mathbb{R}_+$ is the *social welfare weight* on a given individual, interpreted as the social value of giving that individual one unit of money. Individuals observe their own types,[6] but neither r nor λ of any given individual are observed by the designer.

There is a unit mass of objects, with each object characterized by a one-dimensional quality $q \in Q \subseteq [0, 1]$, where Q is a compact set. The assumption of unit mass is without loss of generality: If there is only a mass $\mu < 1$ of objects, we can always add an extra mass $1 - \mu$ of "null" objects with $q = 0$ because receiving an object with quality 0 in our model is equivalent to not receiving an object at all. Let F denote the cumulative distribution function of q, that is, $F(q)$ is the total mass of objects of quality equal to or less than q.

If an agent with willingness to pay r is assigned a good with quality q in exchange for a monetary transfer t, that agent's utility is $rq - t$; if that agent has a social welfare weight λ, then the contribution of that individual to the social welfare function will be $\lambda(rq - t)$.

Note that our framework incorporates a few strong assumptions about the environment.

[5] We do not define notation for the joint distribution because it will not be needed, other than through conditional expectations and some marginal distributions that we introduce next.

[6] As we show, it does not matter whether an individual can observe her own social welfare weight λ.

First, we assume that agents' utility is quasi-linear in money. Second, we assume that agents differ only in their "intensity" of preferences but they agree on the ranking of qualities. Third, each agent's utility only depends on the *expected* outcome—agents are risk neutral.[7] Fourth, social preferences are captured by weights that are exogenous—reflecting an implicit assumption that the designer does not take into account how her chosen allocation impacts social preferences. These simplifying assumptions allow us to derive tight results without imposing any restrictions on the set of mechanisms.

Assignments and mechanisms. An assignment Γ is a collection of $|I|$ measurable functions $\Gamma_i : [\underline{r}_i, \bar{r}_i] \to \Delta(Q)$ with $\Gamma_i(q|\, r)$ interpreted as the probability that an agent in group i with willigness to pay r is assigned an object with quality q or less. The assignment Γ is *feasible* if

$$\Gamma_i(\cdot|\, r) \text{ is a CDF for all } i, \text{ and } r \in [\underline{r}_i, \bar{r}_i]; \tag{2.1}$$

$$\sum_{i\in I} \mu_i \int_{\underline{r}_i}^{\bar{r}_i} \Gamma_i(q|\, r) dG_i(r) \geq F(q), \quad \forall q \in Q. \tag{2.2}$$

Condition (2.2) states that the distribution of *assigned* qualities is first-order stochastically dominated by the distribution of *available* qualities. The condition reflects the availability of free disposal—a decrease in quality can be achieved by randomizing between a given quality and quality 0. Because the utility of agents only depends on the expected quality, it will be convenient to denote

$$Q^{\Gamma_i}(r) = \int_0^1 q d\Gamma_i(q|\, r);$$

we write simply $Q_i(r)$ if the reference to the underlying assignment Γ_i is clear.

To describe feasible mechanisms, we rely on the Revelation Principle. A direct mechanism $(\Gamma_i,\, t_i)_{i\in I}$ asks agents to report their willingness to pay r, assigns objects according to $\Gamma_i(q|\, r)$ in group i, and charges agents according to the transfer function $t_i(r)$. As it will turn out, we do not have to include the social welfare weight λ in the agent's report because no incentive-compatible mechanism can improve the social welfare function by trying to elicit this information from agents (see Claim 2).

Lump-sum payments to agents may or may not be allowed in different applications of our framework. We use the following modeling approach to accommodate all possible cases: There is no hard budget constraint for the designer but the mechanism must use non-negative

[7] While we cannot easily accommodate risk aversion, we can capture some aspects of risk aversion over the assigned quality by defining the agent's utility to be $rv(q) - t$ for some concave function v. In that case, we would define a new random variable $\tilde{q} = v(q)$ with CDF $\tilde{F}$, called "risk-adjusted quality," and apply our results with $\tilde{F}$ as the primitive distribution of quality.

transfers, i.e., $t_i(r) \geq 0$ for all i and r.[8] However, lump-sum payments to agents may happen "outside of the mechanism;" this is captured through the designer's value for generating monetary surplus in the mechanism (in the objective function that we formally introduce in the next subsection). For example, if the value for generating monetary surplus is 0 in the designer's objective, then the constraint of non-negative transfers is binding and means that lump-sum payments are not allowed. However, if the value for generating monetary surplus is equal to the value of giving a lump-sum payment to all agents, then it is *as if* lump-sum payments to all agents were allowed. We comment on other cases later. For incentive-compatible mechanisms, the condition that transfers are non-negative is equivalent to requiring that for each group i, the utility $\underline{U}_i$ of type $\underline{r}_i$ satisfies $\underline{U}_i \leq \underline{r}_i Q^{\Gamma_i}(\underline{r}_i)$.

Formally, a mechanism $(\Gamma_i, t_i)_{i \in I}$ is *feasible* if

- Γ is a feasible assignment, i.e., it satisfies conditions (2.1)–(2.2);
- each agent reports her willingness to pay truthfully:

$$rQ^{\Gamma_i}(r) - t_i(r) \geq rQ^{\Gamma_i}(\hat{r}) - t_i(\hat{r}), \quad \forall i, r, \hat{r}; \tag{2.3}$$

- each agent receives non-negative utility from the mechanism but does not receive a positive money transfer:

$$0 \leq \underline{U}_i \leq \underline{r}_i Q^{\Gamma_i}(\underline{r}_i), \quad \forall i. \tag{2.4}$$

The following lemma follows from standard arguments that extend Myerson (1981).

Claim 1. *A mechanism is feasible if and only if Γ is a feasible assignment, $Q^{\Gamma_i}(r)$ is non-decreasing in r for all i, and $t_i(r)$ satisfies*

$$U_i(r) \equiv rQ^{\Gamma_i}(r) - t_i(r) = \underline{U}_i + \int_{\underline{r}_i}^{r} Q^{\Gamma_i}(\tau) d\tau \tag{2.5}$$

for some $\underline{U}_i \in [0, \underline{r}_i Q^{\Gamma_i}(\underline{r}_i)]$.

Note that (2.5), commonly referred to as the envelope formula, provides an expression for the utility of an agent with willingness to pay r in an incentive-compatible mechanism.

[8] Because agents are buyers in our framework, this constraint on transfers has no impact on the set of implementable allocation rules—it only constrains lump-sum transfers.

The objective function. We assume that the designer maximizes the expectation of a weighted sum of revenue and agents' utilities weighted by their social welfare weights. The following observation—which has been made before in different contexts—implies that our definition of a feasible mechanism is without loss of generality for maximizing this objective.

Claim 2. *The designer cannot increase the expectation of her objective function by using an incentive-compatible mechanism that elicits information about* λ.[9]

Claim 2 is intuitive: Since, conditional on r, λ has no bearing on the individual's preferences, no incentive-compatible revelation mechanism can condition the allocation or payments directly on the reported λ. That is, if the mechanism attempted to elicit λ, then agents would always report whatever λ would lead to the best possible treatment by the mechanism, regardless of their true type. As a consequence, the designer must form beliefs about λ based on the information she *is* able to elicit and observe—that is, r and i. Define

$$\lambda_i(r) \equiv \mathbb{E}[\lambda |\, i,\, r]$$

to be the expectation of λ conditional on i and r, under their joint distribution. To distinguish $\lambda_i(r)$ from the underlying social welfare weight λ, we call $\lambda_i(r)$ the *Pareto weight* on an agent with label i and willingness to pay r. For technical reasons, we assume that $\lambda_i(r)$ is continuous in r for each i. Let

$$\bar{\lambda}_i \equiv \int_{\underline{r}_i}^{\bar{r}_i} \lambda_i(\tau) dG_i(\tau)$$

be the average Pareto weight for group i.

With this, we can write the designer's objective function as

$$\alpha \underbrace{\sum_{i \in I} \mu_i \left(\int_{\underline{r}_i}^{\bar{r}_i} t_i(r) dG_i(r) \right)}_{\text{revenue}} + \underbrace{\sum_{i \in I} \mu_i \left(\int_{\underline{r}_i}^{\bar{r}_i} \lambda_i(r) U_i(r) dG_i(r) \right)}_{\text{social surplus with weights } \lambda_i}, \tag{OBJ}$$

where $\alpha \geq 0$ is the weight on revenue. Let

$$h_i(r) \equiv \frac{1 - G_i(r)}{g_i(r)}$$

[9] The claim follows from analogous results proven by Jehiel and Moldovanu (2001) and Che, Dessein, and Kartik (2013a); see also Dworczak ® Kominers ® Akbarpour for the formulation and proof of the claim in a similar economic context.

denote the inverse hazard rate of G_i, and let

$$J_i(r) \equiv r - \frac{1 - G_i(r)}{g_i(r)}$$

denote the virtual surplus function. It is well known from Myerson (1981) that $h_i(r)$ measures the information rents of an agent with WTP r, while $J_i(r)$ captures the designer's revenue in an incentive-compatible mechanism. Finally, let

$$\Lambda_i(r) \equiv \mathbb{E}_{\tilde{r} \sim G_i}[\lambda_i(\tilde{r}) | \tilde{r} \geq r]$$

be the average Pareto weight attached to agents with willingness to pay above r. A simple calculation then shows an alternative representation of the objective function (OBJ).

Claim 3. *The objective function (OBJ) can be written as*

$$\sum_{i \in I} \mu_i \left(\int_{\underline{r}_i}^{\bar{r}_i} V_i(r) Q^{\Gamma_i}(r) dG_i(r) + \left(\bar{\lambda}_i - \alpha\right) \underline{U}_i \right), \tag{OBJ'}$$

where

$$V_i(r) \equiv \alpha J_i(r) + \Lambda_i(r) h_i(r).$$

The function $V_i(r)$ can be interpreted as the expected social value of allocating a unit of quality to an agent in group i with WTP r in an incentive-compatible mechanism. Note that in the standard paradigm of fully transferable utility, this value function would simply be $J_i(r) + h_i(r) = r$, thus reducing to a measure of allocative efficiency. With $\alpha = 0$ and constant Pareto weights, the value function would reduce to $h_i(r)$, which corresponds to the case of maximizing agent surplus when payments are interpreted as "money burning." In our setting, the value function consists of a weighted sum of virtual surplus $J_i(r)$ (corresponding to revenue) and information rents $h_i(r)$ weighted by the function $\Lambda_i(r)$ representing the Pareto weights. The weight $\Lambda_i(r)$ on the information rent of type r is given by the expected social welfare weight on all agents in group i with a WTP above r—this is a consequence of the envelope formula (2.5), which dictates that in order for the mechanism to remain incentive-compatible, any increase in utility of type r must also be received by all higher types.

Our objective function is quite general but has important limitations within the context of redistribution. Primarily, the approach of using exogenous welfare weights reflects the assumption that the designer takes inequality as given. With this formulation, she cannot express preferences over any inequality created by the mechanism itself. In particular, we do not accommodate quotas that control the overall fairness of the outcome, and are popular in some contexts, such as school choice (see for example Echenique and Yenmez, 2015; Bodoh-Creed and Hickman, 2018).

Interpretation. Claim 2 makes it clear that $\lambda_i(r)$—the Pareto weight—is effectively a primitive of our model. Nevertheless, we introduced the unobserved social welfare weights to emphasize the economic forces that give rise to any particular $\lambda_i(r)$.

The average Pareto weights $\bar{\lambda}_i$ and $\bar{\lambda}_j$ differ to the extent that the labels i and j capture observable information that is correlated with the social welfare weights. For example, if tax data allows the designer to determine the income bracket for each agent, then agents associated with lower income brackets might receive a higher average Pareto weight.

Similarly, dispersion in $\lambda_i(r)$ for any given i should be interpreted as residual correlation between willingness to pay and social welfare weights, conditional on i. For a concrete example, suppose first that no observable information is available, but we elicit the willingness to pay of two individuals, A and B, for a high-quality house in an attractive neighborhood. Agent A is willing to pay $\$500,000$, while agent B is only willing to pay $\$50,000$. While the differences between A and B's willingness to pay may be driven by preferences, they likely also reflect characteristics such as income and opportunity cost of money that could in turn affect the welfare weights. Thus, without observing the characteristics that inform welfare directly, the designer may place a higher Pareto weight on the agent with lower willingness to pay, reflecting a Bayesian belief that this agent is more likely to be poor. Now suppose that the designer additionally has access to tax data, and she knows that agents A and B have the same income. Conditional on that information, the correlation between willingness to pay and welfare weights becomes weaker; willingness to pay originally appeared to be more strongly correlated with welfare weights due to the omission of a relevant variable—income. However, that correlation is likely still negative, as long as other unobserved characteristics—such as health shocks or future job prospects—influence *both* the welfare weights and willingness to pay for a house. More generally, the more informative the label, the less residual correlation one would expect between r and λ. But there are also cases when the correlation can be very weak even in the absence of informative labels. For example, when the good to be allocated is a movie ticket, and agent A is willing to pay \$10, while agent B is only willing to pay \$1, the most likely inference is that agent A is more interested in that movie than agent B—not necessarily that B is very poor or otherwise socially disadvantaged. Summarizing, $\lambda_i(r)$ naturally depends on the strength of the underlying social preferences and the degree to which they are uncovered by the label i, but also on the characteristics of the good, such as the relative importance of personal taste versus ability to pay in determining the willingness to pay for it.

As discussed in the Introduction, we think of α—the weight on revenue—as representing the marginal social value generated by an additional dollar in the designer's (unmodeled) budget.[10] For example, if a local authority generates revenue by running a public housing program, the monetary surplus can be returned to citizens as a tax cut, or used to invest in the construction of new homes. Thus, by varying α, we can analyze how the optimal allocation mechanism changes depending on the best available use of revenue for the designer.

Several special cases are of particular interest. When $\alpha = \bar{\lambda}_i$, a dollar of revenue has the same value to the designer as giving a dollar to a randomly selected agent within group i; similarly, when $\alpha = \bar{\lambda} \equiv \sum_i \mu_i \bar{\lambda}_i$, a dollar of revenue has the same value to the designer as giving a dollar to a randomly selected agent from the whole population. These cases are mathematically equivalent to assuming that the designer uses the revenue to finance lump-sum payments to agents in group i, or all agents, respectively.

The case $\alpha = \bar{\lambda}$ can be thought of as a default specification in which the set of agents represents the entire (local) population, the revenue generated from the mechanism subsidizes the government's budget, and the marginal cost of financing the budget is $\bar{\lambda}$. If the designer uses the additional dollar of revenue to lower taxes (so as to balance the budget), then it is also possible that $\alpha < \bar{\lambda}$ if taxation is progressive. For both of these interpretations, we are implicitly assuming that the designer cannot use the extra dollar of revenue to give lump-sum payments to some "preferred" group i (a group i with $\bar{\lambda}_i > \bar{\lambda}$).

More generally, whenever $\alpha < \bar{\lambda}_i$ for some label i, lump-sum payments to group i are restricted (this is exactly when our assumption of non-negative transfers in the mechanism has bite). This could be, for example, a consequence of political constraints.[11] Another interpretation is that lump-sum payments are allowed but there are frictions (such as administrative costs) that decrease their marginal value below parity. In the extreme case $\alpha = 0$, our model becomes mathematically equivalent to a costly-screening (money-burning) model in which an agent's payment to the designer is more appropriately interpreted as a costly activity (such as standing in a queue) that is socially wasteful.

On the other hand, setting $\alpha > \bar{\lambda}_i$ captures cases in which the designer has a higher value from spending the revenue outside of group i. This could be because label-contingent lump-sum payments are feasible, and there exists another group j with $\bar{\lambda}_j > \bar{\lambda}_i$. Another possibility is that the designer can spend the monetary surplus generated by the mechanism on a socially valuable outside cause (such as infrastructure investment).

[10] In the public finance literature, α is often referred to as the *marginal value of public funds*.

[11] Liscow and Pershing (2022) show, using survey data, that the general population (in the US) express a much stronger political support for in-kind redistribution than for cash transfers.

3 Optimal mechanisms

We identify an optimal mechanism in two steps:

1. First, the objects are allocated "across" groups: F is split into $|I|$ CDFs $F_i^\star$.
2. Then, the objects are allocated "within" groups: For each label i, the objects $F_i^\star$ are allocated optimally according to the expected-quality schedule $Q_i^\star$.

We first explain how to solve the "within" problem, and then use the solution to that problem to solve the "across" problem.

3.1 The "within" problem

In this step, we assume that F_i is the CDF of object qualities that are to be allocated to agents with label i. Formally, we refer to the *within problem for group i* as maximizing (OBJ') subject to feasibility with $I = \{i\}$, $\mu_i = 1$, and $F = F_i$. For a function Ψ, let $\mathrm{co}(\Psi)$ denote the concave closure of Ψ (i.e., the point-wise smallest concave function that bounds Ψ from above) and let $\mathrm{cd}(\Psi)$ denote the concave decreasing closure of Ψ (i.e., the point-wise smallest concave decreasing function that bounds Ψ from above). When i is fixed, we will sometimes abuse terminology slightly by referring to r as the agent's *type*.

We say that there is *assortative* matching among types $r \in [a, b]$, if $Q_i^\star(r) = F_i^{-1}(G_i(r))$ for all $r \in [a, b]$. To account for the possibility that some objects may remain unallocated, we say that the matching is *effectively assortative* when it is assortative for

$$r \in [\inf\{r : Q_i^\star(r) > 0\}, \bar{r}_i].$$

We say that there is *random* matching among types $r \in [a, b]$ if $Q_i^\star(r) = \bar{q}$ for some $\bar{q} \in [0, 1]$ and all $r \in [a, b]$.[12]

Remark 1. Because we have not imposed any assumptions on F (for example, we have not ruled out degenerate distributions of quality), assortative and random matching could coincide (if F_i is constant in the relevant range). In particular, the two concepts do not differ when all types in a given interval are not allocated any objects. The distinction between random and assortative matching can be guaranteed to be meaningful for each group i by assuming that *(i)* $F(0) = 0$, *(ii)* F is continuous, and *(iii)* it is optimal to allocate all objects within group i, which is implied by $\int_{\underline{r}_i}^{r_i} V_i(r) dG_i(r) \geq 0$ for all r_i.

[12] Throughout, $H^{-1}(x)$ denotes the generalized inverse of a right-continuous non-decreasing function H on $[a, b]$: $H^{-1}(x) = \min\{y \in [a, b] : H(y) \geq x\}$, for all $x \leq \max_y\{H(y)\}$.

Theorem 1. *Define*

$$\Psi_i(t) \equiv \int_t^1 V_i(G_i^{-1}(x))dx + \max\{0,\, \bar{\lambda}_i - \alpha\}\, \underline{r}_i\, \mathbf{1}_{\{t=0\}}.$$

The optimal value of the within problem for group i *is given by*

$$\int_0^1 \mathrm{cd}(\Psi_i)(F_i(q))dq.$$

An optimal solution is given by an expected-quality schedule

$$Q_i^\star(r) = \Phi_i^\star(G_i(r))\mathbf{1}_{\{r \geq G_i^{-1}(x_i^\star)\}},$$

where $[0,\, x_i^\star]$ *is the maximal interval on which* $\mathrm{cd}(\Psi_i)$ *is constant, and* $\Phi_i^\star$ *is non-decreasing and satisfies*

$$\Phi_i^\star(x) = \begin{cases} \frac{\int_a^b F_i^{-1}(y)dy}{b-a} & \text{if } x \in [a,\, b] \text{ and } [a,\, b] \text{ is a maximal interval on which } \mathrm{co}(\Psi_i) \text{ is affine} \\ F_i^{-1}(x) & \text{otherwise} \end{cases}$$

for almost all x.[13]

Moreover, it is optimal to set $\underline{U}_i = 0$ *when* $\alpha \geq \bar{\lambda}_i$, *and* $\underline{U}_i = Q_i^\star(\underline{r}_i)\underline{r}_i$ *when* $\alpha \leq \bar{\lambda}_i$.

As mentioned in Section 1.1, the proof of Theorem 1 uses relatively standard techniques known as "generalized ironing" that extend Myerson's methods to richer environments. For completeness, and because several features of our setting (primarily the non-negativity of transfers and the continuous distribution of quantity) require these methods to be adjusted, we present a complete argument in the appendix. In the proof, we work with an arbitrary objective function of the form (OBJ'), not necessarily coming from maximizing a weighted sum of revenue and surplus.

Theorem 1 describes a simple procedure to obtain a closed-form solution to the within-group problem:

1. Compute the function Ψ_i that is a non-linear transformation of the original objective function. A noteworthy feature of Ψ_i is that it incorporates the constraint that transfers are non-negative: Whenever $\bar{\lambda}_i > \alpha$, this constraint must bind, and hence $\underline{U}_i$ is set to the maximal feasible level $Q_i(\underline{r}_i)\underline{r}_i$. In the transformed objective function Ψ_i, this corresponds to an upward jump at 0.

[13] An interval $[a,\, b]$ is a maximal interval on which $\mathrm{co}(\Psi_i)$ is affine if $\mathrm{co}(\Psi_i)$ is affine on $[a,\, b]$ and no interval $[c,\, d] \supsetneq [a,\, b]$ has that property.

2. Compute the concave closure $\text{co}(\Psi_i)$ and the concave decreasing closure $\text{cd}(\Psi_i)$ of Ψ_i.

3. If $\text{co}(\Psi_i) < \text{cd}(\Psi_i)$ over some initial interval $(0, x_i^\star)$, then objects of quality below the $x_i^\star$ quantile of F_i are not allocated (the designer uses the free disposal option), and hence agents with willingness to pay below $r_i^\star = G_i^{-1}(x_i^\star)$ are assigned quality 0. This can only happen if Ψ_i is not decreasing everywhere, which requires $V_i(r)$ to be negative for some r.

4. The remaining object qualities are partitioned into intervals; the remaining agents are partitioned in the order of increasing willingness to pay to match the mass of objects within each interval; whenever $\text{co}(\Psi_i)$ is affine on a (maximal) interval, the matching between types and quality is random within that interval; whenever $\text{co}(\Psi_i)$ is strictly concave on an interval, the matching between types and quality is assortative.

The function Ψ_i plays a key role in determining the properties of the optimal mechanism. To gain intuition, we can use integration by parts and substitution, and obtain that for any $r > \underline{r}_i$,

$$\Psi_i(G_i(r)) = \int_r^{\bar{r}_i} \tau\lambda_i(\tau) dG_i(\tau) + (\alpha - \Lambda_i(r)) r (1 - G_i(r)). \tag{3.1}$$

Thus, the value of Ψ_i at some quantile $x = G_i(r)$, is the value to the designer from selling quality 1 at a price of r.

3.2 The "across" problem

Based on the solution to the within problem for each i separately, we can now formulate the *across problem* as

$$\max_{(F_i)_{i \in I}} \left\{ \sum_{i \in I} \mu_i \int_0^1 \text{cd}(\Psi_i)(F_i(q)) dq \right\} \tag{3.2}$$

$$\text{such that } \sum_{i \in I} \mu_i F_i(q) = F(q), \quad \forall q \in Q. \tag{3.3}$$

Once the optimal $F_i^\star$ are found that solve (3.2)–(3.3), the optimal solution within each group i is described by Theorem 1.

Our second technical result describes a solution procedure for the across problem. Let $\overline{V}_i(x) \equiv |\text{cd}(\Psi_i)'(x)|$ denote the (absolute value of the) slope of $\text{cd}(\Psi_i)$ at quantile x. Intuitively, $\overline{V}_i(x)$ represents the "ironed" social value that—unlike the social value $V_i(G_i^{-1}(x))$—is guaranteed to be non-decreasing in x.

Theorem 2. *There exists a non-decreasing non-negative function $V^{\min}(q)$ such that for all i and q, the optimal solution $(F_i^\star)_{i\in I}$ to (3.2)–(3.3) satisfies*

$$\begin{cases} F_i^\star(q) = 0 & \text{if } \overline{V}_i(0) > V^{\min}(q), \\ F_i^\star(q) = 1 & \text{if } \overline{V}_i(1) < V^{\min}(q), \\ F_i^\star(q) \text{ solves } \overline{V}_i(F_i^\star(q)) = V^{\min}(q) & \text{otherwise.} \end{cases}$$

Moreover, $V^{\min}(q) = \min_{i:\, F_i^\star(q)<1} \{\overline{V}_i(F_i^\star(q))\}$.

Theorem 2 gives rise to a simple procedure for allocating goods of different qualities across groups. The algorithm allocates the objects by gradually increasing the CDFs $F_i^\star$, in the order of increasing (ironed) marginal social values $\overline{V}_i(\cdot) = |\mathrm{cd}(\Psi_i)'(\cdot)|$. The function $V^{\min}(q)$ keeps track of the running minimum over these values across groups. Starting from the lowest quality, we increase the CDF $F_i^\star$ for group i with the smallest $\overline{V}_i$ at 0 (in the case where there are several such groups, the proof of Theorem 2 describes how to break the ties). At any q, we increase the CDF of group(s) i with the lowest $\overline{V}_i$ at $F_i^\star(q)$. That is, only groups i with $\overline{V}_i(F_i^\star(q)) = V^{\min}(q)$ are allocated objects with quality q. When some $F_i^\star(q)$ reaches 1, we stop increasing the CDF for that group.

The procedure described in the preceding paragraph is a greedy algorithm in that it allocates quality levels sequentially, from lowest to highest, and the allocation of the given level of quality only depends on the ranking of marginal social values across the groups, evaluated at the "current" allocation. A greedy algorithm is optimal because, for all i, the ironed marginal value $\overline{V}_i(\cdot)$ changes monotonically with the allocation to group i. This is a consequence of the fact that the optimal within-group allocation concavifies the value function Ψ_i, as shown in Theorem 1. Section 5 illustrates the greedy procedure in a simple numerical example, and gives a graphical interpretation of the greedy procedure.

The proof of Theorem 2 is in the appendix. Intuitively, we solve the program (3.2)–(3.3) by first considering a relaxed problem in which the constraint that $F_i(q)$ is a CDF is dropped, and later verifying that there exists a solution to the relaxed program that is feasible. The index $V^{\min}(q)$ is the Lagrange multiplier on the resource constraint (3.3) for the relaxed problem.

4 Economic implications

We now discuss the main economic implications of our framework. We first focus on the within-group problem, and emphasize the circumstances under which using a non-market allocation becomes optimal. Then, we turn attention to how the structure of the optimal within-group allocation affects the split of qualities across different groups. In the next section, we illustrate and build upon this analysis with a parametric example.

4.1 When to use a non-market mechanism?

4.1.1 Label-revealed inequality

Our first result shows that a non-market allocation becomes optimal when a certain group has a high average Pareto weight and the good is desired by all agents in that group—in a sense that we make precise next.

We say that the good is *universally desired in group* i if $\underline{r}_i > 0$, that is, if the willingness to pay of agents in group i is bounded away from 0.[14] The main example of universally desired goods are essential goods (such as housing or basic health care) that everyone has a need for. An implicit assumption behind this interpretation is that each agent has at least some ability to pay, so that a good fails to be universally desired only if some agents have no intrinsic value for it.[15]

Proposition 1 (Label-revealed inequality)**.** *If the average Pareto weight* $\bar{\lambda}_i$ *in group* i *is strictly larger than the weight on revenue* α*, and the good is universally desired in group* i*, then there exists* $r_i^\star > \underline{r}_i$ *such that the optimal allocation within group* i *is random at a price of* 0 *for all types* $r \leq r_i^\star$*.*

Proposition 1 states that it is always optimal to allocate some objects randomly to the lowest-WTP agents at a price of 0 if *(i)* the designer cares more about the surplus of an *average* agent within group i than about revenue, and *(ii)* the good is universally desired. The first assumption is likely to hold when label i is targeted for preferential treatment or affirmative action, but making direct monetary transfers to group i is not feasible. In the natural case $\alpha = \bar{\lambda}$, we have $\alpha < \bar{\lambda}_i$ if the label i is associated with a group of agents that the designer cares about more than about the average agent in the population.

For intuition, note that for a fixed allocation, when $\alpha < \bar{\lambda}_i$, the designer would like to minimize the transfers that agents pay. The non-negative transfers condition prevents the designer from giving a monetary transfer to agents directly, and implementing assortative matching requires prices to be increasing. Consider, instead, providing the lowest qualities

[14] Using the lower bound $\underline{r}_i$ to define universally desired goods makes our results cleaner, but the gist of the assumption is that G_i is concentrated on values of r above $\underline{r}_i$. That is, our results that assume universally desired goods continue to hold for distributions that attach a small enough mass to $r \in [0, \underline{r}_i]$.

[15] If there are some agents with literally no income, then their willingness to pay for *any* good is zero, and our framework does not correctly account for their welfare since it places no social value on their allocation, regardless of λ.

for free to all agents with willingness to pay below some cutoff r_i. When—and only when—the good is universally desired, this policy leads to an increase in the utility of the lowest-WTP agents. Increasing the utility of lowest-WTP agents, in itself, does not necessarily constitute an improvement in the designer's objective because the designer need not be directly concerned about the welfare of those agents; indeed, the only assumption we made is about the *average* Pareto weight. However—and this is the key observation—providing the goods for free to agents with low willingness to pay also allows the designer to lower prices (and hence increase utility) for higher types. This comes at a cost: Providing the goods for free precludes any screening in the corresponding region, reducing allocative efficiency; in particular, the highest-WTP agents in the free-allocation region must necessarily receive below-efficient quality. However, it can be shown that the reduction in allocative efficiency is always second-order relative to the benefits when the region of random matching is small (see Appendix B for a formal argument).

The optimal mechanism determines the size of the random-allocation interval by trading off a decrease in prices against a decrease in allocative efficiency. Thus, it is often the case that random matching at the bottom of the distribution of willingness to pay coincides with assortative matching at the top of the distribution. In the next section, we illustrate how the random-allocation region varies with the primitives of the model using a parametric example.

Here, we are instead interested in the circumstances under which the trade-off is resolved towards a fully random allocation. This type of in-kind redistribution is quite common in practice: the good is allocated for free to those satisfying certain verifiable eligibility criteria (which are captured by the label i in our model). Because the price is set to 0, some sort of rationing becomes necessary (which in practice may take the form of an explicit or implicit lottery). To rule out trivial cases, we assume that F_i—the pool of quality levels available to group i—is non-degenerate.[16]

Proposition 2 (Optimality of free provision)**.** *A necessary condition for a fully random allocation to be optimal within group i is that*

$$\alpha \bar{r}_i \leq \int_{\underline{r}_i}^{\bar{r}_i} r \lambda_i(r) dG_i(r). \tag{4.1}$$

Condition (4.1) becomes sufficient if $V_i(r) = \alpha J_i(r) + \Lambda_i(r) h_i(r)$ is quasi-convex.

The key condition (4.1) is derived from a hypothetical scenario in which the designer has only one (infinitesimal) unit of the object with quality 1 to allocate: For full randomization

[16] When F_i is a degenerate (Dirac delta) distribution, there is no difference between assortative and random matching.

to be optimal, it must be that the value of revenue from selling that object at a maximal price to the highest willingness-to-pay agent is smaller than the value of allocating this object uniformly at random at a price of 0. This necessary condition becomes sufficient under a regularity condition on $V_i(r)$.

Overall, Proposition 2 provides some support for in-kind distribution, but only if certain restrictive conditions are met. First, the designer must not be able to target a direct cash transfer to the "eligible" agents. Indeed, a direct consequence of inequality (4.1) is that optimality of full randomization requires that the average Pareto weight $\bar{\lambda}_i$ in group i be strictly higher than the weight on revenue α. In particular, if lump-sum transfers to group i are feasible for the designer, then a fully random allocation *cannot* be optimal. Second, the weight on revenue (measuring how effectively it can be used for other purposes) should be relatively small. Third, such schemes are more likely to be optimal for universally desired goods; note that if the Pareto weights are non-increasing, then the right-hand side of (4.1) is bounded above by $(1/2)(\bar{r}_i + \underline{r}_i)\bar{\lambda}_i$, and thus when $\underline{r}_i = 0$, the average Pareto weight must be at least twice as large as the weight on revenue. However, when $\underline{r}_i$ is large, it may suffice that $\bar{\lambda}_i$ is only slightly above α. Moreover, holding fixed $\bar{\lambda}_i > \alpha$, (4.1) is satisfied in the limit as the support of willingness to pay shrinks to a point. Thus, optimality of free provision is more likely for goods for which heterogeneity in tastes is limited.

An interesting corollary of Proposition 2 is that it is *never* optimal to allocate goods fully at random to the target population at a constant strictly positive price—even though such an allocation mechanism is feasible for universally desired goods. This is because a fully non-market allocation can only be justified if the average Pareto weight strictly exceeds the weight on revenue; but if that is the case, then the unique optimal price is 0.[17]

4.1.2 WTP-revealed inequality

So far, we have focused on cases in which random allocation is optimal because the average Pareto weight on group i exceeds the weight on revenue α. We have also argued that it matters that the good is universally desired. Next, we ask whether the use of random allocation is limited to these cases. The following result provides a negative answer, by fully characterizing when the designer should resort to randomization for at least some agents in group i.

[17] While it is tempting to criticize some existing social programs that charge a relatively small price and rely on some form of rationing, it is important to emphasize that the result only applies if the price is below *everyone's* WTP (an empirical statement); besides, small prices could be imposed for reasons other than raising revenue and screening that our model does not capture (for example, due to moral-hazard concerns).

Proposition 3 (WTP-revealed inequality)**.** *Suppose that (i) the average Pareto weight $\bar{\lambda}_i$ in some group i does not exceed the weight on revenue α, or (ii) the good is not universally desired for group i. Then, every optimal mechanism provides random allocation to agents in group i with willingness to pay in some (non-degenerate) interval if and only if the social value function $V_i(r) = \alpha J_i(r) + \Lambda_i(r) h_i(r)$ is not non-decreasing.*

Proposition 3 can also be viewed as providing necessary and sufficient conditions for a market solution to be optimal. The first assumption rules out the circumstances that lead to the conclusion of Proposition 1; we already know that assortative matching cannot be optimal in that case. When the designer can give a lump-sum payment to agents in group i or the good is not universally desired, optimality of fully assortative matching reduces to checking the monotonicity of the function $\alpha J_i(r) + \Lambda_i(r) h_i(r)$, which is the weighted sum of revenue and social-welfare-weighted information rents of the agents.

Assuming differentiability, we can provide further economic intuition: Random matching will be used for a subset of agents as long as for some r, we have

$$\alpha + \Lambda_i'(r) h_i(r) + (\Lambda_i(r) - \alpha) h_i'(r) < 0.$$

Suppose that the inverse hazard rate is non-increasing—an assumption that is satisfied by many commonly used distributions and implies that effectively assortative matching maximizes revenue. Then, random matching will be optimal for agents with willingness to pay close to r if either *(i)* the average Pareto weight on types above r is sufficiently greater than the weight on revenue, or *(ii)* the Pareto weights are declining sufficiently quickly with r. That last condition can be interpreted as saying that, conditional on i, willingness to pay is strongly negatively correlated with the unobserved social welfare weights; this is more likely to be true when the label i is not very informative of the agents' underlying weights (e.g., when the label does not include any information about the agent's income). Intuitively, if Pareto weights are declining with willingness to pay around r, the designer would like to give more rents to agents with types just below r. This can be achieved by making the allocation random in a (small) interval around r. Indeed, compared to assortative matching, agents near the left end of that interval will now receive a higher expected quality and hence—by the envelope formula (2.5)—a higher expected utility. This modification generally reduces both allocative efficiency and revenue, and hence the redistributive motive must be strong enough to justify random matching as optimal.

Proposition 3 is related to results in the literature. First, when there is only one label, quality is binary ($q = 0$ or $q = 1$), and $\alpha = \bar{\lambda}$ (the revenue is redistributed as a lump-sum payment), our setting reduces to the one-sided version of the model of DKA, who show that competitive pricing (which is a special case of assortative matching) may fail to be optimal when the Pareto weights have large dispersion. Under the assumption of non-increasing inverse hazard rate, non-increasing Pareto weights, and $\alpha \geq \bar{\lambda}_i$, a simple calculation based on Proposition 3 shows that assortative matching is optimal in our framework when $\alpha \geq \max_r\{\lambda_i(r) - \Lambda_i(r)\}$. Thus, when revenue can be used more flexibly, non-optimality of a market allocation requires both a high dispersion and a high level of the Pareto weights. Second, Proposition 3 relates to results known from the analysis of the costly screening model, in which transfers are replaced by "money-burning" (corresponding to the case $\alpha = 0$). Among others, Hartline and Roughgarden (2008), Condorelli (2012), and Chakravarty and Kaplan (2013) showed that the assortative allocation maximizes unweighted agent surplus when the inverse hazard rate is non-decreasing.[18] Proposition 3 extends this condition to the case when surplus is weighted by the Pareto weights: it is required that $\Lambda_i(r)h_i(r)$—the product of the inverse hazard rate at r and the average Pareto weight on all types above r—is non-decreasing. At the same time, Proposition 1 implies that the conclusion of Proposition 3 is true *only* under the assumption that the good is *not* universally desired (that is, only if $\underline{r}_i = 0$—an assumption that is made in the aforementioned papers).

4.2 How to allocate objects based on labels?

So far we have focused on allocation within individual groups. We now focus on what the insights of Section 4.1, combined with Theorem 2, tell us about the allocation of objects *across* the groups. We begin by characterizing the structure of $\mathrm{supp}(F_i^\star)$—the set of object qualities allocated to group i—in simple cases in which the optimal allocation takes the same form in all groups.

Proposition 4 (Across-group allocation with random matching)**.** *Suppose that it is optimal to use a (fully) random allocation in each group $i \in I$. Relabel the groups so that lower $i = 1, \ldots, |I|$ corresponds to lower $\int_{\underline{r}_i}^{\bar{r}_i} \tau\lambda_i(\tau)dG_i(\tau)$. Then, there exists an optimal mechanism in which* $\mathrm{supp}(F_i^\star) = [q_i, q_{i+1}] \cap \mathrm{supp}(F)$*, for some* $\{q_i\}_{i=1}^{|I|+1}$ *with* $\min \mathrm{supp}(F) = q_1 \leq q_2 \leq \ldots \leq q_{|I|} \leq q_{|I|+1} = \max \mathrm{supp}(F)$*.*

Proposition 4 states that when all groups receive a random allocation (the conditions for optimality of such an allocation are given in Proposition 2), the optimal across-group allocation is particularly simple: groups can be ordered, and groups higher in the ranking receive uniformly higher qualities. Intuitively, under fully random allocation, the designer's

[18] Similar conditions were obtained as early as the work of McAfee and McMillan (1992) in a setting where bidders collude but cannot share payments among each other; then, bidding in the auction becomes equivalent to burning utility (see also Bauer, 2022 for a related model and result).

marginal value from allocating a unit of quality to group i is equal to $\int_{\underline{r}_i}^{\bar{r}_i} \tau\lambda_i(\tau)dG_i(\tau)$ and does not depend on the previously allocated qualities. Thus, in the implementation of the greedy algorithm from Section 3, the designer maximizes overall welfare by first allocating the lowest qualities to the group with the lowest marginal social value, then allocating the lowest of the remaining qualities to the group with the second-lowest marginal social value, and so on. This is in sharp contrast to the optimal across-group allocation when the market mechanism is used within each group.

Proposition 5 (Across-group allocation with assortative matching)**.** *Suppose that it is optimal to use effectively assortative matching in each group* $i \in I$*. Relabel the groups so that lower* $i = 1, \ldots, |I|$ *corresponds to lower* $\bar{r}_i$*. Then, there exists an optimal mechanism in which* $\text{supp}(F_i^\star) = [\underline{q}_i, \bar{q}_i] \cap \text{supp}(F)$*, for some* $\{\underline{q}_i, \bar{q}_i\}_{i=1}^{|I|}$ *with* $\bar{q}_1 \leq \bar{q}_2 \leq ... \leq \bar{q}_{|I|} = \max \text{supp}(F)$*. Moreover, if the good is not universally desired for any group, then* $\underline{q}_i = \min \text{supp}(F)$ *for all* $i \in I$*.*

When assortative matching is used within groups (as is optimal under the conditions given in Proposition 3), we should in general expect non-trivial overlaps in the quality levels allocated to different groups. In the special case that all groups have the same support of willingness to pay and the good is not universally desired, it is in fact optimal for all groups to receive the same range of qualities. Intuitively, under assortative matching, the marginal value of allocating an object to a given group i varies with how many objects have already been allocated. This is because higher-WTP agents generate more value for the designer than lower-WTP agents within the same group. Overlaps in qualities across groups i and j occur whenever the highest-WTP agent within group i generates more social value then the lowest-WTP agent within group j, and vice versa.

The social value generated by the lowest quality allocated to a given group (which is received by the lowest-WTP agent under assortative matching) is often 0; in fact, this is always the case for goods that are not universally desired.[19] If the good is not universally desired for any group, then it is optimal for the designer to allocate the lowest-quality goods to all groups.

A similar observation explains the conclusion about the highest qualities. The marginal value of allocating the highest-quality object is determined by the value $\alpha\bar{r}_i$ generated by the highest-WTP agent in group i. Thus, the ranking of the upper bounds on WTP determines

[19] Note a subtle difference: It is of course also true that agents with willingness to pay 0 generate no value for the designer when objects are allocated randomly. However, precisely because of the randomness, the marginal value from the perspective of the object is positive. In contrast, under assortative matching, certain qualities are allocated to the lowest-WTP agents deterministically, and hence can have marginal value 0 to the designer.

which groups receive the highest-quality objects. It may seem surprising that these values do not depend on the Pareto weights. In particular, if $\bar{r}_i$ is only slightly higher than $\bar{r}_j$, then group i receives at least some higher-quality objects even if the designer puts no weight on the welfare of agents in group i, and a high weight on the welfare of agents in group j. To see why, note that the utility of the highest type $\bar{r}_i$ is pinned down by the allocation to lower types $r < \bar{r}_i$ within her group (see the envelope formula (2.5)). In particular, the utility of the highest-WTP agent in an incentive-compatible mechanism does not depend on the quality of the object that she receives—higher quality only translates into a higher price. This implies that the allocation at the top of the distribution only affects the designer's revenue, and hence the highest-quality object is allocated to the group i with the highest upper bound on willingness to pay $\bar{r}_i$.

Finally, we analyze the structure of the across-group allocation when groups differ in their internal allocation. To deliver the main insight in the sharpest possible form, we focus on the case of two groups with opposite modes of allocation. We consider a more general situation in the context of our parametric example in the next section.

Proposition 6 (Intermediate quality to a random-matching group)**.** *Suppose that* $|I| = 2$ *and that it is optimal to use effectively assortative matching in group* $i = 0$ *and fully random matching in group* $i = 1$*. Then, there exist* $\underline{q} \leq \bar{q}$ *such that* $\text{supp}(F_1^\star) = [\underline{q},\, \bar{q}] \cap \text{supp}(F)$*, and* $\text{supp}(F_0^\star) = ([0,\, \underline{q}] \cup [\bar{q},\, 1]) \cap \text{supp}(F)$*. Moreover, assuming a non-degenerate distribution of quality,* $\bar{q} < 1$ *if* $\alpha \bar{r}_0 > \int_{\underline{r}_1}^{\bar{r}_1} \tau \lambda_1(\tau) dG_1(\tau)$*; and* $\underline{q} > 0$ *when* $\alpha \underline{r}_0 < \int_{\underline{r}_1}^{\bar{r}_1} \tau \lambda_1(\tau) dG_1(\tau)$*.*

In practice, when a certain group of eligible agents receives goods for free, the quality of those goods tends to be lower than the quality in the "market" (this lower quality may also take the form of rationing, i.e., some agents receiving quality $q = 0$). Proposition 6 indicates that this is typically not optimal when eligibility is verifiable (via the label). Instead, under permissive conditions (for example, when the good is not universally desired for group 0), the goods the designer chooses to provide for free are those of *intermediate* quality.

For intuition, note that when a random-allocation mechanism is used for group 1, the designer's payoff depends only on the *expected* quality allocated to that group. In contrast, when assortative matching is used, the designer's payoff depends on the *dispersion* in quality. The latter observation is particularly intuitive in the context of revenue maximization: A revenue-maximizing seller chooses to decrease the allocation of the low types in order to lower the information rents of the high types (Myerson, 1981). In fact, it will often be optimal not to allocate some objects in group 0, in which case the marginal value of quality is 0 up to some point (in the greedy algorithm described in Theorem 2). At the same time, the marginal value of quality allocated to agents with high willingness to pay in group 0 may be large—especially if $\bar{r}_0$ is high. Thus, the designer allocates both the lowest- and highest-quality objects to group 0, leaving the intermediate-quality objects for group 1.

5 Illustrative example

In this section, we analyze an extended parametric example that illustrates and expands on the insights presented in Section 4. We additionally use the example to showcase a graphical solution method based on the techniques developed in Section 3, providing mathematical intuition behind our results.

5.1 The setup

Suppose that in each group i, willingness to pay is distributed uniformly on $[\underline{r}_i,\, \underline{r}_i + 1]$, where $\underline{r}_i \geq 0$. Pareto weights are given by $\lambda_i(r) = \bar{\lambda}_i(\gamma_i + 1)(\underline{r}_i + 1 - r)^{\gamma_i}$, for some $\gamma_i > -1$. Then, group i is completely characterized by the the triple $(\underline{r}_i,\, \bar{\lambda}_i,\, \gamma_i)$. The lower bound on WTP $\underline{r}_i \geq 0$ controls whether (and to what degree) the good is universally desired. The average Pareto weight $\bar{\lambda}_i$ measures—in relation to the fixed weight on revenue α—the strength of the designer's redistributive preference towards group i.[20] The parameter γ_i controls the dispersion in Pareto weights within group i: When $\gamma_i = 0$, the weights are equal; when $\gamma_i < 0$, they are increasing in WTP; and when $\gamma_i > 0$, they are decreasing in WTP, with the limiting case $\gamma_i \to \infty$ corresponding to the Rawlsian objective (positive weight only attached to the agent with the lowest utility within group i).

5.2 Optimal within-group allocation

We first consider the optimal within-group allocation, fixing $i \in I$ and the CDF of quality F_i allocated to group i (assumed non-degenerate). Under our assumptions, the function $\Psi_i(x)$, as defined in Theorem 1, is first convex and then concave in the interval $(0, 1]$, potentially with a jump at 0 (see Figure 5.1); it follows that the concave decreasing closure $\mathrm{cd}(\Psi_i)(x)$ lies above the function $\Psi_i(x)$ precisely in some interval of the form $[0,\, x_i^\star]$, where $x_i^\star \in [0,\, 1]$. Thus, the structure of the optimal within-group mechanism follows immediately from Theorem 1, and takes a particularly simple form.

Result. There exists $x_i^\star \in [0,\, 1]$ such that the optimal mechanism allocates qualities randomly to agents with $r \leq G_i^{-1}(x_i^\star)$, and assortatively to agents with $r > G_i^{-1}(x_i^\star)$.

[20] Note that by construction, $\bar{\lambda}_i$ is equal to the expected value of $\lambda_i(r)$ with respect to the distribution of WTP r in group i.

Economically, our first result shows that in-kind redistribution to the lowest-WTP agents may coexist with assortative matching at the "top of the distribution." This is a manifestation of the trade-off between redistribution and efficiency. Note that both fully random allocation and fully assortative matching are special cases, with $x_i^\star = 1$ and $x_i^\star = 0$, respectively. While we do not expect the simple structure to remain optimal in general (beyond the example, the optimal allocation may "alternate" between random and assortative matching multiple times), we expect the forces that determine $x_i^\star$—the fraction of objects allocated by a non-market mechanism—to be more robust. In the remainder of this subsection, we focus on studying how the optimal cutoff $x_i^\star$ depends on the primitives of the model.

When lump-sum payments are not available

We first assume that $\bar{\lambda}_i \geq \alpha$, so that label i identifies agents treated preferentially by the designer but a direct lump-sum transfer may not be feasible (when $\bar{\lambda}_i > \alpha$). To find $x_i^\star$, we solve the equation $\Psi_i(x) - \Psi_i'(x)x = \Psi_i(0)$, using the fact that the concave decreasing closure $\text{cd}(\Psi_i)$ coincides with Ψ_i at 0, and is tangent to Ψ_i at $x_i^\star$ (if $x_i^\star < 1$). While a closed-form solution is not available in general when Ψ_i has a jump at 0, we can calculate $x_i^\star$ in the special case when there is no dispersion in the Pareto weights.[21]

Result. When $\bar{\lambda}_i \geq \alpha$ and $\gamma_i = 0$, the fraction of objects allocated randomly in the optimal mechanism is

$$x_i^\star = \begin{cases} \sqrt{\frac{2\underline{r}_i(\bar{\lambda}_i - \alpha)}{2\alpha - \bar{\lambda}_i}} & \text{if that expression is well-defined and below } 1, \\ 1 & \text{otherwise.} \end{cases}$$

In line with Proposition 1, the non-market allocation is used for the lowest qualities of the good when the good is universally desired ($\underline{r}_i > 0$) and the designer attaches a strictly higher weight to group i than to revenue ($\bar{\lambda}_i > \alpha$). By the definition of Ψ_i in Theorem 1, these two assumptions together are equivalent to Ψ_i exhibiting a jump at 0. Then, even if the function Ψ_i is concave on $(0, 1]$, the concave decreasing closure $\text{cd}(\Psi_i)$ lies above Ψ_i in some non-degenerate interval $[0, x_i^\star]$ (see the top left panel of Figure 5.1). Moreover, the number of objects allocated randomly is increasing in the size of the jump, and hence increasing in both $\underline{r}_i$ and $\bar{\lambda}_i$.

Next, we ask when a fully non-market solution is optimal ($x_i^\star = 1$). This possibility is illustrated in the right top panel of Figure 5.1: The jump of Ψ_i at 0 is so large that the concave decreasing closure $\text{cd}(\Psi_i)$ lies everywhere above Ψ_i. Since the second assumption in

[21] This special case does not significantly constrain our analysis in terms of economic insight because the intuition behind Proposition 1 does not rely on dispersion in Pareto weights.

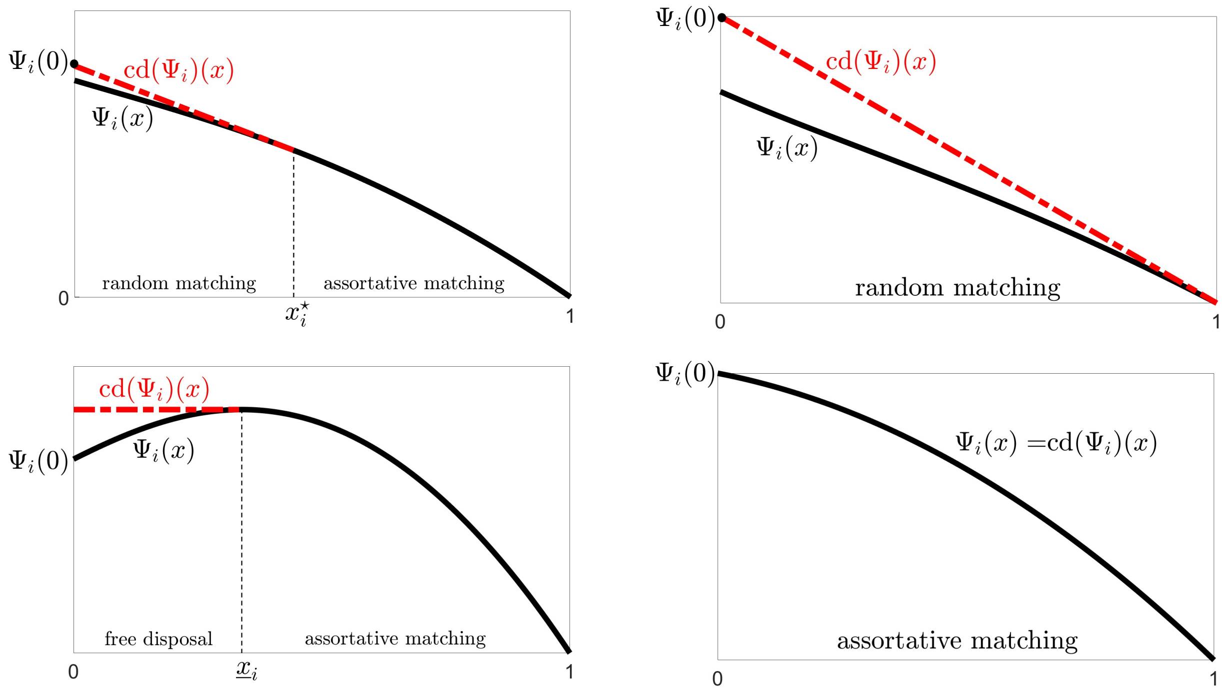

Figure 5.1: Illustration of the within-group solution: The function Ψ_i (solid line) and its concave decreasing closure $\text{cd}(\Psi_i)$ (dashed line when different from Ψ_i)

Proposition 2 always holds in our example, condition (4.1) is both necessary and sufficient.

Result. Fully random matching is optimal ($x_i^\star = 1$) if and only if

$$\bar{\lambda}_i \geq \alpha \frac{\underline{r}_i + 1}{\underline{r}_i + \frac{1}{2+\gamma_i}}.$$

In particular, $\bar{\lambda}_i > \alpha$ is necessary for full randomization to be optimal, while a higher $\underline{r}_i$ makes the condition easier to satisfy (in particular, $\bar{\lambda}_i > \alpha$ becomes sufficient when $\underline{r}_i$ is high enough); this is intuitive because higher $\bar{\lambda}_i - \alpha$ and $\underline{r}_i$ increase the size of the jump of Ψ_i at 0. Perhaps more surprisingly, a high concentration of weights on the lowest-WTP agents (high γ_i) makes the condition more difficult to satisfy. The reason is that—holding the average weight fixed—higher γ_i implies lower weights on agents with high willingness to pay. Then, the motive to maximize revenue dominates the motive to maximize welfare for high types, and the designer optimally uses assortative matching at the top of the distribution. This intuition suggests that free provision of goods can only be optimal when the target population is relatively uniform, perhaps because the label identifying the group is highly informative of the characteristics that determine the social welfare weights.

When lump-sum payments are available

Next, we consider the case $\alpha \geq \bar{\lambda}_i$, when group i is not preferentially treated (or a label-contingent lump-sum payment is feasible), as in the setting of Proposition 3.

We first observe that when some agents in group i have low willingness to pay ($\underline{r}_i$ is low) and the weight on revenue α is large, some objects may be discarded in the optimal mechanism. In our formulation, free disposal of qualities below the $\underline{x}_i$-quantile is optimal when $\Psi_i(x)$ is increasing in $[0, \underline{x}_i]$ (see the bottom left panel in Figure 5.1). The cutoff $\underline{x}_i$ is positive if and only if $\alpha(\underline{r}_i - 1) + \bar{\lambda}_i < 0$. The reason for this is familiar from Myerson (1981): discarding some objects raises revenue. Since the revenue motive for discarding objects is a well-known property, for the remainder of this section, we focus on cases when free disposal is not used.

When lump-sum payments to group i are feasible, Ψ_i does not have a jump at 0, and we can solve for the cutoff $x_i^\star$ explicitly. The cutoff is positive if and only if the function Ψ_i is convex in some initial interval; otherwise, the function Ψ_i is concave everywhere, and hence it coincides with its concave closure, making fully assortative matching optimal (see the bottom right panel of Figure 5.1).

Result. When $\alpha \geq \bar{\lambda}_i$, the fraction of objects allocated randomly in the optimal mechanism is

$$x_i^\star = 1 - \min\left\{1, \left(\frac{2\alpha}{\bar{\lambda}_i(\gamma_i + 1)}\right)^{\frac{1}{\gamma}}\right\}.$$

It follows that fully assortative matching is optimal if and only if $2\alpha \geq \bar{\lambda}_i(\gamma_i + 1)$. When $2\alpha < \bar{\lambda}_i(\gamma_i + 1)$, the lowest-quality objects are allocated for free. Consistent with our discussion in Section 4, the use of a non-market mechanism is supported both by the levels of Pareto weights in group i and their dispersion. However, there is a sense in which dispersion plays a more important role. First, sufficiently high dispersion (as controlled by γ_i) is sufficient for making it optimal to allocate the lowest-quality objects for free. Second, some degree of dispersion (with higher weights on agents with low WTP) is necessary under our assumption that $\alpha \geq \bar{\lambda}_i$ (in fact, we need $\gamma_i > 1$). In particular, when there is no dispersion in Pareto weights, assortative matching is always optimal.

An interesting property of the optimal mechanism in our example is that—whenever label-contingent lump-sum payments are feasible—assortative matching is used at the top of the distribution of WTP; in Appendix C, we show that this is a general property, at least as long as Pareto weights are non-increasing in WTP. Intuitively, under these assumptions, the revenue-maximizing motive dominates the welfare-maximizing motive for agents with high WTP.

Finally, we ask how the level of inequality revealed by WTP (dispersion of Pareto weights) influences the fraction of objects allocated by a non-market mechanism. It turns out that the relationship is non-monotonic, with the use of non-market allocation maximized at an intermediate level of inequality.

Result. When $\alpha \geq \bar{\lambda}_i$, the fraction of objects allocated randomly in the optimal mechanism is zero ($x_i^\star = 0$) both when $\gamma_i = 0$ and (in the limit) when $\gamma_i \to \infty$. The fraction of objects allocated randomly is maximized at the unique positive γ_i that solves the equation

$$\frac{\exp\left(\frac{\gamma_i}{\gamma_i+1}\right)}{(\gamma_i+1)} = \frac{\bar{\lambda}_i}{2\alpha}.$$ [22]

The non-monotonicity exhibited in the example is a general property, as we show in Appendix C. For intuition, consider the case when $\gamma_i \to \infty$ (so that we approach the Rawlsian objective). Even though the optimal mechanism uses random matching for the lowest types, the randomization region actually vanishes as the weights become increasingly skewed. The *average* Pareto weight is fixed at a level below the weight on revenue, so as the weight on the lowest-WTP agents increases, the weight on all higher-WTP agents converges to 0. Thus, the motive to maximize revenue eventually dominates for almost all agents, which makes assortative matching increasingly attractive, and optimal in the limit.[23]

5.3 Optimal across-group allocation

In this subsection, we use the properties of the optimal within-group solution to solve a simple instance of an optimal across-group allocation. We suppose that there are two groups, respectively labeled $i = 0$ and $i = 1$, where the main distinction is that $\bar{\lambda}_1 > \alpha > \bar{\lambda}_0$. That is, group $i = 1$ is poorer, disadvantaged, or for some other reason treated preferentially by the designer. Revenue is used to finance a lump-sum payment to all agents. Additionally, we assume that $\underline{r}_0 \geq \underline{r}_1 > 0$, that is, the good is universally desired, and group 0 has higher WTP. The distribution of quality F has a strictly positive density on $[0, 1]$.

We begin with a simple result that illustrates Proposition 4.

Result. Suppose that the designer uses a fully random allocation in both groups. Then, in

[22]When $\alpha = \bar{\lambda}_i$, the fraction of objects allocated randomly is maximized when γ_i is slightly above 3, and is equal to approximately 20%. As α grows, the maximal fraction of objects allocated randomly declines, while the level of inequality required to achieve it grows. For example, when $\alpha = 2\bar{\lambda}_i$, the maximal fraction is around 10% and is achieved when γ_i is slightly below 9.

[23]While the designer could maximize the welfare of the lowest-WTP agent by giving her a random quality for free, this would necessarily decrease revenue to 0. Since $\alpha \geq \bar{\lambda}_i$, even though the weight on some individuals diverges to ∞, the revenue motive dominates the welfare motive in expectation over all types.

the optimal across-group allocation, the group i with higher $\bar{\lambda}_i \left(\underline{r}_i + \frac{1}{\gamma_i+2}\right)$ receives uniformly higher quality.

The straightforward comparative static in the preceding result is that higher $\bar{\lambda}_i$ makes it more likely that group i receives higher-quality objects. The term $\underline{r}_i$ is a sufficient statistic for the distribution of WTP in our parametric example, and it captures the fact that higher WTP means that allocating an object is more valuable. The $\frac{1}{\gamma_i+2}$ term measures the importance of the distribution of Pareto weights within a group. Group i is more likely to receive priority over the other group if the Pareto weights within that group are more skewed towards agents with high WTP. This is because agents with highest WTP necessarily receive the highest utility among all agents in their group; if they also have a high Pareto weight, then the designer can generate more social value from that group.

From now on, we assume that $2\alpha < \bar{\lambda}_0(\gamma_0+1)$, so that an effectively assortative matching is optimal within group 0. In group 1, meanwhile, we know that it is optimal to allocate qualities below the $x_0^\star$-quantile for free at random, and the remaining qualities assortatively.

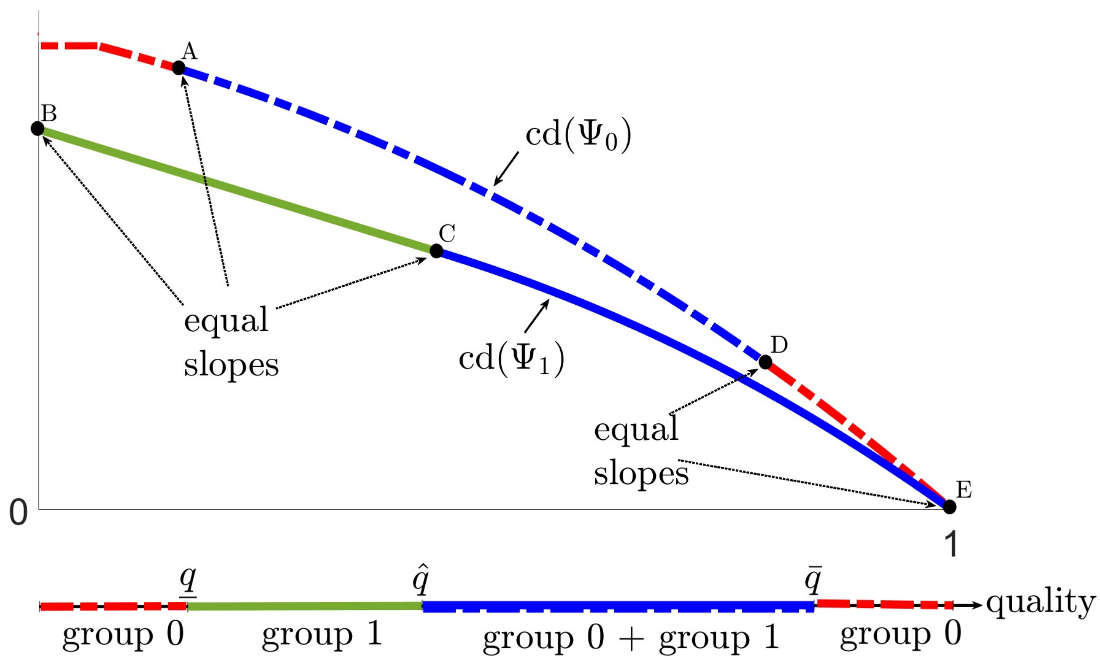

Figure 5.2: Illustration of the across-group solution: The graph depicts $\text{cd}(\Psi_i)$ for both groups $i = 0, 1$, while the line below depicts the resulting allocation of quality.

Figure 5.2 illustrates, qualitatively, how the greedy algorithm from Theorem 2 produces an optimal across-group allocation. Qualities are allocated from lowest to highest (as indicated in the bottom line in the figure), with the "next" level of quality allocated to the group with a smaller (absolute value of the) slope of the function $\text{cd}(\Psi_i)$. Intuitively, the slope of

$\mathrm{cd}(\Psi_i)$ at x captures the marginal value of quality given that a mass x of qualities has been already allocated according to the optimal within-group mechanism.

For intuition, it is helpful to think of α being large relative to $\bar{\lambda}_0$, so that we can approximately treat the designer as maximizing revenue in group 0. In particular, when $\alpha(\underline{r}_0 - 1) + \bar{\lambda}_0 < 0$, the slope of $\mathrm{cd}(\Psi_0)$ is zero in some initial interval, and the optimal mechanism for group 0 uses free disposal. Then, the lowest-quality objects are allocated to group 0 since they are optimally discarded anyway. More generally, the marginal value of allocation to low-WTP agents is low when revenue is the dominant motive for the designer. Allocation of qualities to group 0 continues until the slope of $\mathrm{cd}(\Psi_0)$ equalizes with the slope of $\mathrm{cd}(\Psi_1)$ at 0, as indicated by point A in Figure 5.2.

The next "batch" of qualities is allocated entirely to group 1 (see the green region in Figure 5.2). This is because the lowest qualities in group 1 are allocated randomly; as a result, the marginal value of allocation is constant in the mass of allocated objects in the random-allocation region, as reflected by the constant slope of $\mathrm{cd}(\Psi_1)$ in the interval from B to C.

As soon as assortative matching "kicks in" in group 0 (which happens at point C), the designer optimally splits qualities across both groups. For any quality level q in this region, the fraction of objects allocated to each group is between 0 and 1, and such that the marginal values are always equalized across the groups.

Finally, for the configuration in Figure 5.2, the highest qualities are allocated to group 0. The reason is that the slope of $\mathrm{cd}(\Psi_1)$ at 1 is equal to the slope of $\mathrm{cd}(\Psi_0)$ at some interior point (point D in the figure); this implies that when the designer optimally allocates the last unit of quality to group 1, there are agents in group 0 who generate even higher marginal value—they must receive the highest quality under assortative matching.

The following result summarizes our reasoning, and gives conditions under which the various regions we described are non-degenerate.

Result. There exist cutoffs $0 \le \underline{q} < \hat{q} < \bar{q} \le 1$ such that, in the optimal across-group allocation, objects of quality $q \le \underline{q}$ are allocated to group 0 (with potentially some objects discarded), objects of quality $q \in [\underline{q}, \hat{q}]$ to group 1, objects of quality $q \in [\hat{q}, \bar{q}]$ to both groups (in the sense that each q in this interval is shared by both groups), and objects of quality $q \ge \bar{q}$ to group 0. Additionally, $\underline{q} > 0$ when $\alpha(\underline{r}_0 - 1) + \bar{\lambda}_0 < 0$; and $\bar{q} < 1$ when $\underline{r}_0 > \underline{r}_1$.

Finally, to illustrate Proposition 6, we suppose that a fully random allocation is optimal for group 1. Optimality of fully random matching means that $\mathrm{cd}(\Psi_1)$ is affine, and its slope is constant. In Figure 5.2, this would correspond to the line segment $\overline{\mathrm{BC}}$ being "stretched" to the entire interval [0, 1], and the curved segment $\widehat{\mathrm{CE}}$ removed. In this case, the greedy algorithm allocates all qualities to group 1 in one "batch." In contrast, group 0 receives extremal qualities.

6 Market design implications

The optimal mechanism in our framework is always a combination of (i) *random matching*, which can be seen as a form of in-kind redistribution, and (ii) *assortative matching*, corresponding to the allocation that would arise in a competitive market equilibrium. Random matching can be optimal only when there is enough dispersion in the welfare weights to merit the allocative distortion—and even then, for random matching to be optimal, the designer needs to be able to identify sufficient information about the inequalities in agents' unobserved social welfare weights to be able to target the redistribution properly. The designer can observe, directly or through the mechanism, the label and the willingness to pay. Those give rise to two distinct paths for in-kind redistribution to be optimal:

1. *Label-revealed inequality:* If some label i identifies a group of agents that have a higher welfare weight on average than the weight on revenue α, then Proposition 1 shows that in-kind redistribution becomes optimal when the good being allocated is universally desired.

   Food stamp programs serve as an illustration. Group i can be defined by a set of verifiable eligibility criteria—such as low income—that are strongly correlated with what society associates with those most in need. For various reasons, it might be impractical, politically infeasible, or costly to give monetary transfers to group i, so that $\alpha < \bar{\lambda}_i$ may hold. Then, since food (defined broadly enough) is a universally desired good, in-kind redistribution can be justified by our Proposition 1. We can furthermore ask whether the simple form that many of these programs take—providing an undifferentiated food stamp free of charge—is optimal. Condition (4.1) in Proposition 2 is sufficient (and almost necessary) for optimality of providing a constant quality at a zero price. This condition is more likely to hold when the dispersion in willingness to pay is low.[24] This may indeed be the case in the context of food aid: Food stamp programs tend to be addressed to relatively poor households which do not vary significantly in their ability to pay. Moreover, when the recipients have discretion in choosing individual food items, differences in willingness to pay due to dietary preferences should also be small. Thus, our framework provides a justification for allocating the same food stamp free of charge to everyone who is eligible.

[24] In fact, as $\bar{r}_i$ approaches $\underline{r}_i$, condition (4.1) reduces to the requirement that $\alpha < \bar{\lambda}_i$.

In contrast, consider the example of public housing programs. In some countries, as many as a third of households are eligible for some form of housing assistance, implying that the ability (and hence willingness) to pay of some recipients could be quite large.[25] In such cases, in light of Proposition 2, a fully random allocation is unlikely to be optimal. A superior solution, based on Proposition 1 and illustrated in Section 5, is to provide the lowest-quality houses at a minimal price in a lottery, and use a price gradient for granting access to higher-quality housing. Using a low price for the lottery ensures that even those who opt for higher quality can be charged a below-market price. At the same time, the price gradient ensures a more efficient allocation, and raises more revenue.

2. *WTP-revealed inequality:* The second rationale for using in-kind redistribution in our framework (Proposition 3) is when willingness to pay reveals information about the welfare weights. In general, differences in willingness to pay may reflect both differences in idiosyncratic preferences as well as differences in ability to pay. For markets in which there is strong negative correlation between willingness to pay and the welfare weights, the designer can use the information revealed by agents' behavior to specifically target those individuals within a group who are likely to have a high welfare weight.

   For example, a patient who has a low willingness to pay for an important medical treatment is more likely to be poor, and thus to have a high expected welfare weight. In such cases, the designer can redistribute by introducing a reduced-price lottery for low-quality health care (e.g., providing health care services with higher waiting times) in order to separate low- and high-WTP agents—and subsidize the former via a reduced price.

Our framework also identifies two distinct forces supporting the use of market mechanisms even in the presence of redistributive concerns:

1. *The revenue motive:* As predicted by Proposition 2, assortative matching will be used for at least some agents as long as the weight on revenue α exceeds the average Pareto weight $\bar{\lambda}_i$ in a given group i. Moreover, by Proposition 3, for (Myerson-)regular distributions of willingness to pay, the fraction of objects allocated using the market mechanism increases with α. A high α occurs naturally when revenue is a driving objective unto itself (e.g., when the marketplace owner is a private, for-profit institution). However, the weight on revenue can also be high in public contexts in which

[25] van Dijk (2019) notes that "34% of Dutch households, 26% of Austrian households, and 19% of French households live in subsidized housing."

it is possible to subsidize selected groups of agents via direct lump-sum transfers, or when the designer uses the revenue to fund an outside cause that is socially valuable.

For example, consider a government designing an auction to allocate goods such as spectrum licenses or oil and gas leases to firms. Because the social value of a dollar funding the government budget is probably higher than the marginal social value of giving a dollar to the firms participating in the auction, assortative matching is typically optimal, perhaps with some restriction on supply to further increase revenue.

The same force behind optimality of assortative matching applies in any situation in which direct label-specific lump-sum payments are feasible (so that $\alpha \geq \bar{\lambda}_i$ for any group i). For example, if it is feasible to give cash transfers to those eligible for public housing (perhaps in the form of tax credits), then there is an argument against using lotteries to allocate public housing—we can do better by allocating assortatively at least at the top of the distribution of willingness to pay, and using the resulting revenue to fund monetary transfers to all eligible agents.

2. *The efficiency motive:* Assortative matching is optimal for maximizing the efficiency of the allocation—and this force works in favor of a market allocation even when the weight on revenue α is strictly below the average Pareto weight $\bar{\lambda}_i$. Efficiency becomes the dominant force when Pareto weights do not vary too much with willingness to pay, conditional on the label i. Indeed, Proposition 3 implies that a fully assortative matching becomes optimal when $\alpha \geq \max_r\{\lambda_i(r) - \Lambda_i(r)\}$, which can be true even for very low α when there is little dispersion in $\lambda_i(r)$.

   Low dispersion in $\lambda_i(r)$ can arise in two cases: (i) when the designer does not have strong redistributive preferences to begin with (there is little dispersion in the unobserved welfare weights) or, more interestingly, (ii) when willingness to pay is not correlated with the underlying welfare weights, conditional on the label. Observation (ii) helps explain why a market allocation is desirable for most goods and services even when the designer has strong preferences for redistribution. Agents' needs are unlikely to be strongly correlated with willingness to pay for goods that are relatively cheap (affordable, at least in small quantities, to most people) and whose value depends heavily on tastes. Additionally, the residual correlation between willingness to pay and the unobserved welfare weights decreases when more information becomes available in the form of labels (see the interpretation of the model in Section 2). For example, if a country provides free health care to eligible citizens, it becomes less likely that the low willingness to pay in the non-eligible group reflects adverse social or economic circumstances, since these circumstances would likely be partly captured by the label.

Hence, using in-kind redistribution to address label-reveled inequality should often be expected to coexist with a market allocation to the populations that are not being targeted for redistribution.

7 Concluding remark

Focusing on an objective function that assigns arbitrary welfare weights to market participants sets this work apart from the standard mechanism design paradigm. Indeed, while the mechanism design literature has developed an impressive framework for designing revenue-maximizing auctions and allocatively efficient mechanisms, there has been far less focus on how to use those same tools to understand the ways in which the structure of optimal mechanisms responds to redistributive goals. Our paper is thus one of relatively few attempts thus far using mechanism design to give guidance to real-world market designers about how to optimally structure market-level redistributive systems. We hope to see more work devoted to this problem.

References

Aidt, T. and F. Giovannoni (2011): "Critical decisions and constitutional rules," Social Choice and Welfare, 219–268.

Akbarpour, M. ⓡ E. Budish ⓡ P. Dworczak ⓡ S. D. Kominers (2023): "An economic framework for vaccine prioritization," forthcoming: Quarterly Journal of Economics.

Akerlof, G. A. (1978): "The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning," American Economic Review, 68, 8–19.

Alatas, V., R. Purnamasari, M. Wai-Poi, A. Banerjee, B. A. Olken, and R. Hanna (2016): "Self-targeting: Evidence from a field experiment in Indonesia," Journal of Political Economy, 124, 371–427.

Ashlagi, I., F. Monachou, and A. Nikzad (2020): "Optimal Dynamic Allocation: Simplicity through Information Design," Working Paper.

Atkinson, A. B. and J. E. Stiglitz (1976): "The design of tax structure: direct versus indirect taxation," Journal of Public Economics, 6, 55–75.

Bauer, S. (2022): "Buyers' Welfare Maximizing Auction Design," forthcoming: International Journal of Game Theory.

Besley, T. and S. Coate (1991): "Public provision of private goods and the redistribution of income," The American Economic Review, 81, 979–984.

Blackorby, C. and D. Donaldson (1988): "Cash versus kind, self-selection, and efficient transfers," The American Economic Review, 691–700.

Bodoh-Creed, A. L. and B. R. Hickman (2018): "College assignment as a large contest," Journal of Economic Theory, 175, 88 – 126.

Brooks, B. and S. Du (2021): "Optimal Auction Design With Common Values: An Informationally Robust Approach," Econometrica, 89, 1313–1360.

Chakravarty, S. and T. R. Kaplan (2013): "Optimal allocation without transfer payments," Games and Economic Behavior, 77, 1 – 20.

Che, Y.-K., W. Dessein, and N. Kartik (2013a): "Pandering to Persuade," American Economic Review, 103, 47–79.

Che, Y.-K., I. Gale, and J. Kim (2013b): "Assigning resources to budget-constrained agents," Review of Economic Studies, 80, 73–107.

Condorelli, D. (2012): "What money can't buy: Efficient mechanism design with costly signals," Games and Economic Behavior, 75, 613 – 624.

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

Cremer, H. and F. Gahvari (1997): "In-kind transfers, self-selection and optimal tax policy," European Economic Review, 41, 97–114.

Currie, J. and F. Gahvari (2008): "Transfers in cash and in-kind: Theory meets the data," Journal of Economic Literature, 46, 333–383.

Diamond, P. A. and J. A. Mirrlees (1971): "Optimal taxation and public production I: Production efficiency," American Economic Review, 61, 8–27.

Dworczak, P. and G. Martini (2019): "The Simple Economics of Optimal Persuasion," Journal of Political Economy, 127, 1993–2048.

Dworczak, P. ⓡ S. D. Kominers ⓡ M. Akbarpour (2021): "Redistribution through markets," Econometrica, 89, 1665–1698.

Echenique, F. and M. B. Yenmez (2015): "How to control controlled school choice," American Economic Review, 105, 2679–2694.

Fan, X., Y.-J. Chen, and C. S. Tang (2021): "Allocating Scarce Resources in the Presence of Private Information and Bargaining Power," HKUST Business School Research Paper No. 2021-021.

Gahvari, F. and E. Mattos (2007): "Conditional cash transfers, public provision of private goods, and income redistribution," American Economic Review, 97, 491–502.

Gentzkow, M. and E. Kamenica (2016): "A Rothschild-Stiglitz Approach to Bayesian Persuasion," American Economic Review, 106, 597–601.

Guesnerie, R. and K. Roberts (1984): "Effective policy tools and quantity controls," Econometrica, 59–86.

Hartline, J. D. and T. Roughgarden (2008): "Optimal mechanism design and money burning," in Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing, 75–84.

Jehiel, P. and B. Moldovanu (2001): "Efficient Design with Interdependent Valuations," Econometrica, 69, 1237–1259.

Kang, M. and C. Z. Zheng (2020): "Pareto Optimality of Allocating the Bad," Working Paper.

——— (2022): "Optimal Design for Redistributions among Endogenous Buyers and Sellers," Economic Theory, 1141–1180.

Kang, Z. Y. (2020a): "Markets for goods with externalities," Working Paper.

——— (2020b): "Optimal public provision of private goods," Working Paper.

Kleiner, A., B. Moldovanu, and P. Strack (2021): "Extreme Points and Majorization: Economic Applications," Econometrica, 89, 1557–1593.

Kolotilin, A. (2018): "Optimal Information Disclosure: A Linear Programming Approach," Theoretical Economics, 13, 607–636.

Liscow, Z. D. and A. D. Pershing (2022): "Why Is So Much Redistribution In-Kind and Not in Cash? Evidence from a Survey Experiment," forthcoming: National Tax Journal.

McAfee, R. P. and J. McMillan (1992): "Bidding Rings," American Economic Review, 82, 579–599.

Muir, E. V. and S. Loertscher (2022): "Monopoly pricing, optimal randomization, and resale," Journal of Political Economy, 130.

Müller, A. and D. Stoyan (2002): Comparison Methods for Stochastic Models and Risks, Wiley.

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

Nichols, A. L. and R. J. Zeckhauser (1982): "Targeting transfers through restrictions on recipients," The American Economic Review, 72, 372–377.

Pathak, P., G. Persad, T. Sönmez, and M. Utku Ünver (2022): "Reserve system design for allocation of scarce medical resources in a pandemic: Some perspectives from the field," Oxford Review of Economic Policy, 38, 924–940.

Pathak, P. A., T. Sönmez, M. U. Ünver, and M. B. Yenmez (2020): "Fair Allocation of Vaccines, Ventilators and Antiviral Treatments: Leaving No Ethical Value Behind in Health Care Rationing," Working Paper.

Reuter, M. and C.-C. Groh (2020): "Mechanism Design for Unequal Societies," Working Paper.

Saez, E. and S. Stantcheva (2016): "Generalized social marginal welfare weights for optimal tax theory," American Economic Review, 106, 24–45.

Schmidt, H., P. Pathak, T. Sönmez, and M. U. Ünver (2020): "Covid-19: how to prioritize worse-off populations in allocating safe and effective vaccines," British Medical Journal, 371.

Toikka, J. (2011): "Ironing without control," Journal of Economic Theory, 146, 2510 – 2526.

van Dijk, W. (2019): "The socio-economic consequences of housing assistance," Working Paper.

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

Appendix

A Proofs of the results in the main text

A.1 Proof of Theorem 1

Because the within-group problem can be solved for each group i separately, we fix $i \in I$, and drop the subscripts i to simplify notation. We prove the theorem under the assumption that the designer maximizes a general objective function of the form

$$\int_{\underline{r}}^{\bar{r}} V(r) Q^{\Gamma}(r) dG(r) + v\,\underline{U} \tag{A.1}$$

for some upper semi-continuous function $V : [\underline{r},\, \bar{r}] \to \mathbb{R}$ and some constant $v \in \mathbb{R}$.

Fixing the CDF of available quality F, we call an expected quality schedule $Q : [\underline{r},\, \bar{r}] \to [0,\, 1]$ *feasible* if $Q = Q^{\Gamma}$ for some $\Gamma : [\underline{r},\, \bar{r}] \to \Delta(Q)$, and there exist transfers t such that $(\Gamma,\, t)$ is a feasible mechanism. Given two CDFs $F,\, G : [a,\, b] \to [0,\, 1]$, we say that F is a *mean-preserving spread (MPS)* of G if

$$\int_a^t F(x)dx \geq \int_a^t G(x)dx, \quad \forall t \in [a,\, b],$$

with equality for $t = b$. We say that F *first-order stochastically dominates* G if $F(x) \leq G(x)$ for all $x \in [a,\, b]$. The following lemma describes all feasible expected quality assignments, assuming no free disposal.

Lemma 1. *If F is the CDF of available qualities, then $Q(r)$ is a feasible assignment of expected qualities (with no free disposal) if and only if $Q(r) = \Phi(G(r))$, where $\Phi : [0,\, 1] \to [0,\, 1]$ is a non-decreasing, left-continuous, mean-preserving spread of F^{-1}.*

Proof. Since $F(q)$ is a CDF, we can apply Strassen's Theorem (see Theorem 3.4.2(a) of Müller and Stoyan, 2002): A CDF $\bar{F}(q)$ is a distribution of posterior means of a random variable distributed according to F if and only if F is a mean-preserving spread of $\bar{F}$. Moreover, by the usual argument, the IC constraint (2.3) implies that the assignment of expected qualities must be non-decreasing. This monotonicity condition uniquely pins down $Q(r)$ given $\bar{F}$ and G: $\bar{F}(q)$ is the (normalized) mass of objects of quality q or less available to agents; this mass must be allocated to agents with willingness to pay r or lower; therefore, for any q, there exists r such that $\bar{F}(q) = G(r)$, and it follows that

$$Q(r) = \bar{F}^{-1}(G(r)).$$

Finally, we claim that a function Φ is equal to $\bar{F}^{-1}$ for some feasible $\bar{F}$ if and only if Φ satisfies the conditions of the lemma. That is,

$$\begin{aligned}&\bar{F} \text{ is a CDF on } [0,\,1] \text{ and } F \text{ is a MPS of } \bar{F} \iff \\ &\qquad \bar{F}^{-1} : [0,\,1] \to [0,\,1] \text{ is non-decreasing, left-continuous, and } \bar{F}^{-1} \text{ is a MPS of } F; \end{aligned} \tag{A.2}$$

this follows from Lemma 1 found in Brooks and Du (2021). □

The proof of Lemma 1 can be understood through its connection to information design: We can treat F as the prior distribution of a random variable X (quality); Strassen's Theorem implies that a distribution $\bar{F}$ of posterior means of X can be induced from the prior F (under some signal when X is treated as a state variable) if and only if F is a mean-preserving spread of $\bar{F}$. Hence, in our assignment problem, mean-preserving contractions of the distribution F describe all feasible distributions of expected quality. Moreover, incentive-compatibility constraints imply that there is a unique assignment of expected qualities to types because the assignment must be monotone in the willingness to pay r.

Because the function $\Phi(q)$ from Lemma 1 is left-continuous, its value at 0 is not pinned down. This is a reflection of the fact that the designer's expected payoff from the mechanism does not depend on the allocation for a measure-zero set of types, in particular, on the allocation of type $\underline{r}$. However, the allocation for type $\underline{r}$, $Q(\underline{r})$, appears in the constraint defining the non-negative transfers condition. This constraint is most permissive when $Q(\underline{r})$ is set to its maximal feasible level which is $Q(\underline{r}^+)$ (since Q must be non-decreasing). (Here, and hereafter, we denote $f(x^+) = \lim_{y \searrow x} f(y)$.) Because it is convenient to keep Φ left-continuous also at 0, we will extend the function Φ by assuming that $\Phi(x) = 0$ for all $x \le 0$, and then the non-negative transfers condition becomes $\underline{U} \le \underline{r}\Phi(0^+)$.

Given Lemma 1, we can write the problem of maximizing (A.1) under no-free-disposal as

$$\max_{\Phi} \left\{ \int_{\underline{r}}^{\bar{r}} V(r)\Phi(G(r))dG(r) + \max\{0,\, v\}\,\underline{r}\,\Phi(0^+) \right\}$$

subject to

$$\Phi \text{ is a MPS of } F^{-1}.$$

Indeed, notice that when $v \le 0$, it is optimal to choose $\underline{U}$ as low as possible, and hence $\underline{U} = 0$ in the optimal mechanism ($\underline{U} \ge 0$ by individual rationality). In contrast, when $v > 0$, the non-negative transfers condition implies that it is optimal to set $\underline{U}$ to its maximal feasible level $\underline{r}\,\Phi(0^+)$.

Integration by parts and by substitution yields

$$\int_{\underline{r}}^{\bar{r}} V(r)\Phi(G(r))dG(r) = \int_0^1 \left(\int_t^1 V(G^{-1}(x))dx \right) d\Phi(t).$$

Whenever we write $\int f(x)d\Phi(x)$ for some measurable function f, we mean the Lebesgue integral with respect to the σ-additive measure μ_Φ defined by $\mu_\Phi([a, b]) = \Phi(b^+) - \Phi(a)$, in particular, $\mu_\Phi(\{a\}) = \Phi(a^+) - \Phi(a)$. Under this convention, and recalling that $\Phi(x) = 0$ for $x \le 0$, we can also write

$$\Phi(0^+) = \int_0^1 \mathbf{1}_{\{t=0\}} d\Phi(t).$$

Then, we can write (A.1) as

$$\int_0^1 \left(\int_t^1 V(G^{-1}(x))dx + \max\{0, v\}\,\underline{r}\,\mathbf{1}_{\{t=0\}} \right) d\Phi(t).$$

Using the definition of Ψ from Theorem 1, we conclude that the objective function is $\int_0^1 \Psi(x)d\Phi(x)$; problems of this form admit an easy-to-describe solution.

Lemma 2. *Consider the problem*

$$\max_{\Phi:\,\Phi \text{ is a MPS of} \Phi_0} \left\{ \int_0^1 \Psi(x)d\Phi(x) \right\},$$

where $\Psi(x)$ is an upper semi-continuous function and Φ_0 is given. The value of the problem is $\int_0^1 co(\Psi)(x)d\Phi_0(x)$, and the solution is given by

$$\Phi^\star(x) = \begin{cases} \frac{\int_a^b \Phi_0(x)dx}{b-a} & \text{if } x \in [a, b] \text{ and } [a, b] \text{ is a maximal interval on which } co(\Psi) \text{ is affine,} \\ \Phi_0(x) & \text{otherwise,} \end{cases}$$

for almost all x.

Proof. For any Φ, we have

$$\int_0^1 \Psi(x)d\Phi(x) \le \int_0^1 \mathrm{co}(\Psi)(x)d\Phi(x).$$

Moreover, the function on the right hand side of the inequality is maximized at $\Phi = \Phi_0$

because $\text{co}(\Psi)(x)$ is a concave function. It follows that the value of the problem in the lemma is bounded by $\int_0^1 \text{co}(\Psi)(x)d\Phi_0(x)$. We show that this upper bound can be achieved. Consider the candidate solution $\Phi^\star(x)$ from the statement of the lemma. First, this function is feasible (by Gentzkow and Kamenica, 2016). Moreover, $\text{supp}(\Phi^\star) \subseteq \{x : \Psi(x) = \text{co}(\Psi)(x)\}$, and on that set, $\Phi^\star = \Phi_0$. Indeed, whenever $\Psi(x) < \text{co}(\Psi)(x)$, x must lie in the interior of an interval in which $\text{co}(\Psi)(x)$ is affine, and hence, by definition, $\Phi^\star(x)$ is constant in that region. Thus, $\int_0^1 \Psi(x)d\Phi^\star(x) = \int_0^1 \text{co}(\Psi)(x)d\Phi_0(x)$. □

The form of the solution is consistent with the concurrent findings of Kleiner et al. (2021), who derive general properties of extreme points that emerge as solutions to problems of the form considered in the lemma. The maximization problem in Lemma 2 can also be seen as analogous to a Bayesian persuasion problem in which the designer's preferences over posterior beliefs depend only on the posterior mean (see Kolotilin, 2018, and Dworczak and Martini, 2019) with a key difference: The MPS condition is flipped, requiring the solution Φ to be a mean-preserving spread (rather than a mean-preserving contraction) of the prior Φ_0. This makes the problem very easy to solve by finding a concave closure of the objective function.

Lemmas 1 and 2 immediately imply that the value of the maximization problem under no-free-disposal is given by

$$\int_0^1 \text{co}(\Psi)(x)dF^{-1}(x) = \int_0^1 \text{co}(\Psi)(F(q))dq,$$

where the equality follows from integration by substitution. Moreover, a solution is given by $Q^\star(r) = \Phi^\star(G(r))$, where $\Phi^\star$ is described in Lemma 2.

Next, we modify the solution to allow for free disposal. Allowing for free disposal is equivalent to allowing for "downward" first-order stochastic dominance shifts in the distribution of expected quality allocated to agents. That is, $Q(r)$ is a feasible expected-quality schedule with free disposal if $Q(r) = \bar{\Phi}(G(r))$ for some $\bar{\Phi} \leq \Phi$, where Φ is a mean-preserving spread of F^{-1} (see Lemma 1). Note that $\bar{\Phi}$ dominates Φ in the FOSD order because the FOSD relation is reversed by taking the inverse of the CDFs (and both $\bar{\Phi}$ and Φ are inverses of the CDFs of the expected quality). Therefore, to derive the optimal expected-quality schedule under free disposal from the corresponding solution without free disposal, it is enough to solve an optimization problem of the following form:

Lemma 3. *Consider the problem*

$$\max_{\Phi} \int_0^1 co(\Psi)(x)d\Phi(x)$$

subject to

$$\Phi(x) \leq \Phi^\star(x),$$

where $\Phi^\star(x)$ *is the solution given in Lemma 1. The value of the problem is* $\int_0^1 \mathrm{cd}(\Psi)(x)d\Phi^\star(x)$, *and the solution is given by*

$$\Phi^{\star\star}(x) = \Phi^\star(x)\mathbf{1}_{\{x \geq x^\star\}}$$

for almost all x, *where* $[0, x^\star]$ *is the maximal interval on which the concave decreasing function* $\mathrm{cd}(\Psi)$ *is constant.*

Proof. By definition of $x^\star$, the function $\mathrm{cd}(\Psi)(x)$ is constant and equal to $\mathrm{co}(\Psi)(x^\star)$ on $[0, x^\star]$ and coincides with $\mathrm{co}(\Psi)(x)$ otherwise. On one hand, we have for any feasible Φ,

$$\int_0^1 \mathrm{co}(\Psi)(x)d\Phi(x) \leq \int_0^1 \mathrm{cd}(\Psi)(x)d\Phi(x) \leq \int_0^1 \mathrm{cd}(\Psi)(x)d\Phi^\star(x),$$

where the first inequality follows from the fact that $\mathrm{co}(\Psi) \leq \mathrm{cd}(\Psi)$, and the second follows from the fact that $\mathrm{cd}(\Psi)$ is non-increasing and Φ dominates $\Phi^\star$ in the FOSD order. On the other hand, if we define $\Phi^{\star\star}$ as in the statement of the lemma, then we have

$$\begin{aligned}\int_0^1 \mathrm{co}(\Psi)(x)d\Phi^{\star\star}(x) &= \int_0^{x^\star} \mathrm{co}(\Psi)(x)d\Phi^{\star\star}(x) + \int_{x^\star}^1 \mathrm{co}(\Psi)(x)d\Phi^{\star\star}(x) \\ &= \mathrm{co}(\Psi)(x^\star)\Phi^\star(x^\star) + \int_{x^\star}^1 \mathrm{cd}(\Psi)(x)d\Phi^\star(x) = \int_0^1 \mathrm{cd}(\Psi)(x)d\Phi^\star(x)\end{aligned}$$

by the properties of $\mathrm{co}(\Psi)$, $\mathrm{cd}(\Psi)$, and $\Phi^{\star\star}(x)$. Thus, $\Phi^{\star\star}$ achieves the upper bound and hence is a solution to the problem described in Lemma 3. □

With Lemma 3, Theorem 1 follows directly from Lemma 1: The value of the problem is

$$\int_0^1 \mathrm{cd}(\Psi)(x)d\Phi^\star(x) = \int_0^1 \mathrm{cd}(\Psi)(x)dF^{-1}(x) = \int_0^1 \mathrm{cd}(\Psi)(F(q))dq,$$

where the last equality follows from integration by substitution. The optimal solution is given by an expected-quality schedule

$$Q^\star(r) = \Phi^{\star\star}(G(r)) = \Phi^\star(G(r))\mathbf{1}_{\{G(r) \geq x^\star\}} = \Phi^\star(G(r))\mathbf{1}_{\{r \geq G^{-1}(x^\star)\}},$$

where $\Phi^\star$ is described in Lemma 1. Finally, the choice of the optimal $\underline{U}$ was described in the discussion leading up to Lemma 2.

A.2 Proof of Theorem 2

We solve the program (3.2)–(3.3) by solving a relaxed problem in which the constraint that $F_i(q)$ is a CDF is dropped, and then verifying that the solution of the relaxed program is feasible. The relaxed program is to solve for the optimal $F_i(q)$ for every $q \in Q$ separately:

$$\max_{0 \le x_i \le 1} \left\{ \sum_{i \in I} \mu_i \mathrm{cd}(\Psi_i)(x_i) \right\} \tag{A.3}$$

$$\text{s.t.} \sum_{i \in I} \mu_i \, x_i = F(q). \tag{A.4}$$

This program can be solved using standard Lagrangian techniques (constraint qualification holds trivially in our problem). Fix $q \in Q$. There exists a Lagrange multiplier,[26] which we denote by $L(q)$, such that the optimal $x_i^\star$ maximizes $\sum_{i \in I} \mu_i[\mathrm{cd}(\Psi_i)(x_i) - L(q)x_i]$ while satisfying the constraint (A.4). Because the Lagrangian is concave, the first-order condition is both necessary and sufficient. Let $X_i^\star(q)$ be the set of points satisfying the first-order condition: $X_i^\star(q) = \{x : \mathrm{cd}(\Psi_i)'(x) = L(q)\}$ whenever this set is non-empty, and otherwise $X_i^\star(q) = \{0\}$ if $\mathrm{cd}(\Psi_i)'(0) < L(q)$ and $X_i^\star(q) = \{1\}$ if $\mathrm{cd}(\Psi_i)'(1) > L(q)$. By the preceding argument, we know that there exists a selection $x_i^\star \in X_i^\star(q)$ such that (A.4) holds. Moreover, because each $\mathrm{cd}(\Psi_i)$ is concave and continuous, we know that each $X_i^\star(q)$ is a closed interval (potentially a singleton).

To prove the theorem, it remains to show that there exists a selection $F_i^\star(q)$ from each $X_i^\star(q)$ that is non-decreasing (then, it can be modified on a measure-zero set of points to make it into a CDF; notice that it is guaranteed by the constraint (A.4) that each $F_i^\star$ is 0 at 0 and 1 at 1).

Because the constraint in (A.4) is increasing in q, it follows that the Lagrange multiplier $L(q)$ is a non-increasing function of q. Moreover, the sets $X_i^\star(q)$ are non-decreasing in the strong set order by concavity of $\mathrm{cd}(\Psi_i)$. Define a vector function

$$C(q, \alpha) = \left[(1-\alpha)\min X_1^\star(q) + \alpha \max X_1^\star(q), \ldots, (1-\alpha)\min X_{|I|}^\star(q) + \alpha \max X_{|I|}^\star(q)\right].$$

By definition, for each q, $\sum_i C_i(q, 0) \le F(q)$ while $\sum_i C_i(q, 1) \ge F(q)$. By continuity, there exists $\alpha^\star(q)$ such that $\sum_i C_i(q, \alpha^\star(q)) = F(q)$ (moreover, the values of $C_i(q, \alpha^\star(q))$ are uniquely pinned down, even if $\alpha^\star(q)$ is not). We can now define $F_i^\star(q)$ as $C_i(q, \alpha^\star(q))$. By direct inspection and the strong-set order property of $X_i^\star(q)$, each $F_i^\star(q)$ is non-decreasing, which finishes the proof once we set $V^{\min}(q) = -L(q)$.

[26] In case there are multiple Lagrange multipliers, we pick the largest one.

A.3 Proofs of the results in Section 4

Proof of Proposition 1. The proof is immediate from Theorem 1. The assumptions of Proposition 1 ensure that there is an upward jump in Ψ_i at 0, and therefore $\mathrm{cd}(\Psi_i)(x)$ must be affine on $[0, x]$, for some small enough x. (Of course, when $\mathrm{cd}(\Psi_i)(x)$ is constant for small x, it is possible that types $r \leq r_i^\star$ do not receive any objects; however, this is still random matching according to our definition; see also Remark 1.) □

Proof of Proposition 2. When F_i is a non-degenerate distribution, by Theorem 1, full randomization is optimal if and only if $\mathrm{cd}(\Psi_i)$ is affine, which is true if and only if

$$\Psi_i(x) \leq (1-x)\Psi_i(0) + x\Psi_i(1)$$

for all $x > 0$. We have

$$\Psi_i(0) = \max\{0,\, \alpha - \bar{\lambda}_i\}\underline{r}_i + \int_{\underline{r}_i}^{\bar{r}_i} \tau\lambda_i(\tau)dG_i(\tau).$$

Using the fact that $\Psi_i(1) = 0$, we can write the condition as, for all $r > \underline{r}_i$,

$$\Psi_i(G_i(r)) \leq (1 - G_i(r)) \left[\max\{0,\, \alpha - \bar{\lambda}_i\}\underline{r}_i + \int_{\underline{r}_i}^{\bar{r}_i} \tau\lambda_i(\tau)dG_i(\tau)\right]. \tag{A.5}$$

To see that this implies $\alpha < \bar{\lambda}_i$, note that by diving both sides by $1 - G_i(r)$, using the expression (3.1), and taking the limit as $r \to \bar{r}_i$, we get

$$\alpha\bar{r}_i \leq \max\{0,\, \alpha - \bar{\lambda}_i\}\underline{r}_i + \int_{\underline{r}_i}^{\bar{r}_i} \tau\lambda_i(\tau)dG_i(\tau) < \max\{0,\, \alpha - \bar{\lambda}_i\}\underline{r}_i + \bar{r}_i\bar{\lambda}_i. \tag{A.6}$$

Thus, if $\alpha \geq \bar{\lambda}_i$, we would get $(\alpha - \bar{\lambda}_i)\bar{r}_i < (\alpha - \bar{\lambda}_i)\underline{r}_i$ which is a contradiction. Using this observation to simplify (A.6), we obtain the necessary condition (4.1).

Finally, suppose that $\alpha J_i(r) + \Lambda_i(r)h_i(r)$ is quasi-convex. This implies that Ψ_i is first convex and then concave on $(0, 1]$. The necessary condition implies that $\Psi_i(0) \geq \Psi_i(1) - \Psi'(1)$. Together, these two facts imply that $\Psi_i(x) \leq (1 - x)\Psi_i(0)$, for all x. □

Proof of Proposition 3. The first assumption guarantees that Ψ_i does not have a jump at 0, and hence is a continuous function. Then, Theorem 1 implies that effectively assortative matching is optimal if and only if $V_i(r)$ is non-decreasing in r. (Strictly speaking, this conclusion does not follow from the statement of Theorem 1 when $V_i(r)$ is constant on some intervals, since in this case $\mathrm{co}(\Psi_i)$ may have affine parts. However, the proof of Theorem 1 makes clear that on intervals $[a, b]$ on which $\text{co}(\Psi_i)$ is affine *and* $\text{co}(\Psi_i) = \Psi_i$, the designer is indifferent between random and assortative matching, and hence an assortative matching is also optimal.) $\square$

Proof of Proposition 4. The conclusion is immediate from Theorem 2. When there is fully random matching in group i, the function $\text{cd}(\Psi_i)$ is affine, and thus its slope is constant, equal to $\Psi_i(0)$ (since $\Psi_i(1) = 0$). By Proposition 2, fully random matching requires that $\bar{\lambda}_i > \alpha$, and under this inequality, we have that $\Psi_i(0) = \int_{\underline{r}_i}^{\bar{r}_i} r\lambda_i(r)dG_i(r)$. $\square$

Proof of Proposition 5. By the assumption that effectively assortative matching is optimal, we must have $\text{cd}(\Psi_i)(x) = \Psi_i(x)$, except possibly for $x \leq x_i^\star$ if $\text{cd}(\Psi_i)(x)$ is constant on $[0, x_i^\star]$. By direct calculation (and using the fact that for bounded-support, positive-density distributions, the inverse hazard rate is 0 at the upper bound), we obtain $\Psi_i'(1) = -\alpha\bar{r}_i$. The conclusion follows directly from Theorem 2 and the observation that Ψ_i has a continuous derivative (by the assumptions that $g_i(r)$ and $\lambda_i(r)$ are continuous, and that $g_i(r)$ is strictly positive so that $h_i(r)$ is also continuous). When $\underline{r}_i = 0$, we have that

$$\Psi_i'(0) = -\alpha\left(\underline{r}_i - \frac{1}{g_i(\underline{r}_i)}\right) - \bar{\lambda}_i\frac{1}{g_i(\underline{r}_i)} > 0,$$

and hence Ψ_i is increasing in the neighborhood of 0. Thus, $\text{cd}(\Psi_i)$ is constant in some initial interval, and hence has a zero slope. When $\underline{r}_i = 0$ for all i, by Theorem 2, all groups are allocated the lowest-quality objects. $\square$

Proof of Proposition 6. The first part of the proposition follows immediately from Theorem 2 by observing that the slope of $\text{cd}(\Psi_1)$ is constant, while the (absolute value of the) slope of $\text{cd}(\Psi_0)(q)$ is increasing in q.

We prove the second part. For $\bar{q} < 1$, we need that $|\text{cd}(\Psi_0)'(1)| > |\text{cd}(\Psi_1)'(1)| = \Psi_1(0)$, since then Theorem 2 implies that the highest qualities are allocated to group 0. This yields the condition $\alpha\bar{r}_0 > \int_{\underline{r}_1}^{\bar{r}_1} \tau\lambda_1(\tau)dG_1(\tau)$. For $\underline{q} > 0$, we need that $|\text{cd}(\Psi_0)'(0)| < |\text{cd}(\Psi_1)'(0)| = \Psi_1(0)$, since then Theorem 2 implies that lowest qualities are allocated to group 0. Since group 0 features effectively assortative matching, either $|\text{cd}(\Psi_0)'(0)| = 0$ or $\text{cd}(\Psi_0)'(0) = \Psi_0'(0) > -\alpha\underline{r}_0$. Thus, we obtain the condition $\alpha\underline{r}_0 < \int_{\underline{r}_1}^{\bar{r}_1} \tau\lambda_1(\tau)dG_1(\tau)$. $\square$

Online Appendix

B A precise intuition for Proposition 1

In this appendix, we present a more precise intuition for Proposition 1. Consider Figure B.1 (we drop the subscript i to simplify notation). Suppose that the expected quality schedule $Q(r)$ is strictly increasing. Recall that utility of type r can be expressed as $U(\underline{r})+\int_{\underline{r}}^{r} Q(\tau)d\tau$. Under the assumption that $\bar{\lambda} > \alpha$, the designer wants to minimize prices subject to the constraint $t(r) \geq 0$, and hence we can assume $t(\underline{r}) = 0$ and $U(\underline{r}) = Q(\underline{r})\underline{r}$.

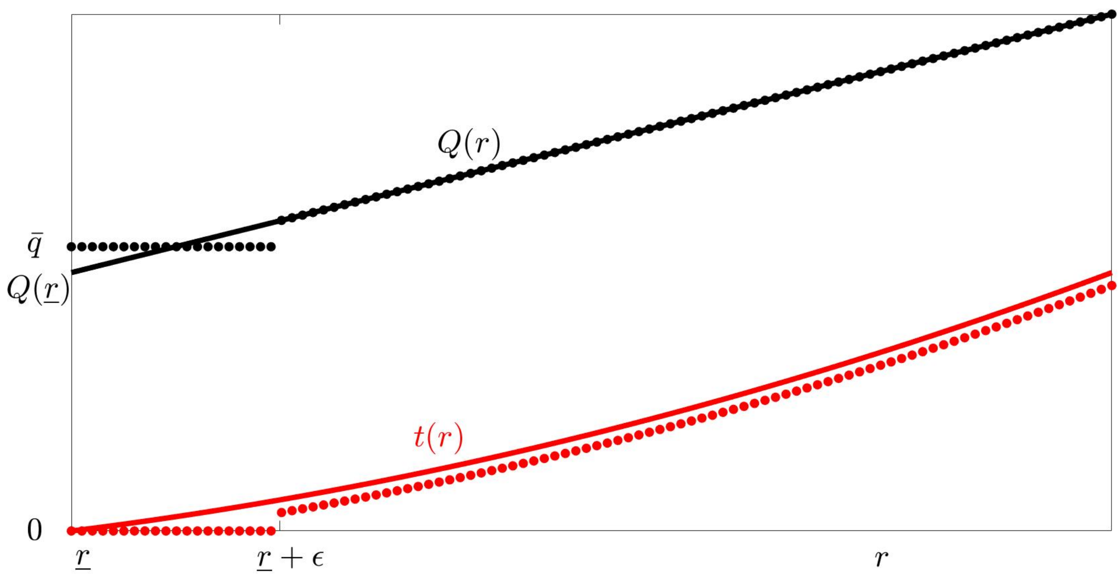

Figure B.1: An expected quality schedule $Q(r)$ and the corresponding payment rule $t(r)$ (solid lines). Dotted lines indicate the perturbation of the mechanism $(Q,\, t)$.

We show that the designer's objective is increased by a perturbation of the mechanism that allocates objects at random and for free to some small set of types $[\underline{r},\, \underline{r}+\epsilon]$. Indeed, such a perturbation allows the designer to decrease prices for everyone else (as long as $\underline{r} > 0$) while only causing a second-order distortion to allocative efficiency. Let q^{ϵ} denote the expected quality of objects allocated to types $[\underline{r},\, \underline{r}+\epsilon]$ under Q, and let $Q^{\epsilon}(r) = q^{\epsilon}$ for $r \leq \underline{r}+\epsilon$ and $Q^{\epsilon}(r) = Q(r)$ otherwise.[27] Setting $t^{\epsilon}(\underline{r}) = 0$ yields $U^{\epsilon}(\underline{r}) = q^{\epsilon}\underline{r}$. The associated change in utility for type r equals

$$\left[U^{\epsilon}(\underline{r}) + \int_{\underline{r}}^{r} Q^{\epsilon}(\tau)d\tau\right] - \left[U(\underline{r}) + \int_{\underline{r}}^{r} Q(\tau)d\tau\right] = (q^{\epsilon} - Q(\underline{r}))\underline{r} + \int_{\underline{r}}^{\min\{r,\underline{r}+\epsilon\}} (q^{\epsilon} - Q(\tau))d\tau.$$

[27]We can assume that $q^{\epsilon} > 0$; if $q^{\epsilon} = 0$, then there is nothing to prove since the allocation is already constant in some interval around the lowest type.

The first term is first-order in ϵ and captures the increase due to the increase in utility for the lowest type $\underline{r}$ (which happens as long as that type values the increase in quality, that is, $\underline{r} > 0$). For types above $\underline{r} + \epsilon$, this increase in utility is achieved via a price discount, which is possible when the allocation of type $\underline{r} + \epsilon$ is decreased (since this relaxes the IC constraints for all higher types). The second term is second-order in ϵ and captures the welfare effects of the distortion in allocation.

C Additional results for Section 4

C.1 Assortative matching at the top

In Sections 4 and 5, we emphasized that random matching of objects to the lowest-WTP agents often coincides with assortative matching "at the top of the distribution." Here, we formalize this observation.

We say that there is *assortative matching at the top* if the mechanism allocates the highest-quality objects assortatively to agents with willingness to pay r above some threshold. For the result, we assume that Pareto weights are non-increasing, i.e., that willingness to pay is negatively correlated with the unobserved welfare weights.

Proposition 7. *If $\lambda_i(r)$ is non-increasing in r, and $\alpha \geq \bar{\lambda}_i$, any optimal mechanism features assortative matching at the top within group i.*

The result is intuitive: Non-increasing Pareto weights along with the assumption that the weight on revenue is weakly larger than the average Pareto weight, imply that the weight on revenue is larger than the weight on the utility of agents with high willingness to pay. Since assortative matching is optimal for revenue maximization at the top of the distribution (the so-called "no distortion at the top" result), it dominates random matching for high enough r.

C.2 Non-monotonicity in the use of non-market mechanisms

In the context of the parametric example in Section 5, we have shown that the fraction of objects allocated randomly is non-monotone in the degree of inequality in Pareto weights; we now show that this conclusion obtains more generally, at least when $\alpha \geq \bar{\lambda}_i$.

Proposition 8. *Suppose that $J_i'(r) \geq \underline{J}_i$, for all r and some constant $\underline{J}_i > 0$. Consider any sequence of within-group-i problems indexed by $n \in \mathbb{N}$, differing only in the specification of Pareto weights λ_i^n. Assume that, for all n, $\lambda_i^n(r)$ is non-increasing in r, and $\bar{\lambda}_i^n \leq \alpha$. If either*

- *for all r and $\epsilon > 0$, $|\lambda_i^n(r) - \bar{\lambda}_i^n| < \epsilon$ for large enough n, or*
- *for all $r > \underline{r}$ and $\epsilon > 0$, $\lambda_i^n(r) < \epsilon$ for large enough n,*

then any convergent sequence of optimal allocations converges point-wise to effectively assortative matching.

Proposition 8 predicts that a market mechanism is optimally used in the two boundary cases when the objective function approaches *(i)* allocative efficiency, and *(ii)* the Rawlsian objective (all weight on the agent with lowest welfare). Combined with Proposition 3, this means that a non-market mechanism is used when there is significant inequality in Pareto weights, but not when the inequality takes the extreme form of all weight attached to the "poorest" agent. In this sense, the use of non-market allocations is non-monotone in the degree of inequality between agents.

Intuitively, there is a trade-off between the revenue and efficiency motives (both of which are forces behind assortative matching) and the WTP-revealed inequality (a force behind random matching). When willingness to pay is relatively uninformative about agents' needs, the revenue and efficiency motives dominate. As willingness to pay becomes increasingly informative, the designer may opt for a partially random allocation to identify those most in need through the mechanism. Eventually, when only a small fraction of agents receive an increasingly high welfare weight, the use of random allocation becomes negligible because the revenue motive and efficiency motives dominate for all remaining agents.

C.3 Who gets the highest-quality object?

Finally, we generalize the observation from Sections 4 and 5 about the allocation of the highest-quality object across groups that feature assortative matching at the top. We show that as long as there is assortative matching at the top in each group (as predicted, for example, by Proposition 7), the allocation of the highest-quality objects depends only on the ranking of the upper limits of willingness to pay.

Proposition 9. *Suppose that, for each group i, it is optimal to have assortative matching at the top. Relabel the groups so that lower $i = 1, \dots, |I|$ corresponds to lower $\alpha\bar{r}_i$. Then, there exists an optimal mechanism in which $\mathrm{supp}(F_i^\star) = [\underline{q}_i,\, \bar{q}_i] \cap \mathrm{supp}(F)$ for some $\{\underline{q}_i,\, \bar{q}_i\}_{i=1}^{|I|}$ with $\bar{q}_1 \le \bar{q}_2 \le \dots \le \bar{q}_{|I|} = \max \mathrm{supp}(F)$.*

Proposition 5 establishes the same conclusion as Proposition 9 under the assumption that matching is (effectively) assortative in each group i. Proposition 9 shows that a much weaker condition is needed—namely, that matching is assortative *at the top* in each group i.

D Proofs the results in Appendix C

Proof of Proposition 7. Suppose that there is random matching at the top, that is, $\Psi_i(x)$ is affine for $x \in [\underline{x}, 1]$ for some $\underline{x}$, and take $\underline{x}$ so that this is the maximal random-allocation region. There are two cases to consider. If $\underline{x} = 0$, then we have to rule out that $\Psi_i(0) \geq \Psi_i(1) - \Psi_i'(1)$; if $\underline{x} > 0$, then it suffices to rule out that $\Psi_i'(\underline{x}) \geq \Psi_i'(1)$ (if $\underline{x} > 0$ is the beginning of the maximal interval of random matching, then the slope of $\mathrm{cd}(\Psi_i)$ at $\underline{x}$ must be equal to the slope of Ψ_i, and that slope must be larger than the slope of Ψ_i at 1 since $\mathrm{cd}(\Psi_i) \geq \Psi_i$ with an equality at 1). Because G_i has bounded support, its inverse hazard rate is 0 at the upper bound; thus, $\Psi_i'(1) = -\alpha\bar{r}_i$. Thus, the first possibility can be ruled out if

$$\alpha\bar{r}_i > (\alpha - \bar{\lambda}_i)\underline{r}_i + \int_{\underline{r}_i}^{\bar{r}_i} \tau\lambda_i(\tau)dG_i(\tau) \iff \alpha(\bar{r}_i - \underline{r}_i) > -\bar{\lambda}_i\underline{r}_i + \int_{\underline{r}_i}^{\bar{r}_i} \tau\lambda_i(\tau)dG_i(\tau). \quad \text{(D.1)}$$

But we have

$$\int_{\underline{r}_i}^{\bar{r}_i} \tau\lambda_i(\tau)dG_i(\tau) - \bar{\lambda}_i\underline{r}_i < \bar{r}_i \int_{\underline{r}_i}^{\bar{r}_i} \lambda_i(\tau)dG_i(\tau) - \bar{\lambda}_i\underline{r}_i = \bar{\lambda}_i(\bar{r}_i - \underline{r}_i);$$

thus, (D.1) can be ruled out by the hypothesis that $\alpha \geq \bar{\lambda}_i$. The second possibility can be ruled out if, for all r,

$$\alpha r - (\alpha - \Lambda_i(r))h_i(r) < \alpha\bar{r},$$

which clearly holds as long as $\alpha \geq \Lambda_i(r)$ which is true by the fact that $\lambda_i(r)$ is non-increasing and $\bar{\lambda}_i \leq \alpha$. □

Proof of Proposition 8. Throughout the proof, we drop the dependence on n from the notation.

Consider the first case first. We know that $J_i'(r) = 1 - h_i'(r) \geq \underline{J}_i$, so $h_i'(r) \leq 1 - \underline{J}_i$. By Proposition 3, since we know that $\alpha \geq \bar{\lambda}_i$ for all n, to prove that assortative matching is optimal, it is enough to prove that the second derivative of Ψ_i is non-positive. The sign of the second derivative of Ψ_i is opposite to the sign of the following expression:

$$\alpha + \Lambda_i(r) - \lambda_i(r) + \underbrace{(\Lambda_i(r) - \alpha)}_{\leq 0} h_i'(r) \geq \alpha - 2\epsilon + (\Lambda_i(r) - \alpha)(1 - \underline{J}_i) = -2\epsilon + \Lambda_i(r) + (\alpha - \Lambda_i(r))\underline{J}_i \geq 0,$$

for all $2\epsilon < \alpha \min\{1, \underline{J}_i\} \leq \Lambda_i(r) + (\alpha - \Lambda_i(r))\underline{J}_i$. Thus, by taking ϵ satisfying that last condition, and n large enough, we conclude that the solution to the problem is assortative matching (this conclusion is stronger than that of Proposition 3 in that assortative matching is exactly optimal for n large enough).

Now consider the second case. By the same calculation as before, for any $x > 0$, there exists a large enough n so that Ψ_i is strictly concave on $[x, 1]$. This means that if there is a random-allocation region that does not vanish in the limit as $n \to \infty$, then it must take the form of $[x_0, x_1]$ with $x_0 \to 0$ and $x_1 > \underline{x} > 0$ as $n \to \infty$, where $\underline{x}$ does not depend on n. (Intuitively, while Ψ_i is concave on $[x, 1]$ for any x if n is large enough, it could be the case that the concave closure of Ψ_i is supported at a point x_0 that converges to 0, and some other point—bounded away from 0—that lies in the region where Ψ_i is concave.) We show that this leads to a contradiction.

First, it is convenient to decompose

$$\Psi_i(x) = \underbrace{\int_x^1 J_i(G_i^{-1}(x))dx}_{\Psi_i^R} + \underbrace{\int_x^1 \Lambda_i(G_i^{-1}(x))h_i(G_i^{-1}(x))dx}_{\Psi_i^W}.$$

Our strategy is to show that, for large enough n, Ψ_i^R is strictly concave (with a second derivative bounded away from 0), while Ψ_i^W and its derivative are arbitrarily small, and thus they cannot change the shape of Ψ_i in the limit.

Note that there exists $m > 0$ such that

$$(\Psi_i^R)''(x) = -\frac{J_i'(G_i^{-1}(x))}{g_i(G_i^{-1}(x))} < -m < 0,$$

by assumption that the derivative of J_i is lower bounded, and that the density g_i is continuous on its support (so it has an upper bound). Also note that for any $\epsilon > 0$, and x such that $G_i^{-1}(x) < \epsilon$, for large enough n, we have

$$\Psi_i^W(x) \leq \Psi_i^W(0) = \int_0^1 \left(\int_{G_i^{-1}(x)}^{\bar{r}_i} \lambda_i(\tau)dG_i(\tau) \right) dr \leq \bar{\lambda}_i G_i^{-1}(x) + \epsilon(\bar{r}_i - \underline{r}_i) \leq \epsilon \cdot M,$$

where the second-to-last inequality uses the assumption that Pareto weights are below ϵ for large enough n, and M is some constant. By the same assumption, for any $\epsilon > 0$, $x > 0$, and large enough n, we have

$$|(\Psi_i^W)'(y)| \leq |-\Lambda_i(G_i^{-1}(y))h_i(G_i^{-1}(y))| \leq \epsilon,$$

for any $y \geq x$.

We are ready to obtain the desired contradiction. A necessary condition for $\text{cd}(\Psi_i)$ to be affine on $[x_0, x_1]$ is that

$$\Psi_i(x_1) - \Psi_i'(x_1)(x_1 - x_0) - \Psi_i(x_0) \leq 0. \tag{D.2}$$

Note, however, that

$$\Psi_i^R(x_1) - (\Psi_i^R)'(x_1)(x_1 - x_0) - \Psi_i^R(x_0) = -\int_{x_0}^{x_1} y(\Psi_i^R)''(y)dy \geq \frac{1}{2}m(x_1 - x_0)^2. \tag{D.3}$$

Since $x_1 \geq \underline{x} > 0$ for all n, and $\underline{x}$ does not depend on n, this expression is bounded away from 0. Yet, by the inequalities established above on Ψ_i^W and $(\Psi_i^W)'$, we have

$$\left|(\Psi_i(x_1) - \Psi_i'(x_1)(x_1 - x_0) - \Psi_i(x_0)) - \left(\Psi_i^R(x_1) - (\Psi_i^R)'(x_1)(x_1 - x_0) - \Psi_i^R(x_0)\right)\right| \leq \epsilon \cdot \tilde{M},$$

for some constant $\tilde{M}$. For large enough n, we can take ϵ small enough so that $\epsilon \cdot \tilde{M} < \frac{1}{2}m(x_1 - x_0)^2$ which is inconsistent with (D.2) and (D.3), a contradiction. □

Proof of Proposition 9. The proof follows from the proof of Proposition 5 by noting that the relevant part of the proof only uses the assumption that matching is assortative at the top in group i; indeed, assortative matching at the top suffices to conclude that $\mathrm{cd}(\Psi_i)(x) = \Psi_i(x)$ for large enough x. □