Efficient Random Assignment with Cardinal and Ordinal Preferences: Online Appendix

James C. D. Fisher

December 11, 2018

1 Introduction

This document collects several results, which supplement those in the main text. Specifically:

- In Section 2, we formally define ex-post Pareto efficiency and show that every interim efficient mixed assignment is ex-post Pareto efficient.
- In Section 3, we establish that a player has the same payoff to sending a report to the Simple Mechanism in the replication extension and in the corresponding T-replication case.
- In Section 4, we extend the baseline model to allow players to have correlated true values, as well as arbitrary (and correlated) knowledge of their preferences. We establish that our results are robust to this extension.
- In Section 5, we study the robustness of the Simple Mechanism with respect to its implicit requirement that the designer know the structure of preferences and show how one can use historical data to close key gaps.

2 Ex-Post Pareto Efficiency

We have in mind that an ex-post Pareto efficient (mixed) assignment $\tilde{\boldsymbol{\phi}}$ is one where no player can be made better off without harming another after all uncertainty about true values is resolved. More precisely, let $\{\boldsymbol{v}_i = (v_{\eta i}, \ldots, v_{Ki})\}_{i\in\mathcal{N}}$ be the players' true values, then a (mixed) assignment $\tilde{\boldsymbol{\phi}}$ is **ex-post Pareto efficient** if, for each pure assignment ϕ in its support, there is no alternative pure assignment ϕ' such that $v_{\phi'(i)i} \geq v_{\phi(i)i}$ for all $i \in \mathcal{N}$ and $v_{\phi'(i)i} > v_{\phi(i)i}$ for some $i \in \mathcal{N}$. Such an assignment is also "ex-post individually rational" in the sense that that each player weakly prefers her assignment to η after all uncertainty about true values is resolved, i.e., $v_{\phi'(i)i} \geq v_{i\eta}$ for every ϕ' in the support of $\tilde{\boldsymbol{\phi}}$.

Our main result is the following.

Proposition OA1. Interim Pareto Efficiency Implies Ex-Post Pareto Efficiency.
If a mixed assignment $\tilde{\boldsymbol{\phi}}$ is interim efficient, then it is also ex-post Pareto efficient.

To prove the proposition, we need to establish that the consistency relationship is convex.

Lemma OA1. Convexity of the Consistency Relation.
Let $\preceq \in \mathcal{B}'$ and let $\{\boldsymbol{x}(k)\}_{k=1}^{L}$ be a finite collection of valuations in $\mathbb{R}^{K+1}$ that are consistent with $\preceq$. If $\{\lambda(k)\}_{k=1}^{L}$ is a finite collection of non-negative weights such that $\sum_{k=1}^{L} \lambda(k) = 1$, then $\sum_{k=1}^{L} \boldsymbol{x}(k)\lambda(k)$ is consistent with $\preceq$.

Proof. Straightforward. We argue by induction on L. This is trivial if $L = 1$, so we begin with $L = 2$. Let $\boldsymbol{x}(1) = (x_\eta, x_1, \ldots, x_K)$ and $\boldsymbol{x}(2) = (x'_\eta, x'_1 \ldots, x'_K)$, and let $\lambda \in [0, 1]$. Let $\boldsymbol{x}^\lambda = \boldsymbol{x}(1)\,\lambda + \boldsymbol{x}(2)\,(1-\lambda) = (x^\lambda_1, \ldots, x^\lambda_J)$. We need to show that $\boldsymbol{x}^\lambda$ is consistent with $\preceq$. Consider two elements o and o' of $\mathcal{O}'$, and observe that $x^\lambda_o = \lambda x_o + (1-\lambda)x'_o$ and $x^\lambda_{o'} = \lambda x_{o'} + (1-\lambda)x'_{o'}$. If $o \preceq o'$, then $x_o \leq x_{o'}$ and $x'_o \leq x'_{o'}$ by consistency, so $x^\lambda_o \leq x^\lambda_{o'}$. If $o \npreceq o'$, then $x_o > x_{o'}$ and $x'_o > x'_{o'}$, so $x^\lambda_o > x^\lambda_{o'}$, i.e., $x^\lambda_o \nleq x^\lambda_{o'}$. Thus, $x^\lambda_o \leq x^\lambda_{o'} \iff o \preceq o'$ and $x^\lambda_o \nleq x^\lambda_{o'} \iff o \npreceq o'$, so $\boldsymbol{x}^\lambda$ is consistent with $\preceq$.

We now suppose that the lemma is true $L-1$ and show that it is true at L. Let $\boldsymbol{y} = \sum_{k=1}^{L-1} \boldsymbol{x}(k)\lambda(k)$. If either $\lambda(L) = 0$ or $\lambda(L) = 1$, then the lemma is trivially true at L. So, we take $\lambda(L) \in (0, 1)$. By the induction hypothesis, $\boldsymbol{z} = \sum_{k=1}^{L-1} \boldsymbol{x}(k)\frac{\lambda(k)}{1-\lambda(L)}$ is consistent with $\preceq$ because $\sum_{k=1}^{L-1} \frac{\lambda(k)}{1-\lambda(L)} = 1$. Since $x(L)$ is also consistent with $\preceq$, we have $x(L)\lambda(L) + (1-\lambda(L))\boldsymbol{z} = \sum_{k=1}^{L} \boldsymbol{x}(k)\lambda(k)$ is consistent with $\preceq$ by the argument in the first paragraph. □

Corollary OA1. Consistency of Expectations.
For each player i, let θ_i denote her type and $\boldsymbol{v}^\dagger_i$ denote her expected true values per equation (2.2) of the main text. If $\theta_i \in \mathcal{B}'$, then $\boldsymbol{v}^\dagger_i$ is consistent with θ_i.

Proof. Follows immediately from Lemma OA1 since, for each player i, $\boldsymbol{v}^\dagger_i$ is the convex combination of points in $I(\theta_i)$, each of which is consistent with $\boldsymbol{v}_i$. □

For each joint vector of true values $\tilde{\boldsymbol{v}} = \{\boldsymbol{v}_i\}_{i\in\mathcal{N}} \in \mathcal{V}^N$, let $\mathcal{W}(\tilde{\boldsymbol{v}}) = (\mathcal{W}_1(\tilde{\boldsymbol{v}}), \ldots, \mathcal{W}_N(\tilde{\boldsymbol{v}})) \subset \Theta^N$ denote the set types that may be chosen by nature with strictly positive possibility for the players. That is, $\mathcal{W}_i(\tilde{\boldsymbol{v}})$ consists of $\boldsymbol{v}_i$ and the order $\preceq_i$ that is consistent with $\boldsymbol{v}_i$.

Proof of Proposition OA1. Straightforward. Let $\{\boldsymbol{v}_i\}_{i\in\mathcal{N}} = \tilde{\boldsymbol{v}}$ be the players' true values, let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_N) \in \mathcal{W}(\tilde{\boldsymbol{v}})$ be their types, and let $\{\boldsymbol{v}^\dagger_i\}_{i\in\mathcal{N}}$ be their expected given these types (see equation (2.2) of the main text). For each player i, let $\boldsymbol{v}_i = (v_{\eta i}, v_{1i}, \ldots, v_{Ni})$ and let $\boldsymbol{v}^\dagger_i = (v^\dagger_{\eta i}, v^\dagger_{1i}, \ldots, v^\dagger_{Ni})$.

We argue by contradiction. Suppose that $\tilde{\boldsymbol{\phi}}^\star$ is interim efficient but not ex-post efficient. Then there is a pure assignment ϕ' in the support of $\tilde{\boldsymbol{\phi}}^\star$ such that $v_{\phi'(i)i} \geq v_{\phi(i)i}$ for each $i \in \mathcal{N}$ and $v_{\phi'(i)i} > v_{\phi(i)i}$ for some $i \in \mathcal{N}$. We'll show that this implies $v^\dagger_{\phi'(i)i} \geq v^\dagger_{\phi(i)i}$ for each $i \in \mathcal{N}$ and $v^\dagger_{\phi'(i)i} > v^\dagger_{\phi(i)i}$ for some $i \in \mathcal{N}$. Given this, we have $\sum_{i\in\mathcal{N}} u_i(\phi'|\theta_i) > \sum_{i\in\mathcal{N}} u_i(\phi|\theta_i)$, i.e., ϕ is interim inefficient at $\boldsymbol{\theta}$. Since an interim efficient mixed assignment only randomizes over interim efficient pure assignments, it follows that $\tilde{\boldsymbol{\phi}}^\star$ is not interim efficient. Thus, we obtain the desired contradiction.

For each player i, it remains to show (i) that $v_{\phi'(i)i} \geq v_{\phi(i)i}$ implies $v^\dagger_{\phi'(i)i} \geq v^\dagger_{\phi(i)i}$ and (ii) that $v_{\phi'(i)i} > v_{\phi(i)i}$ implies $v^\dagger_{\phi'(i)i} > v^\dagger_{\phi(i)i}$. If $\theta_i \in \mathcal{V}$, then we automatically have (i) and (ii) since i knows her true values after learning her type. If $\theta_i \in \mathcal{B}'$, then let $\preceq_i = \theta_i$. The consistency of $\preceq_i$ and $\boldsymbol{v}_i$ gives that $v_{\phi'(i)i} \geq v_{\phi(i)i}$ implies $\phi(i) \preceq_i \phi'(i)$. Since Corollary OA1 gives $\boldsymbol{v}_i^\dagger$ is consistent with $\preceq_i$, we also have that $v_{\phi'(i)i} \geq v_{\phi(i)i}$ implies $v^\dagger_{\phi'(i)i} \geq v^\dagger_{\phi(i)i}$, i.e., (i) obtains. As to (ii): if $v_{\phi'(i)i} > v_{\phi(i)i}$, then consistency gives $\phi(i) \preceq_i \phi'(i)$ and $\phi'(i) \npreceq_i \phi(i)$ and thus Corollary OA1 implies $v^\dagger_{\phi'(i)i} > v^\dagger_{\phi(i)i}$. □

3 Replication Payoff Equivalency

In this section, we show that a player's payoff to sending report r_i to the Simple Mechanism in the replication extension is the same as her payoff to sending $h(r_i)$ to the Simple Mechanism in the corresponding T-replication case. (Recall that h is the bijection that maps types between the replication extension to the corresponding T-replication case; see the Appendix of the main text for details.)

Lemma OA2. Payoff of the Replication Extension and the T-Replication Case.
Let $T = \psi(N)$, then $U_i^{M_S}(r_i|\theta_i) = \ddot{U}_i^{M_S}(h(r_i)|h(\theta_i))$ for each player i, each report $r_i \in \Theta$, and each $\theta_i \in \Theta$.

Proof. (Largely algebraic.) Consider player i. Since $T = \psi(N)$, the set of possible assignments

$$\ddot{\mathcal{O}}'_T = \{\eta, 1_1, 1_2, \dots, 1_T, 2_1, \dots, 2_T, \dots, K_1, \dots, K_T\}$$

is the same in both environments and thus so too is the set of possible assignments. We thus denote the common set of assignments by Φ and index it $j = 1, \dots, |\Phi|$.

For the replication extension, we have

$$\begin{aligned}
U_i^{M_S}(r_i|\theta_i) &= \sum_{\boldsymbol{\theta}_{-i} \in \Theta^{N-1}} u_i(M_S(r_i, \boldsymbol{\theta}_{-i})|\theta_i) \Pr(\boldsymbol{\theta}_{-i}) \\
&= \sum_{\boldsymbol{\theta}_{-i} \in \Theta^{N-1}} \sum_{j=1}^{|\Phi|} u_i(\phi_j|\theta) \tilde{\phi}_j(r_i, \boldsymbol{\theta}_{-i}) \Pr(\boldsymbol{\theta}_{-i}) \\
&= \sum_{\boldsymbol{\theta}_{-i} \in \Theta^{N-1}} \sum_{j=1}^{|\Phi|} \sum_{o_l \in \ddot{\mathcal{O}}'_T} v^\dagger_{o_l} \mathbb{I}(\phi_j(i) = o_l) \tilde{\phi}_j(r_i, \boldsymbol{\theta}_{-i}) \Pr(\boldsymbol{\theta}_{-i}),
\end{aligned}$$

where $v^\dagger_{o_l}$ denotes i's expected true value to each $o_l \in \ddot{\mathcal{O}}'_T$, $\mathbb{I}(\cdot)$ is the classic indicator function, $\tilde{\phi}_j(r_i, \boldsymbol{\theta}_{-i})$ is the probability of each pure assignment ϕ_j under the Simple Mechanism in the replication extension, and $\Pr(\boldsymbol{\theta}_{-i})$ is the probability of the type-vector $\boldsymbol{\theta}_{-i}$ in the replication extension. For the T-replication case, we have

$$\ddot{U}_i^{M_S}(h(r_i)|h(\theta_i)) = \sum_{\boldsymbol{\theta}_{-i}\in\ddot{\Theta}^{N-1}} \sum_{j=1}^{|\Phi|} \sum_{o_l\in\ddot{\mathcal{O}}'_T} \ddot{v}^{\dagger}_{o_l}\mathbb{I}(\phi_j(i)=o_l)\ddot{\tilde{\phi}}_j(h(r_i),\boldsymbol{\theta}_{-i})\ddot{\Pr}(\boldsymbol{\theta}_{-i}),$$

where $\ddot{v}^{\dagger}_{o_l}$ denotes i's expected true value to each $o_l \in \ddot{\mathcal{O}}'_T$, $\ddot{\tilde{\phi}}_j(h(r_i),\boldsymbol{\theta}_{-i})$ is the probability of each pure assignment ϕ_j under the Simple Mechanism in the T-replication case, and $\ddot{\Pr}(\boldsymbol{\theta}_{-i})$ is the probability of the type vector $\boldsymbol{\theta}_{-i}$ in the T-replication case.

Three key facts follow:

Fact 1 We have $v^{\dagger}_{o_l} = \ddot{v}^{\dagger}_{o_l}$ for each $o_l \in \ddot{\mathcal{O}}'_T$.

Proof. Follows from the construction of the T-replication case and the bijective nature of h. The algebra is straight-forward and omitted. △

Fact 2 We have that $\Pr(\boldsymbol{\theta}_{-i}) = \ddot{\Pr}(h(\boldsymbol{\theta}_{-i}))$ for each $\boldsymbol{\theta}_{-i} \in \Theta^{N-1}$.

Proof. Follows from the construction of the T-replication case and the bijective nature of h. The algebra is straight-forward and omitted. △

Fact 3 When players report $\boldsymbol{r} = (r_1,\ldots,r_N) \in \Theta^N$ in the replication environment, then the Simple Mechanism randomizes over the same assignments as when they report $h(\boldsymbol{r}) \in \ddot{\Theta}^N$ in the T-replication case. Thus, $\tilde{\phi}_j(\boldsymbol{r}) = \ddot{\tilde{\phi}}_j(h(\boldsymbol{r}))$ for each $j = 1,\ldots,|\Phi|$ (since h is bijective).

Proof. To see this, consider player k, who makes report r_k in the replication extension and $h(r_k)$ in the T-replication case. There are two cases: $r_k \in \mathcal{V}$ or $r_k \in \Theta\backslash\mathcal{V}$.[1] If the former, then $r_k = (v_\eta, v_1, \ldots, v_K)$ and

$$h(r_k) = (\eta, \underbrace{v_1,\ldots,v_1}_{T\text{ copies}}, \underbrace{v_2,\ldots,v_2}_{T\text{ copies}}, \ldots, \underbrace{v_K,\ldots,v_K}_{T\text{ copies}})$$

by construction. Thus, the Simple Mechanism in the replication environment and the Simple Mechanism in the T-replication case both estimate that k has the same payoff.

If $r_k \in \Theta\backslash\mathcal{V}$, then $h(r_k) \in \ddot{\Theta}\backslash\ddot{\mathcal{V}}$. Hence, the Simple Mechanism estimates k's

[1]Recall $\mathcal{V}$ and Θ are the true value and type spaces in the replication extension, while $\ddot{\mathcal{V}}$ and $\ddot{\Theta}$ are the true value and type spaces in the T-replication case.

expected true values in the replication extension are

$$(\hat{v}^\dagger_\eta, \hat{v}^\dagger_1, \ldots, \hat{v}^\dagger_K) = \sum_{\boldsymbol{v} \in I(r_k)} \boldsymbol{v} \frac{f_{\boldsymbol{v}}(\boldsymbol{v})}{\sum_{\boldsymbol{v} \in I(r_k)} f_{\boldsymbol{v}}(\boldsymbol{v})},$$

where $I(r_k) = \{\boldsymbol{v} \in \mathcal{V} | \boldsymbol{v} \text{ is consistent with } r_k\}$, and in the T-replication case, the Simple Mechanism estimates i's expected true values are

$$(\hat{v}^\ddagger_\eta, \hat{v}^\ddagger_{1_1}, \ldots, \hat{v}^\ddagger_{2_1}, \ldots, \hat{v}^\ddagger_{K_T}) = \sum_{\boldsymbol{v} \in \ddot{I}(h(r_k))} \boldsymbol{v} \frac{f_{\ddot{\boldsymbol{v}}}(\boldsymbol{v})}{\sum_{\boldsymbol{v} \in \ddot{I}(h(r_k))} f_{\ddot{\boldsymbol{v}}}(\boldsymbol{v})},$$

where $\ddot{I}(h(r_k)) = \{\boldsymbol{v} \in \ddot{\mathcal{V}} | \boldsymbol{v} \text{ is consistent with } h(r_k)\}$. Since r_k and $h(r_k)$ encode the same ordinal preference (and only differ in the fact that the latter is defined over additional copies of the objects), we have that $\boldsymbol{v} \in I(r_k)$ if and only if $h(\boldsymbol{v}) \in \ddot{I}(h(r_k))$. Combining this with the fact that $f_{\ddot{\boldsymbol{v}}}(\boldsymbol{v}) = f_{\boldsymbol{v}}(h^{-1}(\boldsymbol{v}))$, where h^{-1} is the inverse of h, implies $\sum_{\boldsymbol{v} \in I(r_k)} f_{\boldsymbol{v}}(\boldsymbol{v}) = \sum_{\boldsymbol{v} \in \ddot{I}(h(r_k))} f_{\ddot{\boldsymbol{v}}}(\boldsymbol{v})$. Hence,

$$\begin{aligned}(\hat{v}^\ddagger_\eta, \hat{v}^\ddagger_{1_1}, \ldots, \hat{v}^\ddagger_{2_1}, \ldots, \hat{v}^\ddagger_{K_T}) &= \sum_{\boldsymbol{v} \in \ddot{I}(h(r_k))} \boldsymbol{v} \frac{f_{\ddot{\boldsymbol{v}}}(\boldsymbol{v})}{\sum_{\boldsymbol{v} \in \ddot{I}(h(r_k))} f_{\ddot{\boldsymbol{v}}}(\boldsymbol{v})}, \\ &= \sum_{\boldsymbol{v} \in I(r_k)} h(\boldsymbol{v}) \frac{f_{\boldsymbol{v}}(\boldsymbol{v})}{\sum_{\boldsymbol{v} \in I(r_k)} f_{\boldsymbol{v}}(\boldsymbol{v})}.\end{aligned}$$

Since

$$h(v_\eta, v_1, \ldots, v_K) = (\eta, \underbrace{v_1, \ldots, v_1}_{T \text{ copies}}, \underbrace{v_2, \ldots, v_2}_{T \text{ copies}}, \ldots, \underbrace{v_K, \ldots, v_K}_{T \text{ copies}}),$$

we have

$$\begin{aligned}&\sum_{\boldsymbol{v} \in I(r_k)} h(\boldsymbol{v}) \frac{f_{\boldsymbol{v}}(\boldsymbol{v})}{\sum_{\boldsymbol{v} \in I(r_k)} f_{\boldsymbol{v}}(\boldsymbol{v})} \\ &= (\sum_{\boldsymbol{v}=(v_\eta, v_1, \ldots, v_K) \in I(r_k)} v_\eta \frac{f_{\boldsymbol{v}}(\boldsymbol{v})}{\sum_{\boldsymbol{v} \in I(r_k)} f_{\boldsymbol{v}}(\boldsymbol{v})}, ((\sum_{\boldsymbol{v}=(v_\eta, v_1, \ldots, v_K) \in I(r_k)} v_o \frac{f_{\boldsymbol{v}}(\boldsymbol{v})}{\sum_{\boldsymbol{v} \in I(r_k)} f_{\boldsymbol{v}}(\boldsymbol{v})})_{\times T})_{o \in \mathcal{O}}).\end{aligned}$$

It is evident that the first term is $\hat{v}^\dagger_\eta$ and that the term inside the double parentheses is $\hat{v}^\dagger_o$. It follows that $\hat{v}^\dagger_\eta = \hat{v}^\ddagger_\eta$ and, for each $l \in \{1, \ldots, T\}$, that $\hat{v}^\dagger_1 = \hat{v}^\ddagger_{1_l}$, $\hat{v}^\dagger_2 = \hat{v}^\ddagger_{2_l}$, ..., and $\hat{v}^\dagger_K = \hat{v}^\ddagger_{K_l}$. Thus, the Simple Mechanism in the replication environment and the Simple Mechanism in the T-replication case both estimate that k has the same payoff.

Since k was a generic player and r_k a generic report, it follows that after getting reports

$\boldsymbol{r} = (r_1, \ldots, r_N)$ the Simple Mechanism in the replication extension has the same objective function as the Simple Mechanism in the T-replication case does after getting reports $h(\boldsymbol{r})$. Hence, the two mechanisms randomize over the same set of pure assignments. $\triangle$

In light of these facts, we have,

$$\begin{aligned}
U_i^{M_S}(r_i|\theta_i) &= \sum_{\boldsymbol{\theta}_{-i}\in\Theta^{N-1}} \sum_{j=1}^{|\Phi|} \sum_{o_l\in\ddot{O}'} v_{o_l}^{\dagger}\mathbb{I}(\phi_j(i)=o_l)\tilde{\phi}_j(r_i,\boldsymbol{\theta}_{-i})\Pr(\boldsymbol{\theta}_{-i}) \\
&= \sum_{\boldsymbol{\theta}_{-i}\in\Theta^{N-1}} \sum_{j=1}^{|\Phi|} \sum_{o_l\in\ddot{O}'} \ddot{v}_{o_l}^{\dagger}\mathbb{I}(\phi_j(i)=o_l)\ddot{\tilde{\phi}}_j(h(r_i),h(\boldsymbol{\theta}_{-i}))\ddot{\Pr}(h(\boldsymbol{\theta}_{-i})) \\
&= \sum_{\boldsymbol{\theta}_{-i}\in\ddot{\Theta}^{N-1}} \sum_{j=1}^{|\Phi|} \sum_{o_l\in\ddot{O}'} \ddot{v}_{o_l}^{\dagger}\mathbb{I}(\phi_j(i)=o_l)\ddot{\tilde{\phi}}_j(h(r_i),\boldsymbol{\theta}_{-i})\ddot{\Pr}(\boldsymbol{\theta}_{-i}) = \ddot{U}_i^{M_S}(h(r_i)|h(\theta_i)),
\end{aligned}$$

where the last line is due to re-indexing. The proposition then follows from the fact that all players with the same type have the same payoff. $\square$

4 Partially Known and Correlated Preferences

In this section, we develop a generalization of our model that allows players to have correlated true values, as well as arbitrary partial (and correlated) knowledge of their preferences – e.g., a player may know she values object o'' twice as much as object o', while only knowing that o' is strictly preferred to object o. We establish that all of the results from the main text are robust to this extension.

Economic Environment

We generalize the baseline model of the main text by changing the way nature draws true values and passes information about them to the players. Specifically, nature first draws a vector of true values for all of the players $\bar{\boldsymbol{v}} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N) \in \mathcal{V}^N$ according to the strictly positive probability mass function $f_{\bar{\boldsymbol{v}}}(\bar{\boldsymbol{v}})$, where $f_{\bar{\boldsymbol{v}}} : \mathcal{V}^N \to (0,1]$ such that $\sum_{\bar{\boldsymbol{v}}\in\mathcal{V}^N} f_{\bar{\boldsymbol{v}}}(\bar{\boldsymbol{v}}) = 1$. Subsequently, nature draws a vector of types $\boldsymbol{t} = (t_1, \ldots, t_N) \in \mathcal{T}^N$ according to the conditional probability mass function $f_{\boldsymbol{t}|\bar{\boldsymbol{v}}}(\boldsymbol{t}|\bar{\boldsymbol{v}})$, where $\mathcal{T}$ is the finite set of all possible types and $f_{\boldsymbol{t}|\bar{\boldsymbol{v}}} : \mathcal{T}^N \to [0,1]$ such that $\sum_{\boldsymbol{t}\in\mathcal{T}^N} f_{\boldsymbol{t}|\bar{\boldsymbol{v}}}(\boldsymbol{t}|\bar{\boldsymbol{v}}) = 1$ for all $\bar{\boldsymbol{v}} \in \mathcal{V}^N$. We assume that, for each $\boldsymbol{t} \in \mathcal{T}^N$, we have $f_{\boldsymbol{t}|\bar{\boldsymbol{v}}}(\boldsymbol{t}|\bar{\boldsymbol{v}}) > 0$ for some $\bar{\boldsymbol{v}} \in \mathcal{V}^N$. Next nature shows each player i her type t_i – a player's type is her "signal" about her true values. (As in the baseline model of the main text, types are private information.) The probability that i has type t_i given true values $\bar{\boldsymbol{v}}$, is $f_{t|\bar{\boldsymbol{v}}}(t_i|\bar{\boldsymbol{v}}) = \sum_{\boldsymbol{t}\in A(t_i)} f_{\boldsymbol{t}|\bar{\boldsymbol{v}}}(\boldsymbol{t}|\bar{\boldsymbol{v}})$, where $A(t_i) = \{(t'_1, \ldots, t'_i, \ldots, t'_N) \in \mathcal{T}^N | t'_i = t_i\}$ is the set of type vectors where i has type t_i, the probability i has type t_i is $f_t(t_i) = \sum_{\bar{\boldsymbol{v}}\in\mathcal{V}^N} f_{t|\bar{\boldsymbol{v}}}(t_i|\bar{\boldsymbol{v}}) f_{\bar{\boldsymbol{v}}}(\bar{\boldsymbol{v}})$, and the probability of $\boldsymbol{t}$ is $f_{\boldsymbol{t}}(\boldsymbol{t}) = \sum_{\bar{\boldsymbol{v}} \in \mathcal{V}^N} f_{\boldsymbol{t}|\bar{\boldsymbol{v}}}(\boldsymbol{t}|\bar{\boldsymbol{v}}) f_{\bar{\boldsymbol{v}}}(\bar{\boldsymbol{v}})$. Notice that f_t and $f_{\boldsymbol{t}}$ are strictly positive.

Given type t_i, the vector of players' expected true values is

$$(\boldsymbol{v}_1^\dagger, \ldots, \boldsymbol{v}_i^\dagger, \ldots, \boldsymbol{v}_N^\dagger) = \sum_{\bar{\boldsymbol{v}} \in \mathcal{V}^N} \bar{\boldsymbol{v}} \frac{f_{\bar{\boldsymbol{v}}}(\bar{\boldsymbol{v}}) f_{t|\bar{\boldsymbol{v}}}(t_i|\bar{\boldsymbol{v}})}{f_t(t_i)}. \tag{4.1}$$

Thus, upon observing her type t_i, player i's vector of expected true values is $(v_{\eta i}^\dagger, v_{1i}^\dagger, \ldots, v_{oi}^\dagger, \ldots, v_{Ki}^\dagger) = \boldsymbol{v}_i^\dagger$. We represent this vector as a payoff function $u_i(\cdot|t_i)$ by setting $u_i(o|t_i) = v_{oi}^\dagger$ for each $o \in \mathcal{O}'$.

This extension is equivalent to the baseline model of the main text when true values are independently and identically distributed and each player has the same chance of observing her true values. Specifically, for $\alpha \in (0,1)$, then the models are equivalent when $f_{\bar{\boldsymbol{v}}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i, \ldots, \boldsymbol{v}_N) = \prod_{i=1}^N f_{\boldsymbol{v}}(\boldsymbol{v}_i)$ for some probability mass function $f_{\boldsymbol{v}}(\boldsymbol{v})$, (ii) $\mathcal{T} = \Theta$, (iii) $\alpha \in (0,1)$, and (iii) $f_{\boldsymbol{t}|\bar{\boldsymbol{v}}}(t_1, \ldots, t_i, \ldots, t_N|\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i, \ldots, \boldsymbol{v}_N) = \prod_{i=1}^N f_{t|\boldsymbol{v}}(t_i|\boldsymbol{v}_i)$ where

$$f_{t|\boldsymbol{v}}(t|\boldsymbol{v}) = \begin{cases} \alpha & t = \boldsymbol{v} \\ 1-\alpha & t = c(\boldsymbol{v}) \\ 0 & \text{else.} \end{cases}$$

(Recall that $c(\boldsymbol{v})$ gives the element of $\mathcal{B}'$ that is consistent with $\boldsymbol{v}$.) Equivalency also holds for the cases of $\alpha = 1$ or $\alpha = 0$; the details are omitted for expositional simplicity.

We need to modify our definitions to account for the correlation of types. We do this now for interim efficiency and ϵ-Bayesian incentive compatibility; our other definitions extend similarly. Regarding interim efficiency, we write $u_i(\cdot|\boldsymbol{t})$ for player i's (expected) payoff given the joint type vector $\boldsymbol{t} \in \mathcal{T}$. Specifically, let

$$(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i, \ldots, \boldsymbol{v}_N) = \sum_{\bar{\boldsymbol{v}} \in \mathcal{V}^N} \bar{\boldsymbol{v}} \frac{f_{\bar{\boldsymbol{v}}}(\bar{\boldsymbol{v}}) f_{\boldsymbol{t}|\bar{\boldsymbol{v}}}(\boldsymbol{t}|\bar{\boldsymbol{v}})}{f_{\boldsymbol{t}}(\boldsymbol{t})}$$

be the players' expected payoffs given $\boldsymbol{t}$. Let $(v_{\eta i}, v_{1i}, \ldots, v_{Ki}) = \boldsymbol{v}_i$, then $u_i(\cdot|\boldsymbol{t})$ is defined by $u_i(o|\boldsymbol{t}) = v_{oi}$ for each $o \in \mathcal{O}'$. We say that the (mixed) assignment $\tilde{\boldsymbol{\phi}}^\star$ is **interim (utilitarian) efficient** if it maximizes social surplus, i.e., if it solves $\max_{\tilde{\boldsymbol{\phi}} \in \Delta\Phi} \sum_{i \in \mathcal{N}} u_i(\tilde{\boldsymbol{\phi}}|\boldsymbol{t})$.

To state our definition of ϵ-Bayesian incentive compatibility, it helps to specify how the distribution of players' true values and types changes as N changes. To these ends, let $f_{\bar{\boldsymbol{v}}}^N(\boldsymbol{t})$ be the probability mass function of values when there are N players and let $f_{\boldsymbol{t}|\bar{\boldsymbol{v}}}^N(\boldsymbol{t}|\bar{\boldsymbol{v}})$ be the conditional probability mass function of types when there are N players. Then,

$f^N_{\boldsymbol{t}}(\boldsymbol{t}) = \sum_{\bar{\boldsymbol{v}} \in \mathcal{V}^N} f^N_{\boldsymbol{t}|\bar{\boldsymbol{v}}}(\boldsymbol{t}|\bar{\boldsymbol{v}}) f^N_{\bar{\boldsymbol{v}}}(\bar{\boldsymbol{v}})$.

We write $U_i^M(r_i|t_i)$ for i payoff to making report r_i to mechanism M when her type is t_i when the other $N-1$ players report truthfully. That is,

$$U_i^M(r_i|t_i) = \sum_{\boldsymbol{t}_{-i} \in \mathcal{T}^{N-1}} u_i(M(r_i, \boldsymbol{t}_{-i})|t_i, \boldsymbol{t}_{-i}) \Pr^{N-1}(\boldsymbol{t}_{-i}|t_i),$$

where $\boldsymbol{t}_{-i} = (t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_i)$ and $\Pr^{N-1}(\boldsymbol{t}_{-i}|t_i) = f^{N-1}_{\boldsymbol{t}}(t_i, \boldsymbol{t}_{-i}) / \sum_{\boldsymbol{t}' \in A(t_i)} f^{N-1}_{\boldsymbol{t}}(\boldsymbol{t}')$.

Given $\epsilon > 0$. A mechanism M is **ϵ-Bayesian incentive compatible** if it is the case that, for each player i and each type $t_i \in \mathcal{T}$, i does not gain more than ϵ from lying strategically when everyone else tells the truth, i.e., if $\max_{r_i \in \mathcal{T}} U_i^M(r_i|t_i) \leq U_i^M(t_i|t_i) + \epsilon$.

Disregarding Cardinal Information

For each player i and each $t_i \in \mathcal{T}$, let $q(t_i)$ give the unique order in $\mathcal{B}(\mathcal{O}')$ that is consistent with i's expected true values after learning her type is t_i, i.e., consistent with $\boldsymbol{v}_i^\dagger$ defined via equation (4.1). Let $q(t_1, t_2, \ldots, t_N) = (q(t_1), q(t_2), \ldots, q(t_N))$. We say that a mechanism M **disregards cardinal information** if, for each $\boldsymbol{r} \in \mathcal{T}^N$, we have $M(\boldsymbol{r}) = M(q(\boldsymbol{r}))$. In other words, a mechanism disregards cardinal information if it makes the same assignment when players report their types or the orders that are consistent *with their expected values* (after they observe their own types). This definition generalizes the one used in the main text since Corollary OA1 gives a player's expected true values are always consistent with her ordinal preference in the baseline model of the main text.

The next assumption ensures that (i) it is always possible for players to learn their true values and so is analogous to the $\alpha > 0$ assumption of the main text.

Assumption OA1. We have $\mathcal{V} \subset \mathcal{T}$ and that $f_{t|\boldsymbol{v}}(\boldsymbol{v}|\boldsymbol{v}) > 0$ for all $\boldsymbol{v} \in \mathcal{V}$.

Proposition OA2. Disregarding Cardinal Information and Inefficiency.
Let Assumptions 1 and OA1 hold and let M be a mechanism that disregards cardinal information, then there is a $\bar{N} > 0$ such that M is interim inefficient when there are at least $\bar{N}$ players.

Proof. Analogous to the Proof of Proposition 1 since Assumption OA1 allows us to initially endow all players with the same, known true values and then Assumption 1 allows us to perturb one player's true values to render M's assignment inefficient. The formal argument is omitted. □

It follows that mechanisms that only use ordinal information are subject to the exact same inefficiency problems as they are in the baseline model. This motivates our study of a generalization of the Simple Mechanism.

Generalized Simple Mechanism

We begin with a description of the Generalized Simple Mechanism.

The Generalized Simple Mechanism.

Suppose the players report $\boldsymbol{r} = (r_1, \ldots, r_N) \in \mathcal{T}^N$. Then, the mechanism returns the mixed assignment $M_{GS}(\boldsymbol{r})$, which is constructed as follows. First, estimate expected true values,

$$(\hat{\boldsymbol{v}}_1, \ldots, \hat{\boldsymbol{v}}_N) = \sum_{\bar{\boldsymbol{v}} \in \mathcal{V}^N} \bar{\boldsymbol{v}} \frac{f_{\bar{\boldsymbol{v}}}(\bar{\boldsymbol{v}}) f_{\boldsymbol{t}|\bar{\boldsymbol{v}}}(\boldsymbol{r}|\bar{\boldsymbol{v}})}{f_{\boldsymbol{t}}(\boldsymbol{r})}.$$

For each player i, represent $\hat{\boldsymbol{v}}_i$ as a function $\hat{u}_i$ by setting $\hat{u}_i(o) = \hat{v}_o$ for each $o \in \mathcal{O}'$, where $(\hat{v}_\eta, \ldots, \hat{v}_K) = \hat{\boldsymbol{v}}_i$. Second, compute

$$\sigma = \arg\max_{\phi \in \Phi} \sum_{i \in \mathcal{N}} \hat{u}_i(\phi(i)).$$

Third, set $M_{GS}(\boldsymbol{r})$ to assign weight $1/|\sigma|$ to each pure assignment in σ. ○

As the next results show, the Generalized Simple Mechanism retains all of the features of the Simple Mechanism. We say that the true value and type distributions are **exchangeable** if (i) $f_{\bar{\boldsymbol{v}}}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N) = f_{\bar{\boldsymbol{v}}}(\boldsymbol{v}_{\rho(1)}, \ldots, \boldsymbol{v}_{\rho(N)})$ for any permutation $\rho(1), \ldots, \rho(N)$ of the original indices and (ii) $f_{\boldsymbol{t}|\bar{\boldsymbol{v}}}(t_1, \ldots, t_N|\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N) = f_{\boldsymbol{t}|\bar{\boldsymbol{v}}}(t_{\rho'(1)}, \ldots, t_{\rho'(N)}|\boldsymbol{v}_{\rho'(1)}, \ldots, \boldsymbol{v}_{\rho'(N)})$ for any permutation $\rho'(1), \ldots, \rho'(N)$ of the original indices for each $\bar{\boldsymbol{v}}$. It is easily seen that (i) exchangeability allows for correlation between the true values and types and that (ii) exchangeability is implied when (i') true values are independently and identically distributed and (ii') each player has the same chance of observing her true values, as in the main text. Exchangeability ensures that two players with the same type have the same (expected) true values.

Proposition OA3. The Generalized Simple Mechanism is Interim Efficient and Symmetric. *We have $M_{GS}(\boldsymbol{t})$ is interim efficient. In addition, when the distributions of true values and types are exchangeable, the $M_{GS}(\boldsymbol{t})$ is also symmetric.*

Proof. Analogous to the Proof of Proposition 2 in the main text and omitted. □

To state our analogue of Proposition 3, we need the following assumption. Recall, from the Proof of Lemma A2, that a **type *E*** player is one who observes their true values, has the lowest value for η, and has the highest value for every other object.

Assumption OA2. Two parts:

(i) We have $\mathcal{V}$ is a product of finite grids.

(ii) The probability that there are a finite number of type E players goes to zero as N goes to infinity. That is, $\lim_{N\to\infty} \Pr^N(E_l) \to 0$ for all $0 \le l < \infty$, where $E_l = \{\boldsymbol{t} | \boldsymbol{t}$ specifies l players have type $E\}$ and $\Pr^N(E_l) = \sum_{\boldsymbol{t}\in E_l} f^N_{\boldsymbol{t}}(\boldsymbol{t})$.

Part (i) is standard, together with Assumption OA1 it allows for the possibility of type E players. Part (ii) ensures that the number of type E increases as N increases (with probability one). It is easily seen that part (ii) holds when true values are independently and identically distributed and each player has the same chance of observing her true values, as in the main text.

Proposition OA5. The General Simple Mechanism is ϵ-Bayesian Incentive Compatible.
Let Assumptions OA1 and OA2 hold, then, for every $\epsilon > 0$, there is an $\bar{N} > 0$ such that the Generalized Simple Mechanism is ϵ-Bayesian incentive compatible when there are at least $\bar{N}$ players.

Proof. Analogous to the Proofs of Lemma A1 and Proposition 3 since there type E players per Assumptions OA1 and OA2 who increase competitive pressure as N increases. The only major changes, besides a shift in notation, are in the Proof of Lemma A1 where we need to use part (ii) of Assumption OA2 instead of the closed form expression for $\Pr(E_l)$ computed in the main text. However, because we only use this closed form to show $\lim_{N\to\infty} \Pr(E_l) = 0$ for all l, we can easily make this substitution. □

Since an analogue of Proposition 3 holds, an analogue of Proposition 4 holds as well.

Proposition OA6. ϵ-Bayesian Incentive Compatibility and Replication
Let Assumptions OA1 and OA2 hold. For every $\epsilon > 0$, when the linking function ψ is arbitrary there may not exist a symmetric, interim efficient, and ϵ-Bayesian compatible mechanism. However, there is a linking function $\bar{\psi}$ and an integer $\bar{N} > 0$ such that the Generalized Simple Mechanism is ϵ-Bayesian incentive compatible when there are at least $\bar{N}$ players.

Proof. The first part of the proposition follows from Example A of the main text. The second part of the proposition follows from Proposition OA5: using this proposition, we are able to establish a series of thresholds N_1, N_2, ... above which the Simple Mechanism is ϵ-Bayesian incentive compatible with 1, 2, 3, ... replications of objects respectively. Thus, we can use analogues of the arguments used in the Proof of Proposition 4 to establish the existence of the requisite linking function and the proposition. The details are omitted. □

5 Limited Knowledge of Preferences

The Simple Mechanism implicitly requires that the designer have knowledge of structure of preferences. Yet, in practice, the designer may not know the distribution of true value $f_{\boldsymbol{v}}$ or even the set of true values $\mathcal{V}$, and so finds it difficult to implement the Simple Mechanism. In this section, we show how to address this problem via the use of historical data. We first focus on the case where the designer is ignorant of $f_{\boldsymbol{v}}$, but knows $\mathcal{V}$. Then, we address the case where the designer is ignorant of both $f_{\boldsymbol{v}}$ and $\mathcal{V}$. For simplicity, we work in the baseline model of the main text.

Ignorance of the Distribution of True Values

Suppose the data from previous assignments consists of the players' reports and the objects they obtained. Specifically, there are data for W previous assignments consisting of N players each. We denote an assignment w, we denote a player who was part of this assignment i_w, and we denote the mechanism that made the assignment M_w. Let r_{i_w} denote i_w's report to M_w and let o_{i_w} denote the object she received.[2]

We make two major "identification" assumptions. First, $\alpha \in (0,1)$ so that all types occur with strictly positive probability. Second, for each assignment w, the mechanism M_w is (approximately) incentive compatible, so players do not lie to M_w. (It is easily seen that mechanisms like Random Serial Dictatorship meet the latter requirement.)

For each $\boldsymbol{v} \in \mathcal{V}$, $A(\boldsymbol{v}) = \{i_w | r_{i_w} = \boldsymbol{v}\}$ be the set of players who report values $\boldsymbol{v}$ and let

$$\hat{f}_{\boldsymbol{v}}(\boldsymbol{v}) = \frac{|A(\boldsymbol{v})|}{\sum_{\boldsymbol{v}' \in \mathcal{V}} |A(\boldsymbol{v}')|}$$

be the fraction of players who report values $\boldsymbol{v}$ among those who report a cardinal preference.

Lemma OA3. Convergence in Probability.
For each $\boldsymbol{v} \in \mathcal{V}$, we have

$$\hat{f}_{\boldsymbol{v}}(\boldsymbol{v}) \xrightarrow{p} f_{\boldsymbol{v}}(\boldsymbol{v}) \text{ as } W \to \infty.$$

Proof. Since types are truthfully reported, $|A(\boldsymbol{v})|$ is the number of players in our sample of $W\,N$ players with type $\boldsymbol{v}$. Since types are independent and identically distributed across the players in our data and all types occur with strictly positive probability, $|A(\boldsymbol{v})|/(W\,N) \xrightarrow{p} \alpha f_{\boldsymbol{v}}(\boldsymbol{v})$ by the Weak Law of Large Numbers. Likewise, $\sum_{\boldsymbol{v}' \in \mathcal{V}} |A(\boldsymbol{v}')|/(W\,N) \xrightarrow{p} \alpha$. Thus, the ratio converges in probability to $f_{\boldsymbol{v}}(\boldsymbol{v})$. $\square$

We modify the Simple Mechanism by using $\hat{f}_{\boldsymbol{v}}$ in place of $f_{\boldsymbol{v}}$. Creatively, we dub this edited mechanism the **Modified Simple Mechanism**; the next lemma summarizes its key properties.

Proposition OA7. Properties of the Modified Simple Mechanism.

[2]Data like this is available for settings where people choose objects periodically – as is the case with college dorm rooms, school courses, various public assistance programs, and so on.

The Modified Simple Mechanism is (i) symmetric and (ii) ϵ-Bayesian incentive compatible for every $\epsilon > 0$, when $\mathcal{V}$ is the product of finite grids and N is finitely large. In addition, the Modified Simple Mechanism becomes interim efficient with probability one as $W \to \infty$.

Proof. The arguments for symmetry and ϵ-Bayesian incentive compatibility are standard and are omitted. The last sentence is a direct consequence of Lemma OA3. By Lemma OA3, as $W \to \infty$, the Modified Simple Mechanism's objective function converges in probability to the expected social surplus function. Hence, the mechanism becomes interim efficient, and so is approximately interim efficient with sufficient historical data. $\square$

Ignorance of the Set and Distribution of True Values

While the above discussion assumes the designer knows $\mathcal{V}$, this is often not the case. Fortunately, under the same assumptions as above, when W is large, the designer will observe $\mathcal{V}$ via the cardinal reports in the data with probability one: simply, all points in $\mathcal{V}$ are eventually endowed to players who observe their cardinal preferences by our first identification assumption and the designer, in turn, observes these points by our second identification assumption. Thus, once W is large, $\mathcal{V}$ is known and the previous analysis applies.