UTMD Working Paper
The University of Tokyo
Market Design Center

UTMD-104

No Screening is More Efficient with Multiple Objects

Shunya Noda
University of Tokyo

Genta Okada
University of Tokyo

First Draft: August 20, 2024
This Version: December 14, 2025

No Screening is More Efficient with Multiple Objects*

Shunya Noda† Genta Okada‡

First Draft: August 20, 2024 This Version: December 14, 2025

Abstract

We study efficient mechanism design for allocating multiple heterogeneous objects. The aim is to maximize the residual surplus, the total value generated from an allocation minus the costs of screening. We discover a robust trend indicating that no-screening mechanisms, such as serial dictatorship with exogenous priority order, tend to perform better as the variety of goods increases. We analyze the underlying reasons by characterizing asymptotically efficient mechanisms in a stylized environment. We also apply an automated mechanism design approach to numerically derive efficient mechanisms and validate the trend in general environments. Building on these implications, we propose the *register-invite-book system* (RIB) as an efficient system for scheduling vaccination against pandemic diseases.

Keywords: Money Burning, Multi-dimensional Types, Assignment Market, Pandemic Vaccine, Extreme Value Theory, Automated Mechanism Design

---

*Authors are listed in alphabetical order. We are grateful to Itai Ashlagi, Yu Awaya, Nima Haghpanah, Yuichiro Kamada, Michihiro Kandori, Fuhito Kojima, Junpei Komiyama, Yukio Koriyama, Satoshi Kurihara, Fumio Ohtake, Wojciech Olszewski, Parag Pathak, Satoru Takahashi, Alex Teytelboym, Yuichi Yamamoto, Frank Yang, and all participants of the conference celebrating the 65th birthday of Michihiro Kandori (Tokyo), Summer Workshop on Economic Theory 2024 (Otaru), 7th International Workshop on Matching Under Preferences (Oxford), Japanese Economic Association 2024 Autumn Meeting (Fukuoka), Tokyo Conference on Market Design 2025 (Tokyo), Econometric Society World Congress 2025 (Seoul), and a seminar at Tokyo University of Science for helpful comments. This work has been supported by JST PRESTO Grant Number JPMJPR2368 and JST ERATO Grant Number JRMJER2301, Japan. All remaining errors are our own.

†Contact Author. Graduate School of Economics, the University of Tokyo 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan. E-mail: shunya.noda@e.u-tokyo.ac.jp

‡Graduate School of Economics, the University of Tokyo 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan. E-mail: ogenta0628@g.ecc.u-tokyo.ac.jp

1 Introduction

Policymakers often face the challenge of allocating scarce resources under conditions of information asymmetry. The goal is to allocate goods to those with the highest valuations, but when tractable and socially costless screening devices such as monetary transfers are unavailable, screening can become costly. For example, first-come-first-served (FCFS) mechanisms give high-value agents an incentive to claim the good immediately after distribution begins, wasting the effort of acting early as ordeals. Conversely, lottery-based mechanisms prevent wasted effort but risk misallocating goods. Maximizing social welfare (defined as *residual surplus*, the agents' total payoffs from the realized allocation minus wasted screening costs) requires balancing allocative efficiency against wasted screening costs.

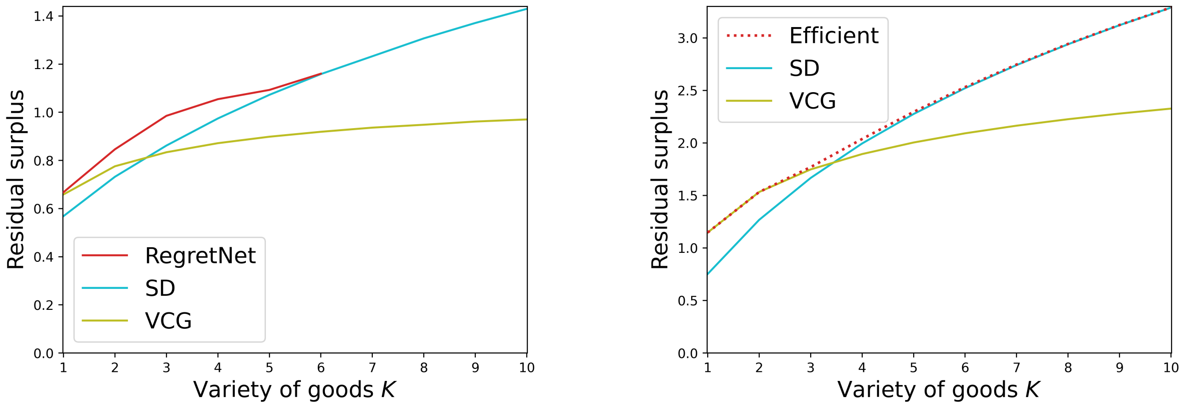
Despite its importance, the literature has largely solved problems only in single-dimensional settings with homogeneous goods (Hartline and Roughgarden, 2008). Indeed, real-world problems often involve multiple heterogeneous goods. The scheduling of vaccine appointments during the COVID-19 pandemic was a prominent example—reservation slots, each of which specifies when, where, and which vaccine will be administered, are highly heterogeneous goods, and potential recipients have diverse preferences over slots. In such settings, agents' effort influences not only whether they secure a good but also which good they obtain. Since the efficient allocation of a large number of heterogeneous goods has not been sufficiently established in the literature, economists were unable to provide effective guidance on vaccine appointment scheduling. Consequently, various uncoordinated methods were adopted worldwide, often leading to confusion.[1]

This paper explores efficient screening mechanisms for allocating heterogeneous objects. Agents have multi-dimensional preferences over objects but only demand one item. They can signal their preferences through costly effort, which serves as a form of payment. The central planner aims to maximize residual surplus.[2] Two prominent candidates are *serial dictatorship with exogenous order* (SD), where agents sequentially choose their most preferred available goods without paying screening costs, and the *Vickrey-Clarke-Groves mechanism* (VCG), which achieves allocative efficiency at the expense of high screening costs. Between these extremes, the planner has many options for balancing screening costs and allocative efficiency. Characterizing optimal mechanisms in this multi-dimensional setting is widely believed to be challenging.

This study uncovers a robust trend: as the variety of goods increases, no-screening mechanisms like SD tend to become more efficient. Figure 1a illustrates this result. We consider a scenario where $2K$ agents each demand a single unit among K heterogeneous objects, and each agent i's

[1] As detailed in Section 5.2, numerous regions that adopted FCFS, including Florida, Massachusetts, Washington, D.C., and many local governments in Japan, experienced severe competition for reservation slots. By contrast, several countries and regions that adopted other systems (such as an invitation-based system) experienced far fewer such disruptions.

[2] The environment is the assignment market originally studied by Koopmans and Beckmann (1957) and Shapley and Shubik (1971), but our objective differs in that it is the residual surplus, where monetary transfers are not zero-sum but burned.



(a) Finite Market (Weibull with Parameter 0.8)

(b) Continuous Market (Weibull with Parameter 0.6)

Figure 1: The relationship between the variety of goods and the per-capita residual surplus achieved by SD, VCG, and the (numerically) efficient mechanism.

value for object k, v_k^i, is drawn i.i.d. from a Weibull distribution with parameter 0.8. The horizontal axis of the figure represents the value of K, and the vertical axis shows the residual surplus per agent. Under this distribution, when $K = 1$ (i.e., the single-good case), the optimality of VCG has been proven by Hartline and Roughgarden (2008). However, for $K \geq 3$, SD outperforms VCG, and when $K \geq 6$, the performance gap between SD and the mechanism designed by RegretNet (Dütting et al., 2019), an automated mechanism design method for discovering efficient mechanisms, becomes negligible. As K increases further, the performance of SD continues to improve, whereas the performance gains of VCG taper off. This paper's findings suggest that such trends will likely hold in a wide range of environments.

Formally, this paper examines the impact of multiple goods on efficient mechanisms through theoretical analysis, computational experiments, and real-world applications. First, we use extreme value theory to analytically characterize behavior in a stylized environment, showing that the performance of no-screening mechanisms improves as the number of goods increases. Second, we employ automated mechanism design to explore efficient mechanisms numerically, in more general settings, confirming the robustness of our findings. Third, we consider vaccine appointment scheduling during COVID-19, demonstrating that the *register-invite-book system* (RIB), an SD-based implementation, effectively distributes heterogeneous reservation slots.

We begin with a stylized large market with a continuum of agents and objects, called a continuous i.i.d. market. Agents consume at most one object, and objects are divided into $K < \infty$ types with equal capacities. Each agent's value for each object type is drawn i.i.d. from a marginal distribution. We first demonstrate that this setup reduces the problem to a single-dimensional environment, where each agent's value corresponds to the largest order statistic—the highest value among all K options (for a formal statement of this reduction including underlying assumptions, see Section 3 and Theorem 1).

We analyze the reduced single-dimensional environment and the distribution of the largest order statistic to characterize efficient mechanisms in the stylized environment. First, we prove that in a single-dimensional environment, a no-screening mechanism is efficient if and only if the value distribution satisfies the *new better than used in expectation* (NBUE) property. NBUE is a weaker condition than *increasing hazard rate* (IHR), shown as a sufficient condition by Hartline and Roughgarden (2008). Second, we prove that, for any marginal value distribution, as the variety of goods, K, increases, the distribution of the largest order statistic tends to satisfy NBUE and IHR, implying that no-screening mechanisms become more likely to be efficient as the number of goods grows. Third, using extreme value theory, we analyze the limit as the variety of goods K approaches infinity. We show that in this limit, the no-screening mechanism is efficient for a broad class of marginal distributions, and even when it is not, it is practically close to being efficient.

Figure 1b illustrates the observation for the continuous market. Parallel to the setting of Figure 1a, there are K types of objects with an equal mass and agents whose mass is twice that of the objects. Each agent i's value for each object type k, v_k^i, is independently drawn from a Weibull distribution with parameter 0.6. We observe that the performance of SD improves as K increases. For such a continuous market, we can also prove as a theorem that whenever the marginal value distribution satisfies the von Mises condition with parameter 0 (von Mises, 1936), such as the Weibull distribution (with any parameter), the performance ratio of SD to the exact efficient mechanism converges to one. (The setting and results will be reviewed in Section 3.9.)

Analytical and even numerical analyses for finite markets with heterogeneous objects have been known to be extremely challenging. Nevertheless, we leverage deep-learning techniques for automated mechanism design, which have recently been developed with a primary focus on revenue maximization (Dütting et al., 2019), to derive mechanisms for maximizing residual surplus. Consistent with the results from the analysis of stylized environments, we show that in these finite markets, as the number of objects (or types of objects) increases, the performance of SD improves, as depicted in Figure 1a.

We further conduct numerical analyses of cases with correlated values, which are practically important but challenging for the application of extreme value theory.[3] Correlation between values assigned to different objects by a single agent (within-agent correlation) brings the situation closer to a single-dimensional (i.e., homogeneous-good) environment, thereby weakening the effect that the multiplicity of objects improves the performance of no-screening mechanisms. However, the correlation between values assigned to a single object by different agents (between-agent correlation) enhances the relative performance of no-screening mechanisms, as in such cases, the screening costs outweigh the allocative benefits. While these types of correlations can either amplify or mitigate the effects of increasing the variety of goods, they do not qualitatively impact the policy implication

[3] According to a review by Majumdar et al. (2020), except for cases with weakly correlated values—where the correlation length is finite and a renormalization group-like argument can reduce the problem to that of uncorrelated values—the behavior of the extreme value statistic in the limit is largely unknown.

that the relative performance of SD improves as the variety of goods increases.

Our findings have significant implications for vaccine distribution during pandemics. During COVID-19, many authorities used FCFS systems, leading to stress and inefficiency. By contrast, SD-based systems, used in some (local) governments, including British Columbia and Singapore, effectively allocated reservation slots without inducing wasted effort. We build on these successful examples as well as the analytical and numerical results to establish the register-invite-book system (RIB) for vaccine distribution. RIB maximizes residual surplus by eliminating screening while maintaining the practical convenience of FCFS.

2 Related Literature

Maximizing residual surplus parallels designing mechanisms with monetary transfers, such as auctions that maximize welfare while discarding payments. This problem, often called *money burning*, has been studied extensively. For environments with homogeneous goods, (Hartline and Roughgarden, 2008; Yoon, 2011; Condorelli, 2012; Chakravarty and Kaplan, 2013) provide characterizations of efficient mechanisms. More recent works, such as Akbarpour et al. (2023b) and Tokarski (2025a), combine wasteful effort (ordeals) with other screening devices, such as payments or damages, to propose mechanisms that screen not only for the need for the good but also simultaneously for other private characteristics, including social welfare weights and the cost of completing the ordeal.

However, characterizing optimal mechanisms for multi-dimensional types with multiple objects remains challenging. While revenue maximization has received considerable attention (e.g., Manelli and Vincent, 2006; Pavlov, 2011; Haghpanah and Hartline, 2021; Giannakopoulos and Koutsoupias, 2018; Daskalakis et al., 2015), no complete characterization exists, even in simple scenarios. Similarly, maximizing residual surplus in multi-object environments is complex. The existing literature provides only partial characterizations, and typically only in restricted environments. For example, Tokarski (2025b) shows that when distributing two goods using deterministic mechanisms, one can improve residual surplus by deliberately degrading the value of the goods. Fotakis et al. (2016) and Goldner and Lundy (2024) propose approximation mechanisms with performance guarantees. Although our characterization is also confined to a stylized environment, our results suggest that when the number and variety of goods are large, the efficient mechanism may become relatively simple and practically implementable.

Reflecting the analytical difficulty of designing mechanisms for environments with multi-dimensional types, methods for automated mechanism design have also advanced substantially (Sandholm, 2003; Conitzer and Sandholm, 2004; Dütting et al., 2019). To our knowledge, our study is the first to apply these methods to residual surplus maximization. It is also worth noting that our study provides practical insights, whereas revenue-maximizing mechanisms are often impractical for real-world implementation because they frequently have complex structures or are highly sensitive to the environment.

Serial dictatorship (SD) mechanisms are well-studied for their desirable properties. Pycia and Troyan (2024) demonstrate that the random serial dictatorship in sequential form is the only mechanism that satisfies Pareto efficiency, symmetry, and obvious strategy-proofness (Li, 2017). Bade (2015) demonstrate SD's robustness when agents can acquire preference information, and Hakimov et al. (2023) confirm its empirical effectiveness in school choice. Additionally, SD achieves at least $1 - 1/e$ of the first-best matching size (Krysta et al., 2014; Noda, 2020) and prevents scalping (Hakimov et al., 2021). Combining these advantages with our findings, SD is a promising option for allocating many heterogeneous goods without payments.

Finally, research on medical resource allocation during crises has grown rapidly. Studies have analyzed incentives for production (Ahuja et al., 2021; Castillo et al., 2021; Athey et al., 2022), capacity design (Noda, 2018), priorities (Gans, 2022; Akbarpour et al., 2023a), and reserves (Pathak et al., 2022, 2024). Our paper complements this literature by focusing on the downstream process of vaccine distribution, offering practical insights for future public health crises.

3 Continuous Market

We first consider a large market where a continuum of agents demands a continuum of goods, which can be classified into a finite variety of types. By imposing a few simplifying assumptions, this market can be reduced to a well-characterized single-dimensional environment, allowing us to analytically derive the efficient mechanism. We demonstrate that as the goods variety increases, the relative performance of the no-screening mechanism improves. Then, we characterize the properties of the efficient mechanism in the limit, using the insights from extreme value theory.

3.1 Model

There is a unit mass of a continuum of agents $I \coloneqq [0, 1]$ with a unit demand. Each object has a type $k \in K$, where K is a finite set of object types. The mass of type-k objects is denoted by $m_k \in \mathbb{R}_+$. We denote the total mass of objects by $\bar{m} \coloneqq \sum_{k \in K} m_k$.

Each agent is specified by their valuation vector, $\mathbf{v} = (v_k)_{k \in K} \in V \subset \mathbb{R}_+^K$, where v_k denotes this agent's valuation for a type-k object. For notational simplicity, we omit the index specifying the agent's identity; thus, $\mathbf{v}$ represents the valuation vector of a single agent, not their profile. Agents can choose a payment (effort cost) level, and the central planner can observe it. Since efforts (ordeals) are used just like payments, we refer to them simply as "payments," whereas we emphasize that all payments are burned without enriching the central planner as revenue. Accordingly, when this agent obtains object k with probability $x_k \in [0, 1]$ for each k while making a payment of p, her payoff is $\sum_{k \in K} v_k x_k - p$.

Throughout this section, we consider direct mechanisms that determine an allocation and payment based only on the agent's own report and the distribution of the reported type profiles.[4] This restriction has two effects. First, we can drop an index representing the agent's identity because an agent's identity does not alter the agent's outcome. Second, in a continuous market, there is no aggregate uncertainty, and thus the distribution of reported types remains constant in equilibrium. Accordingly, we can also omit the other agents' type reports (or their distribution) from the arguments of the mechanism. A mechanism $\mathcal{M}$ is comprised of an allocation rule $\mathbf{x} : V \to [0,1]^K$ and a payment rule $p : V \to \mathbb{R}_+$, where $\mathbf{x}(\mathbf{v})$ and $p(\mathbf{v})$ are the respective allocation and payment when the agent has a valuation $\mathbf{v}$. Let $F : V \to [0,1]$ be the cumulative distribution function of agents' valuations. Our main measure of social welfare, *residual surplus* from a mechanism $\mathcal{M} = (\mathbf{x}, p)$ is defined as

$$RS(\mathcal{M}) \coloneqq \int_V \left(\sum_{k \in K} v_k x_k(\mathbf{v}) - p(\mathbf{v}) \right) dF(\mathbf{v}). \tag{1}$$

In our setting, because payments are burned, the objective function (1) differs from the *gross surplus*, social welfare considered in a setting with monetary transfers:

$$GS(x) \coloneqq \int_V \sum_{k \in K} v_k x_k(\mathbf{v}) dF(\mathbf{v}).$$

An allocation rule $\mathbf{x}$ is *allocatively efficient* if it maximizes the gross surplus.

Below, we list several constraints a mechanism must satisfy. A mechanism $\mathcal{M}$ satisfies the *resource constraint* if, for each k, there is at most mass m_k of agents who receive object k:

$$\int_V x_k(\mathbf{v}) dF(\mathbf{v}) \leq m_k \text{ for all } k \in K. \tag{2}$$

A mechanism $\mathcal{M}$ is *strategy-proof* if there is no gain from misreporting:

$$\sum_{k \in K} v_k x_k(\mathbf{v}) - p(\mathbf{v}) \geq \sum_{k \in K} v_k x_k(\mathbf{v}') - p(\mathbf{v}') \text{ for all } \mathbf{v}, \mathbf{v}' \in V. \tag{3}$$

A mechanism $\mathcal{M}$ is *individually rational* if no agent obtains a negative payoff:

$$\sum_{k \in K} v_k x_k(\mathbf{v}) - p(\mathbf{v}) \geq 0 \text{ for all } \mathbf{v} \in V. \tag{4}$$

A mechanism $\mathcal{M}$ satisfies the *unit demand* condition if an agent obtains at most one object:

$$\sum_{k \in K} x_k(\mathbf{v}) \leq 1 \text{ for all } \mathbf{v} \in V. \tag{5}$$

A mechanism $\mathcal{M}$ is *efficient* if it maximizes the residual surplus RS subject to all the above

[4] This property is obtained in the continuous limit of a sequence of mechanisms satisfying Liu and Pycia's (2016) regularity condition.

constraints: the resource constraint, strategy-proofness, individual rationality, and unit demand.

3.2 Additional Restrictions for Analytical Tractability

For theoretical analyses, we make several additional simplifying assumptions to enable a closed-form characterization of an efficient mechanism. Specifically, we assume that the set of all possible valuations V can be written as $V = [0, \bar{v}]^K$, where $\bar{v} \in \mathbb{R} \cup \{+\infty\}$, and there exists a *marginal value distribution* $G : [0, \bar{v}] \to [0, 1]$ such that the distribution function of valuation vectors, F, can be written as a product of the G function: For all $\mathbf{v} = (v_k)_{k \in K}$, we have $F(\mathbf{v}) = \prod_{k \in K} G(v_k)$. Throughout this paper, we assume that G has full support and is twice continuously differentiable. The first and second-order derivative of G is written as g and g', respectively. We further assume that all object types have the same capacity; i.e., we set $m_k = \bar{m}/K$ for all objects $k \in K$. We focus on the case of $\bar{m} \in (0, 1)$ to make the problem non-trivial. We refer to this environment as a *continuous i.i.d. market*, which is parametrized by its marginal value distribution G, the cardinality of object types, $|K|$, and the capacity parameter $\bar{m}$. Slightly abusing notation, we denote K to represent its cardinality $|K|$ unless it causes confusion.[5]

We restrict our attention to symmetric mechanisms. We say that a mechanism is *neutral* if it treats all objects symmetrically, without using their a priori labels.

Definition 1 (Neutrality)**.** A mechanism $\mathcal{M} = (\mathbf{x}, p)$ is *neutral* if for all permutations of objects $\sigma : K \to K$, we have

$$\mathbf{x}(\sigma(\mathbf{v})) = \sigma(\mathbf{x}(\mathbf{v})),$$
$$p(\sigma(\mathbf{v})) = p(\mathbf{v}).$$

Since all objects are assumed to be symmetric, there always exists an efficient mechanism that satisfies neutrality.

Proposition 1. In a continuous i.i.d. market, there exists an efficient and neutral mechanism.

The proof is similar to that of Theorem 1 of Rahme et al. (2021).

The final assumption we impose for the theoretical analysis is ex post efficiency. A mechanism is *ex post (Pareto) efficient* if the allocation resulting from the mechanism satisfies the property that no group of agents can simply exchange objects (without monetary transfers) and still make a mutual gain.

Definition 2 (Ex Post Efficiency)**.** An allocation rule x is *ex post efficient* if there is no allocation rule x' that satisfies the following four conditions: (i) the resource constraint, (ii) the unit demand condition, (iii) $\sum_{k \in K} v_k x'_k(\mathbf{v}) \geq \sum_{k \in K} v_k x_k(\mathbf{v})$ for all $\mathbf{v} \in V$, and (iv) there exists $V' \subset V$ such that $F(V') > 0$ and $\sum_{k \in K} v_k x'_k(\mathbf{v}) > \sum_{k \in K} v_k x_k(\mathbf{v})$ for all $\mathbf{v} \in V'$.

[5] Appendix A.2 studies the case of asymmetric capacity.

In a continuous i.i.d. market, any finitely many type-level exchanges can be represented as a measurable allocation rule that redistributes an arbitrarily small but positive mass of agents while respecting feasibility and unit-demand constraints. Consequently, the absence of any such rule that yields a pointwise improvement over x on a set of positive measure is equivalent to the absence of profitable trades at the realized allocation. This establishes the correspondence between the intuitive "no mutually beneficial exchanges" condition and Definition 2.

Although it is not clear a priori whether an optimal residual-surplus-maximizing mechanism must satisfy ex post efficiency, restricting attention to ex post efficient mechanisms is well-motivated for several reasons. First, in many practical environments, agents can reallocate resources through informal swaps or secondary markets. If a mechanism generates allocations that are ex post inefficient, such off-mechanism trades will naturally occur, potentially undermining the intended incentive properties or creating inequities among agents depending on their ability to coordinate such trades. Second, ex post inefficiency is difficult to defend on normative grounds: if a group of agents could all be made weakly better off through simple exchanges that leave all other agents unaffected, it is hard to justify the original allocation as a policy recommendation.

We also note that, even without imposing neutrality or ex post efficiency, our automated mechanism design computations consistently output mechanisms that satisfy ex post efficiency up to numerical tolerances. This suggests that ex post efficiency emerges endogenously in optimal residual-surplus–maximizing designs (at least in a certain class of environments).

3.3 Reduction to a Single-Dimensional Environment

Imposing neutrality and ex post efficiency removes the complexity arising from multidimensional types and thereby greatly simplifies the analysis of the optimal mechanism. To achieve ex post efficiency, we should allocate objects in such a way that agents receive their favorite one whenever possible. In a continuous i.i.d. market, both the popularity and scarcity of objects are equal, and neutrality requires the mechanism to allocate objects symmetrically. Accordingly, under a mechanism that is both neutral and ex post efficient, each agent faces only two possibilities: receiving their favorite object or receiving nothing at all.

Theorem 1. In a continuous i.i.d. market, if $\mathcal{M} = (x, p)$ is neutral and ex post efficient, then there exists $V' \subset V$ such that $F(V') = 1$ and $x_k(\mathbf{v}) > 0$ implies $v_k = \max_{l \in K} v_l$ for all $\mathbf{v} \in V'$.

Proofs are shown in Appendix B. This theorem holds only in the continuous i.i.d. market. When the popularity or scarcity of objects is asymmetric, some agents may, of course, end up receiving unpopular or abundant objects, even if those are not their favorite ones. In finite markets, objects symmetric ex ante may become asymmetric ex post, depending on realized preferences, causing some agents to receive their second or lower choices. However, in broad environments, agents tend to receive highly preferred objects when both goods and preferences are diverse. Theorem 1 captures this insight in its most extreme yet tractable form. As we later show through numerical analysis, the qualitative implications derived from the analysis of the continuous i.i.d. market carry over to more general environments.

Due to Theorem 1, since each agent can only obtain their favorite object, valuations for other objects can be disregarded. Thus, the problem can be reduced to a simpler environment with single-dimensional types, which we call the *reduced market*. We denote each agent's valuation for the "single object" in the reduced market as v, and let G_K be its distribution. The value v is defined as the largest order statistic of valuations in the original market: $v := \max_{k\in K} v_k$. Thus, we can derive its distribution G_K using the marginal value distribution G as $G_K(v) := (G(v))^K$. We refer to G_K as the *reduced value distribution*. The "single object" in the reduced market has capacity $\bar{m}$, as it is the total capacity of the original environment. Slightly abusing the notation, we use (x, p) to represent the allocation rule and the payment rule, whereas now $x : [0, \bar{v}] \to [0, 1]$ and $p : [0, \bar{v}] \to \mathbb{R}_+$ maps each agent's value for the favorite object v to the probability that she obtains her favorite good $x(v)$ and her payment $p(v)$, respectively.

The residual surplus maximization problem for the reduced market is expressed as follows.

$$
\begin{aligned}
&\max_{(x,p)} \int_0^{\bar{v}} (vx(v) - p(v))\, dG_K(v) \\
\text{s.t.}\quad & \int_0^{\bar{v}} x(v) dG_K(v) = \bar{m}, && \text{(Resource Constraint)} \\
& vx(v) - p(v) \geq vx(v') - p(v') \text{ for all } v, v' \in [0, \bar{v}], && \text{(Strategy-Proofness)} \\
& vx(v) - p(v) \geq 0 \text{ for all } v \in [0, \bar{v}], && \text{(Individual Rationality)} \\
& x(v) \leq 1 \text{ for all } v \in [0, \bar{v}]. && \text{(Unit Demand)}
\end{aligned}
$$

The solution is an efficient mechanism for a reduced market. It is straightforward to construct the efficient mechanism in the original continuous i.i.d. market based on that solution. Letting (x^r, p^r) be an efficient mechanism of the reduced problem, an efficient mechanism of the continuous i.i.d. market, $(\mathbf{x}^o, p^o)$, is given by $x^o_k(\mathbf{v}) = x^r(\max_{l\in K} v_l)$ if $k = \arg\max_{l\in K} v_l$ and $x^o_k(\mathbf{v}) = 0$ otherwise,[6] and $p^o(\mathbf{v}) = p^r(\max_{k\in K} v_k)$. Since $k = \arg\max_l v_l$ occurs with probability $1/K$, we have $\int_V x^o_k(\mathbf{v}) dF(\mathbf{v}) = \bar{m}/K$ for every k, implying that the original resource constraint is satisfied.

3.4 Mechanisms

No Screening/Serial Dictatorship For a reduced market, a *no-screening mechanism* is defined as a mechanism that uniformly assigns the right to obtain the most preferred object with probability $\bar{m} \in (0, 1)$, requiring no payment. That is, $x(v) = \bar{m}$ and $p(v) = 0$ for all $v \in [0, \bar{v}]$.

[6] For simplicity of exposition, we ignore the case of ties, i.e., $\arg\max_{k\in K} v_k$ has multiple elements. Since the probability of a tie is zero, any tie-breaking rule can be applied.

The residual surplus achieved under the no-screening mechanism is

$$RS(\mathcal{M}_{SD}) = \bar{m}\mathbb{E}_{v\sim G_K}[v] = \bar{m}\int_0^{\bar{v}} v dG_K(v).$$

The revenue equivalence theorem of Myerson (1981) implies that the gross surplus $\bar{m}\mathbb{E}[v]$ cannot be improved further without burning payments.

This result is achieved in a continuous i.i.d. market if we apply *serial dictatorship (SD)*, in which agents take their most preferred objects sequentially following an exogenous priority order. In SD, the first $\bar{m}$ agents of the priority order, facing all goods available, obtain their favorite goods. Exactly when mass $\bar{m}$ agents finish choosing, all goods are simultaneously exhausted, and the remaining $1-\bar{m}$ agents receive nothing.[7]

Note that, besides SD, various no-screening mechanisms (e.g., the random favorite mechanism discussed in Appendix A.2) yield an equivalent reduced mechanism to SD in a continuous i.i.d. market.

Full Screening/VCG For a reduced market, a *full-screening mechanism* is defined as a mechanism that identifies agents' values to achieve allocative efficiency, even at the cost of screening. In the reduced market, the allocatively efficient allocation rule is to assign the right to obtain the most preferred object with probability 1 to the agents with the highest value for a mass $\bar{m}$ units, while not allocating the good to any other agents; i.e., $x(v) = 1$ if $v \geq q$ and $x(v) = 0$ otherwise, where $q \coloneqq G_K^{-1}(1-\bar{m})$ is the $(1-\bar{m})$-quantile of the reduced value distribution G_K. To implement this allocation rule, the mechanism should require paying a price of q for allocated agents. The residual surplus achieved under VCG is

$$RS(\mathcal{M}_{VCG}) = \bar{m}\mathbb{E}_{v\sim G_K}[v-q|v>q] = \int_q^{\bar{v}} (v-q)\, dG_K(v).$$

The revenue equivalence theorem of Myerson (1981) implies that to achieve the gross surplus $\bar{m}\mathbb{E}[v|v \geq q]$, each allocated agent needs to burn q; thus, the total expected burned payment is $\bar{m}q$.

For general environments, including finite environments and environments with continuous agents and objects but non-i.i.d. values, the *Vickrey-Clarke-Groves mechanism* (VCG), which requires each agent to pay their externality, achieves allocative efficiency. Furthermore, Green and Laffont (1977, 1979); Holmström (1979) show that, under certain conditions, VCG is the unique direct mechanism that achieves allocative efficiency even in environments with multi-dimensional types.

Direct mechanisms such as VCG are not practically used in money-burning problems. However,

[7]Formally, the outcome of SD in a continuous environment is specified by an eating mechanism. While we skip it because this paper does not intensively analyze the general continuous multidimensional environment, interested readers should refer to Noda (2018). The asymptotic equivalence of (random) SD and an eating mechanism is established by Che and Kojima (2010).

when indirect mechanisms are employed, the resulting equilibrium allocations often coincide with or resemble those achieved by direct mechanisms. For example, during the COVID-19 pandemic, FCFS was used to allocate vaccine reservation slots. Under this mechanism, agents who claim after the distribution begins receive, in order of arrival, their most preferred goods among those still available. If we formulate this as a problem where agents freely choose their payment (effort cost) to claim earlier, then (i) agents who pay a higher payment are more likely to obtain goods, and (ii) payments are made regardless of whether the good is obtained. In a reduced single-dimensional market, this incentive structure is equivalent to an all-pay auction. In the reduced environment, an all-pay auction implements the same allocation rule and expected utilities as the VCG by revenue equivalence. Furthermore, whereas FCFS might yield a different outcome from VCG in general environments, VCG, which achieves allocative efficiency, represents the opposite extreme of a no-screening mechanism and serves as a useful benchmark.[8]

3.5 Characterizations of Efficient Mechanisms

Myerson (1981) proves that (i) an allocation rule can comprise a strategy-proof mechanism if and only if x is *monotonic* (i.e., $x(v') \geq x(v)$ for all $v' \geq v$), and (ii) if an allocation is nondecreasing, strategy-proofness uniquely identifies the payment rule (up to constant). Using this approach, Hartline and Roughgarden (2008) show that the expectation of the residual surplus is equal to the expectation of the *virtual valuation for utility* ϑ under strategy-proofness (see their Lemma 2.6):

$$\mathbb{E}[vx(v) - p(v)] = \int_0^{\bar{v}} (vx(v) - p(v))\, dG_K(v) = \int_0^{\bar{v}} \vartheta(v; G_K)x(v)dG_K(v) = \mathbb{E}[\vartheta(v; G_K)x(v)].$$

where

$$\vartheta(v; G_K) \coloneqq \frac{1 - G_K(v)}{g_K(v)}.$$

In the context of revenue maximization, we exploit the relation of $\mathbb{E}[p(v)] = \mathbb{E}[(v - \vartheta(v; G_K))x(v)]$ to characterize an optimal mechanism. The value, $v - \vartheta(v; G_K)$, is defined as the virtual valuation *for payment*. Since this paper's focus is on residual-surplus maximization, we rather refer to $\vartheta(v; G_K)$ as the virtual valuation. Note that $r(v; G_K) = 1/\vartheta(v; G_K) = g_K(v)/(1 - G_K(v))$ is the *hazard rate* function of the distribution G_K.

The problem reduces to the maximization of $\mathbb{E}[\vartheta(v)x(v)]$ subject to the resource constraint, individual rationality, the unit demand condition, and monotonicity of x. To do so, we apply a procedure called "ironing" (Myerson, 1981) to obtain a nondecreasing *ironed virtual value* function $\bar{\vartheta}(v; G_K)$. Then, to derive an efficient mechanism, starting from higher values of v, we greedily

[8] In the finite market studied in Section 4, we assume finitely many agents and objects, unit demand, and quasi-linear utility. This setting coincides with the assignment markets by Koopmans and Beckmann (1957); Shapley and Shubik (1971). In this environment, it is also known from Crawford and Knoer (1981); Demange et al. (1986) that dynamic auctions, which gradually raise the prices of over-demanded goods, can find the allocation that maximizes gross surplus.

increase the allocation probabilities $x(v)$ whereas assigning equal allocation probabilities within intervals where $\bar{\vartheta}(v; G_K)$ is constant, until the resource constraint $\mathbb{E}[x(v)] \leq \bar{m}$ becomes binding.

Below, we establish a primitive condition on the reduced value distribution G_K that is necessary and sufficient for a no-screening mechanism to be efficient. Specifically, we prove that the condition is G_K satisfying the *new better than used in expectation* (NBUE) property. To the best of our knowledge, this paper is the first to characterize the equivalence between NBUE and constant ironed virtual value, thereby establishing conditions for the optimality of no screening.

Definition 3. A distribution function G satisfies the *new better than used in expectation property (NBUE)* if for all $t \in (0, \bar{v})$, we have the following:

$$\mathbb{E}_G[v] \geq \mathbb{E}_G[v - t | v > t]. \tag{6}$$

NBUE, like the hazard rate, originates from the survival analysis literature. Consider the lifetime v of a product, following a distribution G. The left-hand side of (6) represents the mean residual life of a new product $\mathbb{E}[v]$, whereas the right-hand side is the mean residual life conditional on survival up to age t, $\mathbb{E}[v - t|v > t]$. NBUE requires that, for any t, a new product lasts longer in expectation than a used product. Note that NBUE is weaker than the *increasing hazard rate* (IHR) condition, which requires the hazard rate $r(\cdot; G)$ is nondecreasing. This is because IHR requires that for all v, the product becomes increasingly likely to fail at v as time passes, whereas NBUE only requires that the new product lasts longer in expectation.[9]

Theorem 2. The following two statements hold.

(i) If the reduced value distribution G_K is NBUE, then for all $\bar{m} \in (0, 1)$, a no-screening mechanism is efficient.

(ii) If the reduced value distribution G_K is not NBUE, then for all $\bar{m} \in (0, 1)$, a no-screening mechanism is not efficient.

The proof proceeds by carefully tracking the mathematical steps involved in the ironing procedure. It is straightforward to see that NBUE is necessary for no screening to be efficient. Under no screening, allocating one good randomly to an agent yields a residual surplus of $\mathbb{E}[v]$ (which is the left-hand side of (6)). In contrast, if we use a take-it-or-leave-it offer requiring a payment t, agents accept it if and only if their value v is larger than t. By making this offer to $1/\Pr(v > t)$ agents, the planner can allocate the good to one agent on average, generating a residual surplus of $\mathbb{E}[v - t|v > t]$ (which is the right-hand side of (6)). Accordingly, if NBUE is violated, the planner can outperform the no-screening mechanism by (partly) using take-it-or-leave-it offers. Theorem 2

[9] We can easily construct examples of distributions that are NBUE but not IHR, by considering a distribution whose hazard rate is generally increasing as a trend, yet has small intervals where it decreases.

demonstrates not only this necessity but also the sufficiency: when NBUE holds, no mechanism can outperform no screening, for any capacity parameter $\bar{m}$.[10]

Conversely, the following theorem demonstrates that the *decreasing hazard rate* (DHR) property of the reduced value distribution G_K, that is, its hazard rate $r(\cdot; G_K)$ is nonincreasing, is essential for the efficiency of full-screening mechanisms.

Theorem 3. The following two statements hold.

(i) If the reduced value distribution G_K has a DHR, then for all $\bar{m} \in (0, 1)$, a full-screening mechanism is efficient.

(ii) If the reduced value distribution G_K does not have a DHR, then there exists $\bar{m} \in (0, 1)$ such that a full-screening mechanism is not efficient.

As previously described, an efficient mechanism is obtained by increasing $x(v)$ from the highest values of v. If, at the moment the resource constraint $\mathbb{E}[x(v)] \leq \bar{m}$ becomes binding, the current v does not belong to an interval where the ironed virtual value $\bar{\vartheta}(v; G_K)$ is flat, then full screening is efficient. Thus, the absence of such intervals—equivalently, a nondecreasing virtual value—is a sufficient condition for full screening efficiency.[11] This condition corresponds exactly to the distribution G_K satisfying DHR. Conversely, if G_K does not satisfy DHR, there must exist intervals where the ironed virtual value is flat, implying the existence of some $\bar{m}$ for which full screening is inefficient. Note that full screening may still be efficient even when the ironed virtual value is flat in the low-v region. Thus, DHR is not a necessary condition for full screening to be efficient for every $\bar{m}$.

3.6 Effects from Increasing Variety

In this section, we show that as the variety of objects K increases, the reduced value distribution G_K is more likely to satisfy conditions such as NBUE or IHR, which ensure the efficiency of no-screening mechanisms.

First, we can show that NBUE, the necessary and sufficient condition for no screening to be efficient, becomes easier to satisfy as the variety of objects K increases.

Theorem 4. Suppose that G_K satisfies NBUE. Then, G_{K+1} also satisfies NBUE.

The intuition behind Theorem 4, based on the terminology of survival analysis, is as follows. While $G(v)$ represents the probability that a single product fails by time v, its largest order statistic

[10] As their Corollary 2.11, Hartline and Roughgarden (2008) provide IHR as a tractable sufficient condition for no screening to be efficient in a finite market. However, since IHR is stronger than NBUE, IHR is not necessary. Alternatively, while IHR is equivalent to a nonincreasing virtual value, this is only a loose sufficient condition for a nonincreasing (constant) *ironed* virtual value, a property equivalent to NBUE.

[11] For a finite market, Hartline and Roughgarden (2008) proposes this sufficient condition as Corollary 2.12.

$G_K(v)$ represents the probability that all K products fail by time v. When calculating the mean residual life of a product, we are guaranteed that all K products are initially functioning. In contrast, when calculating the mean residual life after t periods of use, $\mathbb{E}[v|v > t]$, the condition $v > t$ merely indicates that the largest order statistic $v = \max\{v_1, \ldots, v_K\}$ exceeds t, ensuring only that at least one product remains functioning at time t. Even when conditioning on the event "at least one product survives," the number of surviving products tends to fall as time t increases. Consequently, when the variety of goods K becomes larger, the mean residual life increases substantially for $t = 0$, but this effect is smaller for large t. This makes the NBUE condition more likely to hold when we have larger K.

The following result is immediate from Theorems 2 and 4.

Corollary 1. Suppose that a no-screening mechanism is efficient in a continuous i.i.d. market $(G, K, \bar{m})$. Then, a no-screening mechanism is also efficient in a continuous i.i.d. market $(G, K + 1, \bar{m})$.

Second, we can also show that as the variety of goods K increases, the region in which the reduced value distribution has an IHR expands.

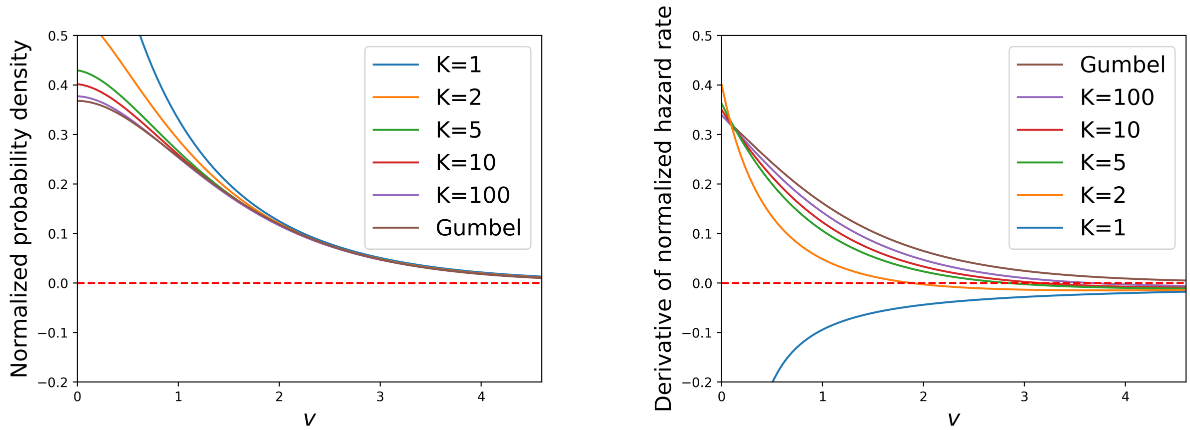
Theorem 5. Let $r'(\cdot; G_K)$ be the derivative of $r(\cdot; G_K)$ with respect to v. Then, $r'(v; G_{K+1}) \geq 0$ if $r'(v; G_K) \geq 0$. Furthermore, for all $v \in [0, \bar{v}]$, there exists K_0 such that for all $K > K_0$, $r'(v; G_K) > 0$.

Again, we describe the intuition using the terminology of survival analysis. The hazard rate of G_K indicates "the probability density of all K products failing exactly at time v, conditional on at least one product surviving until then." For all K products to fail at time v, exactly one product must survive until just before v, with the remaining $K - 1$ products having already failed. This condition is rarely satisfied with $K > 1$ when v is small but becomes increasingly likely as v grows larger. Therefore, increasing the number of products K significantly reduces the hazard rate at small v, while having only a limited effect at large v. As a result, a larger K makes G_K more likely to satisfy the IHR condition.

Theorem 5 implies that, as the variety of goods increases, DHR becomes harder to satisfy, making full screening less likely to be an efficient mechanism.

Corollary 2. For every marginal distribution function G, there exists K and $\bar{m}$ such that full screening is inefficient for a continuous i.i.d. market $(G, K, \bar{m})$.

Remark 1. These analyses, based on the properties of distributions for largest order statistics, strictly apply only to the continuous i.i.d. market, where agents have no chance of obtaining goods other than their top choices. However, even when this strong assumption does not hold, if the variety of goods is large and agent preferences are diverse, each agent typically receives one of their highly preferred goods. In such cases, the distributions of these relevant order statistics



(a) Normalized probability density (b) Derivative of the normalized hazard rate

Figure 2: The behavior of the distribution of the normalized largest order statistic $(a_K v + b_K)$ when the value distribution is the Weibull distribution with shape parameter 0.9. The sequence (a_K, b_K) is shown in Appendix A.1.

should closely resemble that of the largest order statistic. Indeed, numerical simulations show qualitatively similar results even in finite markets (see Section 4).

The effect of increasing the variety of goods K on making G_K satisfy NBUE or IHR is substantial. Figure 2 illustrates the case of a Weibull distribution with parameter 0.9, which is globally DHR. As shown in Panel 2b, the derivative of the hazard rate is negative everywhere when $K = 1$, but becomes positive over a large region even with $K = 2$. At this point, a no-screening mechanism already outperforms a full-screening mechanism. As K increases further, this region continues to expand, and in the limit $K \to \infty$, the distribution of the largest order statistic G_K converges to the Gumbel distribution, which is globally IHR.

3.7 Extreme Value Theory

We present the large-market results based on the asymptotic properties of the reduced value distribution G_K in the limit of $K \to \infty$. The limiting behavior of the largest order statistic has been extensively analyzed in the literature of extreme value theory. We utilize these results to characterize efficient mechanisms in the large-market limit.

Definition 4. Let G, H be distribution functions. We say that H is an *extreme value distribution* and that G lies in its *domain of attraction* of H if G and H are nondegenerate and there exist a sequence of constants $(a_K, b_K)_{K=1}^{\infty}$ with $a_K > 0$ for all K, such that $\hat{G}_K(w) \coloneqq G_K(a_K w + b_K) \to H(w)$ as $K \to \infty$ for all w at which H is continuous.

For notational simplicity, we define a normalized reduced value distribution $\hat{G}_K$ and allocation rule $\hat{x}$ as $\hat{G}_K(w) \coloneqq G_K(a_K w + b_K)$ and $\hat{x}(w) \coloneqq x(a_K w + b_K)$. Since $a_K > 0$, the residual surplus maximization for $\hat{G}_K$ is equivalent to that for G_K.

Table 1: Extreme Value Distributions of Common Probability Distributions

Limit	Marginal Value Distribution
Gumbel	Exponential, Normal, Gamma, Log-Normal, Logistic, Weibull, Gumbel, etc.
Fréchet	Pareto, Cauchy, t, F, Burr, Log-Gamma, Zipfian, Fréchet, etc.
Reverse-Weibull	Uniform, Beta, Reverse-Weibull, etc. (Unimportant for our analysis)

When the number of object types K becomes sufficiently large, the shape of the reduced value distribution G_K becomes less dependent on the shape of the marginal value distribution, G. Consequently, the limit distribution can only be either Gumbel, Fréchet, or reverse-Weibull distribution.

Proposition 2 (Fisher and Tippett (1928); Gnedenko (1943))**.** Only the Gumbel, Fréchet, and reverse-Weibull distributions,

$$\Lambda(w) = \exp(-\exp(-w)), \quad -\infty < w < \infty \tag{Gumbel}$$

$$\Phi_\alpha(w) = \exp(-w^{-\alpha}), \quad w \geq 0, \alpha > 0, \tag{Fréchet}$$

$$\Psi_\alpha(w) = \begin{cases} \exp(-(-w)^\alpha) & w < 0 \\ 1 & w \geq 0, \end{cases} \quad \alpha > 0 \tag{reverse-Weibull}$$

can be an extreme-value distribution.

For the proof of Proposition 2, see Theorems 1.1.3 and 1.1.6 of De Haan and Ferreira (2006).[12] Theorem 1.1.6 of De Haan and Ferreira (2006) additionally implies that, whenever a distribution function G belongs to the domain of attraction of Gumbel, $\hat{G}_K(w) \coloneqq G_K(a_K w + b_K) \to \Lambda(w)$ with

$$a_K = s(G^{-1}(1 - 1/K)),\ b_K = G^{-1}(1 - 1/K),\ \text{where } s(v) = \frac{\int_v^{\bar{v}} (1 - G(v'))dv'}{1 - G(v)}. \tag{7}$$

Furthermore, whenever a distribution function G belongs to the domain of attraction of Fréchet with shape parameter α, $\hat{G}_K(w) \coloneqq G_K(a_K w + b_K) \to \Phi_\alpha(w)$ with

$$a_K = G^{-1}(1 - 1/K),\ b_K = 0 \tag{8}$$

Throughout the paper, when we refer to $(a_K, b_K)_{K=1}^\infty$, these constants are defined by (7) or (8).

We omit the condition for reverse-Weibull because the convergence occurs only if $\bar{v} < +\infty$ and for such a case, we do not need extreme value theory for characterizing an efficient mechanism in the limit (see Theorem 6).

Proposition 2 does not state that the distribution of the normalized maximum converges.[13]

[12] Whereas De Haan and Ferreira (2006) employ the representation of generalized extreme value distribution, as they introduce in pp.9-10, it is equivalent to Proposition 2.

[13] For example, the binomial, Poisson, and geometric distributions do not have an extreme distribution.

Nevertheless, it is also known that many continuous distribution functions have an extreme value distribution.[14] Table 1 lists extreme value distributions of common probability distributions. The limit is identical as long as distributions have the same right-tail property, even when their overall shape differs significantly.

The extreme value theory literature has also characterized the conditions under which a marginal value distribution G belongs to the domain of attraction of each extreme value distribution. However, this paper focuses not on the convergence of the distribution function G, but on the derivative of the hazard rate, i.e.,

$$r'(w;\hat{G}_K) = \frac{\hat{g}'_K(w)(1-\hat{G}_K(w)) + (\hat{g}_K(w))^2}{\left(1-\hat{G}_K(w)\right)^2}.$$

For convergence of $r'(\cdot;\hat{G}_K)$, we also need a convergence of $\hat{g}_K$ and $\hat{g}'_K$, which are the first and second derivatives of $\hat{G}_K$. The following is the necessary and sufficient condition for it.

Definition 5. Let G and H be distribution functions. We say that G lies in the *twice differentiable domain of attraction* of H if G and H are nondegenerate and twice differentiable for all sufficiently large w and $\hat{G}_K^{(l)}(w) \to H^{(l)}(w)$ as $K \to \infty$ uniformly for all w in any finite interval for all $l = 0, 1, 2$.

Definition 6 (von Mises condition)**.** A distribution function G satisfies the *von Mises condition with parameter* γ if G is twice differentiable for all sufficiently large v and

$$\vartheta'(v;G) = \left(\frac{1-G(v)}{g(v)}\right)' \to \gamma \text{ as } v \to \infty. \tag{9}$$

We denote the set of all distribution functions satisfying (9) by $\mathbf{VMC}(\gamma)$.

The von Mises condition (9) is originally established by von Mises (1936) as a tractable sufficient condition for convergence of a distribution function. Later, Pickands III (1986) proves that (9) is a necessary and sufficient condition for convergence of the second derivative.

Proposition 3 (Theorem 5.2 of Pickands III (1986))**.**

(i) G is in the twice differentiable domain of attraction of the Gumbel distribution Λ if and only if $G \in \mathbf{VMC}(0)$.

(ii) G is in the twice differentiable domain of attraction of the Fréchet distribution with shape parameter α, Φ_α, if and only if $G \in \mathbf{VMC}(1/\alpha)$.

3.8 Large Market Limit

We characterize efficient mechanisms in a large market limit of the continuous i.i.d. market.

[14] Pickands III (1975) mentions that *"Most 'textbook' continuous families of distribution functions are discussed in Gumbel (1958). They all lie in the domain of attraction of some extremal distribution function."* (page 119)

Bounded Support We start with the case of bounded support, i.e., $\bar{v} < +\infty$. In this case, no screening is efficient with sufficiently large K, while full screening is conversely the worst. When K is large, almost every agent has at least one good that they value close to $\bar{v}$. Therefore, even with no screening, the gross surplus approaches $\bar{v}$, and there is little need to screen the intensity of preferences. On the other hand, if screening is conducted, almost all the surplus is burned due to competition.

Theorem 6. Suppose that $\bar{v} < +\infty$. Then, for any $\epsilon > 0$ and $\delta > 0$, there exists $K_0 > 0$ such that for all $K > K_0$, we have both $RS(\mathcal{M}_{SD}) > \bar{m}(1-\delta)(\bar{v}-\epsilon)$ and $RS(\mathcal{M}_{VCG}) < \bar{m}\epsilon$.

Gumbel Case Next, we consider the case where the marginal value distribution G satisfies $\mathbf{VMC}(0)$, and thus belongs to the twice differentiable domain of attraction of the Gumbel distribution Λ. The hazard rate function of Gumbel is given by

$$r(w;\Lambda) = \frac{\lambda(w)}{1-\Lambda(w)} = \frac{\exp(-w)}{\exp(\exp(-w))-1},$$

and thus the Gumbel distribution Λ has an IHR. Accordingly, if the reduced value distribution is exactly Gumbel, then an efficient mechanism is a no-screening mechanism.

Proposition 3 ensures that whenever $G \in \mathbf{VMC}(0)$, $\hat{G}_K(w)$ converges to $\Lambda(w)$ up to the second derivative, uniformly in any finite interval of w. Accordingly, for any finite interval $[\underline{c}, \bar{c}]$, the hazard rate becomes increasing in $[\underline{c}, \bar{c}]$ with sufficiently large K (see Lemma 1 in the Appendix). Furthermore, by the standard arguments in optimal mechanism design, the efficient allocation rule is guaranteed to be constant over the interval where the hazard rate is nondecreasing. (For clarity, we formally present and prove this result as Lemma 2 in the Appendix.)

Accordingly, for any interval $[\underline{c}, \bar{c}]$, there exists K_0 such that for all $K > K_0$, an efficient allocation rule becomes constant in the interval: $\hat{x}(w) = \bar{x}$ for $w \in [\underline{c}, \bar{c}]$. Furthermore, as $\hat{G}(\bar{c}) - \hat{G}(\underline{c}) \to \Lambda(\bar{c}) - \Lambda(\underline{c})$, by taking small $\underline{c}$, large $\bar{c}$, and large K, we can make $\hat{G}_K(\bar{c}) - \hat{G}_K(\underline{c})$ arbitrarily close to one. Since (i) the allocation is flat on the interval $[\underline{c}, \bar{c}]$, (ii) this interval can be taken arbitrarily large to carry probability mass arbitrarily close to one, and (iii) the allocation must be flat within the interval, the resource constraint forces this constant allocation level to be arbitrarily close to $\bar{m}$, with any residual allocation absorbed by the negligible outside mass. In this sense, an efficient mechanism conducts no screening asymptotically.

Theorem 7. If $G \in \mathbf{VMC}(0)$, then for all $\epsilon > 0$, there exists K_0 such that for all $K > K_0$, there exists a constant $\bar{x}_K \in [0,1]$ such that an efficient allocation rule $\hat{x}$ for (G, K) satisfies $\Pr(\hat{x}(w) = \bar{x}_K) > 1-\epsilon$. Furthermore, $\bar{x}_K \in ((\bar{m}-\epsilon)/(1-\epsilon), \bar{m}/(1-\epsilon))$.

Fréchet Case Finally, we consider the case where the marginal value distribution G satisfies $\mathbf{VMC}(1/\alpha)$ for some $\alpha > 0$. When $\alpha \in (0,1]$, the right tail of Fréchet, Φ_α, is extremely heavy, and the distribution does not have a mean; thus, we focus on the case of $\alpha > 1$, where agents' expected payoffs are well-defined. The hazard rate function of Fréchet with parameter α is given by

$$r(w;\Phi_\alpha) = \frac{\phi_\alpha(w)}{1-\Phi_\alpha(w)} = \frac{\alpha w^{-\alpha-1}\exp(-w^{-\alpha})}{1-\exp(-w^{-\alpha})}.$$

We say that a distribution has an *increasing and decreasing hazard rate (IDHR)* if there exists w^* such that the hazard rate is increasing for $w < w^*$ and decreasing for $w > w^*$. We can verify that Fréchet has an IDHR.[15]

As the reduced value distribution has neither IHR nor DHR, we apply Myerson's (1981) ironing to obtain an efficient mechanism. We can show that there exists a threshold $w^{**} > w^*$ such that the efficient mechanism fully screens agents' preferences for $w > w^{**}$, whereas it conducts no screening for $w < w^{**}$.[16] If the supply $\bar{m}$ is insufficient to allocate all agents with $w > w^{**}$, then only those with the highest values will win an item. Otherwise, an efficient mechanism first allocates goods with probability one to agents with values $w > w^{**}$. The remaining goods are then distributed equally among the other agents without payment.

Proposition 4. When the reduced value distribution function $\hat{G}_K$ has an IDHR, there exists $w^{**} > w^*$ with which an efficient allocation rule $\hat{x}$ is represented as follows:

(i) If $1-\hat{G}_K(w^{**}) \geq \bar{m}$, then

$$\hat{x}(w) = \begin{cases} 1 & \text{for } w \in [\hat{G}_K^{-1}(1-\bar{m}), +\infty), \\ 0 & \text{for } w \in [0, \hat{G}_K^{-1}(1-\bar{m})). \end{cases}$$

(ii) If $1-\hat{G}_K(w^{**}) < \bar{m}$, then

$$\hat{x}(w) = \begin{cases} 1 & \text{for } w \in [w^{**}, +\infty), \\ \dfrac{\bar{m}-(1-\hat{G}_K(w^{**}))}{\hat{G}_K(w^{**})} & \text{for } w \in [0, w^{**}). \end{cases}$$

Figure 3a shows the values of $\Phi_\alpha(w^*)$ and $\Phi_\alpha(w^{**})$ for various shape parameters α.[17] As α becomes larger, the right tail of Fréchet decays, and in the limit of $\alpha \to \infty$, the distribution

[15]For Fréchet with parameter α, w^* is the unique solution of

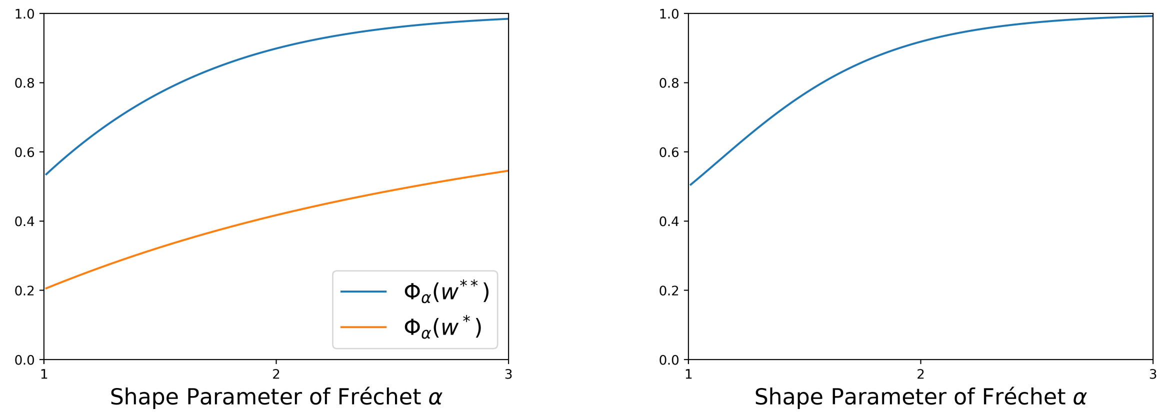
$$\frac{w^{-\alpha}}{1-\exp(-w^{-\alpha})} = \frac{\alpha+1}{\alpha}$$

with $w > 0$.

[16]The equation (28) in Appendix B characterizes w^{**}.

[17]For Fréchet with parameter α, w^{**} is a solution of

$$(1-\exp(-w^{-\alpha}))w + \Gamma\left(\frac{\alpha-1}{\alpha}, w^{-\alpha}\right) - \frac{1}{\alpha}w^{\alpha+1}(1-\exp(-w^{-\alpha})) = 0, \tag{10}$$



(a) Region in which the extreme value distribution has an IHR ($\Phi_\alpha(w^*)$) and the efficient mechanism uses a constant allocation ($\Phi_\alpha(w^{**})$)

(b) Performance of the no screening mechanism relative to the efficient one

Figure 3: How the no-screening mechanism is close to the efficient mechanism in the large market limit under the Frećhet case

converges to Gumbel. Accordingly, as α increases, both $\Phi_\alpha(w^*)$ and $\Phi_\alpha(w^{**})$ increase and converge to one. The convergence of $\Phi_\alpha(w^*)$ is relatively slow. By contrast, the fraction of agents with $w > w^{**}$ diminishes rapidly even with relatively small α, implying that under Fréchet with relatively small α, an efficient mechanism conducts no screening for most agents. For example, when $\alpha = 3$ and $\bar{m} = 0.5$, an efficient mechanism requires a substantial amount of money burning for obtaining a good with probability 1, which only agents with the highest 1.5% of values will choose. The remaining units, which can cover 48.5% of the population, are distributed to the rest of 98.5% of agents with no screening, just as SD would do. In addition, Figure 3b illustrates that the performance of SD rapidly approaches that of the efficient mechanism as the shape parameter α increases.

Finally, we characterize the efficient mechanism when $G \in \mathbf{VMC}(1/\alpha)$ and K is large but finite. The condition $G \in \mathbf{VMC}(1/\alpha)$ implies that for any finite interval $[\underline{c}, \bar{c}]$, if K is sufficiently large, $\hat{G}_K$ is close to Fréchet Φ_α within that interval. By choosing $\underline{c}$ close to zero and $\bar{c}$ larger than w^{**}, and accordingly taking a large K, the efficient allocation rule $\hat{x}$ for $\hat{G}_K$ becomes constant over most of the interval $[0, w^{**}]$.

Theorem 8. If $G \in \mathbf{VMC}(1/\alpha)$, then for all $\epsilon > 0$, there exists $K_0 \in \mathbb{Z}_{++}$ such that for all $K > K_0$, there exists a constant $\bar{x}_K \in [0,1]$ such that an efficient allocation rule $\hat{x}$ for (G, K) satisfies $\Pr(\hat{x}(w) = \bar{x}_K) \in (\Phi_\alpha(w^{**} - \epsilon), \Phi_\alpha(w^{**} + \epsilon))$, where w^{**} is a solution of the equation (10).

where $\Gamma(s,x) = \int_x^\infty t^{s-1} e^{-t} dt$ is the upper incomplete gamma function. See Appendix B.10 for the derivation of (10).

3.9 Finite Variety (with Continuum Copies)

Figure 1b in the Introduction illustrates how residual surplus changes with the number of object types K under SD (or no screening), VCG (or full screening), and the efficient mechanism. This visualization is based on a continuous i.i.d. market, where the total mass of objects $\bar{m}$ is $1/2$, and the marginal value distribution G is Weibull with parameter 0.6. The performance of each mechanism is derived analytically using reduction to a single-dimensional environment.

The marginal value distribution has a DHR; thus, when $K = 1$, i.e., the single-good case, full screening is efficient. However, since the distribution also belongs to $\mathbf{VMC}(0)$, in the limit of $K \to \infty$, the no-screening mechanism becomes asymptotically efficient. For intermediate, finite K, the no-screening mechanism outperforms full screening for $K \geq 4$ under this setting. Furthermore, beyond this point, the performance of the no-screening mechanism rapidly converges to that of the efficient mechanism. This observation suggests that the key policy implications of our study, the superiority of SD in multi-object allocation without money, likely hold for practical settings.

4 Finite Market

This section considers finite markets, where a finite number of objects are allocated to a finite number of agents (i.e., no continuum copies of agents and objects). The analysis of finite markets is widely believed to be considerably more challenging than analyzing continuous markets. In this study, instead of characterizing efficient mechanisms analytically, we adopted the deep learning-based approach developed by Dütting et al. (2019), originally for numerically constructing revenue-maximizing auction mechanisms.

4.1 Model

We consider a finite market with a finite set of agents I and a finite set of object types K. Slightly abusing notation, I and K also represent the cardinality of those sets. The endowment of each object type k is $m_k \in \mathbb{Z}_{++}$. We assume that the goods are scarce: $I > \sum_{k \in K} m_k$. Agents' utilities are identical to that of the continuous market: When agent i obtains object k with probability x_k^i for each k while making a payment of p^i, her payoff is $\sum_{k \in K} v_k^i x_k^i - p^i$.

Let $V^i \subset \mathbb{R}^K$ be the set of all possible valuation vectors for agent i, and $\mathbf{v}^i$ be its generic element. We define the set of all valuation profiles as $V := \prod_{i \in I} V^i$, and $\mathbf{v} = (\mathbf{v}^i)_{i \in I}$ as its generic element. Here, we abandon equal treatment of equals and a direct mechanism $\mathcal{M}$ is comprised of an allocation rule $\mathbf{x} : V \to [0,1]^{I \times K}$ and a payment rule $\mathbf{p} : V \to \mathbb{R}_+^I$, where $\mathbf{x}^i(\mathbf{v})$ and $p^i(\mathbf{v})$ are the respective allocation and payment for agent i under valuation profile $\mathbf{v}$. Let F be the probability distribution that the valuation profile $\mathbf{v}$ follows.

The (per capita) *residual surplus* from a mechanism $\mathcal{M} = (\mathbf{x}, \mathbf{p})$ is

$$RS(\mathcal{M}) := \int_V \frac{1}{I} \sum_{i \in I} \left(\sum_{k \in K} v_k^i x_k^i(\mathbf{v}) - p^i(\mathbf{v}) \right) dF(\mathbf{v}). \tag{11}$$

A mechanism $\mathcal{M}$ satisfies the *resource constraint* if

$$\sum_{i \in I} x_k^i(\mathbf{v}) \leq m_k \text{ for all } k \in K, \mathbf{v} \in V. \tag{12}$$

A mechanism $\mathcal{M}$ is *strategy-proof* if

$$\sum_{k \in K} v_k^i x_k^i(\mathbf{v}^i, \mathbf{v}^{-i}) - p^i(\mathbf{v}^i, \mathbf{v}^{-i}) \geq \sum_{k \in K} v_k^i x_k^i(\hat{\mathbf{v}}^i, \mathbf{v}^{-i}) - p^i(\hat{\mathbf{v}}^i, \mathbf{v}^{-i}) \text{ for all } i \in I, \mathbf{v}^i, \hat{\mathbf{v}}^i \in V^i, \mathbf{v}^{-i} \in V^{-i}. \tag{13}$$

A mechanism $\mathcal{M}$ is *individually rational* if

$$\sum_{k \in K} v_k^i x_k^i(\mathbf{v}) - p^i(\mathbf{v}) \geq 0 \text{ for all } i \in I, \mathbf{v} \in V. \tag{14}$$

A mechanism $\mathcal{M}$ satisfies the *unit demand* condition if

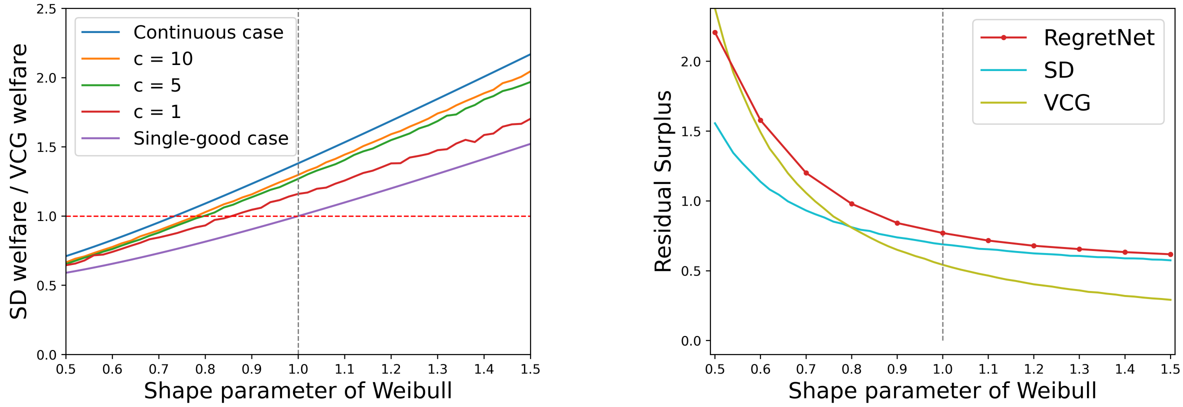
$$\sum_{k \in K} x_k^i(\mathbf{v}) \leq 1 \text{for all } i \in I, \mathbf{v} \in V. \tag{15}$$

A mechanism $\mathcal{M}$ is *efficient* if it maximizes $RS(\mathcal{M})$ subject to (12), (13), (14), and (15).

4.2 RegretNet

To numerically derive an efficient mechanism for finite markets, we adopt *RegretNet*, a deep learning-based method for designing revenue-maximizing auction mechanisms proposed by Dütting et al. (2019). In RegretNet, a mechanism, which is a function that takes a valuation profile as input and returns an allocation and a payment profile, is represented by a neural network, which can approximate a wide range of functions using a relatively small number of parameters. In Dütting et al. (2019), the loss function used for learning consists of (i) the negated empirical expected revenue of the seller and (ii) a penalty term for violating strategy-proofness. RegretNet uses the augmented Lagrange method to find a mechanism (i.e., the parameters) minimizing such loss.

Residual-surplus maximization and revenue maximization have different objective functions but share the same constraints. Therefore, we can derive an efficient mechanism by replacing the objective function without making other modifications.



(a) The ratio of the residual surplus achieved by SD and VCG, $(RS(\mathcal{M}_{SD})/RS(\mathcal{M}_{VCG}))$

(b) The performance of SD, VCG, and the mechanism developed by RegretNet for $c = 5$

Figure 4: Comparison of the performance of SD, VCG, and the numerically efficient mechanism in a finite market

4.3 I.I.D. Case

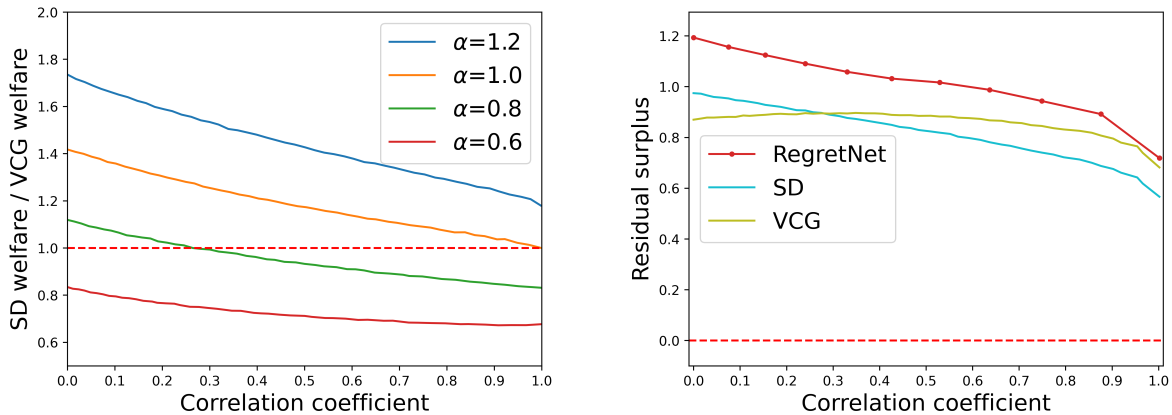
We examine the extent to which the continuous-market model well approximates a finite market. We first consider a market with $4c$ agents and two object types ($K = 2$) each with $c \in \mathbb{Z}_{++}$ capacities. Each agent i's value for object k, v_k^i follows Weibull, i.i.d.

In Figure 4a, we generate 10,000 samples and compare the ratio of the residual surplus achieved by SD and VCG, i.e., $RS(\mathcal{M}_{SD})/RS(\mathcal{M}_{VCG})$ for $c = 1, 5, 10$. We also plot (i) the ratio in a continuous market with a unit mass of agents and two objects with $m_k = 0.25$ unit of endowments (i.e., $\bar{m} = 0.5$), which corresponds to a finite market in the limit of $c \to \infty$, and (ii) the ratio in a finite market with two agents and one object, which corresponds to the single-good case.

As characterized by Hartline and Roughgarden (2008), SD outperforms VCG if and only if $\alpha > 1$, for the single-good case. In the continuous case, SD achieves larger social welfare than VCG as long as the shape parameter is larger than roughly 0.7, given $K = 2$. We observe a similar pattern for finite markets: Even with $c = 1$, the threshold at which SD starts to outperform VCG is strictly smaller than 1.0, and as c increases, the market becomes closer to the continuous case.

Figure 4b shows the residual surplus of SD, VCG, and the numerically efficient mechanism returned by RegretNet for the case of $c = 5$. As we consider the case of $K = 2$, where the variety of goods is minimal, RegretNet strictly outperforms both SD and VCG under some parameters, particularly around $\alpha = 0.8$, where SD surpasses the performance of VCG. However, based on the analysis of the continuous market, this gap is expected to diminish as K and c increase.[18]

[18] In Appendix A.3, we illustrate the mechanism returned by RegretNet for the case of $c = 1$ and $\alpha = 0.7$ and $\alpha = 0.8$. Furthermore, in Appendix A.4, we present figures corresponding to Figure 4b for $c = 1$, single-good, and continuous cases.



(a) The performance ratio between SD and VCG: $(RS(\mathcal{M}_{SD})/RS(\mathcal{M}_{VCG}))$

(b) The performance of SD, VCG, and RegretNet given $\alpha = 0.8$

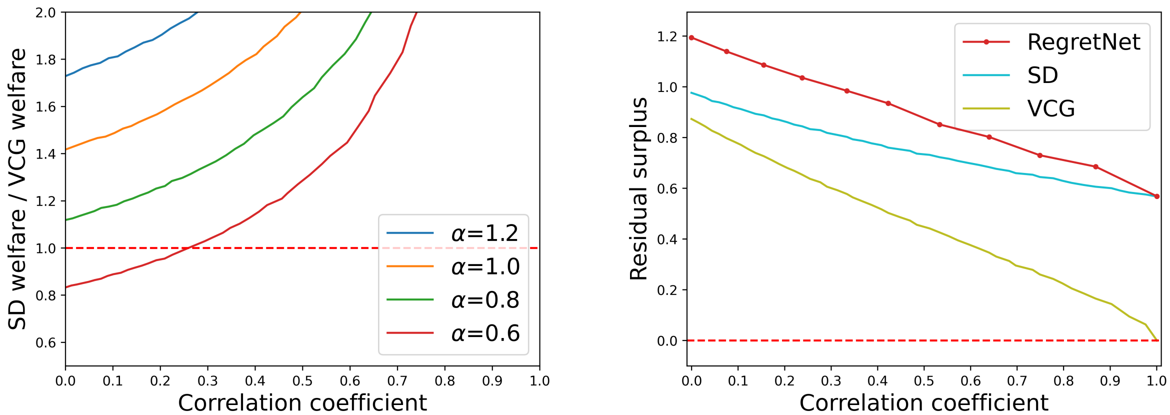
Figure 5: The performance of mechanisms under within-agent correlation. The horizontal axis represents $\mathrm{Corr}(v_k^i, v_l^i)$.

4.4 Correlation

Next, we relax the i.i.d. assumption and investigate cases where values are correlated. Throughout this section, we consider an environment with 8 agents and 4 objects. We employ a *Gaussian copula*, constructed by applying a probability integral transform to a multivariate normal distribution. This allows us to generate multiple joint distributions F of $\mathbf{v} = (v_k^i)_{k \in K, i \in I}$ whose marginal distribution of each v_k^i are identical while the correlation between v_k^i and v_l^j for $(i,k) \neq (j,l)$ varies. Using these joint distributions, we analyze how the performance of SD, VCG, and RegretNet changes as the strength of within-agent correlation and between-agent correlation varies. The detailed method is described in Appendix A.5.

Within-Agent Correlation We first evaluate the effect of the *within-agent correlation*, the correlation between v_k^i and v_l^i for two distinct objects $k \neq l$ (i.e., some agents need any object more strongly than other agents). Using the Gaussian copula, we generate joint distribution F whose marginal distribution is Weibull while the correlation coefficients between each pair of v_k^i and v_l^i, $\mathrm{Corr}(v_k^i, v_l^i)$, vary. For $i \neq j$, $\mathrm{Corr}(v_k^i, v_l^j) = 0$ for all $k, l \in K$. By construction, when $\mathrm{Corr}(v_k^i, v_l^i) = 0$, values are drawn i.i.d., and when $\mathrm{Corr}(v_k^i, v_l^i) = 1$, $v_k^i = v_l^i$ holds for all $k, l \in K$ with probability one, i.e., goods are perfectly homogeneous.

Figure 5a shows the performance ratio of SD and VCG, whereas Figure 5b shows the residual surplus achieved by SD, VCG, and RegretNet for the case of $\alpha = 0.8$. With every parameter α, as the within-agent correlation becomes stronger, the relative performance of VCG improves. This suggests that within-agent correlation attenuates the performance improvement of the no-screening mechanism as the variety of goods increases. When the within-agent correlation is perfect, the prob-



(a) The performance ratio between SD and VCG: $(RS(\mathcal{M}_{SD})/RS(\mathcal{M}_{VCG}))$

(b) The performance of SD, VCG, and RegretNet given $\alpha = 0.8$

Figure 6: The performance of mechanisms under between-agent correlation. The horizontal axis represents $\mathrm{Corr}(v_k^i, v_k^j)$.

lem essentially has single-dimensional types. When the within-agent correlation is intermediate, the results are also intermediate between the single-dimensional type and the multi-dimensional i.i.d. types. However, it is important to note that even with within-agent correlation, the multiplicity of objects improves the performance of SD relative to VCG.

Between-Agent Correlation Next, we evaluate the effect of the *between-agent correlation*, the correlation between v_k^i and v_k^j for two distinct agents $i \neq j$ (i.e., some objects are more popular than other objects). When such a correlation becomes stronger, the environment approaches a common value scenario, where goods with high common value generate high value regardless of who receives them. In such cases, the performance of SD improves, and SD becomes the first-best in the case of pure common values. In contrast, adopting a mechanism that involves screening, such as VCG, would lead to intense competition for popular goods. Consequently, the common value would be entirely lost due to money burning. The following theorem formalizes this observation.

Proposition 5. Consider a profile of marginal distributions $(F_k)_{k \in K}$, where $\mathbb{E}_{v \sim F_k}[v] = \mu_k > 0$ and $\mathrm{Var}_{v \sim F_k}(v) = \sigma_k^2 < +\infty$. Consider a sequence of joint distribution functions F^ρ such that for all ρ, (i) for every agent $i \in I$ and object $k \in K$, the marginal distribution of v_k^i is F_k, and (ii) for every distinct agents $i, j \in I$ and object $k \in K$, $\mathrm{Corr}(v_k^i, v_k^j) \geq \rho$. Then, as $\rho \to 1$, we have (i) $RS(\mathcal{M}_{SD}; F^\rho)/RS(\mathcal{M}_{FB}^\rho; F^\rho) \to 1$, and (ii) $RS(\mathcal{M}_{VCG}; F^\rho) \to 0$, where $\mathcal{M}_{FB}^\rho$ is a mechanism that maximizes residual surplus (11) subject to the resource constraint (12) and the unit demand condition (15).

Figure 6a shows the performance ratio of SD and VCG, whereas Figure 6b shows the residual surplus achieved by SD, VCG, and RegretNet for the case of $\alpha = 0.8$. The simulation setting parallels that for the within-agent correlation cases, whereas we vary $\mathrm{Corr}(v_k^i, v_k^j)$ instead of $\mathrm{Corr}(v_k^i, v_l^i)$. We set $\mathrm{Corr}(v_k^i, v_l^j) = 0$ for all $k \neq l$. When $\mathrm{Corr}(v_k^i, v_k^j) = 0$, values are drawn i.i.d., and when $\mathrm{Corr}(v_k^i, v_k^j) = 1$, we have pure common values. With every parameter α, as the between-agent correlation becomes stronger, the SD and VCG performance ratios increase, indicating that the relative performance of SD improves. Furthermore, in the case of perfect correlation, the residual surplus under VCG becomes zero.

5 Application: Pandemic Vaccination

5.1 Problem

The COVID-19 pandemic significantly disrupted global economic and social welfare, prompting rapid vaccine development and deployment. By the end of 2021, around 9.14 billion COVID-19 vaccinations had been administered worldwide (Mathieu et al., 2021), marking one of history's largest allocation problems.

COVID-19 vaccines developed during the pandemic were clinically confirmed to be effective in preventing severe illness and were anticipated to reduce infections and limit transmissions. Efficient vaccine allocation was crucial for safeguarding healthcare workers and reducing hospital burdens. However, vaccine production occurred gradually, while initial demand was immediately high, creating early vaccine scarcity. Although vaccines are mostly homogeneous, the reservation slots, defined by "when, where, and which vaccine," are heterogeneous, and people have diverse preferences; thus, it is precisely the type of large-scale multi-object allocation problem studied in this paper. Managing these reservations effectively is inherently a market design challenge.

Rapid vaccine development outpaced mechanism design analyses, resulting in logistical inefficiencies and public frustration. Future pandemics remain likely due to globalization, yet improved vaccine technologies suggest rapid development will continue. Therefore, proactively researching optimal mass vaccination logistics, including mechanisms with or without screening, is essential for future preparedness.

5.2 Lessons from the COVID-19 Pandemic

During the early rollout of COVID-19 vaccines, many countries and regions relied on FCFS. These systems turned every available slot into a scarce heterogeneous good and induced high-intensity competition among potential recipients. In early 2021, several counties in Florida opened public vaccination sites with no reservation system. In Lee County, people paid substantial effort costs (such as camping overnight and queuing for hours) to secure doses until officials acknowledged the chaos and announced a shift to a reservation system (Fox 4 News, 2021). Even if a reservation system is in place, similar confusion arises when reservations are accepted on an FCFS basis. In other Florida counties that initially used an FCFS-based reservation system, as our theory predicts, applicants rushed the system immediately after reservations opened and all slots were exhausted within minutes, which in turn led to overwhelmed call centers. Consequently, several local governments quickly abandoned FCFS (Coleman, 2021). Similar problems arose in Massachusetts. When the state opened eligibility to older adults, the system experienced delays and outages due to surges in access, leaving many residents unable to make reservations and frustrated (NBC Boston, 2021). In Washington, D.C., when an online portal was launched for newly eligible residents, it immediately crashed under heavy traffic, and phone lines were jammed. The incident led to public apologies from officials and a review of the system (Barthel et al., 2021). Reservation surges, along with system and call center failures and the resulting public dissatisfaction, were widely observed across many regions in Japan as well (e.g., Asahi Shimbun, 2021a,b; Yomiuri Shimbun, 2021).

In environments with many heterogeneous reservation slots, our theory implies that such screening-induced races are particularly wasteful, which motivates replacing FCFS with no-screening mechanisms, and several local governments indeed introduced alternative mechanisms. Minnesota introduced a random lottery for seniors over 65. Residents had a 24-hour window to preregister, and names were drawn at random; those selected received notice that directed them to complete a booking, while the others remained in a queue for future lotteries (Olson, 2021). Local press reported that the lottery was perceived as fairer and less stressful than the prior FCFS system (Stolle, 2021). San Luis Obispo County in California, after facing similar FCFS failures, announced plans to introduce a lottery (Pitcher, 2021). A lottery-based mechanism was also introduced in Kakogawa City, Japan, as a direct-mechanism version of random serial dictatorship (Tada, 2021). There, the high communication cost inherent to direct mechanisms—stemming from the requirement to submit a full preference order—was one reason for later reverting to FCFS.

British Columbia, Canada, introduced a "Get Vaccinated System." In this reservation system, residents registered once, and the government later sent invitations to book an appointment as supply and age eligibility permitted (Government of British Columbia, 2021). In Singapore, the government created an online pre-registration form and generated a list of prospective recipients; when vaccine slots became available, those on the list received SMS invitations with booking links (Liew, 2022). These can be seen as real-world cases where the authorities used a reservation system recommended in this paper.

5.3 FCFS

FCFS allows participants to choose their favorite option among currently available slots. FCFS is widely used in real-world allocation problems due to its various advantages. First, it is simple in that participants can easily figure out an optimal action. Once participants access the system, they can instantly see that it is optimal to choose their most preferred option among those displayed.[19] Second, it has low communication costs: participants need only indicate their most preferred avail-

[19]However, under FCFS, agents face complex strategic considerations in deciding how much effort to spend on early system access, depending on their valuation vector and the severity of competition.

able option, without revealing more detailed information such as a full preference order. Third, reservations are confirmed immediately: participants do not need to hold multiple time slots as potential appointment times.

These features are often not satisfied by mechanisms derived as optimal in mechanism design research. For example, VCG requires participants to report their entire valuation vector, and running an equivalent ascending auction also entails substantial communication costs. The random favorite mechanism, which can outperform SD in cases such as asymmetric capacities, demands complex strategic considerations from participants (see Appendix A.2). Even SD, when implemented as a direct mechanism, requires the revelation of the full preference order. To retain the same practical advantages as FCFS, it must be implemented in sequential form.

However, under scarcity, the drawbacks of FCFS become more pronounced. FCFS is a screening mechanism that prioritizes participants who exert more effort to access the system earlier after reservations open. Our analysis indicates that such mechanisms are inefficient for allocating many heterogeneous goods (reservation slots). Indeed, as surveyed in Section 5.2, there are many cases where FCFS caused reservation surges and disruptions, leading to a switch to alternative mechanisms. We also note that FCFS unfairly advantages those with better Internet access or personal assistance, exacerbating inequities in vaccine access.

5.4 Register-Invite-Book System (RIB)

We propose the *register-invite-book system* (RIB), which we consider the most suitable mechanism for handling pandemic vaccine appointments. This mechanism is similar to SD in sequential form and was effectively used for COVID-19 vaccine distribution in places including British Columbia, Canada, and Singapore. RIB preserves the practical advantages of FCFS while minimizing competition and preventing the waste of screening costs.

As the name suggests, RIB first requires participants to complete *registration*. By registering, participants provide contact information such as an email address and input information relevant to their priority (e.g., age, occupation, underlying health condition). At this stage, no information about participants' preferences over slots needs to be entered. The system sorts the registered participants using methods other than FCFS, such as age or a lottery. The system sequentially sends an *invitation* to the next participant in line, while controlling the timing based on current slot availability (i.e., sends invitations more rapidly if many slots are remaining). Upon receiving an invitation, participants access the system and select their preferred slot from available options to complete their *booking* on an FCFS basis. However, since invitations are sent in small batches, only a small number of invitees can book at any given time.

The final stage of RIB retains the desirable practical properties of FCFS—simplicity, low communication cost, and immediate confirmation. However, the incentive for participants to exert effort to secure a slot is low because invitations are sent gradually, and few slots can be lost by minor delays. Therefore, RIB is approximately SD implemented in sequential form, sorting agents by an exogenous priority order and allowing them to choose their most preferred remaining slot in turn. RIB thus combines the practical advantages of FCFS with the theoretically desirable properties of SD, particularly the tendency to be an efficient mechanism when many heterogeneous objects are allocated, as discussed in this paper.

In the application of vaccine distribution, RIB has other noteworthy advantages. First, it is easy to prioritize specific groups of people. Those at higher risk, those prioritized for reasons of fairness, and those prioritized to prevent the spread of infection are likely to emerge for many diseases and vaccines, as discussed by Pathak et al. (2022, 2024); Akbarpour et al. (2023a). Prioritization criteria, such as infection risk or fairness, can matter even more than willingness to pay or exert effort. In RIB, policymakers can freely adjust the order in which invitations are sent to agents in the queue based on their observable attributes. Moreover, to distribute reservation slots among agents who belong to the same priority groups, SD implemented RIB saves the screening cost.

Second, RIB minimizes communication costs in the case of gradually supplied goods. Other mechanisms typically suffer from increased communication costs—for instance, when FCFS is used and new slots are added every midnight, participants have to repeatedly access the system every night until they successfully make a reservation, which imposes unnecessary effort on participants and increased server load.[20] Implementing SD as a direct mechanism would require creating a new lottery notification each time vaccines are supplied and having participants submit their preferences for new dates, which incurs high unmodeled costs. In contrast, with RIB, participants only need to access the system twice, for registration and booking. This property greatly reduces communication costs, even compared with FCFS.

Finally, RIB allows participants to predict when they might secure a reservation slot. Under FCFS, the competition for slots occurs each time new slots are released, and therefore, participants cannot predict their chances of securing a slot. This uncertainty persists until a reservation is secured, requiring participants to keep potential dates open and causing mental stress. In contrast, with RIB, the system can continuously disclose how many people with higher priorities are waiting for invitations, allowing participants to estimate when their turn might come. Although the exact timing can be difficult to predict due to potential new registrations and uncertainties in vaccine supply, RIB provides a much clearer and more predictable framework than FCFS.

Overall, RIB combines the practical simplicity of FCFS with the efficiency of SD. This makes it a state-of-the-art mechanism particularly suited for large-scale, multi-object vaccine distribution scenarios encountered during pandemics.

[20]Such a system was indeed used in New York City during the COVID-19 pandemic, and people indeed checked the system right after midnight (Krasnoff, 2021).

6 Concluding Remarks

We studied the efficient money-burning mechanism for allocating multiple heterogeneous goods, where agents take costly actions to signal the magnitude of their values. While the structure of residual-surplus maximizing mechanisms in multi-object allocation has remained largely unexplored, our findings reveal a clear trend: as variety increases, no-screening mechanisms like SD tend to become more efficient.

We first explored this trend by analyzing a stylized environment with a continuum of agents and objects, assuming values for all object types are drawn i.i.d., and the capacity of each object type is uniform. Agents tend to acquire the goods they prefer more. As the variety of goods increases, the distribution of the value for the "most preferred goods" increasingly favors no-screening mechanisms. Using extreme value theory, we showed that in the limit of a large variety, the efficient mechanism involves little to no screening.

Second, we applied the automated mechanism design method established by Dütting et al. (2019) to numerically explore efficient money-burning mechanisms. The results demonstrated that several strong assumptions imposed in the stylized environment—such as a very large variety of goods, a continuum of agents and objects, and uniform capacities—are not essential for the trend that SD tends to be efficient in multi-object allocations. This suggests that SD can perform well in practical settings. Given the various excellent properties of SD discussed in prior studies, our findings position SD as a promising option.

Finally, we emphasize broader implications for future research and policy design. Policymakers often face the challenge of allocating goods without monetary transfers, and our results suggest that imposing burdensome administrative procedures to ensure that only those who truly need goods and services receive them may often be unnecessary. This insight extends to diverse allocation problems beyond vaccine distribution. Furthermore, while the immediate urgency of COVID-19 vaccine distribution has subsided, preparing effective allocation mechanisms remains a critical issue for future pandemics. The lessons from this research highlight the importance of developing and implementing mechanisms like RIB or exploring other efficient no-screening mechanisms during periods of preparedness rather than crisis.

References

Abdulkadiroğlu, A., Y.-K. Che, and Y. Yasuda (2011): "Resolving Conflicting Preferences in School Choice: The "Boston Mechanism" Reconsidered," *American Economic Review*, 101, 399–410.

Ahuja, A., S. Athey, A. Baker, E. Budish, J. C. Castillo, R. Glennerster, S. D. Kominers, M. Kremer, J. Lee, C. Prendergast, C. M. Snyder, A. Tabarrok, B. J.

Tan, and W. Wiecek (2021): "Preparing for a Pandemic: Accelerating Vaccine Availability," *AEA Papers and Proceedings*, 111, 331–35.

Akbarpour, M., E. Budish, P. Dworczak, and S. D. Kominers (2023a): "An Economic Framework for Vaccine Prioritization," *Quarterly Journal of Economics*, 139, 359–417.

Akbarpour, M., P. Dworczak, and F. Yang (2023b): "Comparison of Screening Devices," in *Proceedings of the 24th ACM Conference on Economics and Computation*, 60.

Asahi Shimbun (2021a): "Vaccine Reservations Overwhelming Phone Lines: Concerns of Nationwide Confusion," *Asahi Shimbun*, in Japanese. Accessed: 2024-08-06.

——— (2021b): "Yokohama City Halts Vaccine Reservations: Overwhelming Access, Insufficient Response," *Asahi Shimbun*, in Japanese. Accessed: 2024-08-06.

Athey, S., J. C. Castillo, E. Chaudhuri, M. Kremer, A. Simoes Gomes, and C. M. Snyder (2022): "Expanding Capacity for Vaccines against COVID-19 and Future Pandemics: A Review of Economic Issues," *Oxford Review of Economic Policy*, 38, 742–770.

Bade, S. (2015): "Serial Dictatorship: The Unique Optimal Allocation Rule when Information is Endogenous," *Theoretical Economics*, 10, 385–410.

Barthel, M., N. Delgadillo, and M. Austermuhle (2021): "D.C. Vaccine Appointment System Crashes As Thousands More People Become Eligible," DCist.

Castillo, J. C., A. Ahuja, S. Athey, A. Baker, E. Budish, T. Chipty, R. Glennerster, S. D. Kominers, M. Kremer, G. Larson, J. Lee, C. Prendergast, C. M. Snyder, A. Tabarrok, B. J. Tan, and W. Wiecek (2021): "Market Design to Accelerate COVID-19 Vaccine Supply," *Science*, 371, 1107–1109.

Chakravarty, S. and T. R. Kaplan (2013): "Optimal Allocation without Transfer Payments," *Games and Economic Behavior*, 77, 1–20.

Che, Y.-K. and F. Kojima (2010): "Asymptotic Equivalence of Probabilistic Serial and Random Priority Mechanisms," *Econometrica*, 78, 1625–1672.

Coleman, E. (2021): "One State is Using Eventbrite to Schedule Covid-19 Vaccine Appointments," Online news report.

Condorelli, D. (2012): "What Money Can't Buy: Efficient Mechanism Design with Costly Signals," *Games and Economic Behavior*, 75, 613–624.

Conitzer, V. and T. Sandholm (2004): "Self-interested Automated Mechanism Design and Implications for Optimal Combinatorial Auctions," in *Proceedings of the 5th ACM Conference on Electronic Commerce*, 132–141.

Crawford, V. P. and E. M. Knoer (1981): "Job Matching with Heterogeneous Firms and Workers," *Econometrica*, 437–450.

Daskalakis, C., A. Deckelbaum, and C. Tzamos (2015): "Strong Duality for a Multiple-Good Monopolist," in *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, 449–450.

De Haan, L. and A. Ferreira (2006): *Extreme Value Theory: An Introduction*, vol. 3, Springer.

Demange, G., D. Gale, and M. Sotomayor (1986): "Multi-item Auctions," *Journal of Political Economy*, 94, 863–872.

Dütting, P., Z. Feng, H. Narasimhan, D. Parkes, and S. S. Ravindranath (2019): "Optimal Auctions through Deep Learning," in *International Conference on Machine Learning*, PMLR, 1706–1715.

Fisher, R. A. and L. H. C. Tippett (1928): "Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample," in *Mathematical proceedings of the Cambridge philosophical society*, Cambridge University Press, vol. 24, 180–190.

Fotakis, D., D. Tsipras, C. Tzamos, and E. Zampetakis (2016): "Efficient Money Burning in General Domains," *Theory of Computing Systems*, 59, 619–640.

Fox 4 News (2021): "Lee County Responds to Criticism of Vaccine Distribution," Online news report.

Fu, H. (2017): "Notes on Myerson's Revenue Optimal Mechanisms," Lecture Note.

Gans, J. S. (2022): "Optimal Allocation of Vaccines in a Pandemic," *Oxford Review of Economic Policy*, 38, 912–923.

Giannakopoulos, Y. and E. Koutsoupias (2018): "Duality and Optimality of Auctions for Uniform Distributions," *SIAM Journal on Computing*, 47, 121–165.

Gnedenko, B. (1943): "Sur La Distribution Limite Du Terme Maximum D'Une Serie Aleatoire," *Annals of Mathematics*, 44, 423–453.

Goldner, K. and T. Lundy (2024): "Simple Mechanisms for Utility Maximization: Approximating Welfare in the I.I.D. Unit-Demand Setting," Working Paper.

Government of British Columbia (2021): "COVID-19 Immunizations," Provincial government announcement.

Green, J. and J.-J. Laffont (1977): "Characterization of Satisfactory Mechanisms for the Revelation of Preferences for Public Goods," *Econometrica*, 427–438.

——— (1979): *Incentives in Public Decision-Making*, Elsevier North-Holland.

Gumbel, E. J. (1958): *Statistics of Extremes*, Columbia University Press.

Haghpanah, N. and J. Hartline (2021): "When Is Pure Bundling Optimal?" *Review of Economic Studies*, 88, 1127–1156.

Hakimov, R., C.-P. Heller, D. Kübler, and M. Kurino (2021): "How to Avoid Black Markets for Appointments with Online Booking Systems," *American Economic Review*, 111, 2127–51.

Hakimov, R., D. Kübler, and S. Pan (2023): "Costly Information Acquisition in Centralized Matching Markets," *Quantitative Economics*, 14, 1447–1490.

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*, New York, NY, USA: Association for Computing Machinery, STOC '08, 75–84.

Holmström, B. (1979): "Groves' Scheme on Restricted Domains," *Econometrica*, 1137–1144.

Koopmans, T. C. and M. Beckmann (1957): "Assignment Problems and the Location of Economic Activities," *Econometrica*, 53–76.

Krasnoff, B. (2021): "How to Score a COVID-19 Vaccine Appointment," The Verge.

Krysta, P., D. Manlove, B. Rastegari, and J. Zhang (2014): "Size Versus Truthfulness in the House Allocation Problem," in *Proceedings of the fifteenth ACM conference on Economics and Computation*, 453–470.

Li, S. (2017): "Obviously Strategy-Proof Mechanisms," *American Economic Review*, 107, 3257–3287.

Liew, M. E. (2022): "How Singapore Vaccinated Nearly 90 Percent of its Population in a year," GovInsider.

Liu, Q. and M. Pycia (2016): "Ordinal Efficiency, Fairness, and Incentives in Large Markets," Working Paper.

Majumdar, S. N., A. Pal, and G. Schehr (2020): "Extreme Value Statistics of Correlated Random Variables: A Pedagogical Review," *Physics Reports*, 840, 1–32.

Manelli, A. M. and D. R. Vincent (2006): "Bundling as an Optimal Selling Mechanism for a Multiple-Good Monopolist," *Journal of Economic Theory*, 127, 1–35.

Mathieu, E., H. Ritchie, E. Ortiz-Ospina, M. Roser, J. Hasell, C. Appel, C. Giattino, and L. Rodés-Guirao (2021): "A Global Database of COVID-19 Vaccinations," *Nature Human Behaviour*, 5, 947–953.

Myerson, R. B. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–73.

NBC Boston (2021): "Massachusetts Vaccination Website Crash: What Went Wrong?" Online news report.

Noda, S. (2018): "Large Matchings in Large Markets with Flexible Supply," Working Paper.

——— (2020): "Size versus truncation robustness in the assignment problem," *Journal of Mathematical Economics*, 87, 1–5.

Olson, J. (2021): "Minnesota Seniors Get 24 hours to Sign up for Vaccine Lottery," The Minnesota Star Tribune.

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 (2024): "Fair Allocation of Vaccines, Ventilators and Antiviral Treatments: Leaving No Ethical Value Behind in Healthcare Rationing," *Management Science*, forthcoming.

Pavlov, G. (2011): "Optimal Mechanism for Selling Two Goods," *B.E. Journal of Theoretical Economics*, 11, 1–35.

Pickands III, J. (1975): "Statistical Inference Using Extreme Order Statistics," *Annals of Statistics*, 119–131.

——— (1986): "The Continuous and Differentiable Domains of Attraction of the Extreme Value Distributions," *Annals of Probability*, 14, 996–1004.

Pitcher, T. (2021): "Vaccine Lottery Registration System Soon to be Implemented in SLO County," Mustang News.

Pycia, M. and P. Troyan (2024): "The Random Priority Mechanism is Uniquely Simple, Efficient, and Fair," Working Paper.

Rahme, J., S. Jelassi, J. Bruna, and S. M. Weinberg (2021): "A Permutation-Equivariant Neural Network Architecture for Auction Design," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 35, 5664–5672.

Sandholm, T. (2003): "Automated Mechanism Design: A New Application Area for Search Algorithms," in *International Conference on Principles and Practice of Constraint Programming*, Springer, 19–36.

Shapley, L. S. and M. Shubik (1971): "The Assignment Game I: The Core," *International Journal of Game Theory*, 1, 111–130.

Stolle, M. (2021): "Random Lottery Works, say Local Residents, but Lack of Vaccine Slows the Process," Post Bulletin.

Tada, I. (2021): "Regarding the Web Reservation Lottery Application Form for COVID-19 Vaccines," In Japanese. Accessed: 2024-08-06.

Tokarski, F. (2025a): "Equitable Screening," Working Paper.

——— (2025b): "Screening with Damages and Ordeals," Working Paper.

von Mises, R. (1936): "La Distribution De La Plus Grande De n Valeurs," *Revue de Mathématique de l'Union Interbalkanique*, 1, 141–160.

Wilson, R. (1985): *Game-theoretic Analyses of Trading Processes*, Institute for Mathematical Studies in the Social Sciences, Stanford University.

Yomiuri Shimbun (2021): "Vaccine Reservations Flood Municipalities: Strengthened Response, Increased Phone Lines, In-Person Appointments," *Yomiuri Shimbun Tokyo Morning Edition*, page 23, April 27, in Japanese.

Yoon, K. (2011): "Optimal Mechanism Design when Both Allocative Inefficiency and Expenditure Inefficiency Matter," *Journal of Mathematical Economics*, 47, 670–676.

Appendix

A Supplementary Results

A.1 Property of the Weibull Distribution

This section details the example of the Weibull distribution with shape parameter 0.9, which is illustrated in Figure 2.

The Weibull distribution with shape parameter $\alpha > 0$ is given by

$$G(v) = 1 - \exp(-v^{\alpha}),$$

and its probability density function is

$$g(v) = \alpha v^{\alpha-1} \exp(-v^{\alpha}).$$

Note that Weibull has an IHR if $\alpha \geq 1$ and a DHR if $\alpha \leq 1$. When $\alpha = 1$, it is identical to a standard exponential distribution.

The inverse hazard rate is given by

$$\vartheta(v; G) = \frac{1 - G(v)}{g(v)} = \frac{1}{\alpha} v^{1-\alpha}$$

and thus

$$\vartheta'(v; G) = \frac{1-\alpha}{\alpha} v^{-\alpha} \to 0 \text{ as } v \to \infty,$$

implying that the Weibull distribution satisfies the von Mises condition for all $\alpha > 0$, implying that the Weibull distribution belongs to the twice differentiable domain of attraction of Gumbel.

When G is the Weibull distribution with shape parameter α, from (7), we have

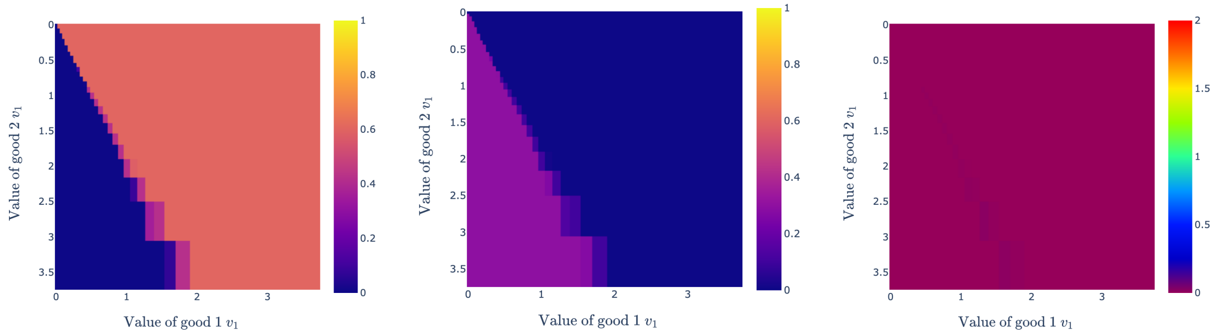
$$a_K = \frac{K}{\alpha} \Gamma\left(\frac{1}{\alpha}, \log K\right), \; b_K = (\log K)^{\frac{1}{\alpha}}.$$

The probability density of the normalized largest order statistic, which is illustrated in Figure 2a is given by $\hat{g}_K(w) = a_K g_K(a_K w + b_K)$. The normalized hazard rate is given by

$$r(w; \hat{G}_K) = \frac{a_K g_K(a_K w + b_K)}{1 - G_K(a_K w + b_K)},$$

and its derivative, illustrated as Figure 2b, is given by

$$r'(w; \hat{G}_K) = \frac{a_K^2 g'_K(a_K w + b_K)(1 - G_K(a_K w + b_K)) + a_K^2 g_K^2(a_K w + b_K)}{(1 - G_K(a_K w + b_K))^2}.$$



(a) $x_1(v_1, v_2)$ (b) $x_2(v_1, v_2)$ (c) $p(v_1, v_2)$

Figure A.1: An efficient mechanism for the case of asymmetric capacity. The optimal solution can be achieved with RF.

A.2 Asymmetric Capacity

In Section 3, for analytical tractability, we assumed that all object types have an equal capacity. This assumption is not always realistic; for example, there are only two weekend days compared to five weekdays. In such environments, an ex post efficient mechanism may assign non-favorite objects when an agent's favorite object is scarce, making it difficult to reduce the problem to a single-dimensional environment as discussed in Section 3.2. This complexity makes it challenging to characterize an efficient mechanism.

In this section, we solve a linear program (LP) to numerically derive an efficient mechanism and compare its structure with that of the efficient mechanism in a continuous i.i.d. market. The residual surplus function RS and constraints (2), (3), (4), and (5) are all linear in the allocation and payment rule, (x, p), and thus this problem is an LP. To use a solver, we discretize the valuation space $V \subset \mathbb{R}_+^K$ by dividing each dimension into n intervals with equal probability weights to produce n^K grid points, each of which is specified as a K-tuple of lower endpoints. Throughout this paper, we take $n = 30$ intervals. We consider the case of $K = 2$ and assume that the value for each object is drawn from a standard exponential distribution independently. If both objects had an equal capacity, then SD would be efficient. Here, we instead assume that $m_1 = 0.4$ and $m_2 = 0.1$.

Figure A.1 illustrates an optimal solution to the LP for the case of asymmetric capacity. For all figures, the horizontal axis represents the value of object 1 (v_1), and the vertical axis represents the value of object 2 (v_2). The colors of panels (a), (b), and (c) indicate the probability of allocating objects 1 and 2 ($x_1(v_1, v_2), x_2(v_1, v_2)$), and the payment ($p(v_1, v_2)$), respectively.

The efficient mechanism is neither SD nor VCG. The mechanism offers two options: Receiving object 1 with high probability or receiving object 2 with low probability, without requiring payments. Following the terminology of Goldner and Lundy (2024), we call this mechanism the

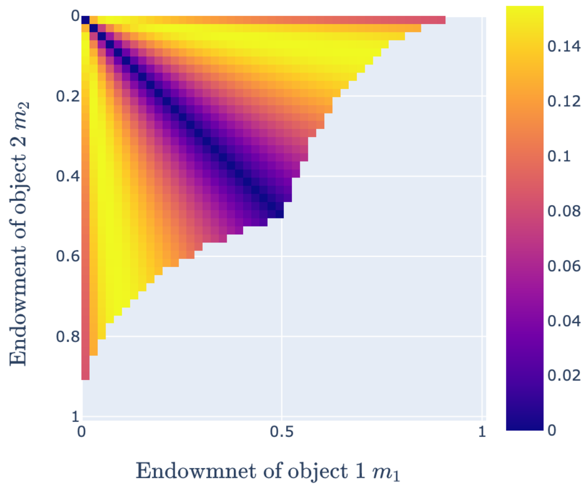


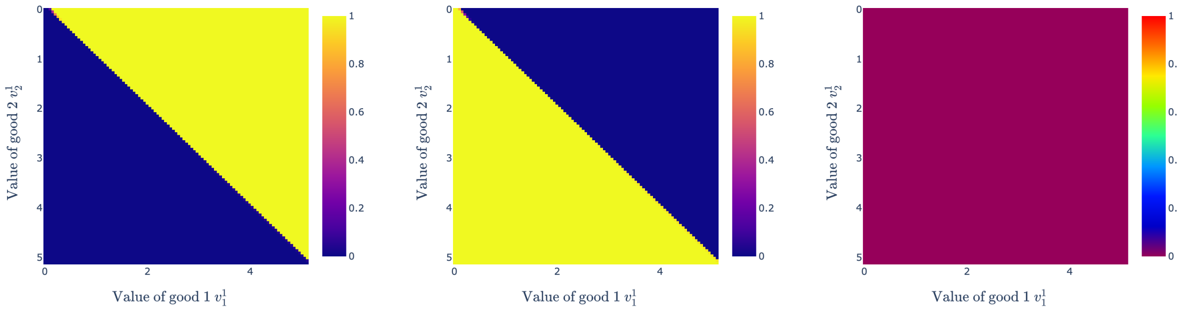
Figure A.2: Percentage difference of the residual surplus achieved by RF and SD $(RS(\mathcal{M}_{RF}) - RS(\mathcal{M}_{SD}))/RS(\mathcal{M}_{RF})$.

random-favorites mechanism (RF).[21] By requiring which object to claim before drawing a lottery, this mechanism partially screens for agents' preference for their favorite object relative to the other one—only agents who strongly prefer object 2 over object 1 desire to claim object 2, given that object 1 is much easier to obtain. By contrast, under SD, claiming object 2 when it is available is risk-free, and therefore, all agents who prefer object 2 over object 1 obtain object 2 whenever possible. Consequently, RF outperforms SD with asymmetric capacities.

While RF outperforms SD in the residual surplus, it comes with several drawbacks. While RF can be implemented as a strategy-proof mechanism in a continuous market, to determine the allocation probability of each option, we need information about the value distribution, which is often considered challenging in the market design literature (Wilson, 1985). By contrast, SD is detail-free. Furthermore, RF cannot satisfy strategy-proofness in a finite market, even if the market is large.[22] This is because the object an agent should claim depends on which objects other agents are claiming. In addition, RF cannot satisfy the practically desirable properties that FCFS and the sequential-form SD meet (discussed in Section 5).

[21] Goldner and Lundy (2024) study an (ex ante) symmetric environment, and therefore, each agent i claims $k \in \arg\max v_k^i$ as their favorite object in equilibrium. This property does not necessarily hold with asymmetric capacity because the probability of being assigned depends on which object to claim.

[22] Several studies on matching theory (e.g., Abdulkadiroğlu et al., 2011) have demonstrated that non-strategy-proof mechanisms may outperform strategy-proof mechanisms in equilibrium because their equilibrium strategies can reflect agents' preference intensity. Note also that RF may not satisfy even a weaker notion of truthfulness, such as Bayesian incentive compatibility (BIC). While Goldner and Lundy (2024) proves that RF satisfies BIC when each object k has a unit capacity and each value v_k^i is drawn i.i.d., these assumptions are rarely satisfied practically.



(a) $x_1^1(v_1^1, v_2^1, 2, 1)$ (b) $x_2^1(v_1^1, v_2^1, 2, 1)$ (c) $p^1(v_1^1, v_2^1, 2, 1)$

Figure A.3: Allocation and payment for agent 1 from an efficient mechanism learned by RegretNet. The marginal value distribution is the Weibull distribution with parameter 0.8. Agent 2's value is fixed to $(v_1^2, v_2^2) = (2, 1)$.

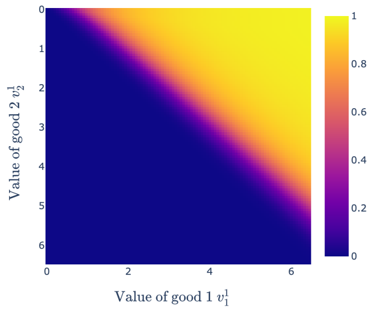
To evaluate the tradeoff between social welfare and the simplicity of the mechanism, we compare the performance of SD and RF. The allocation probability of RF should satisfy a market-clearing condition. That is, given the allocation probability, agents should optimally choose one option. This determines the demand for each object, and together with the resource constraint, the demand determines the allocation probability. The initial allocation probability should be consistent with the last one. For each pair of endowments (m_1, m_2), there exists a unique RF that satisfies the market-clearing condition, and its performance has a closed-form representation when values follow an exponential distribution. The derivation is presented in Appendix B.11.

Figure A.2 displays the percentage difference of residual surplus between RF and SD (i.e., $(RS(\mathcal{M}_{RF}) - RS(\mathcal{M}_{SD}))/RS(\mathcal{M}_{RF})$) for various endowments, (m_1, m_2). For any endowments, RF performs as good as or better than SD. Indeed, when $m_1 = m_2$, RF and SD return an identical allocation and the performance difference is zero. The percentage difference is at most 15%, implying that the gain from using RF is not excessively large.

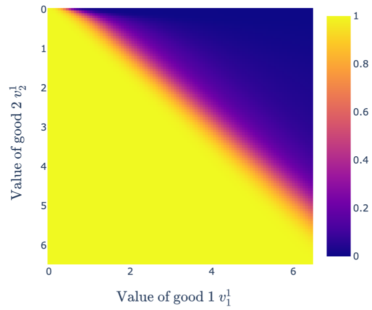
A.3 Illustration of Mechanisms Returned by RegretNet

Figures A.3 and A.4 show the approximately efficient mechanisms learned by RegretNet. The horizontal and vertical axes represent v_1^1 and v_2^1, agent 1's value for objects 1 and 2. Agent 1's allocation probability and the amount of agent 1's payment are represented by color. While RegretNet returns the full shape of the mechanism, for the illustration's sake, we fix agent 2's valuation to $(v_1^2, v_2^2) = (2, 1)$ to draw these figures.

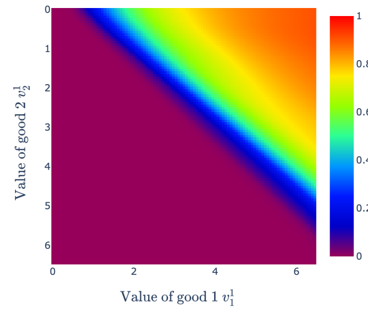
Figure A.3 shows the case for a Weibull distribution with parameter 0.8. The mechanism allocates agent 1 to their favorite object, without payment and regardless of agent 2's value, implying that the mechanism is SD, which prioritizes agent 1 over agent 2.



(a) $x_1^1(v_1^1, v_2^1, 2, 1)$



(b) $x_2^1(v_1^1, v_2^1, 2, 1)$



(c) $p^1(v_1^1, v_2^1, 2, 1)$

Figure A.4: Allocation and payment for agent 1 from an efficient mechanism learned by RegretNet. The marginal value distribution is the Weibull distribution with parameter 0.7. Agent 2's value is fixed to $(v_1^2, v_2^2) = (2, 1)$.

Figure A.4 shows the case for a Weibull distribution with parameter 0.7. The value distribution has a thicker tail, and the learned mechanism allocates the contested object with a price of the other agent's value—since agent 2 prefers object 1 over object 2 by $v_1^2 - v_2^2 = 2 - 1 = 1$, agent 1 obtains object 1 only if her excess value $v_1^1 - v_2^1$ is larger than 1, and agent 1 has to pay the price of 1 in such cases. This is how VCG allocates objects.

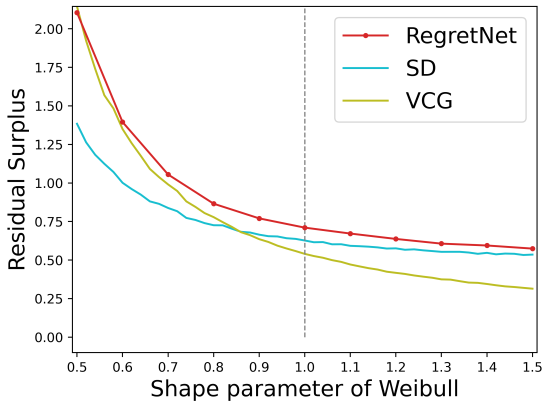
A.4 Performance of SD, VCG, and the (numerically) efficient mechanisms

Figure A.5 illustrates the residual surplus achieved by SD, VCG, and the (numerically) efficient mechanism across four cases: $c = 1$, $c = 5$, the single-good case, and the continuous case. (Note that the $c = 5$ case is identical to Figure 4b in the main text but is included here for comparison.)

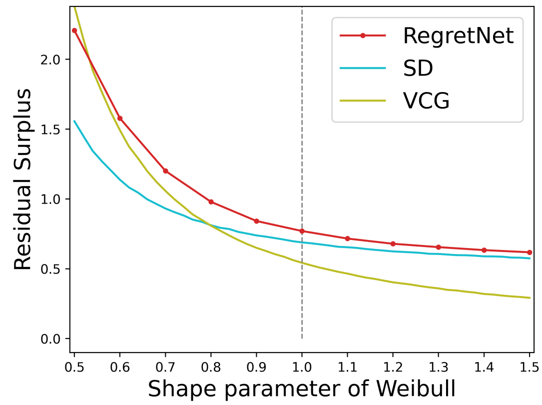
For the single-good and continuous market cases, the efficient mechanism is derived analytically using the results from Hartline and Roughgarden (2008) and Section 3. In contrast, for the $c = 1$ and $c = 5$ cases, the mechanism is obtained numerically using RegretNet.

In the single-good case, the efficient mechanism precisely aligns with the better-performing mechanism between SD and VCG across all values of α. In the continuous case, the same pattern holds visually, although a closer numerical inspection reveals that the efficient mechanism can outperform both SD and VCG by a small but strictly positive margin. We also note that, in the continuous case, where there is minimal but non-zero heterogeneity among goods (i.e., $K = 2$), the threshold value of the Weibull parameter shifts from 1.0 to below 0.8. As K increases, this threshold approaches $\alpha = 0$.

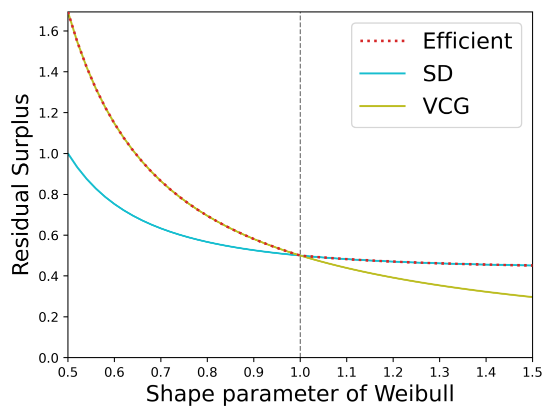
The case with $c = 1$ exhibits characteristics that lie between those of the single-good case and the $c = 5$ case. Even for $c = 1$, the effect of having multiple goods are clearly evident, as reflected



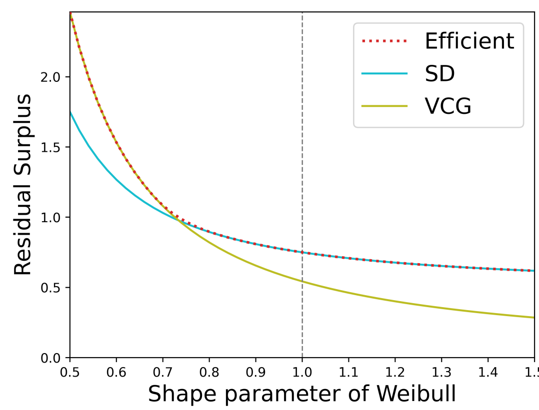
(a) $c = 1$



(b) $c = 5$



(c) Single-good case



(d) Continuous case

Figure A.5: The performance of SD, VCG, and the (numerically) efficient mechanism

in the observed performance difference.

A.5 Simulation Settings for the Case of Correlated Values

This section illustrates how we construct the joint distribution F for the case of correlated values. Let $\Sigma \in [-1,1]^{(I\times K)\times(I\times K)}$ be a $(I \times K)$-dimensional correlation matrix. The *Gaussian copula* $C_\Sigma : [0,1]^{I\times K} \to [0,1]$ is defined as

$$C_\Sigma(\mathbf{w}) = N_\Sigma(N^{-1}(w_1^1), \ldots, N^{-1}(w_K^I)), \tag{16}$$

where N_Σ is the joint cumulative distribution function of an $(I\times K)$-dimensional multivariate normal distribution with mean vector zero and covariance matrix Σ, and N^{-1} is the inverse cumulative distribution function of a (univariate) standard normal.

By construction of C_Σ, for all (i,k), the marginal distribution of w_k^i is a uniform distribution over $[0,1]$. Accordingly, by defining

$$F(\mathbf{v}) = C_\Sigma(G(v_1^1), \ldots, G(v_K^I)), \tag{17}$$

We can construct a joint cumulative distribution F of $\mathbf{v}$ such that for all (i,k), the marginal value distribution of v_k^i is G. Furthermore, by imposing symmetry on the covariance matrix Σ used in the construction of the copula—such as keeping within-agent correlation constant or between-agent correlation constant—the symmetry will be preserved in the covariance matrix of the output joint distribution F. Although covariance and correlation coefficients are not necessarily preserved during the transformation process, the relative strength of the correlations is maintained. Therefore, if stronger correlations are set in the covariance matrix Σ, the resultant joint distribution F will also exhibit stronger correlations. As a result, by varying the covariance matrix Σ, we can generate a joint distribution F with the desired strengths of within-agent correlation and between-agent correlation.

In the numerical analysis, instead of directly computing the joint distribution F using (16) and (17), we generate random variables $\mathbf{v}$ following the intended joint distribution F by generating $\mathbf{u} = (u_1^1, \ldots, u_K^I)$ following a multivariate normal distribution with mean 0 and covariance matrix Σ, and then transforming it as $\mathbf{v} = (G^{-1}(N(u_1^1)), \ldots, G^{-1}(N(u_K^I)))$.

We use the following covariance matrices Σ for our simulations. For the analysis to evaluate the effect of the within-agent correlation, we employ (i) $\mathrm{Cov}(u_k^i, u_k^i) = 1$ for all $(i,k) \in I \times K$, (ii) $\mathrm{Cov}(u_k^i, u_l^i) = c$ for all $i \in I$, for all distinct $k, l \in K$, for some $c \in [0,1]$, and (iii) $\mathrm{Cov}(u_k^i, u_l^j) = 0$ otherwise. For the analysis to evaluate the effect of the between-agent correlation, we employ (i) $\mathrm{Cov}(u_k^i, u_k^i) = 1$ for all $(i,k) \in I \times K$, (ii) $\mathrm{Cov}(u_k^i, u_k^j) = c$ for all $k \in K$, for all distinct $i, j \in I$, for some $c \in [0,1]$, and (iii) $\mathrm{Cov}(u_k^i, u_l^j) = 0$ otherwise.

Finally, we describe how we calculate the correlation coefficient of the generated distribution F, which is represented on the horizontal axis of Figures 5 and 6. The method used is similar for both cases, thus we will only describe the within-agent correlation case. Noting that all v_k^i share the same marginal distribution, the correlation coefficient between v_k^i and v_l^i for $k \neq l$ is defined as follows:

$$\mathrm{Corr}(v_k^i, v_l^i) = \frac{\mathrm{Cov}(v_k^i, v_l^i)}{\mathrm{Var}(v_k^i)}. \tag{18}$$

Here, $\mathrm{Var}(v_k^i)$ is known since we specify the marginal distribution as Weibull. However, the joint distribution F derived using the Gaussian copula does not have a tractable closed-form representation, thus $\mathrm{Cov}(v_k^i, v_l^i)$ cannot be derived in closed form. Therefore, we estimate it by sampling. Specifically, let the sample size be T, and denote the value of v_k^i in the t-th sample as $v_k^i(t)$. The sample covariance $S(i,k,l)$ between v_k^i and v_l^i is expressed as follows:

$$S(i,k,l) = \frac{1}{T} \sum_{t=1}^{T} \left(v_k^i(t) - \mu \right) \left(v_l^i(t) - \mu \right),$$

where $\mu := \mathbb{E}[v_k^i(t)]$. Note that we can calculate μ from the marginal value distribution G that each $v_k^i(t)$ follows, and thus there is no need to compute its sample mean. Due to the symmetry of the covariance matrix Σ used to generate $\mathbf{u}$, $\text{Cov}(v_k^i, v_l^i) = \text{Cov}(v_{\hat{k}}^i, v_{\hat{l}}^i)$ holds as long as $k \neq l$ and $\hat{k} \neq \hat{l}$. Using this property, we aggregate the samples of within-agent correlation across different (i, k, l) to obtain the following sample covariance:

$$\widehat{\text{Cov}}(v_k^i, v_l^i) = \frac{1}{I} \frac{2}{K(K-1)} \sum_{i=1}^{I} \sum_{k=1}^{K} \sum_{l=k+1}^{K} S(i, k, l). \tag{19}$$

In the simulations, we used a sample size of $T = 100,000$ for each setting. By substituting the sample covariance obtained from (19) into (18), we derive the correlation coefficients plotted on the horizontal axis of Figure 5.

B Proofs

B.1 Proof of Theorem 1

Proof. We show the contrapositive. Suppose that there exists $V^* \subset V$ with $F(V^*) > 0$ such that for each $\mathbf{v} \in V^*$, there exists $k \in K$ such that $k \notin \arg\max_{l \in K} v_l$ but $x_k(\mathbf{v}) > 0$. Since G is a continuous distribution, $F(\{\mathbf{v} \in [0, \bar{v}]^K : v_k = v_l \text{ for some } k \neq l\}) = 0$. Hereafter, we ignore any $\mathbf{v}$ such that $\arg\max_{l \in K} v_l$ is not a singleton.

For each $\mathbf{v} \in V^*$, we denote $k^+(\mathbf{v}) := \arg\max_{l \in K} v_l$, and define k^- as follows:

$$K^-(\mathbf{v}) = \{l \in K : x_l(\mathbf{v}) > 0 \text{ and } v_l < v_{k^+(\mathbf{v})}\},$$

$$k^-(\mathbf{v}) = \arg\min_{l \in K^-(\mathbf{v})} v_l.$$

That is, $k^-(\mathbf{v})$ is defined as the good allocated to a value-$\mathbf{v}$ agent with a positive probability, while it is not a value-$\mathbf{v}$ agent's favorite object. When there are multiple objects satisfying this condition, we pick the least preferred one.

For each pair $(a, b) \in K^2$ with $a \neq b$, define

$$V_{a,b} := \{\mathbf{v} \in V^* : k^-(\mathbf{v}) = a, k^+(\mathbf{v}) = b\}.$$

Then,

$$F(V^*) = \sum_{a \neq b} F(V_{a,b}).$$

Since the number of pairs is finite, there exists a pair (k^-, k^+) such that $F(V_{k^-,k^+}) > 0$. Fix such a pair and write $V^{**} = V_{k^-,k^+}$. For every $\mathbf{v} \in V^*$, $k^-(\mathbf{v}) = k^-$ and $k^+(\mathbf{v}) = k^+$.

We define a mapping $T : V \to V$ by

$$Tv_{k^+} = v_{k^-}, Tv_{k^-} = v_{k^+}, Tv_k = v_k \text{ for all } k \in K \setminus \{k^+, k^-\}.$$

By neutrality, we have $x_{k^+}(\mathbf{v}) = x_{k^-}(T\mathbf{v})$, $x_{k^-}(\mathbf{v}) = x_{k^+}(T\mathbf{v})$, and $x_k(\mathbf{v}) = x_k(T\mathbf{v})$ for all $k \in K \setminus \{k^+, k^-\}$. We define an alternative allocation rule $\mathbf{x}'$ by

$$\begin{aligned} x'_{k^+}(\mathbf{v}) &= x_{k^+}(\mathbf{v}) + x_{k^-}(\mathbf{v}) \text{ for all } \mathbf{v} \in V^{**}, \\ x'_{k^-}(\mathbf{v}) &= 0 \text{ for all } \mathbf{v} \in V^{**}, \\ x'_{k^+}(\mathbf{v}) &= 0 \text{ for all } \mathbf{v} \in TV^{**}, \\ x'_{k^-}(\mathbf{v}) &= x_{k^+}(\mathbf{v}) + x_{k^-}(\mathbf{v}) \text{ for all } \mathbf{v} \in TV^{**}, \\ \mathbf{x}'(\mathbf{v}) &= \mathbf{x}(\mathbf{v}) \text{ for all } \mathbf{v} \notin V^{**} \cup TV^{**}. \end{aligned}$$

By construction, $\sum_{k \in K} x'_k(\mathbf{v}) = \sum_{k \in K} x_k(\mathbf{v}) \leq 1$, implying that $\mathbf{x}'$ satisfies the unit demand condition. Furthermore, since we have $F(\mathbf{v}) = F(T\mathbf{v})$ in a continuous i.i.d. market, the resource constraint is also satisfied. Finally, all agents with valuation $\mathbf{v} \in V^{**} \cup TV^{**}$ prefers $\mathbf{x}'$ over $\mathbf{x}$, whereas $F(V^{**}) > 0$ by assumption. Accordingly, $\mathbf{x}$ is not ex post efficient. □

B.2 Proof of Theorem 4

Proof. By direct calculation, we can verify that a cumulative distribution function F is NBUE if and only if

$$\int_t^{\bar{v}} \frac{F(t)(1 - F(v))}{1 - F(t)} dv \leq \int_0^t (1 - F(v)) dv \tag{20}$$

for all $t \in [0, \bar{v}]$. We will show that if (20) is satisfied with $F = G^K$, then (20) is also satisfied with $F = G^{K+1}$.

Since the right-hand side of (20) with $F = G^{K+1}$ is larger than that with $F = G^K$, it suffices to show that the left-hand side of (20) with $F = G^{K+1}$ is smaller than that with $F = G^K$. Thus, the desired equation is

$$\int_t^{\bar{v}} \left[\frac{G^K(t)(1 - G^K(v))}{1 - G^K(t)} - \frac{G^{K+1}(t)(1 - G^{K+1}(v))}{1 - G^{K+1}(t)} \right] dv \geq 0.$$

To this end, we will prove

$$\frac{G^K(t)(1 - G^K(v))}{1 - G^K(t)} - \frac{G^{K+1}(t)(1 - G^{K+1}(v))}{1 - G^{K+1}(t)} \geq 0 \tag{21}$$

for all $v \in [t, \bar{v})$.

For notational simplicity, let $a = G(t)$ and $b = G(v)$. Then, (21) can be expressed as

$$\frac{a^K(1-b^K)}{1-a^K} - \frac{a^{K+1}(1-b^{K+1})}{1-a^{K+1}} \geq 0.$$

This inequality is equivalent to

$$(1-b^K)(1-a^{K+1}) - a(1-b^{K+1})(1-a^K) \geq 0,$$

or

$$(1+b+\cdots+b^{K-1})(1+a+\cdots+a^K) - a(1+b+\cdots+b^K)(1+a+\cdots+a^{K-1}) \geq 0. \tag{22}$$

Let $A = 1+a+\cdots+a^{K-1}$, and $B = 1+b+\cdots+b^{K-1}$. Then, (22) becomes

$$B(1+aA) - a(1+bB)A \geq 0,$$

or equivalently,

$$(B-aA) + aAB(1-b) \geq 0. \tag{23}$$

It follows from $a = G(t)$, $b = G(v)$, and $v \geq t$ that $a, b \in [0,1]$ and $B \geq A \geq aA \geq 0$. Accordingly, (23) is indeed satisfied, as desired. □

B.3 Proof of Theorem 2

We apply the construction of the ironed virtual value established in Section 6 of Myerson (1981) to a distribution function G_K.

Definition 7 (Ironed Virtual Value)**.** The ironed virtual value $\bar{\vartheta}$ is constructed as follows:

1. For $q \in [0,1]$, define $h(q) = \vartheta(G_K^{-1}(q); G_K)$.
2. Define $H(q) = \int_0^q h(r)dr$.
3. Define I as the convex hull of H, the largest convex function bounded above by H for all $q \in [0,1]$.
4. Define $i(q)$ as the derivative of $I(q)$, where defined, and extend to all of $[0,1]$ by right continuity.
5. $\bar{\vartheta}(v; G_K) = i(G_K(v))$.

Proof of Theorem 2. The ironed virtual value $\bar{\vartheta}$ becomes constant if and only if $I(q)$ is linear. Since I is the convex hull of H, $I(1) = H(1)$ must be the case; if $I(1) > H(1)$, I is not bounded above by H, and if $I(1) < H(1)$, I is not the largest convex function bounded above by H. Accordingly, $I(q)$

is linear if and only if $I(q) = H(1) \cdot q$. It is easy to see that such a function is the pointwise largest among all convex functions satisfying $I(0) = H(0) = 0$ and $I(1) = H(1)$. Thus, such a function I is the convex hull of H if and only if I is bounded above by H, i.e., $I(q) = H(1)q \leq H(q)$ for all $q \in (0, 1)$.

From the definition of H, we have

$$\begin{aligned} H(q) &:= \int_0^q \vartheta(G_K^{-1}(r); G_K) dr \\ &= \int_0^{G_K^{-1}(q)} \vartheta(v'; G_K) g_K(v') dv' \\ &= \int_0^{G_K^{-1}(q)} (1 - G_K(v')) dv'. \end{aligned}$$

Accordingly, $H(1)q \leq H(q)$ is equivalent to

$$\int_0^{G_K^{-1}(q)} (1 - G_K(v)) dv \geq q \int_0^{\bar{v}} (1 - G_K(v)) dv.$$

Define $t = G_K^{-1}(q)$. Then, the above equation can be rewritten as

$$\int_0^t (1 - G_K(v)) dv \geq G_K(t) \int_0^{\bar{v}} (1 - G_K(v)) dv,$$

or equivalently,

$$\begin{aligned} &\int_0^t (1 - G_K(v)) dv \geq G_K(t) \mathbb{E}[v] \\ \Leftrightarrow &\int_t^{\bar{v}} (1 - G_K(v)) dv \leq (1 - G_K(t)) \mathbb{E}[v] \\ \Leftrightarrow &\frac{1}{1 - G_K(t)} \int_t^{\bar{v}} (v - t) dG_K(v) \leq \mathbb{E}[v] \\ \Leftrightarrow &\mathbb{E}[v - t | v > t] \leq \mathbb{E}[v], \end{aligned}$$

as desired. □

B.4 Proof of Theorem 5

Proof. Recall that

$$r(v; G_K) = \frac{g_K(v)}{1 - G_K(v)} = \frac{K g(v) (G(v))^{K-1}}{1 - (G(v))^K}.$$

By direct calculation, we can show that

$$r'(v; G_K) = \frac{K (G(v))^{K-2}}{(1 - (G(v))^K)^2} \left(K (g(v))^2 - ((g(v))^2 - g'(v) G(v))(1 - (G(v))^K) \right).$$

Thus, $r'(v; G_K) \geq 0$ if and only if

$$K(g(v))^2 \geq ((g(v))^2 - g'(v)G(v))(1 - (G(v))^K). \tag{24}$$

We show that, for any fixed v, if (24) holds with K, then (24) also holds with $K+1$. This claim is trivial if $g'(v) \geq 0$. In the following, we consider the case of $g'(v) < 0$.

Suppose that (24) holds with K. Then, we have

$$K(g(v))^2 \geq ((g(v))^2 - g'(v)G(v))(1 - (G(v))^K).$$

and thus

$$(K+1)(g(v))^2 \geq ((g(v))^2 - g'(v)G(v))\frac{K+1}{K}(1 - (G(v))^K).$$

Therefore, it suffices to show that

$$\frac{K+1}{K}(1 - x^K) \geq 1 - x^{K+1}$$

for all $x \in [0,1]$, or equivalently,

$$(K+1)(1 - x^K) - K(1 - x^{K+1}) \geq 0. \tag{25}$$

It follows from $1 - x^K = (1-x)(x^{K-1} + x^{K-2} + \cdots + x + 1)$ that (25) holds if

$$(K+1)(x^{K-1} + x^{K-2} + \cdots + x + 1) - K(x^K + x^{K-1} + \cdots + x + 1) \geq 0,$$

or equivalently,

$$x^{K-1} + \cdots + x + 1 \geq Kx^K. \tag{26}$$

Indeed, (26) holds for all $x \in [0,1]$ because $x^k \geq x^K$ holds for all $k = 0, 1, \ldots, K-1$, as desired.

We show that for all $v \in [0, \bar{v}]$, there exists K_0 such that for all $K > K_0$, $r'(v; G_K) > 0$. The right-hand side of (24) is bounded above by $((g(v))^2 - g'(v)G(v))$. Since we assume that G has full support, i.e., $g(v) > 0$ for all $v \in (0, \bar{v})$, there exists K_0 such that

$$K_0(g(v))^2 > (g(v))^2 - g'(v)G(v).$$

Clearly, for all $K > K_0$, we have (24), implying that $r'(v; G_K) > 0$. □

B.5 Proof of Theorem 6

Proof. Take $\epsilon > 0$ and $\delta > 0$ arbitrarily. Let $\bar{\delta} = \min\{\delta, (1 - \bar{m})/2\}$. Since $\bar{v} = \sup \text{support}(G) < +\infty$, we have $G(\bar{v} - \epsilon) < 1$. Let K_0 be the smallest integer K satisfying $G_K(\bar{v} - \epsilon) := (G(\bar{v} - \epsilon))^K < \bar{\delta}$, i.e., $K_0 = \lfloor \log \bar{\delta} / \log G(\bar{v} - \epsilon) \rfloor + 1$. By construction of K_0, for any $K > K_0$, we have $G_K(\bar{v} - \epsilon) < \bar{\delta}$, or equivalently, $\Pr(v \in (\bar{v}-\epsilon, \bar{v}]) > 1-\bar{\delta}$.

Given such K_0, for any $K > K_0$,

$$\begin{aligned} RS(\mathcal{M}_{SD}) &= \bar{m}\int_0^{\bar{v}} v dG_K(v) \\ &> \bar{m}\int_{\bar{v}-\epsilon}^{\bar{v}} (\bar{v}-\epsilon) dG_K(v) \\ &= \bar{m}\Pr(v \in (\bar{v}-\epsilon, \bar{v}])(\bar{v}-\epsilon) \\ &> \bar{m}(1-\bar{\delta})(\bar{v}-\epsilon) \\ &\geq \bar{m}(1-\delta)(\bar{v}-\epsilon). \end{aligned}$$

Next, we evaluate $RS(\mathcal{M}_{VCG})$ with the same K_0. For any $K > K_0$, we have $G_K(\bar{v}-\epsilon) < \bar{\delta} < 1-\bar{m}$. Accordingly, we have $q \coloneqq G_K^{-1}(1-\bar{m}) > \bar{v}-\epsilon$. Accordingly,

$$\begin{aligned} RS(\mathcal{M}_{VCG}) &= \int_q^{\bar{v}} (v-q)\, dG_K(v) \\ &< \int_q^{\bar{v}} (\bar{v}-(\bar{v}-\epsilon))\, dG_K(v) \\ &= \epsilon(1-G_K(q)) \\ &= \bar{m}\epsilon. \end{aligned}$$

□

B.6 Proof of Theorem 7

We first introduce two lemmas.

Lemma 1. If $G \in \mathbf{VMC}(0)$, then for any finite interval $[\underline{c}, \bar{c}]$, there exists K_0 such that for all $K > K_0$, we have $\vartheta'(w; \hat{G}_K) < 0$ for $w \in [\underline{c}, \bar{c}]$.

Proof. Recall that $\vartheta(w; \hat{G}_K) \coloneqq (1-\hat{G}_K(w))/\hat{g}_K(w)$.

$$\vartheta'(w; \hat{G}_K) \coloneqq \left(\frac{1-\hat{G}_K(w)}{\hat{g}_K(w)}\right)' = \frac{-(\hat{g}_K(w))^2 - \hat{g}'_K(w)(1-\hat{G}_K(w))}{(\hat{g}_K(w))^2}.$$

Since G has full support and twice differentiable, there exists $\underline{g}$ such that $\hat{g}_K(w) > \underline{g}$ for all $w \in [\underline{c}, \bar{c}]$. Furthermore, $\hat{G}_K, \hat{g}_K, \hat{g}'_K$ converges uniformly to $\Lambda, \lambda, \lambda'$ in $[\underline{c}, \bar{c}]$. Accordingly, for all $w \in [\underline{c}, \bar{c}]$,

$$\vartheta'(w; \hat{G}_K) \to \frac{-(\lambda(w))^2 - \lambda'(w)(1-\Lambda(w))}{(\lambda(w))^2} = \vartheta'(w; \Lambda) < 0,$$

where the convergence is uniform in $[\underline{c}, \bar{c}]$. Hence, there exists K_0 such that for all $K > K_0$, we have $\vartheta'(w; \hat{G}_K) < 0$ for $w \in [\underline{c}, \bar{c}]$. □

Lemma 2. Let $[a,b]$ be an interval on which virtual valuation $\vartheta(\cdot;\hat{G}_K)$ is nonincreasing, and x be the allocation rule of any strategy-proof mechanism. Then, the allocation rule

$$\tilde{x}(w) := \begin{cases} \dfrac{\displaystyle\int_a^b x(w')\hat{g}_K(w')dw'}{\hat{G}_K(b)-\hat{G}_K(a)} & \text{if } w \in [a,b], \\ x(w) & \text{otherwise,} \end{cases}$$

is a monotonic allocation rule that satisfies the resource constraint. Furthermore, its residual surplus is no less than that of x.

This lemma is a basic step in deriving revenue-maximizing mechanisms and has been well-known. We include it here for completeness. The proof follows the approach outlined in Fu (2017).

Proof. Since x is an allocation rule of a strategy-proof mechanism, x is monotonic; thus, $\tilde{x}$ is also monotonic.

Next, we show that $\tilde{x}$ satisfies the resource constraint.

$$\begin{aligned}
&\int_{-\infty}^{\infty} \tilde{x}(w)\hat{g}_K(w)dw \\
&= \int_a^b \tilde{x}(w)\hat{g}_K(w)dw + \int_{w\notin[a,b]} x(w)\hat{g}_K(w)dw \\
&= (\hat{G}_K(b)-\hat{G}_K(a))\frac{\int_a^b x(w)\hat{g}_K(w)dw}{\hat{G}_K(b)-\hat{G}_K(a)} + \int_{w\notin[a,b]} x(w)\hat{g}_K(w)dw \\
&= \int_{-\infty}^{\infty} x(w)\hat{g}_K(w)dw.
\end{aligned}$$

Since x satisfies the resource constraint, $\tilde{x}$ also satisfies it.

For $w \notin [a,b]$, x and $\tilde{x}$ are identical and hence generate the same residual surplus. On $[a,b]$, let $\hat{G}_K(\cdot|w \in [a,b])$ denote the conditional distribution of w given $w \in [a,b]$, which has density $\hat{g}_K/(\hat{G}_K(b)-\hat{G}_K(a))$. Then,

$$\begin{aligned}
&\int_a^b \tilde{x}(w)\vartheta(w)\hat{g}_K(w)dw \\
&= \frac{\int_a^b x(w)\hat{g}_K(w)dw}{\hat{G}_K(b)-\hat{G}_K(a)}\int_a^b \vartheta(w)\hat{g}_K(w)dw \\
&= \mathbb{E}_{w\sim\hat{G}_K(\cdot|w\in[a,b])}[x(w)]\mathbb{E}_{w\sim\hat{G}_K(\cdot|w\in[a,b])}[\vartheta(w)](\hat{G}_K(b)-\hat{G}_K(a)) \\
&\geq \mathbb{E}_{w\sim\hat{G}_K(\cdot|w\in[a,b])}[x(w)\vartheta(w)](\hat{G}_K(b)-\hat{G}_K(a)) \\
&= \int_a^b x(w)\vartheta(w)\hat{g}_K(w)dw.
\end{aligned}$$

The inequality is an application of Harris inequality, which states that for any nondecreasing function f and nonincreasing function g on $\mathbb{R}$ and any probability measure on $\mathbb{R}$, we have $\mathbb{E}[fg] \leq \mathbb{E}[f]\mathbb{E}[g]$. $\square$

Proof of Theorem 7. Given for any $\epsilon > 0$, take $\bar{c}$ and $\underline{c}$ to satisfy $1-\epsilon/2 < (\Lambda(\bar{c})-\Lambda(\underline{c}))$. Then, since $\hat{G}_K(w) \to \Lambda(w)$ as $K \to \infty$, there exists K_0 such that for all $K > K_0$, we have $1-\epsilon < \hat{G}_K(\bar{c})-\hat{G}_K(\underline{c})$. Lemmas 1 and 2 imply that there also exists K_0 such that for all $K > K_0$, for all $w \in [\underline{c}, \bar{c}]$, $\hat{x}(w) = \bar{x}_K$ for some $\bar{x}_K \in [0,1]$. Let K_0 be the larger one of these two. Then, clearly we have $\Pr(\hat{x}(w) = \bar{x}_K) > 1-\epsilon$. Furthermore, the resource constraint implies that

$$\bar{m} = \int_{-\infty}^{\infty} \hat{x}(w)\hat{g}_K(w)dw \geq \int_{\underline{c}}^{\bar{c}} \hat{x}(w)\hat{g}_K(w)dw > \bar{x}_K(1-\epsilon),$$

or equivalently, $\bar{x}_K < \bar{m}/(1-\epsilon)$. The resource constraint also implies

$$\bar{m} = \int_{-\infty}^{\infty} \hat{x}(w)\hat{g}_K(w)dw \leq \int_{\underline{c}}^{\bar{c}} \hat{x}(w)\hat{g}_K(w)dw + \int_{w\notin[\underline{c},\bar{c}]} \hat{g}_K(w)dw = (1-p)\bar{x}_K + p,$$

where $p = \Pr(w \notin [\underline{c}, \bar{c}]) < \epsilon$. The above inequality can be rewritten as $\bar{x}_K \geq (\bar{m}-p)/(1-p)$. The right-hand side is larger than $(\bar{m}-\epsilon)/(1-\epsilon)$ because it is decreasing in p and $p < \epsilon$. Accordingly, $\bar{x}_K > (\bar{m}-\epsilon)/(1-\epsilon)$. $\square$

B.7 Proof of Proposition 4

Proof. We first prove the following lemma.

Lemma 3. When the reduced value distribution $\hat{G}_K$ has an IDHR, the ironed virtual value $\bar{\vartheta}$ is given by

$$\bar{\vartheta}(w;\hat{G}_K) = \begin{cases} \vartheta(w^{**};\hat{G}_K) & \text{for } w \in [0, w^{**}] \\ \vartheta(w;\hat{G}_K) & \text{for } w \in (w^{**}, +\infty) \end{cases} \tag{27}$$

where w^{**} is the solution of

$$\frac{\int_0^w \vartheta(w';\hat{G}_K)\hat{g}_K(w')dw'}{\hat{G}_K(w)} = \vartheta(w;\hat{G}_K). \tag{28}$$

Proof. Since $\hat{G}_K$ has an IDHR, H is concave for $[0, \hat{G}_K(w^*))$ and convex for $(\hat{G}_K(w^*), +\infty)$. Accordingly, there exists a unique line passing through the origin and being tangent to H. Let the q-coordinate of the point of tangency be q^{**}. Then, the convex hull I of H is given by

$$I(q) = \begin{cases} \dfrac{H(q^{**})}{q^{**}} \cdot q & \text{for } q < q^{**} \\ H(q) & \text{for } q \geq q^{**}, \end{cases}$$

implying that the ironed virtual value is given by (27), where $w^{**} = \hat{G}_K^{-1}(q^{**})$.

Furthermore, since q^{**} is defined as the point of tangency, it solves $H(q)/q = h(q)$, or equivalently,

$$\frac{\int_0^q \vartheta(\hat{G}_K^{-1}(r); \hat{G}_K)dr}{q} = \vartheta(\hat{G}_K^{-1}(q)).$$

Finally, we obtain (28) by replacing q with $\hat{G}_K(w)$ and applying the substitution $r = \hat{G}_K(w)$ to the integral in the left-hand side. □

By Theorem 2.8 of Hartline and Roughgarden (2008), an efficient mechanism maximizes the value of $\hat{x}(w)$ in regions where $\bar{\vartheta}(w; \hat{G}_K)$ is larger while fulfilling the resource constraint and the unit demand condition. The conclusion is immediate from this. □

B.8 Proof of Theorem 8

Proof. By Theorem 5, for sufficiently large K, there exists $\tilde{c} > 0$ such that $\hat{G}_K$ has an IHR in $[0, \tilde{c})$. Take $\underline{c} < \tilde{c}$ and $\bar{c} > w^{**}$. It follows from $G \in \mathbf{VMC}(1/\alpha)$ and Proposition 3 that $\hat{G}_K, \hat{g}_K, \hat{g}'_K$ converges uniformly to $\Phi_\alpha, \phi_\alpha, \phi'_\alpha$ in $[\underline{c}, \bar{c}]$. Accordingly, for any $\delta' > 0$, for sufficiently large K, $\hat{G}_K$ has an IHR for $[0, w^* - \delta')$.

Since $\vartheta(\cdot; \hat{G}_K)$ is decreasing for $[0, w^* - \delta')$, there exists $w^\dagger > w^* - \delta'$ such that $\hat{x}(w) = \bar{x}_K$ for $w \in [0, w^\dagger)$, for some $\bar{x}_K \in [0, 1]$. For a reduced distribution $\hat{G}$, define $\Theta(w; \hat{G})$ by

$$\Theta(w; \hat{G}) := \frac{\displaystyle\int_0^w \vartheta(w'; \hat{G})\hat{g}(w')dw'}{\hat{G}(w)} - \vartheta(w; \hat{G}).$$

For the same reasons used in the proof of Lemma 3, $w^\dagger$ must solve $\Theta(w; \hat{G}_K) = 0$, and w^{**} must solve $\Theta(w; \Phi_\alpha) = 0$. Since Φ_α has an IDHR, $\Theta(\cdot; \Phi_\alpha)$ is decreasing in the neighborhood of w^{**}. Furthermore, for any $\delta'' > 0$, by taking sufficiently large K, we can make

$$|\Theta(w; \hat{G}_K) - \Theta(w; \Phi_\alpha)| < \delta''.$$

Accordingly, for any $\epsilon > 0$, for sufficiently large K, $w^\dagger \in (w^{**} - \epsilon, w^{**} + \epsilon)$. Together with the fact that $\hat{G}_K(w) \to \Phi_\alpha(w)$ uniformly in $[\underline{c}, \bar{c}]$, we have the desired conclusion. □

B.9 Proof of Proposition 5

Proof. First, it is straightforward to see that, for all ρ,

$$RS(\mathcal{M}_{SD}; F^\rho) \geq \mathbb{E}\left[\sum_{k=1}^K m_k \min_{i \in I} v_k^i\right] = \sum_{k=1}^K m_k \mathbb{E}\left[\min_{i \in I} v_k^i\right],$$

$$RS(\mathcal{M}_{FB}^{\rho}; F^{\rho}) \leq \mathbb{E}\left[\sum_{k=1}^{K} m_k \max_{i\in I} v_k^i\right] = \sum_{k=1}^{K} m_k \mathbb{E}\left[\max_{i\in I} v_k^i\right]$$

Furthermore, since we assume that the goods are scarce (i.e., $I > \sum_{k\in K} m_k$), for each $\mathbf{v}$, there exists an agent $i^*(\mathbf{v}) \in I$ who is not allocated. Under VCG, each agent who obtains object k must pay at least $v_k^{i^*(\mathbf{v})}$. Accordingly,

$$\mathbb{E}[(\text{VCG payment})] \geq \sum_{k=1}^{K} m_k \mathbb{E}\left[v_k^{i^*(\mathbf{v})}\right] \geq \sum_{k=1}^{K} m_k \mathbb{E}\left[\min_{i\in I} v_k^i\right]$$

Therefore,

$$\frac{\sum_{k\in K} m_k \mathbb{E}[\min_i v_k^i]}{\sum_{k\in K} m_k \mathbb{E}[\max_i v_k^i]} \leq \frac{RS(\mathcal{M}_{SD}; F^{\rho})}{RS(\mathcal{M}_{FB}^{\rho}; F^{\rho})} \leq 1,$$
$$0 \leq RS(\mathcal{M}_{VCG}; F^{\rho}) \leq \sum_{k=1}^{K} m_k \left(\mathbb{E}\left[\max_{i\in I} v_k^i\right] - \mathbb{E}\left[\min_{i\in I} v_k^i\right]\right).$$

Accordingly, it suffices to show that for each object k, $\mathbb{E}[\min_{i\in I} v_k^i] \to \mu_k$ and $\mathbb{E}[\max_{i\in I} v_k^i] \to \mu_k$ as $\rho \to 1$.

For each $k \in K$, let H_k^{ρ} denote the distribution of $\max_{i\in I} v_k^i$ and L_k^{ρ} that of $\min_{i\in I} v_k^i$. Using standard integral identities,

$$\mathbb{E}\left[\max_{i\in I} v_k^i\right] = \int_0^{\infty} (1 - H_k^{\rho}(w))dw, \qquad \mathbb{E}\left[\min_{i\in I} v_k^i\right] = \int_0^{\infty} (1 - L_k^{\rho}(w))dw.$$

Let $S_k(v) := I(1 - F_k(v))$. Since $\int_0^{\infty} |S_k(w)|dw = I\mu_k < \infty$, S_k is integrable. Furthermore, we have (i) $1-L_k^{\rho}(w) \leq 1-H_k^{\rho}(w)$ by definition, and (ii) $1-H_k^{\rho}(w) \leq S_k(w)$ by the union bound. Thus, by the dominated convergence theorem, once we show pointwise convergence, $H_k^{\rho}(w) \to F_k(w)$ and $L_k^{\rho}(w) \to F_k(w)$ at all continuity points of F_k, then the dominated convergence theorem implies $\mathbb{E}[\max_i v_j^i] \to \mu_j$ and $\mathbb{E}[\min_i v_j^i] \to \mu_j$.

Finally, we prove pointwise convergence. Let $X_k^i := (v_k^i - \mu_k)/\sigma_k$ so that for all $i \in I$, $k \in K$, and $j \neq i$, we have $\mathbb{E}[X_k^i] = 0$, $\text{Var}(X_k^i) = 1$, and $\text{Cov}(X_k^i, X_k^j) \geq \rho$. Define the differences $D_k^i := X_k^1 - X_k^i$. Then, $\mathbb{E}[D_j^i] = 0$ and $\text{Var}(D_k^i) = 2 - 2\text{Cov}(X_k^1, X_k^i) \leq 2(1-\rho)$.

By the union bound and Chebyshev's inequality,

$$\begin{aligned}
\Pr\left(\max_{i\neq 1} |D_k^i| \geq \varepsilon\right) &\leq \sum_{i\neq 1} \Pr(|D_k^i| \geq \varepsilon) \\
&\leq \sum_{i\neq 1} \frac{\text{Var}(D_k^i)}{\varepsilon^2} \\
&\leq \frac{2(I-1)(1-\rho)}{\varepsilon^2} \xrightarrow[\rho\to 1]{} 0.
\end{aligned}$$

Let $M_k \coloneqq \max_{i \neq 1} |v_k^1 - v_k^i|$. Then,

$$\begin{aligned}
\Pr\left(\max_{i \neq 1} |v_k^1 - v_k^i| \geq \varepsilon\right) &\leq \sum_{i \neq 1} \Pr\left(|v_k^1 - v_k^i| \geq \varepsilon\right) \\
&\leq \sum_{i \neq 1} \Pr\left(|D_k^i| \geq \frac{\varepsilon}{\sigma_k}\right) \xrightarrow[\rho \to 1]{} 0.
\end{aligned}$$

Fix a continuity point v of F_k. Then

$$\begin{aligned}
|H_k^\rho(v) - F_k(v)| &= \Pr\Big(\max_i v_k^i > v \text{ and } v_k^1 \leq v\Big) \\
&\leq \Pr(M_k \geq \varepsilon) + \Pr\big(M_k < \varepsilon \text{ and } v_k^1 \in (v - \varepsilon, v + \varepsilon)\big) \\
&\leq \Pr(M_k \geq \varepsilon) + \big(F_k(v + \varepsilon) - F_k(v - \varepsilon)\big).
\end{aligned}$$

The first term vanishes as $\rho \to 1$, and the second term can be made arbitrarily small by choosing ε. Thus $H_k^\rho(v) \to F_k(v)$. We can prove that $L_k^\rho(v) \to F_k(v)$ in a similar manner. □

B.10 Efficient Mechanism for a Fréchet Case

We transform (28) to derive (10) as follows.

$$\begin{aligned}
0 &= \frac{\int_0^v \vartheta(w; \Phi_\alpha)\phi_\alpha(w)dw}{\Phi_\alpha(v)} - \vartheta(v; \Phi_\alpha), \\
0 &= \frac{\int_0^v (1 - \Phi_\alpha(w))dw}{\Phi_\alpha(v)} - \frac{1 - \Phi_\alpha(v)}{\phi_\alpha(v)}, \\
0 &= \int_0^v (1 - e^{-w^{-\alpha}})dw - \frac{1 - e^{-v^{-\alpha}}}{\alpha v^{-\alpha-1}}, \\
0 &= \left[w - \frac{1}{\alpha}\Gamma\left(-\frac{1}{\alpha}, w^{-\alpha}\right)\right]_0^v - \frac{1 - e^{-v^{-\alpha}}}{\alpha v^{-\alpha-1}}, \\
0 &= v - \frac{1}{\alpha}\Gamma\left(-\frac{1}{\alpha}, v^{-\alpha}\right) - \frac{1 - e^{-v^{-\alpha}}}{\alpha v^{-\alpha-1}}, \\
0 &= v - \frac{1}{\alpha}(-\alpha)\left[\Gamma\left(1 - \frac{1}{\alpha}, v^{-\alpha}\right) - ve^{-v^{-\alpha}}\right] - \frac{1 - e^{-v^{-\alpha}}}{\alpha v^{-\alpha-1}}, \\
0 &= (1 - e^{-v^{-\alpha}})v + \Gamma\left(\frac{\alpha - 1}{\alpha}, v^{-\alpha}\right) - \frac{1}{\alpha}v^{\alpha+1}(1 - e^{-v^{-\alpha}}).
\end{aligned}$$

B.11 Random Favorite (RF) Mechanism

We characterize the RF mechanism for the following setting:

- A continuous i.i.d. market.
- Two object types: $K = 2$.

- The marginal value distribution is standard exponential: $G(v) = 1 - e^{-v}$.
- The capacities of objects 1 and 2 are given by (m_1, m_2), where $m_1 \geq m_2 > 0$ and $m_1 + m_2 \leq 1$.

RF offers two options: without making any payments, (i) obtain object 1 with probability $a \in (0, 1]$ and obtain object 2 with a zero probability, or (ii) obtain object 2 with probability $b \in (0, 1]$ and obtain object 1 with a zero probability. An agent claims object 1 if $av_1 \geq bv_2$, or equivalently, $av_1/b \geq v_2$ and claims object 2 otherwise.

Conditional on v_1, the probability that an agent claims object 2 is $1 - G(av_1/b) = e^{-av_1/b}$. Thus, the total mass of agents who claim object 2 is

$$\int_0^\infty e^{-\frac{a}{b}v_1} e^{-v_1} dv_1 = \frac{b}{a+b}.$$

Accordingly, given (a, b), the demand for object 1 is $a^2/(a+b)$ and the demand for object 2 is $b^2/(a+b)$. The market clearing condition requires that $a^2/(a+b) = m_1$ and $b^2/(a+b) = m_2$. Solving these equations, we obtain $a = m_1 + \sqrt{m_1 m_2}$ and $b = m_2 + \sqrt{m_1 m_2}$ as a unique (a, b) satisfying the agent's optimality condition and the market clearing condition.

The residual surplus achieved by RF is

$$\begin{aligned}
&RS(\mathcal{M}_{RF}) \\
&= a \int_0^\infty v_1 (1 - e^{-\frac{a}{b}v_1}) e^{-v_1} dv_1 + b \int_0^\infty v_2 (1 - e^{-\frac{b}{a}v_2}) e^{-v_2} dv_2 \\
&= a \left(1 - \frac{b^2}{(a+b)^2}\right) + b \left(1 - \frac{a^2}{(a+b)^2}\right) \\
&= (a+b) - \frac{ab(a+b)}{(a+b)^2} \\
&= m_1 + m_2 + 2\sqrt{m_1 m_2} - \sqrt{m_1 m_2} \\
&= m_1 + m_2 + \sqrt{m_1 m_2}.
\end{aligned}$$

Next, we compute the residual surplus achieved by SD. Since $m_1 \geq m_2$ and objects 1 and 2 are equally popular, in the beginning, all agents obtain their favorite objects, and object 2 is exhausted when the first $2m_2 (< m_1 + m_2 < 1)$ unit of agents makes a choice. Afterward, all agents obtain object 1, and when an additional $m_1 - m_2$ unit of agents makes a choice, object 1 is also exhausted. Accordingly, the residual surplus achieved by SD is

$$\begin{aligned}
&RS(\mathcal{M}_{SD}) \\
&= 2m_2 \int_0^\infty v g_2(v) dv + (m_1 - m_2) \int_0^\infty v g(v) dv \\
&= 2m_2 \int_0^\infty v (2e^{-v} - 2e^{-2v}) dv + (m_1 - m_2) \int_0^\infty v e^{-v} dv \\
&= 3m_2 + (m_1 - m_2)
\end{aligned}$$

$$= m_1 + 2m_2.$$

Since we assume $m_1 \geq m_2$, $RS(\mathcal{M}_{RF}) \geq RS(\mathcal{M}_{SD})$ always holds, and the equality holds if and only if $m_1 = m_2$. In Figure A.2, we plot

$$\frac{RS(\mathcal{M}_{RF}) - RS(\mathcal{M}_{SD})}{RS(\mathcal{M}_{RF})} = \frac{\sqrt{m_1 m_2} - \min\{m_1, m_2\}}{m_1 + m_2 + \sqrt{m_1 m_2}}$$

to illustrate the performance difference between RF and SD.