Random redistribution and discrimination*

Stéphane Gauthier[†]
PSE, University of Paris 1 and Institute for Fiscal Studies
Guy Laroque[‡]
Sciences-Po, UCL and Institute for Fiscal Studies

December 15, 2024

Abstract

We analyze random redistribution among risk-averse agents with quasilinear utility. For randomness to yield social benefits, the optimal deterministic redistribution must involve bunching. Differential treatment that is unattainable in the deterministic case due to bunching, becomes feasible in the stochastic case, allowing for discrimination. Randomness in redistribution implies a shift from the downward to an upward incentive structure that makes redistribution goals aligned with incentives. We exhibit an example of a Rawlsian government that should design lotteries directed toward the most risk-averse population rather than the least risk-averse.

JEL classification numbers: C61, D82, D86, H21.

Keywords: redistribution, asymmetric information, bunching, random taxation.

*We are grateful to Francis Bloch, Ben Brooks, Bernard Caillaud, Killian Dengreville, Piotr Dworczak, Marc Fleurbaey, Deniz Kattwinkel, David Martimort, Nicolas Werquin, Floris Zoutman and participants to the Public Finance and Mechanism Design session of the 2023 conference of the Society for the Advancement of Economic Theory in Paris for feedback on our work. The usual disclaimers apply.

[†]stephane.gauthier@univ-paris1.fr; PSE, 48 bd Jourdan, 75014 Paris, France.

[‡]g.laroque@ucl.ac.uk; UCL, Drayton House, 30 Gordon St, London WC1H 0AX, UK.

1 Introduction

In a first-best environment the government observes the innate traits of individuals, which can serve as a basis for the design of redistribution policies. Financial assistance, such as food stamps, housing subsidies or cash benefits can be provided to the low-skilled poor, disabled persons or those with qualifying medical conditions preventing gainful activity. The spectrum of policies available to the government often is more limited when some relevant characteristics of individuals are not publicly known. Authorities then have to spend extra resources to perform a suitable targeting of income support to those in need, implying a balance between efficiency and equity concerns. Mirrlees (1971) formalized the additional costs due to privately known traits by accounting for incentive constraints that ensure self-selection by those in need while deterring unwarranted claims.

A general lesson from this literature is that the burden of asymmetric information typically falls on those in need, rather than those in more favorable situations. Indeed a typical response to asymmetric information is to reduce the amount of assistance, as the lower aid enables the government to target those in need as beneficiaries of assistance while others are discouraged.

Several ideas have been explored to expand redistribution through improved targeting of assistance. Most of them involve some form of ordeal mechanism subjecting vulnerable populations to challenging tasks or stressful conditions, in the spirit of Nichols and Zeckhauser (1982). The government can for instance rely on time-consuming shameful queuing to distribute essential goods to low-income households. It may also implement unnecessarily complex and lengthy application processes to prove eligibility for benefits, or impose additional conditions after admission to continue receiving benefits, such as requiring beneficiaries to regularly send their children to school or undergo health check-ups. The social usefulness of these complementary schemes thus depends on balancing the direct cost borne by the targeted population and the benefits from relaxed incentive constraints associated with the discouragement of undue claims.

In these examples, the ordeal is usually taken as deterministic, i.e., pain comes for sure. However it is known from, e.g., Lang (2017), Ederer, Holden, and Meyer (2018) or Lang (2023), randomization limits gaming social rules by blurring incentives. Vague standards or legal uncertainty may deter firms to undertake strategies detrimental to the society; hospitals, for instance, may be discouraged from selecting healthy but less costly to treat patients if they are not aware of the exact amount of compensation they will receive. In this paper, we are interested into such a form of ordeal, which is to impose random noisy transfers to risk averse recipients.

In public finance, Weiss (1976), Stiglitz (1981), Stiglitz (1982) or Brito, Hamilton, Slutsky, and Stiglitz (1995) have shown that deterministic redistribution sometimes is socially dominated.[1] Income lotteries can be due to random noise in taxes, because of e.g., administration errors, tax evasion coupled with non-comprehensive auditing, or uncertainty about the actual fiscal regime amid frequent tax reforms. In Wijkander (1988) or Dworczak, Kominers, and Akbarpour (2021), lotteries occur in the presence of quotas and rationing in the allocation of certain goods or services, as limits on market transactions lead some agents to engage in trade with strictly interior probabilities (with some risk of being rationed). In the same vein, random labor and before-tax income variations can be induced by the minimum wage and the risk of unemployment; they can also be due to randomness in occupations for students who apply for medical training in the Netherlands and are accepted by draw. But rather than before-tax income, the most explicit randomizations perhaps concern situations where the government instead relies on random after-tax incomes and allocation of consumption goods for redistributive purposes. Tobin (1971) argues that income tests for housing subsidies make support 'available only for an accidentally or arbitrarily selected few' while randomness from housing programs involving rent regulation improves selectivity in access for low-income populations in Weitzman (1977). Similar situations are common when an agency has to allocate scarce resources, e.g., when the distribution of public piped water and energy resources in developing countries involves shortages through rationing, potentially with random interruptions of supply services.

The literature presents a straightforward argument for random allocation to outperform the best deterministic alternative. Following Hellwig (2007), suppose that the government would like to redistribute income to low-skilled in a population of risk averse workers. Redistribution is potentially limited if the government observes neither skill nor the exact amount of labor, as high-skilled might reduce labor effort to enjoy higher transfers. Randomness in the after-tax income of low-skilled is detrimental to their welfare, but this also expands the scope of possible redistribution by discouraging risk-

[1] See also Pavlov (2011), Gauthier and Laroque (2014), Pycia and Unver (2015) or Gauthier and Laroque (2017) for related approaches.

averse high-skilled from relaxing labor effort. We thus expect a deterministic optimum if high-skilled do not suffer much from income noise, which leaves us with little hope for randomized taxes to improve the welfare of the poor, who are usually found more risk averse than the rich.

Although the above argument sounds intuitive, it does not fully accord with a puzzling parametric example in Strausz (2006). In this example, a regulatory authority faces two types of firms with different production technologies. If it offers the first-best option while not observing technologies, then incentive compatibility fails, as low production cost (efficient) firms would mimic high cost (inefficient) firms. Still, the second-best regulatory policy involves a random option designed for the mimicking (efficient) firms, rather than the mimicked (inefficient) firms.

This example suggests that the argument identified thus far for the role of randomization does not completely account for the impact of random noise in the presence of asymmetric information. Our paper provides a related example that combines bunching in the deterministic optimum and a reversal of incentives, shifting from the familiar downward pattern under deterministic redistribution to an upward pattern when random transfers are introduced. Eventually offering random options to the most risk-averse agents may be optimal.

We consider a Rawlsian government that only values the agents with the lowest utility. If incentive compatibility issues could be dealt with the first-order approach that neglects the possibility of bunching, the best deterministic redistribution policy would involve socially disfavored (rich) types envying the option designed for those more socially favored (poor). However, the strong redistribution motive underlying Rawlsian criteria leads to bunching where many different agents, including those socially favored, have to enjoy the same transfers. Incentives then prevent the authority from discriminating recipients. In extreme cases, deterministic redistribution might not even be possible.

The uniform treatment of the agents in the deterministic optimum with bunching is akin to some form of uniform rationing which blurs the pattern of incentives. Every agent may then be seen as both willing to mimic any other agent and envied by the others. We show that randomization then allows the government to exploit new dimensions of individual heterogeneity, e.g., risk aversion, in a way that reverses the pattern of incentives compared to the deterministic case. Randomness makes the agents that the government wants to favor now envying the treatment of those with lower social importance, a feature reminiscent to countervailing incentives. A similar reversal occurs in Strausz (2006), but not in Hellwig (2007) where the same structure of incentives prevails both in the deterministic and stochastic cases. Actually the disappointing outcome for pro-poor policies in Hellwig (2007) relies on the fact that high-skilled types continue to envy low skilled once random noise is introduced into the tax system. Our example shows that this is not a general property.

The gain from making incentives aligned however comes with a cost, as risk averse agents have to face randomness. In a particular parametrization of our model the gain from aligned incentives overcomes the cost, and redistribution should involve a random allocation for the socially favored (lowest utility) agents, though they display the highest risk aversion. This may provide incentive-based justifications for randomness in social assistance, or water or energy distribution.

The paper proceeds as follows. Our setup with random redistribution is described in Section 2. Section 3 characterizes the role played by bunching in the deterministic optimum. Section 4 shows that, in the presence of small random noise in taxes, the direction of incentives can be reversed compared to the deterministic case. Section 5 provides a condition for socially useful randomization in the polar case where bunching prevents any deterministic redistribution. Some properties of random redistribution in this case are discussed in Section 6. The analysis is generalized in Section 7 to partial bunching, and Section 8 presents parametric examples where optimal redistribution involves randomness. Finally, Section 9 concludes.

2 General framework

A government designs a redistribution policy between a continuum of agents in a population of total unit size. Heterogeneity across agents is characterized by θ, a real parameter taking values in $\Theta = \left[\theta^{\inf}, \theta^{\sup}\right]$, which is referred to as the type of the agent. It has cumulative distribution function $F : \Theta \rightarrow [0, 1]$ associated with positive density $f : \Theta \rightarrow \mathbb{R}_{++}$.

The preferences of a type θ agent are represented by the quasilinear utility function

$$u(c, \theta) - y, \tag{1}$$

where c is thought of as a quantity purchased for a payment of y. For instance, c may correspond to the consumption of water or electricity distributed via the supply system, while y represents the usage fee. These variables could also represent benefits and contributions to a regulatory agency.

The function u is increasing and differentiable in c and θ. It is strictly concave in c, implying that every agent is risk-averse with respect to c. The quasilinear formulation in (1) is commonly employed in the presence of asymmetric information because it facilitates addressing incentive issues. However, it implies that the first derivative $u'_c(c, \theta)$ with respect to c, which also represents the marginal rate of substitution between consumption c and payment y, is independent of y, a property that may lack realism.

Formulation (1), where θ enters as an argument of the function u, suggests to interpret type as reflecting some general consumption preferences or tastes. A more concrete interpretation of θ could posit that it represents a private endowment, which may serve as a substitute for purchased consumption. For instance, in the analysis of Cape Town water crisis carried out by Abajian et al. (2024), θ would capture grey water, stored rainwater, or some capacity to extract groundwater, which serve as substitutes for piped water delivered in quantity c. In the same vein, this parameter could reflect additional sources of housing capacity, such as family transfers, that supplement public housing benefits, or various types of solar panels substitutes for publicly provided electricity. In this context, households pay an amount y for consuming a quantity c of piped water or publicly provided electricity, given their endowment θ in these commodities. This interpretation is consistent with the parametric example in Section 8, where $u(c, \theta)$ is specified as $\ln(c+\theta)$, with $c+\theta$ representing the total consumption of water or electricity, namely the sum of publicly provided consumption and its substitutes.

In these example, agents with higher types (i.e., those with greater endowments of private alternatives) likely place a lower value on an incremental amount of the good they buy. Formally, this is captured by the condition that the derivative $u'_c(c, \theta)$ is decreasing in θ for all (c, θ).

Note that this form of the Spence-Mirrlees (single-crossing) condition departs from the main strand of the literature in industrial organization or public finance, where $u'_c(c, \theta)$ is increasing in θ. For instance, in Mussa and Rosen (1978), $u(c, \theta) = \theta c$, where θ represents the valuation of the good by the consumer. In Lollivier and Rochet (2003), an agent with labor productivity θ must exert a labor effort of y/θ to earn a before-tax income y. The agent's utility is $v(c) - y/\theta$, or equivalently, after multiplying by θ, $\theta v(c) - y$. Thus, $u(c, \theta) = \theta v(c)$, and the cross-derivative $u''_{c\theta}(c, \theta) = v'(c) > 0$, implying that individuals who work more place a higher value on additional after-tax income.

The government designs a redistribution policy between the agents. The policy is defined by a menu $(\tilde{c}(\theta), \tilde{y}(\theta))$ of consumption and payment lotteries. Such pairs are referred to as contracts. The menu of contracts is feasible if aggregate consumption falls below aggregate payment,

$$\int_{\Theta} \mathbb{E}[\tilde{c}(\theta) - \tilde{y}(\theta)] \, \mathrm{d}F(\theta) \leq 0. \tag{2}$$

The government is assumed to know the distribution of types, but not to observe the value of θ for every agent, which remains private information to the agent. Indeed Abajian et al. (2024) suggest that substitutes for public piped water are difficult to observe accurately in Cape Town. Therefore the government must also ensure that agents choose the contract designed for them. This is satisfied if the incentive constraints

$$\mathbb{E}[u(\tilde{c}(\theta), \theta) - \tilde{y}(\theta)] \geq \mathbb{E}[u(\tilde{c}(\tau), \theta) - \tilde{y}(\tau)] \tag{3}$$

hold for all (θ, τ) in $\Theta \times \Theta$.

An optimal redistribution policy is a menu $(\tilde{c}(\theta), \tilde{y}(\theta))$ that maximizes the social welfare objective of the government subject to the feasibility constraint (2) and the incentive constraints (3).

We expect wealthier consumers to have better access to private substitutes. A redistributive government may accordingly prioritize individuals with low access, more likely less well-off types. Here, we consider the polar case of a Rawlsian government that only values agents with the lowest utility. In all the paper, these agents appear to have the least substitute type $\theta^{\inf}$. Let $V(\theta) = \mathbb{E}[u(\tilde{c}(\theta), \theta) - \tilde{y}(\theta)]$ denote the expected indirect utility of type θ when she chooses the lottery designed for her. The Ralwsian social objective is

$$V(\theta^{\inf}). \tag{4}$$

A deterministic policy only consists of degenerate lotteries $(c(\theta), y(\theta))$ where every type θ pays $y(\theta)$ with certainty and consumes $c(\theta)$ with certainty (the absence of a tilde mark applies to deterministic options). We are interested into circumstances where some agents face non-degenerate lotteries in the optimal redistribution policy. Given the quasilinear form of agents' utility, there is no role for random payments, as replacing lottery $\tilde{y}(\theta)$ with the sure outcome $y(\theta) = \mathbb{E}[\tilde{y}(\theta)]$ affects neither the constraints nor the social objective. We shall therefore consider $y(\theta)$ deterministic for all types and focus on random consumption.

3 Randomness and uniform rationing

This section shows that a menu of consumption lotteries $(\tilde{c}(\theta))$ can be socially useful only if incentive considerations restrain the government from relying on deterministic discrimination, i.e., the same treatment has to be applied to different types of agents.

We refer to Laffont and Martimort's (2002) optimal 'relaxed' redistribution to formalize the idea of restrained discrimination. The relaxed policy is a menu $(\tilde{c}(\theta), y(\theta))$ maximizing the social objective (4) subject to the feasibility constraint (2) and the necessary first-order conditions for a local truthful report in (3),

$$V'(\theta) = \mathbb{E}\left[u'_\theta(\tilde{c}(\theta), \theta)\right] \tag{5}$$

for all θ.

It coincides with the optimal redistribution policy if (5) is also sufficient to meet the incentive constraints (3). Under our Spence-Mirrlees assumption, this is met if the relaxed policy entails consumption that does not increase with θ. Otherwise, if relaxed consumption increases for some values of θ, optimal redistribution involves bunching with different types receiving the same amount of consumption goods.

The decreasing pattern of relaxed consumption aligns with the social objective by providing more to lower types; full equality would actually occur when $c(\theta) + \theta$ is constant. This pattern echoes the observation that wealthy residents in Cape Town (presumably with better access to private underground water) have consumed less piped water $c(\theta)$ than lower-income households since the 2017 water crisis, making piped water inferior (though we expect total water consumption $c(\theta) + \theta$ to remain normal). However the payment $y(\theta)$ is designed to overcompensate the gain from alignment of consumption. Indeed utility must increase with θ by (5), while the government seeks to favor the lowest type, which highlights a conflict between incentives and redistribution goals.

How the utility of the lowest type relates to the distribution of consumption obtains by summing up (5) over types, to get

$$V(\theta) = V(\theta^{\inf}) + \int_{\Theta} \mathbb{E}\left[u'_{\theta}(\tilde{c}(z), z)\right] \, \mathrm{d}z.$$

Replacing $y(\theta)$ with $\mathbb{E}[u(\tilde{c}(\theta), \theta)] - V(\theta)$ into the feasibility constraint (2) then gives $V(\theta^{\inf})$. After using the integration by parts formula, it writes as

$$V(\theta^{\inf}) = \int_{\Theta} \mathbb{E}[W(\tilde{c}(\theta), \theta)] \, \mathrm{d}F(\theta) \tag{6}$$

where

$$W(c, \theta) = u(c, \theta) - c - m(\theta) u'_{\theta}(c, \theta)$$

represents the virtual contribution of type θ to social welfare when she consumes c with certainty, and $m(\theta) = [1 - F(\theta)] / f(\theta)$ is the Mills ratio.

In the optimal deterministic relaxed redistribution policy, every type θ consumes

$$c^*(\theta) = \arg\max_c W(c, \theta).$$

This provides us with a benchmark for desired discrimination. It can be implemented if $c^*(\theta)$ is non-increasing in θ. The government can then discriminate agents by designing a profile of transfers that make public consumption lower for higher θ substitute types, those with lower social importance, who put less value on additional consumption.

Since then $W\left(c^*(\theta), \theta\right) \geq W(c, \theta)$ for all c and θ, we have $W\left(c^*(\theta), \theta\right) \geq \mathbb{E}[W(\tilde{c}, \theta)]$ for all $\tilde{c}$ and θ. In particular, the inequality holds true if for all θ we set $\tilde{c}$ equal to the lottery $\tilde{c}(\theta)$ that maximizes (6) subject to (3). This yields:

Lemma 1. *A random redistribution policy is socially useless if the optimal deterministic redistribution policy coincides with the optimal deterministic relaxed policy, i.e., the optimal deterministic redistribution policy involves no bunching.*

Proof. The argument given above leads to

$$\int_{\Theta} \mathbb{E}[W(\tilde{c}(\theta), \theta)] \, \mathrm{d}F(\theta) \leq \int_{\Theta} W\left(c^*(\theta), \theta\right) \, \mathrm{d}F(\theta)$$

for every menu $(\tilde{c}(\theta))$. The right-hand side of this inequality gives social welfare in the optimal deterministic redistribution policy in the absence of bunching (this is welfare in an optimal relaxed policy). The left-hand side is an upper bound for social welfare in the optimal redistribution policy in the presence of random noise. This upper bound is achieved if incentive compatibility can be addressed using the first-order approach. This concludes the proof. □

Socially useful randomness in redistribution can be achieved only if the incentive constraints associated with the optimal deterministic policy lead to bunching, where different types of agents face the same contract.

In the context of water distribution, the authority would be compelled to allocate similar amounts of pipe water to recipients with different, but privately known, outside access capacities θ. This can be interpreted as some uniform rationing. In a market where all agents would face the same price, those with greater outside access capacities (θ is high) would purchase less pipe water than others (θ is small), as they value less pipe water from the market (the marginal utility $u'_c(c, \theta)$ decreases with θ). They may accordingly be viewed as being forced to have greater consumption, while the other agents are rationed. This interpretation suggests that randomness with say, random interruptions in water or energy distribution, possibly targeted on specific segments of the population, could serve as a means to approach desired discrimination. This idea is explored in the next section.

4 A reversal of incentives

If type θ faces the deterministic option $(c(\theta), y(\theta))$, the incentive constraints (3) simplify to

$$\begin{aligned} V(\theta) = u(c(\theta), \theta) - y(\theta) &\geq u(c(\tau), \theta) - y(\tau) \\ &= V(\tau) + u(c(\tau), \theta) - u(c(\tau), \tau) \end{aligned}$$

for all θ and τ. The authority has to give the informational rent $u(c(\tau), \theta) - u(c(\tau), \tau)$ to type θ to ensure that they do not mimic type τ. Since u is increasing with type (better access to substitutes allows agents to derive greater utility from a given amount of the publicly provided good), this rent is positive for $\theta > \tau$. The familiar downward pattern of incentives prevails in the deterministic case, with the authority discouraging high types from imitating lower types.

We are going to show that randomness leads to informational rents given to low types, rather than high types. Our argument is more straightforward in cases where incentives imply that all agents are subject to bunching, where every agent consumes c^* and pays y^*, yielding indirect utility $u(c^*, \theta) - y^*$ to type θ (a more general analysis is given in Section 7). The optimal consumption c^* is such that $u(c, \theta^{\inf}) - c$ is maximized for $c = c^*$. If interior, it satisfies

$$u'_c\left(c^*, \theta^{\inf}\right) = 1. \tag{7}$$

This can be implemented in a decentralized setup through the use of a two bracket schedule, where a low (resp. high) unit price applies to consumption below (resp., above) c^*. Since feasibility requires $y^* = c^*$, the Rawlsian government is prevented from any redistribution in the optimal deterministic policy.

Consider now small randomizations in consumption where type θ faces the lottery $\tilde{c}(\theta) = c^* + \tilde{\varepsilon}(\theta)$. We restrict to small noise, with realizations of the random variable $\tilde{\varepsilon}(\theta)$ close to 0. We also require that greater noise comes with greater transfers by setting $\mathbb{E}[\tilde{\varepsilon}(\theta)] = \text{var}[\tilde{\varepsilon}(\theta)] = \lambda v(\theta)$, with λ a positive real number close to 0 and $v(\theta) \geq 0$ a (rescaled) variance bounded from above.[2]

For such lotteries, the (second-order Taylor expansion of the) expected utility of type θ when she chooses the lottery $\tilde{c}(\tau)$ designed for type τ writes

$$\mathbb{E}\left[u\left(c^* + \tilde{\varepsilon}(\tau), \theta\right)\right] \simeq u\left(c^*, \theta\right) + \lambda u'_c\left(c^*, \theta\right)\left(1 - \frac{A\left(c^*, \theta\right)}{2}\right) v(\tau)$$

where

$$A(c, \theta) = -\frac{u''_{cc}(c, \theta)}{u'_c(c, \theta)} > 0$$

is the coefficient of absolute risk aversion of type θ. This shows that randomness allows the government to exploit heterogeneity in both marginal utility $u'_c\left(c^*, \theta\right)$ and risk aversion. The sub-utility $u\left(c^*, \theta\right)$ derived from consumption in the deterministic case is now modified by $\lambda S\left(c^*, \theta\right) v(\tau)$ for type θ,

[2]Our methodology can be applied to any menu where $\mathbb{E}[\tilde{\varepsilon}(\theta)]$ and $\text{var}[\tilde{\varepsilon}(\theta)]$ are linearly related. We did not investigate more general formulations.

with

$$S\left(c^{*},\theta\right)=u_{c}'\left(c^{*},\theta\right)\left(1-\frac{A\left(c^{*},\theta\right)}{2}\right).$$

The quantity $S(c^{*},\theta)$ approximates the extra utility derived by type θ following a unit marginal increase from c^{*} in her consumption of the publicly provided good. It appears as an extra valuation of the consumption lottery under scrutiny. In view of the Spence-Mirrlees assumption and (7), we have

$$S\left(c^{*},\theta\right)\leq u_{c}'(c^{*},\theta^{\inf})+\frac{1}{2}u_{cc}''\left(c^{*},\theta\right)<1.$$

The shape of $S(c,\theta)$ appears otherwise difficult to characterize. However, the interpretation in terms of extra utility suggests that the relevant economic case should involve $S(c,\theta)$ taking positive values. Furthermore, based on the Spence-Mirrlees assumption, we expect it to be decreasing in θ, $S_{\theta}'(c^{*},\theta)<0$. This is indeed the case for the specifications presented below.

Example 1. CRRA preferences. Suppose that $u(c,\theta)=c^{1-\theta}/(1-\theta)$ for $\theta\neq 1$. From (7), the optimal consumption is $c^{*}=1$. At this point,

$$S(c^{*},\theta)=1-\frac{\theta}{2}\text{ and }S_{\theta}'(c^{*},\theta)=-\frac{1}{2}<0.$$

Referring to the upper bound of 2 for the coefficient θ of relative risk aversion obtained by Chetty (2006), we have $S(c^{*},\theta)>0$.

Example 2. CARA preferences. Let $u(c,\theta)=-\exp(-\theta\left(c-\bar{c}\right))/\theta$ for some consumption $\bar{c}\geq 0$. Then (7) gives $c^{*}=\bar{c}$, and both $S(c^{*},\theta)$ and $S_{\theta}'(c^{*},\theta)$ are as in the CRRA case. The estimates of the coefficient of absolute risk aversion θ in Cohen and Einav (2007) are of an order of magnitude of 10^{-2} at most. For such values, $S(c^{*},\theta)>0$.

Example 3. Logarithmic Preferences. Let $u(c,\theta)=\ln\left(c+\theta\right)$. Then (7) gives $c^{*}=1-\theta^{\inf}$. For $\theta^{\inf}$ close enough to 0,

$$S(c^{*},\theta)=\frac{1}{c^{*}+\theta}\left(1-\frac{1}{2}\frac{1}{c^{*}+\theta}\right)>0,$$

and

$$S_{\theta}'(c^{*},\theta)=-\frac{1}{\left(c^{*}+\theta\right)^{2}}\left(1-\frac{1}{c^{*}+\theta}\right)<0.$$

From now onward, based on the insights from these examples, we assume:

Assumption A1. *The extra valuation $S(c^*, \theta)$ of consumption is positive and decreases with θ.*

In the presence of random noise on consumption, the incentive constraints (3) are: for all τ and θ,

$$u(c^*, \theta) + \lambda S(c^*, \theta) v(\theta) - y(\theta) \geq u(c^*, \theta) + \lambda S(c^*, \theta) v(\tau) - y(\tau),$$

or equivalently,

$$\begin{aligned} U(\theta) = \lambda S(c^*, \theta) v(\theta) - y(\theta) &\geq \lambda S(c^*, \theta) v(\tau) - y(\tau) \\ &= U(\tau) + \lambda (S(c^*, \theta) - S(c^*, \tau)) v(\tau). \end{aligned}$$

Incentives are now driven by the component $U(\theta)$ of the overall utility $V(\theta)$. The monotonicity properties of $U(\theta)$ may differ from that of $V(\theta)$, allowing for a reversal of incentives. Formally, the incentive constraints have a simple textbook structure in $v(\theta)$ and $y(\theta)$, so that one can apply the usual arguments to get:

Lemma 2. *Consider a menu where the government uses small random perturbations $\tilde{\varepsilon}(\theta)$ to the optimal deterministic consumption c^* with mean and variance $\lambda v(\theta)$ for some $\lambda \geq 0$ close to 0. The incentive constraints (3) are satisfied if and only if*

$$U'(\theta) = \lambda S'_\theta(c^*, \theta) v(\theta)$$

and $v(\theta)$ is non-increasing for all θ.

Proof. Using the envelope theorem, a necessary first-order condition for a local truthful report is $U'(\theta) = \lambda S'_\theta(c^*, \theta) v(\theta)$ for all θ. The second-order conditions write $S'_\theta(c^*, \theta) v'(\theta) \geq 0$ for all θ where v is differentiable. Finally,

$$\frac{\partial}{\partial \tau} (S(c^*, \theta) v(\tau) - y(\tau)) = \int_\tau^\theta S'_\theta(c^*, z) v'(\tau) \, \mathrm{d}z$$

has the same sign as $\theta - \tau$ since $S'_\theta(c^*, z) < 0$ for all z. It follows that (3) is satisfied for all τ and θ. This concludes the proof. □

Lemma 2 conveys a reversal of incentives implied by the introduction of random noise in redistribution. By Assumption A1, incentives require

$U(\theta)$ to be non-increasing. Note that the overall utility $V(\theta)$ remains increasing ($V'(\theta) = u'_\theta(c^*, \theta) + \lambda S'_\theta(c^*, \theta)v(\theta) \simeq u'_\theta(c^*, \theta) > 0$ for small λ and bounded $v(\theta)$). The reversal is characterized by a shift in the informational rent $\lambda\left(S(c^*, \theta) - S(c^*, \tau)\right) v(\tau)$ given to type θ to prevent mimicking type τ. Unlike the deterministic case, the rent now turns positive for $\tau > \theta$. This upward, rather than downward, pattern aligns incentives with social preferences.

5 Welfare improving randomization

The existing argument for randomization is based on relaxed incentives when random noise is applied to mimickees, requiring these agents to be less risk averse than mimickers (Hellwig, 2007). The reversal of incentives in Lemma 2 instead occurs when $S'_\theta(c^*, \theta) < 0$, which is consistent with a risk aversion decreasing with θ. Actually, $A(c^*, \theta) = 1/(c^* + \theta)$ in Example 3, type θ^{inf} agents are the most risk averse.

Still, type θ^{inf} also faces the greatest noise. The following result provides a condition under which the social gain from aligning incentives outweighs the loss from random noise.

Proposition 1. Optimal random redistribution. *Consider a menu where the government uses random perturbations $\tilde{\varepsilon}(\theta)$ to the optimal deterministic consumption c^* with mean and variance $\lambda v(\theta)$ for some $\lambda \geq 0$ close to 0. Let $v(\theta)$ be non-negative, non-increasing and bounded from above. The random menu improves upon the deterministic optimum if and only if*

$$\int_\Theta \phi(c^*, \theta)v(\theta)\ \mathrm{d}F(\theta) > 0 \tag{8}$$

where

$$\phi(c, \theta) = S\left(c, \theta\right) - 1 - m(\theta)S'_\theta\left(c, \theta\right).$$

A proof is in Appendix A. A heuristic argument for deriving (8) follows from the methodology developed by Saez (2001) for optimal taxation. The Rawlsian government must maximize the resources that can be transferred to type θ^{inf} agents. Once random noise is introduced, the net contribution of a θ agent is $y(\theta) - [c^* + \lambda v(\theta)]$. Therefore, using the definition of $U(\theta)$, the total collected resources can be written

$$\int_{\Theta} [\lambda S(c^*, \theta) v(\theta) - U(\theta) - c^* - \lambda v(\theta)] \, \mathrm{d}F(\theta) \tag{9}$$

Consider a reform that increases the (rescaled) variance $v(\theta)$ of consumption by a small amount $\mathrm{d}v$ for all types between θ and $\theta + \mathrm{d}\theta, \mathrm{d}\theta$ positive close to 0. These types, who are directly concerned by the reform, are in total number $f(\theta)\mathrm{d}\theta$. The argument distinguishes behavioral and mechanical effects of the reform.

The behavioral effect is the change in total collected resources that abstracts from the adjustments needed to meet incentives, i.e., with $U(\theta)$ temporarily maintained fixed at its initial level. It captures the net social cost of the randomization that transits through the noise bearing on the socially favored agents. The payment made by every type θ directly concerned by the reform increases by $\lambda S(c^*, \theta) \, \mathrm{d}v$. The total resources thus increase by $\lambda S(c^*, \theta) \; f(\theta)\mathrm{d}v \; \mathrm{d}\theta$. However, the government takes advantage of the noise to raise average consumption to every such agents, which costs $\lambda \mathrm{d}v$ per agent. Overall the change in total collected resources is $\lambda [S(c^*, \theta) - 1] \, f(\theta)\mathrm{d}v \; \mathrm{d}\theta$. Given $U(\theta)$, the reform yields a lower amount of resources, which represents a net cost for the society (recall that $S(c^*, \theta) < 1$).

This cost has to be compared to the social gain from getting incentives aligned with redistribution tastes. Such a gain obtains when one accounts for the mechanical response of $U(\theta)$ to the reform. By Lemma $2, U'(\theta)$ changes by $\mathrm{d}U'(\theta) = \lambda S'_\theta(c^*, \theta) \, \mathrm{d}v$ for every type directly concerned by the reform. It follows that the utility changes by $\mathrm{d}U = \mathrm{d}U'(\theta)\mathrm{d}\theta = \lambda S'_\theta(c^*, \theta) \, \mathrm{d}v \; \mathrm{d}\theta$ for every type above $\theta + \mathrm{d}\theta$, implying a change in total collected resources equal to $-(1 - F(\theta))\lambda S'_\theta(c^*, \theta) \, \mathrm{d}v \; \mathrm{d}\theta$. Since $S'_\theta \leq 0$, these are indeed additional resources.

The whole change in resources that can be transferred following the introduction of the noise thus is

$$\lambda [S(c^*, \theta) - 1] \, f(\theta)\mathrm{d}v \; \mathrm{d}\theta - (1 - F(\theta))\lambda S'_\theta(c^*, \theta) \, \mathrm{d}v \; \mathrm{d}\theta$$

or equivalently, $\lambda \phi(c^*, \theta) \, f(\theta)\mathrm{d}v \; \mathrm{d}\theta$. The $\phi(c^*, \theta)$ term in Proposition 1 balances the loss in net resources from agents concerned by the reform (their higher payment does not compensate the cost from the additional consumption they receive) and the greater resources allowed by the reduced informational rents given to high types above θ.

6 Shape of random redistribution

The inequality (8) shows that there is no social improvement from randomness if $S'_\theta(c^*, \theta) > 0$, i.e., in the absence of reversal of incentives (recall that $S(c^*, \theta) < 1$).

If the additional resources $\phi(c^*, \theta)$ collected from type θ decreases with θ, then (8) is satisfied if and only if $\phi(c^*, \theta^{\inf}) > 0$. In this case, there exists a threshold type θ^*, $\theta^* > \theta^{\inf}$, such that every type $\theta \leq \theta^*$ should face a random contract.

If $\phi(c^*, \theta)$ instead increases with θ, then the government would like to make consumption random for high types specifically, but this would violate the monotonicity requirement for incentive compatibility that $v(\theta)$ must be non-increasing. Redistribution would then have to involve randomness for all agents. This provision cannot be optimal, however, as agents with the highest types should always face certainty.

We now provide a general characterization of the optimal shape of random transfers, including the special case where $\phi(c^*, \theta)$ increases with θ. It follows from the observation that random redistribution is optimal if ond only if the highest value of the sum in (8) in Proposition 1 is positive for a non-increasing profile of rescaled variance. The form of this problem allows us to rely on the methodology introduced by Myerson (1981) and consider the new function

$$H(c^*, q) = \int_0^q \phi(c^*, F^{-1}(z))\,\mathrm{d}z$$

for every quantile $q \in [0, 1]$ of the type distribution. Then, let G be the concave hull of H (this is the smallest concave function satisfying $G(c^*, q) \geq H(c^*, q)$ for all q), and

$$\bar{\phi}(c^*, \theta) = G'_q(c^*, F(\theta))$$

as the so-called priority rule. We have:

Proposition 2. *Consider a menu where the government uses random perturbations $\tilde{\varepsilon}(\theta)$ to the optimal deterministic consumption c^* with mean and variance $\lambda v(\theta)$ for some $\lambda \geq 0$ close to 0. The random menu improves upon the deterministic optimum if and only if*

$$\bar{\phi}(c^*, \theta^{\inf}) > 0.$$

There exists $\theta^* \geq \theta^{\inf}$ such that the highest amount of extra resources implied by randomization obtains by setting $v(\theta) > 0$ and non-increasing for all $\theta < \theta^*$, and $v(\theta) = 0$ for all $\theta \geq \theta^*$.

The proof mirrors Myerson (1981), Section 6 pp. 68-69, or Condorelli (2012), and thus it is omitted. Actually, a non-increasing variance maximizes the sum in (8) if it maximizes this sum with $\phi(c^*, \theta)$ replaced with $\bar{\phi}(c^*, \theta)$. Proposition 2 then follows from the fact that priority $\bar{\phi}(c^*, \theta^{\inf})$ is non-increasing in θ since $G(c^*, q)$ is concave in q.

The two polar cases with $\phi(c^*, \theta)$ monotone in θ discussed above obtain for $H(q)$ either concave or convex. In the concave case, $\phi(c^*, \theta)$ is decreasing in θ, and thus $\bar{\phi}(c^*, \theta^{\inf}) = \phi(c^*, \theta^{\inf})$. In the convex case, $\phi(c^*, \theta)$ is increasing in θ,

$$\bar{\phi}(c^*, \theta) = H(c^*, 1) = \int_{\Theta} \phi(c^*, z)\, \mathrm{d}z$$

for all θ, and consumption should be random if and only if a policy with $v(\theta) = v > 0$ for all θ could yield an extra amount of total collected resources.

Consider now an arbitrary curvature of $H(q)$. By Proposition 2, we can restrict attention to $v(\theta) = v > 0$ if $\theta \leq \theta^*$, and $v(\theta) = 0$ otherwise. For such policies, (8) is met if and only if there exists a threshold type $\theta^* \in \left[\theta^{\inf}, \theta^{\sup}\right]$ such that[3]

$$\frac{1 - S(c^*, \theta^{\inf})}{1 - S(c^*, \theta^*)} < 1 - F(\theta^*). \tag{10}$$

Thus, agents with the highest types should never be exposed to randomness, since the left-hand side of (10) is positive for all types while its right-hand side is 0 at $\theta^* = \theta^{\sup}$.

The writing (10) provides us with a better understanding of the economic conditions where random noise in redistribution can be useful. First, $S(c^*, \theta^{\inf})$ has to be much higher than $S(c^*, \theta)$ for almost all $\theta > \theta^{\inf}$. Second, most agents should have high types close enough to $\theta^{\sup}$, so that $1 - F(\theta)$ in the right-hand side of (10) remains close to 1 for a large part of the population. Overall, Rawlsian redistribution by means of lotteries requires wide dispersion in the valuation of the good $S(c^*, \theta)$ and low variance in the extra consumption capability θ.

Such a combination is reminiscent of circumstances identified by Weitzman (1977) and Spence (1977), where the price system is a better instrument

[3] A detailed derivation of (10) is in Appendix B.

than rationing to allocate resources. This may not come as a surprise. When (10) is met, the socially favored types $\theta^{\inf}$ value the good much more than the rest of the population, providing the government with strong incentives to allocate them a greater amount of goods. However, this cannot be achieved in a deterministic fashion, which involves uniform rationing. Screening then is difficult to make since it has to be based on small differences in extra capabilities θ. Following Lemma 2, the government can instead base screening on valuation $S(c^*, \theta)$ in the presence of random noise. This allows, to some extent, for the replication of the market allocation, with higher, though stochastic, transfers to those actually rationed, who would have consumed more through the market.

7 Non-uniform partial bunching

So far we have considered the polar case of uniform bunching in the deterministic optimum where every type faces the same option (c^*, y^*). In practice, authorities often rely on schemes with more than just two brackets. Then, provided that bunching operates everywhere, the general form of the optimal deterministic schedule consists of a collection of contracts (c_i^*, y_i^*) assigned to every agent with a type ranging from $\bar{\theta}_i$ to $\bar{\theta}_{i+1}$ $\left(\bar{\theta}_i < \bar{\theta}_{i+1}\right)$. Proposition 3 below extends Proposition 1 to such schedules.

Proposition 3. Non-uniform deterministic bunching. *Suppose that the optimal deterministic redistribution policy consists of* n *different brackets, with every type of agents in* $\left[\bar{\theta}_i^*, \bar{\theta}_{i+1}^*\right)$ *consuming* c_i^* *against payment* y_i^**. There exists a random policy that improves upon the deterministic optimum if*

$$\sum_{i=1}^{n} \int_{\bar{\theta}_i^*}^{\bar{\theta}_{i+1}^*} \phi\left(c_i^*, \theta\right) v(\theta) \, \mathrm{d}F(\theta) > 0$$

for some non-increasing profile of consumption variance $(\mathrm{d}v(\theta))$ *close to* 0*.*

We outline the argument for the two-interval configuration $n = 2$ characterized by an interior threshold type $\bar{\theta}^*$ such that every type $\theta < \bar{\theta}^*$ consumes c_1^* while the remaining higher types $\theta \geq \bar{\theta}^*$ consume c_2^* $(c_1^* > c_2^*)$. This deterministic schedule is dominated if small random noise $\mathrm{d}v(\theta) = \lambda v(\theta)$, $\lambda > 0$ small, on consumption yields a higher amount of total collected resources while the socially favored type $\theta^{\inf}$ agents do not loose, $\mathrm{d}U(\theta^{\inf}) \geq 0$. These resources can then be redistributed to each agent by uniformly lowering payments, without violating incentive constraints.

The utility of every type $\theta < \bar{\theta}^*$ changes by

$$\mathrm{d}U(\theta) = \mathrm{d}U(\theta^{\inf}) + \int_{\theta^{\inf}}^{\theta} \lambda S'_\theta \left(c_1^*, z\right) \ \mathrm{d}v(z) \ \mathrm{d}z$$

so that the total change in utility of these agents can be written, after using the integration by parts formula,

$$\mathrm{d}U(\theta^{\inf}) - \left[1 - F\left(\bar{\theta}^*\right)\right] \mathrm{d}U\left(\bar{\theta}^*\right) + \int_{\theta^{\inf}}^{\bar{\theta}^*} \lambda S'_\theta \left(c_1^*, \theta\right) m(\theta) \ \mathrm{d}v(\theta) \ \mathrm{d}z \ \mathrm{d}F(\theta).$$

Similarly, the utility of every type $\theta \geq \bar{\theta}^*$ changes by

$$\mathrm{d}U(\theta) = \mathrm{d}U(\theta^{\inf}) + \int_{\theta^{\inf}}^{\bar{\theta}^*} \lambda S'_\theta \left(c_1^*, z\right) \ \mathrm{d}v(z) \ \mathrm{d}z + \int_{\bar{\theta}^*}^{\theta} \lambda S'_\theta \left(c_2^*, z\right) \ \mathrm{d}v(z) \ \mathrm{d}z,$$

which now yields a total utility change for these agents equal to

$$\left[1 - F\left(\bar{\theta}^*\right)\right] \mathrm{d}U\left(\bar{\theta}^*\right) + \int_{\bar{\theta}^*}^{\theta^{\sup}} \lambda S'_\theta \left(c_2^*, \theta\right) m(\theta) \ \mathrm{d}v(\theta) \ \mathrm{d}F(\theta).$$

Then, from (9), the additional amount of collected resources implied by random noise writes

$$\int_{\theta^{\inf}}^{\bar{\theta}^*} \lambda \phi \left(c_1^*, \theta\right) \ \mathrm{d}v(\theta) \ \mathrm{d}F(\theta) + \int_{\bar{\theta}^*}^{\theta^{\sup}} \lambda \phi \left(c_2^*, \theta\right) \ \mathrm{d}v(\theta) \ \mathrm{d}F(\theta) - \mathrm{d}U(\theta^{\inf}).$$

Given a profile $(\mathrm{d}v(\theta))$, which Lemma 2 shows must be non-increasing to meet incentive requirements, the highest amount of collected resources that does not hurt type $\theta^{\inf}$ agents obtains by setting $\mathrm{d}U(\theta^{\inf}) = 0$. This yields the condition given in Proposition 3 for $n = 3$, with $\bar{\theta}_1^* = \theta^{\inf}$, $\bar{\theta}_2^* = \bar{\theta}^*$ and $\bar{\theta}_3^* = \theta^{\sup}$.

Remark 1. Partial bunching. Proposition 3 also applies for $\bar{\theta}^*_{n+1} < \theta^{\sup}$, i.e., in the absence of bunching at the top of the distribution. Then, one can set $v(\theta) = 0$ for all types that are not concerned by bunching, $\theta \geq \bar{\theta}^*_{n+1}$. The change in collected resources is

$$\sum_{i=1}^{n} \int_{\bar{\theta}^*_i}^{\bar{\theta}^*_{i+1}} \lambda\phi\left(c^*_i, \theta\right) \mathrm{d}v(\theta)\mathrm{d}F(\theta) - \mathrm{d}U(\theta^{\inf}) + \left[1 - F(\bar{\theta}^*_{n+1})\right] \mathrm{d}U(\bar{\theta}^*_{n+1}).$$

As above, $\mathrm{d}U(\theta^{\inf}) = 0$ maximizes the additional revenue. The perturbation argument guarantees incentive compatibility among types below $\bar{\theta}^*_{n+1}$. To avoid failures of incentives involving types above $\bar{\theta}^*_{n+1}$ one can give $\mathrm{d}U(\bar{\theta}^*_{n+1})$ to every such types. This costs $[1 - F(\bar{\theta}^*_{n+1})]\mathrm{d}U(\bar{\theta}^*_{n+1})$ in terms of tax resources, hence the result in Proposition 3 for this special case.

8 Examples of optimal randomization

We exhibit two specific parametrizations where bunching occurs with deterministic policy tools and randomized contracting is socially useful.

8.1 Multiplicative utility

We first consider a variant of the multiplicative formulation used by, e.g., Lollivier and Rochet (1983), where utility is

$$h(\theta)v(c) - y, \tag{11}$$

with $v(c)$ is increasing concave, and $h(\theta)$ is a (twice differentiable) decreasing function taking positive values.

The virtual contribution $W(c, \theta)$ defined in Section 3 is $h(\theta)v(c) - c - m(\theta)h'(\theta)v(c)$. Consumption $c^*(\theta)$ in the optimal relaxed policy thus satisfies the first-order condition $[h(\theta) - m(\theta)h'(\theta)]\, v'(c) - 1 = 0$ at $c = c^*(\theta)$. This amount locally maximizes the virtual contribution. Differentiating the first-order condition shows that

$$c'(\theta) > 0 \Leftrightarrow 1 - m'(\theta) < m(\theta)\frac{h''(\theta)}{h'(\theta)}.$$

Standard distributions have a decreasing Mills ratio, $m'(\theta) \leq 0$. In the case where θ is exponentially distributed with rate parameter λ, the Mills ratio is $m(\theta) = 1/\lambda$ and the inequality is met if $h''(\theta)/h'(\theta)$ is positive (so that $h(\theta)$ is concave) and high enough. Heterogeneity is low at the bottom, and more pronounced among high types, who differ more significantly from low types. In this case, bunching occurs, and by (7) every agent consumes c^* such that $h(\theta^{\text{inf}})v'(c^*) = 1$.

By Propositions 1 and 2, small random noise on types $\theta \leq \theta^*$ improves upon the deterministic optimum if there is $\theta^* \leq \theta^{\text{sup}}$ satisfying

$$v'(c^*)\left(1 - \frac{A(c^*)}{2}\right)\int_{\theta^{\text{inf}}}^{\theta^*} [h(\theta) - m(\theta)h'(\theta)]\, \mathrm{d}F(\theta) - F(\theta^*) > 0.$$

Note that agents have the same risk aversion, $A(c) = -v''(c)/v'(c)$ does not vary with θ. After applying the integration by parts formula, this inequality rewrites as

$$v'(c^*)\left(1 - \frac{A(c^*)}{2}\right)\left[h(\theta^{\text{inf}}) - [1 - F(\theta^*)]\, h(\theta^*)\right] - F(\theta^*) > 0. \tag{12}$$

By Assumption A1, $A(c^*) < 2$, and so (12) is not satisfied for $\theta^* = \theta^{\text{sup}}$ (we have used $h(\theta^{\text{inf}})v'(c^*) = 1$). Agents with high substitutes should indeed face a deterministic option.

However, (12) can be satisfied at the bottom of the type distribution. The left-hand side of the inequality is 0 for $\theta^* = \theta^{\text{inf}}$. As a result, it is optimal to rely on small random noise at the bottom if the derivative

$$v'(c^*)\left(1 - \frac{A(c^*)}{2}\right)\left[f(\theta^{\text{inf}})h(\theta^{\text{inf}}) - h'(\theta^{\text{inf}})\right] - f(\theta^{\text{inf}}) > 0,$$

a condition that will be met if $f(\theta^{\text{inf}})$ is close enough to 0 (λ is close enough to 0 in the exponential distribution case).

In this example, the process of screening is complicated by the shape of h, since its slope close to 0 implies tiny differences across types. No differential treatment can be implemented using deterministic tools. Randomness among low types is socially beneficial when there are few such agents. Then the efficiency cost of randomness is limited compared to the gains from relaxed incentives.

8.2 Log-utility and weibull distribution

We now consider an example where utility depends on the sum of c and θ, in accordance with the interpretation of θ as a substitute for the consumption good. Preferences are represented by the logarithmic utility function $u(c, \theta) = \ln(c + \theta)$ used in Example 3.

Using (6) the optimal deterministic relaxed redistribution policy maximizes

$$V\left(\theta^{\text{inf}}\right) = \int_{\Theta} \left[\ln(c(\theta) + \theta) - c(\theta) - \frac{m(\theta)}{c(\theta) + \theta}\right] \mathrm{d}F(\theta). \tag{13}$$

The consumption $c^*(\theta)$ that maximizes pointwise $V(\theta^{\text{inf}})$ is the nonnegative root of the first-order condition $(c^*(\theta) + \theta)^2 - (c^*(\theta) + \theta) - m(\theta) = 0$,

$$c^*(\theta) = \frac{1 + (1 + 4m(\theta))^{1/2}}{2} - \theta. \tag{14}$$

Bunching occurs in the deterministic optimum if this quantity is increasing,

$$m'(\theta) > (1 + 4m(\theta))^{1/2}. \tag{15}$$

The inequality can be satisfied for well-chosen log-logistic, Weibull, and variants of Weibull distributions such as generalized or power generalized Weibull commonly used in econometric models for duration data. Our example uses a generalized Weibull distribution (see Dimitrakopoulou, Adamidis, and Loukas (2007) for properties of this distribution). Its cumulative distribution function is

$$F(\theta) = 1 - \exp\left[1 - (1 + \lambda\theta^b)^a\right]$$

for $\theta \geq 0$, with a, b and λ positive parameters. The Mills ratio $m(\theta)$ is increasing for $a < 1$ and $b \leq 1$.

We set $a = 0.5, b = 0.05$ and $s = 0.5$. The condition (15) for bunching is satisfied if and only if $\theta \leq 19.9$, which corresponds to 22.7 percent of the population with the lowest types.

The optimal deterministic policy consists of a single contract (c^*, y^*) offered to every type $\theta \leq \bar{\theta}^*$ while all the other types are assigned the optimal relaxed contract $(c^*(\theta), y^*(\theta))$. The social objective $V(\theta^{\text{inf}})$ thus is

$$\int_0^{\bar{\theta}^*} \left[\ln(c^* + \theta) - c^* - \frac{m(\theta)}{c^* + \theta}\right] \mathrm{d}F(\theta) \tag{16}$$

$$+\int_{\bar{\theta}^*}^{+\infty}\left[\ln(c^*(\theta)+\theta)-c^*(\theta)-\frac{m(\theta)}{c^*(\theta)+\theta}\right]\mathrm{d}F(\theta).$$

Using the expression of $c^*(\theta)$ given in (14), the optimal threshold $\bar{\theta}^*$ is such that

$$c^* = c^*(\bar{\theta}^*) = \frac{1+\left(1+4m(\bar{\theta}^*)\right)^{1/2}}{2} - \bar{\theta}^* \tag{17}$$

To characterize the amount c^*, we apply the integration by parts formula and rewrite the contribution of types below $\bar{\theta}^*$ to the social objective $V(\theta^{\inf})$ in (16) as

$$-\left[1-F(\bar{\theta}^*)\right]\ln\left(c^*+\bar{\theta}^*\right)+\ln c^* - c^*F(\bar{\theta}^*).$$

Solving for the first-order condition for c^* to maximize this contribution, which is a quadratic equation in c^*, the only positive root is

$$c^* = \frac{1-\bar{\theta}^*}{2}+\frac{1}{2}\sqrt{\left(1-\bar{\theta}^*\right)^2+\frac{4\bar{\theta}^*}{F(\bar{\theta}^*)}}.$$

Replacing this expression of c^* into (17) defines the optimal threshold $\bar{\theta}^*$. Numerical computations (see the `R` code in Appendix B) yield $c^* = 4.02$ and $\bar{\theta}^* = 74.87$, with $F(\bar{\theta}^*) = 23.88$ percent.

Let us now account for random consumption with variance $\mathrm{d}v(\theta) = \mathrm{d}v > 0$ for all $\theta \leq \theta^*$, where θ^* is a threshold type below $\bar{\theta}^*$. Higher types face a deterministic option, with $\mathrm{d}v(\theta) = 0$ for all $\theta > \theta^*$. A social improvement obtains if, by Proposition 3,

$$-F(\theta^*) - \frac{1-F(\theta^*)}{c^*+\theta^*}\left(1-\frac{1}{2\left(c^*+\theta^*\right)}\right)+\frac{1}{c^*}\left(1-\frac{1}{2c^*}\right) > 0. \tag{18}$$

The shape of the left-hand side is depicted in Figure 1. It is 0 when evaluated at $\theta^* = 0$, and decreasing for θ^* close enough to 0. This implies that randomness should not concern a too narrow subset of the smallest types. For higher values of θ^*, it is single-peaked, reaching its global maximum of 0.0265 for $\bar{\theta} = 73.83$. It takes positive values for all $\theta^* \in \left[9.76, \bar{\theta}^*\right]$, with $F(9.76) = 7.33$ percent.

Thus, in this example, randomness should apply to at least the bottom 7.33 and at most 23.88 percent of types.

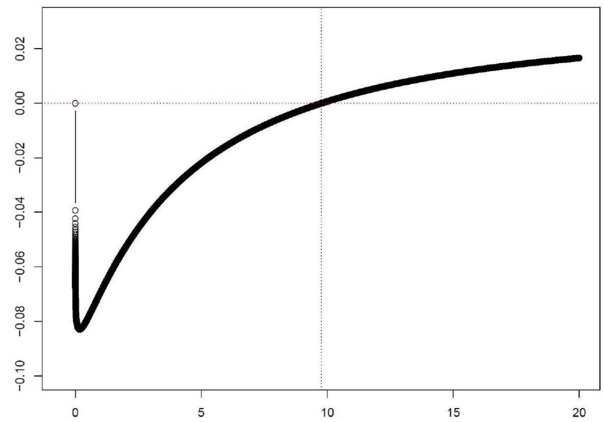

Figure 1: Random redistribution with a generalized Weibull distribution

The figure depicts function of θ^* that appears in the left-hand side of (18). It is drawn for a generalized Weibull distribution with parameters $a = 0.5, b = 0.05$ and $s = 0.5$. The threshold θ^* is on the horizontal axis. See the `R` *code in Appendix B for recovering the figure. It is optimal to expose all agents with type θ below θ^* if the function takes a positive value when evaluated at θ^*. The vertical dotted line at $\bar{\theta} = 9.76$ gives the least value of the threshold such that the function reaches positive values. For readability purposes, the figure does not represent the function for θ^* above 20. The function is actually single-peaked at takes positive values for θ^* below $\bar{\theta}^* = 74.87$, and negative values for higher θ^*, so that it is not optimal to expose all agents to random contracts.*

Relying on the interpretation of (8) as a change in total collected resources, the highest social welfare gain that can be achieved equals $0.0265 \times \lambda v$ USD. Since the less well-off get at most $\ln(c^* + \theta^{\inf}) = \ln(4.02)$USD, this gain represents a $0.0265/\ln(4.02) \times \lambda v$ share of the initial level of welfare. For $\lambda v = 1$, i.e., a one-unit increase in the average transfer (25 percent of the initial consumption), we find a modest welfare gain, with lower bound of 2 percent.

9 Conclusion

Our paper examines the choice between deterministic versus random redistribution. We have shown that the random alternative is preferred only if the best deterministic policy implies a uniform treatment of different types of agents. Randomness then allows the government to exploit new dimensions of individual heterogeneity and implement discriminatory treatment.

The existing literature following Hellwig (2007) suggests that rationing, viewed as implying randomness in the allocation of goods designed for the poor, can be justified as far as these agents display lower risk aversions. Our paper shows that stochastic redistribution can be socially useful even though random noise bears on the most risk averse agents. In this respect, it can be used to justify policies relying on rationing the less well-off part of the population to improve its welfare. This may be relevant in the case of the provision of goods when recipients differ in the availability of substitutes that are difficult to observe.

Two features in our analysis could be worth addressing in further work. First, we considered the case of a Rawlsian planner, which magnifies tensions from redistribution. A continuity argument suggests that the results should remain unaffected for weighted utilitarian redistributive preferences that place greater importance on agents who value consumption more. On the other hand, the occurrence of bunching in the deterministic optimum may be less plausible for weak redistribution motives, e.g., the unweighted (Benthamite) utilitarian social welfare objective, implying low redistribution made deterministically.

A second feature relates to the interplay between the extent of bunching and optimal randomization. Our parametric examples suggest that random contracts apply to a subset of the agents affected by failures of monotonicity requirements while redistribution should remain deterministic at the top of the distribution. It seems plausible that deterministic redistribution is more suitable when a smaller portion of the population is concerned by bunching.

References

Abajian, A., C. Cole, B.K. Jack, K. Meng, and M. Visser (2024): "Dodging Day Zero: Drought, adaptation, and inequality in Cape Town," Working Paper.

Brito, D., J. Hamilton, S. Slutsky, and J. Stiglitz (1995): "Randomization in optimal income tax schedules," Journal of Public Economics, 56, 189-223.

Chetty, R. (2006): "A new method of estimating risk aversion," American Economic Review, 96 (5): 1821-1834.

Cohen, A. and L. Einav (2007): "Estimating risk preferences from deductible choice," American Economic Review 97, 745-788.

Dimitrakopoulou, T., K. Adamidis, and S. Loukas (2007): "A lifetime distribution with an upside-downbathtub-shaped hazard function," IEEE Transactions on Reliability, 56(2), 308-311.

Dworczak, P., S. D. Kominers, and M. Akbarpour (2021): "Redistribution through markets," Econometrica, 89(4), 1665-1698.

Ederer, F., R. Holden, and M. Meyer (2018): "Gaming and strategic opacity in incentive provision," The RAND Journal of Economics, 49(4), 819-854.

Gauthier, S., and G. Laroque (2014): "On the value of randomization," Journal of Economic Theory, 151, 493-507.

Gauthier, S., and G. Laroque (2017): "Redistribution by means of lotteries," Journal of Economic Theory, 169, 707-716.

Hellwig, M. (2007): "The undesirability of randomized income taxation under decreasing risk aversion," Journal of Public Economics, 91, 791-816.

Laffont, J.-J., and D. Martimort (2002): The theory of incentives: The principalagent model, Princeton University Press.

Lang, M. (2017): "Legal uncertainty as a welfare enhancing screen," European Economic Review, 91, 274-289.

Lang, M. (2023): "Stochastic contracts and subjective evaluations," Rand Journal of Economics, 54, 104-134.

Lollivier, S., and J.-C. Rochet (1983): "Bunching and second-order conditions: A note on optimal tax theory," Journal of Economic Theory, 31, 392-400.

Mirrlees, J. (1971): "An exploration in the theory of optimum income taxation," Review of Economic Studies, 38, 175-208.

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

Mussa, M. and S. Rosen (1978): "Monopoly and product quality," Journal of Economic Theory 18, 301-317.

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

Pavlov, G. (2011): "Optimal mechanism for selling two goods," The B.E. Journal of Theoretical Economics, 11, 1-35.

Pycia, M., and U. Unver (2015): "Decomposing random mechanisms," Journal of Mathematical Economics, 61, 21-33.

Saez, E. (2001): "Using elasticities to derive optimal income tax rates," Review of Economic Studies, 68, 205-229.

Slemrod, J. (2019): "Tax compliance and enforcement," Journal of Economic Literature 57, 904-954.

Spence, M., (1977): "Nonlinear prices and welfare," Journal of Public Economics, 8, 1-18.

Stern, N.H. (1982): "Optimum taxation with errors in administration", Journal of Public Economics, 17, 181-211.

Stiglitz, J. (1981): "The allocation role of the stock market," The Journal of Finance, 36, 235-251.

Stiglitz, J. (1982): "Utilitarianism and horizontal equity: the case for random taxation," Journal of Public Economics, 18, 1-33.

Strausz, R. (2006): "Deterministic versus stochastic mechanisms in principal-agent models," Journal of Economic Theory, 128, 306-314.

Tobin, J. (1971): "On limiting the domain of inequality," Journal of Law and Economics, 13, 263-277.

Weiss, L. (1976): "The desirability of cheating incentives and randomness in the optimal income tax," Journal of Political Economy, 84, 1343-1352.

Weitzman, M. (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.

Wijkander, H. (1988): "Equity and efficiency in public sector pricing: A case of stochastic rationing," Econometrica, 56, 1455-1465.

Appendices

A A Proof of Proposition 1

We first express the payment made by type θ as a function of her indirect utility $V(\theta)$,

$$y(\theta) = u(c^*, \theta) + \lambda S(c^*, \theta) v(\theta) - V(\theta).$$

The expression of the indirect utility $V(\theta)$ obtains from using the first-order necessary condition in Lemma 2 for incentive compatibility, which yields

$$U(\theta) = U\left(\theta^{\inf}\right) + \int_{\theta^{\inf}}^{\theta} \lambda S'_{\theta}(c^*, z) v(z) \mathrm{d}z,$$

so that

$$V(\theta) = u(c^*, \theta) + U\left(\theta^{\inf}\right) + \int_{\theta^{\inf}}^{\theta} \lambda S'_{\theta}(c^*, z) v(z) \mathrm{d}z.$$

The feasibility constraint (2) reads

$$\int_{\Theta} \left[c^* + \lambda v(\theta) - y(\theta)\right] \mathrm{d}F(\theta) = 0.$$

Replacing $y(\theta)$ with its expression in terms of $V(\theta)$, with $V(\theta)$ given above, we find

$$\int_{\Theta} \left[c^* + \lambda v(\theta) - \lambda S(c^*, \theta) v(\theta)\right] \mathrm{d}F(\theta)$$

$$+U\left(\theta^{\inf}\right) + \int_{\Theta} \int_{\theta^{\inf}}^{\theta} \lambda S'_{\theta}(c^*, z) v(z) \, \mathrm{d}z \, \mathrm{d}F(\theta) = 0.$$

Using the integration by parts formula,

$$\int_{\Theta} \int_{\theta^{\inf}}^{\theta} S'_{\theta}(c^*, z) v(z) \, \mathrm{d}z \, \mathrm{d}F(\theta) = \int_{\Theta} m(\theta) S'_{\theta}(c^*, \theta) v(\theta) \, \mathrm{d}F(\theta),$$

the feasibility constraint allows us to get the sub-utility $U(\theta)$ driving incentives in the presence of small random tax perturbations for type $\theta^{\inf}$,

$$U\left(\theta^{\inf}\right) = -\int\limits_{\Theta} \left[c^* + \lambda v(\theta) - \lambda S\left(c^*, \theta\right) v(\theta) + m(\theta) \lambda S'_\theta \left(c^*, \theta\right) v(\theta)\right] \mathrm{d}F(\theta).$$

Social welfare is $V\left(\theta^{\inf}\right) = u\left(c^*, \theta^{\inf}\right) + U\left(\theta^{\inf}\right)$, which is actually

$$u\left(c^*, \theta^{\inf}\right) - \int\limits_{\Theta} \left[c^* + \lambda \phi\left(c^*, \theta\right) v(\theta)\right] \mathrm{d}F(\theta),$$

with $\phi\left(c^*, \theta\right)$ defined in Proposition 1.

The expression of social welfare in the absence of noise obtains by letting $v(\theta) = 0$ for all θ. It reduces to $u\left(c^*, \theta^{\inf}\right) - c^*$. This yields condition (8) in Proposition 1 for socially useful random redistribution (recall that $\lambda \geq 0$ for the variance of consumption to be non-negative). This concludes the proof.

B Detailed derivation of (10)

Applying the integration by parts formula, we have

$$\int\limits_{\theta^{\inf}}^{\theta^*} m(\theta) S'_\theta(c^*, \theta) \, \mathrm{d}F(\theta) = \left[1 - F(\theta^*)\right] S(c^*, \theta^*) - S(c^*, \theta^{\inf}) + \int\limits_{\theta^{\inf}}^{\theta^*} S(c^*, \theta) \, \mathrm{d}F(\theta).$$

Therefore,

$$\int\limits_{\theta^{\inf}}^{\theta^*} \phi(c^*, \theta) \, \mathrm{d}F(\theta) = -F(\theta^*) - \left[1 - F(\theta^*)\right] S(c^*, \theta^*) + S(c^*, \theta^{\inf}).$$

This is positive if and only if

$$S(c^*, \theta^{\inf}) - S(c^*, \theta^*) > F(\theta^*) \left[1 - S(c^*, \theta^*)\right].$$

Since $S(c^*, \theta^*) < 1$, this rewrites as

$$\frac{S(c^*, \theta^{\inf}) - S(c^*, \theta^*)}{1 - S(c^*, \theta^*)} > F(\theta^*),$$

which is equivalent to (10).

C R code for Section 8

```
a <- 0.5; b <- 0.05; s <- 0.5
  FF <- function(x) 1 - exp (1-(1+s*x^b)^a)
  ff <- function(x) {
    ff <- (a*(1+s*x^b )^(a-1)*s*b*x^(b-1))
    ff <- ff*exp(1-(1+s*x^b)^a)
    ff
  }
  mm <- function(x) (1-FF(x)) / ff(x)
  mmprime <- function(x) {
    mmp <- - (a*b*(b-1)*s*x^(b-2)*(1+s*x^b)^(a-1))
    temp <- a*b*s*x^(b-1)*(a-1)*s*b*x^(b-1)
    mmp <- mmp - temp*( 1+s*x^b)^(a-2)
    mmp <- mmp / (a*b*s*x^(b-1)*( 1+s*x^b)^(a-1))^2
    mmp
  }

bunch <- function(x) mmprime(x) - (1+4*mm(x))^(1/2)
  xx <- seq (1e-10,1e3,1e-2)
  plot (xx , bunch(xx))
        # bunching occurs for xx such that bunch (xx) is positive
  max (xx[bunch(xx)>=0]); FF(max(xx[bunch(xx)>=0]))

thetabar <- function(x) x+((1-x)^2+4*x/FF(x))^(1/2)-(1+4*mm(x))^(1/2)
  plot (xx, thetabar(xx))
        # threshold below which bunching occurs has thetabar = 0
  min(xx[thetabar(xx)>=0]); FF(min(xx[thetabar(xx)>=0]))
  theta <- min(xx[thetabar(xx)>=0])
  theta

cbar <- (1-theta)/2+((1-theta)^2+4*theta/FF(theta))^(1/2)/2
  cbar

sfnum <- function(x) {
  sf <- (1-FF(x))*(1-1/(2*(cbar+x)))/(cbar+x)
  sf <- -FF(x)-sf+(1-1/(2*cbar))/cbar
  sf
  }
```

```
  sumphi <- function(x) sfnum(x)-sfnum(1e-10) # sumphi is 0 at 1e-10
  xx <- seq (1e-10, theta, 1e-1)
  max(sumphi(xx)); xx[sumphi(xx)==max(sumphi(xx))]
  sumphi(theta)
  min(xx[sumphi(xx)>0]); FF(min(xx[(xx)>0]))
# Figure exported in the main text
xx <- seq(1e-10, 2, 1e-5)
  plot(xx, sumphi(xx), type="b", xlim =c(-0.5,20), ylim =c(-0.1,0.03))
  xx <- seq(2,20,1e-3)
  points(xx, sumphi(xx), type ="b")
  abline(h=0, col="red", lty="dotted")
  abline(v= min(xx[sumphi(xx)>0]), lty ="dotted")
```