**UMUTCAN SALMAN**

**University of Padova**

# THE REVEALED PREFERENCE THEORY OF AGGREGATE OBJECT ALLOCATIONS

**August 2025**

**Marco Fanno Working Papers – 324**

Dipartimento di Scienze Economiche e Aziendali 'Marco Fanno'

Università degli Studi di Padova

# The Revealed Preference Theory of Aggregate Object Allocations*

Umutcan Salman†

August 20, 2025

**Abstract**

We develop a revealed preference framework to test whether an aggregate allocation of indivisible objects satisfies Pareto efficiency and individual rationality (PI) without observing individual preferences. Exploiting the type-based preferences of Echenique et al. (2013), we derive necessary and sufficient conditions for PI-rationalizability. We show that an allocation is PI-rationalizable if and only if its allocation graph is acyclic, and equivalently if its associated bipartite graph contains no alternating cycles. The bipartite representation admits a matroid structure, enabling a simple greedy algorithm to measure the severity of PI violations and identify the minimal set of individual–object assignments whose removal restores rationalizability. Our results yield the first complete revealed preference test for PI in matching markets and provide an implementable tool for empirical applications.

**JEL codes:** C78, D11

**Keywords:** Aggregate object allocation, Pareto Efficiency, Individual Rationality, Revealed preferences, Matroid theory

*I thank Thomas Demuynck for his support and guidance. I also thank participants at the 3rd Italian Junior Theory Workshop. I acknowledge financial support from the PRIN PNRR 2022 *NextGenerationEU* project P2022P5CHH, "Incentivizing participation of compatible pairs in Kidney Paired Exchange Programs."

†Department of Economics and Management-University of Padova

# 1 Introduction

In many assignment environments, individuals are initially allocated objects by a central authority, (e.g., assignment of employees to public institutions or allocation of social houses to low income families) yet subsequent *exchanges* of assignments among individuals take place in a fully decentralized manner. A prominent example in Turkey involves exchanges among public employees, which grants public servants the right to exchange their positions within the same institution but from different locations, provided they have the same job title and their request is approved administratively.[1] The public authorities are not involved in arranging these exchanges. They learn of them only once agreements have been reached and approve if the legal requirements are satisfied.

Similar exchange practices are found in public housing "mutual exchange" schemes in the United Kingdom, where council or housing association tenants swap homes without intervention from the housing authority, provided both landlords give consent.[2] In the UK's National Health Service, doctors and other healthcare staff frequently trade shifts informally before notifying the rota manager for approval, a practice recognized in national rostering guidelines and formalized in local Trust policies.[3] Large-scale public-sector exchanges also occur in other contexts, such as Indian Railways' mutual transfer system.[4]

A common feature of these markets is that the central authority, after the initial allocation, does not observe individual preferences and therefore cannot directly assess whether the observed reallocation after exchanges is Pareto efficient and individually rational. This raises an important policy question: should the authority invest in centralizing the reallocation or exchange process, by implementing a formal matching mechanism, or are decentralized exchanges already delivering efficient outcomes? This question cannot be answered directly without knowledge of individuals' preferences. The central authority may observe their initial preferences, collected before assignments take place. However, exchanges arise precisely because individuals hold different preferences afterwards, which are no

[1]This procedure is called *Becayiş* in Turkish. A practice rooted in Article 73 of Civil Servants Law No 657, Türkiye. For more information about the procedure, see https://tr.wikipedia.org/wiki/Becayiş

[2]For more information about mutual exchanges, you can visit the website: https://england.shelter.org.uk/housing_advice/council_housing_association/tenancy_exchanges

[3]See Good Rostering Guide (2018) published by BMA and NHS Employers in UK.

[4]The platform railwaysmutualtransfer.com has been created for facilitating mutual transfers for railway employees across various zones and divisions. See the platform for more information.

longer observed.

We propose a revealed preference framework to address this problem. Our approach uses only the initial assignments and the observed exchanges to test whether the resulting allocation could have been generated by a Pareto-efficient and individually rational outcome for some preference profile. Therefore the central planner asks the following revealed preference question:

> *Does there exist a preference profile of individuals such that the observed allocation is both Pareto efficient and individually rational?*

If the answer is no for this question, the authority gains empirical justification for considering a centralized reallocation procedure. If not, the evidence suggests that the current decentralized process may already be achieving efficient outcomes, sparing the cost of mechanism redesign.

We jointly refer to Pareto efficiency and individual rationality as *PI*. Without imposing additional structure on the problem, the question of whether a given reallocation is PI becomes essentially meaningless: for any reallocation, the answer would always be "yes". To see why, note that for any reallocation we can always construct a hypothetical preference profile, that is, a complete ordering of all objects for each individual, such that the reallocation is PI. In particular, imagine that in this constructed profile, each individual ranks the object they receive in the reallocation as their most preferred option. Under this artificial preference profile, the allocation is trivially both Pareto efficient (since no mutually beneficial swap exists) and individually rational (since no one receives an object worse than what they "originally" had in the profile). This means that, without further assumptions, every reallocation can be "rationalized" as PI, making the concept uninformative for empirical or theoretical testing.

To generate testable restrictions or conditions that could potentially be rejected by data, we adopt the type-based approach of Echenique et al. (2013). In this framework, as researchers or as a central planner, we partition individuals into types based on observable characteristics (such as age, occupation, or education). We then assume that all individuals within the same type are identical in their preferences over objects. This homogeneity assumption is admittedly strong: it rules out any unobserved preference heterogeneity within types. However, it is precisely this restriction that gives the analysis empirical content. Without such a structure, the revealed preference framework becomes trivial, as any observed allocation pattern could be explained by some contrived set of individual preferences, leaving no scope for falsification.

Our contributions are twofold. We refer to the combination of initial endowments, final allocation, and type formation of individuals as the *aggregate allocation data.*

First, we construct from aggregate allocation data a directed *allocation graph* (Algorithm 1). By individual rationality, if an agent receives an object different from their initial endowment, they must prefer the new object. By the type-based preferences structure, these preferences extend to all agents of the same type. We show that the observed allocation is PI-rationalizable, meaning that there exists a preference profile of types where the observed allocation is PI, *if and only if* the allocation graph contains no cycles (Theorem 2). The absence of cycles captures the idea that there is no sequence of individuals in which every individual strictly prefers the object of the next individual in the sequence, a situation that would contradict Pareto efficiency.

Second, we provide an equivalent characterization using a *bipartite graph*, obtained by translating edges from the allocation graph into two distinct classes. In this representation, PI-rationalizability holds *if and only if* the bipartite graph contains no alternating cycles (Theorem 3). We further show that the collection of alternating-cycle-free edge sets forms a *matroid.* This structure yields a greedy algorithm (Algorithm 2) that tests for PI-rationalizability and, for non-rationalizable aggregate allocations, identifies the minimal set of individual–object assignments whose removal restores rationalizability.

**Related literature.** Since the pioneering work of Gale and Shapley (1962), matching theory has grown into a rich and influential field. Over the past six decades, much of the progress has been on the normative side, designing algorithms to solve real-life allocation problems in ways that satisfy desirable properties. Two of the most celebrated mechanisms, the Top Trading Cycles (TTC) and the Deferred Acceptance Algorithm (DAA), are now standard tools in practical markets: they allocate students to schools, match medical graduates to hospitals, allocate social houses to low income families, and even facilitate kidney exchanges. In all these cases, the primary focus has been on how to design the market, assuming preferences are observed and using them to obtain an allocation in line with the desirable properties with an algorithm.

While the design side of the field has developed extensively, the positive side has received much less attention. Key questions in the context of object reallocation remain largely unanswered. For instance: *What can we say about the outcomes we actually observe? Is the observed allocation Pareto efficient? Could it result from individually rational behavior?* In settings where $n$ individuals each begin with one object and exchanges are allowed, these questions are both natural and important. Without answers, empirical work faces a fundamental obstacle. More clearly, before attempting to estimate utility parameters or other structural features of preferences (identification), one must first establish that the observed outcome is consistent with basic normative properties (characterization) such as Pareto efficiency and individual rationality.

The distinction between characterization and identification is crucial. Characterization provides testable conditions, necessary and sufficient restrictions, that observed data must satisfy to be consistent with equilibrium concepts such as Pareto efficiency and individual rationality. Identification, by contrast, presupposes such consistency and focuses on recovering the underlying structural features of preferences. In this sense, characterization is not only conceptually different but also logically prior to identification. There is little value in estimating preference parameters if the observed allocation cannot satisfy Pareto efficiency and individual rationality under any possible preference profile. Although such characterization steps are standard in other areas of revealed preference analysis, they remain largely unexplored in the study of object reallocations.

A small but growing literature addresses revealed preference in matching markets. In two-sided matching markets, Echenique et al. (2013) develop a revealed preference analysis for stability, the core equilibrium concept in that setting, and Demuynck and Salman (2022) extends the analysis of Echenique et al. (2013).[5] Echenique et al. (2013) provides a graph theoretic characterization for aggregate marriage data. In aggregate object allocations or one-sided matching markets, Tai (2022) performs a revealed preference analysis for strict core stability using the type-based structure of Echenique et al. (2013). They also provide a graph theoretic characterization for aggregate allocations. Despite the fact that strict core stability implies Pareto efficiency and individual rationality, their characterization fails to extend to Pareto efficiency and individual rationality. We establish this formally with a counterexample.

Our work builds on this emerging positive strand of matching literature. We offer a new characterization results for PI-rationalizability in object reallocations. To the best of our knowledge, this study is the first one that performs a revealed preference analysis for object reallocations.

The structure of the paper is as follows. Section 2 introduces the setting and the key equilibrium notions that guide the analysis. Section 3 presents the first

[5] A matching is stable if there is no pair of agents $(i, j)$ such that both $i$ and $j$ prefer each other to their respective matches.

characterization based on the aggregate allocation graph. Section 4 develops an alternative characterization through a bipartite graph representation and further introduces the associated matroid structure together with a greedy algorithm for non-rationalizable reallocations. Section 5 concludes the paper, while all proofs are collected in the Appendix.

## 2 The Framework

Let $I$ be a finite set of individuals and $H$ a finite set of objects. Object $h \in H$ has $q_h \in \mathbb{N}_+$ identical copies, with $q = (q_h)_{h\in H}$ and $\sum_{h\in H} q_h = |I|$. Each individual $i \in I$ has a complete, transitive, and antisymmetric preference relation $P_i$ over $H$; let $P = (P_i)_{i\in I}$ denote the preference profile. We write $hP_ih'$ if $i$ strictly prefers $h$ to $h'$, and $R_i$ for the weak preference relation induced by $P_i$, i.e., $hR_ih'$ iff $hP_ih'$ or $h = h'$.

Preferences $P$ are unobserved. Instead, we observe individual characteristics and partition $I$ into a finite set of types $T = \{t_1, \dots, t_{|T|}\}$ via a type function $\tau : I \to T$. For $t \in T$, let $I^t = \{i \in I : \tau(i) = t\}$. We assume $\exists\, t \in T$ such that $|I^t| \geq 2$. Individuals of the same type share identical preferences. This restriction, standard in the empirical matching literature (e.g., Choo and Siow (2006)), is necessary for testability: without it, any allocation can be rationalized by assuming each individual ranks their assigned object highest.

An allocation is a mapping $x : I \to H$ satisfying

$$|\{i \in I : x(i) = h\}| = q_h \quad \forall h \in H.$$

We denote by $\omega$ the initial endowment allocation. In our setting, we observe $(x, \omega)$ and $(I, H, q, \tau, T)$, but not $P$. We refer to $\langle I, H, q, \tau, T, x, \omega\rangle$ as an *aggregate allocation instance.*

**Definition 1** (Pareto efficiency (PE))**.** *An allocation $x'$* Pareto dominates *$x$ if*

$$x'(i)R_ix(i) \quad \forall i \in I,$$

*and $x'(j)P_jx(j)$ for some $j \in I$. An allocation is* Pareto efficient *if it is not Pareto dominated by any other allocation.*

**Definition 2** (Individual rationality (IR))**.** *An allocation $x$ is* individually rational *if*

$$x(i)R_i\,\omega(i) \quad \forall i \in I.$$

**Definition 3** (PI-rationalizability)**.** *An allocation $x$ is* PI-rationalizable *if there exists a preference profile $P$ such that $x$ is both PE and IR.*

Tai (2022) study strict-core stability in a related framework. An allocation is *strict-core stable* if there is no coalition $S \subseteq I$ where individuals belonging to $S$ can reallocate their initial endowments so that all members are weakly better off and at least one is strictly better off. Every strict-core stable allocation is PI, but in general the reverse might not be true. Moreover, in the framework of our paper the set of strict core stable allocations may be empty. Example 1 shows this.

**Example 1.** *Consider three agents, $i_1$, $i_2$ and $i_3$, and two objects, $h_1$ and $h_2$, with $q_{h_1} = 2$ and $q_{h_2} = 1$. The information about the endowments and preferences of agents are summarized in the table below.*

| *Agent* | *Endowment* | *Preferences* |
|---|---|---|
| $i_1$ | $h_1$ | $h_2 P_{i_1} h_1$ |
| $i_2$ | $h_2$ | $h_1 P_{i_2} h_2$ |
| $i_3$ | $h_1$ | $h_2 P_{i_3} h_1$ |

*In this object allocation example, there exist two PI allocations. Those are:*

$$x_1 = \{(i_1, h_2), (i_2, h_1), (i_3, h_1)\}, \quad x_2 = \{(i_1, h_1), (i_2, h_1), (i_3, h_2)\}.$$

*If we check blocking coalitions in which individuals can exchange their endowments and at least one agent in the coalition is strictly better off while the other members are getting the same object, then $x_1$ is blocked by $\{i_2, i_3\}$ (yielding $x_2$) and $x_2$ is blocked by $\{i_1, i_2\}$ (yielding $x_1$), so the set of strict core stable allocations is empty.*

Tai (2022) is conducting a revealed preference analysis in this setting. The question that their paper asks if there exists a preference profile where the observed aggregate allocation is strict core stable. In other words, they replace the equilibrium concept in our framework with strict core stability. To answer the question, they form a directed graph. In their graph, the set of vertices is composed of the set of agents, e.g. each agent $i$ is a vertex in their graph. There is an edge from an agent $i$ to another $j$ whenever $i$'s assigned object equals $j$'s initial endowment. Their main characterization result is the following:

**Theorem 1** (Tai (2022))**.** *An aggregate allocation is rationalizable as a (strict) core-stable allocation if and only if, in every strictly connected component of the directed graph, all agents of the same type receive the same object.*

We focus on PI allocations, not strict-core stability for two reasons. First, in our framework, the set of strict core stable allocations might be empty. Second, the characterization result of Tai (2022) does not extend to Pareto efficiency and individual rationality together. In Example 2, we show that in an aggregate allocation instance the observed allocation can be PI-rationalizable although it is not rationalizable as strict core stable.

**Example 2.** *Figure 1 shows a directed graph in the style of Tai (2022). Superscripts denote assigned objects; parentheses indicate initial endowments.*

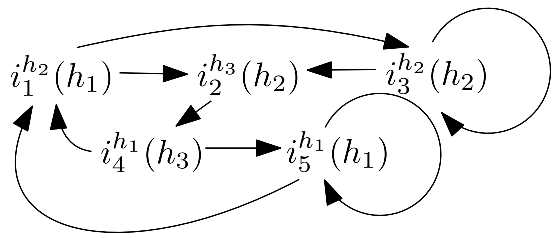


Figure 1: Directed graph representation based on Tai (2022)

*The graph has a single strictly connected component in which agents $i_5$ and $i_1$ (same type) receive different objects, so there exists no preference profile where the allocation is strict-core stable. However, the following preference profile makes it PI:*

| *Agent* | *Type* | *Preferences* |
|---|---|---|
| $i_1$ | $t_1$ | $h_2 P_{i_1} h_5 P_{i_1} h_3 P_{i_1} h_4 P_{i_1} h_1$ |
| $i_2$ | $t_2$ | $h_3 P_{i_2} h_4 P_{i_2} h_2 P_{i_2} h_1 P_{i_2} h_5$ |
| $i_3$ | $t_1$ | $h_2 P_{i_3} h_5 P_{i_3} h_3 P_{i_3} h_4 P_{i_3} h_1$ |
| $i_4$ | $t_3$ | $h_1 P_{i_4} h_2 P_{i_4} h_3 P_{i_4} h_4 P_{i_4} h_5$ |
| $i_5$ | $t_1$ | $h_2 P_{i_5} h_5 P_{i_5} h_3 P_{i_5} h_4 P_{i_5} h_1$ |

*Thus, Tai (2022)'s condition is not necessary for PI-rationalizability.*

## 3 The Aggregate Allocation Graph

To test for PI-rationalizability, we construct a directed *aggregate allocation graph* $AG = (AV, AE)$ with Algorithm 1 from the observed allocation. The construction exploits individual rationality and the type-based preferences structure. If an individual receives an object different from their initial endowment, we infer that they prefer the new object; by the type-based preferences structure, these preferences extend to all members of the same type.

**Algorithm 1** Construction of $AG = (AV, AE)$

**Vertices:** For each $i \in I$, create a vertex $i$ and add it to $AV$.

**Edges:**

**Step 1:** For $i, i' \in I$, if there exists $j \in I^t$ with $t = \tau(i)$, $\omega(j) \neq x(j)$, $x(i) = \omega(j)$, and $x(i') = x(j)$, add edge $(i, i')$ to $AE$.

**Step 2:** For $i, i' \in I$ with $(i, i') \notin AE$, if there is a path $p = \langle i, j_1, \ldots, j_n, i' \rangle$ where $\tau(j) = \tau(i)$ for all $j \in \{j_1, \ldots, j_n\}$, add edge $(i, i')$ to $AE$.

**Step 3:** For $i, i' \in I$ with $(i, i') \notin AE$, if there exists $(i, j) \in AE$ and $x(i') = x(j)$, add edge $(i, i')$ to $AE$.

We write $h \to h'$ if there exists $(i, i') \in AE$ with $x(i) = h$ and $x(i') = h'$. In Algorithm 1, Step 1 encodes individual rationality implied preferences from endowment changes. Step 2 adds the transitive closure within each type. Step 3 propagates preferences across identical copies of an object.

Figure 2 illustrates $AG$ for the instance in Example 2. For example, $i_1$ has endowment $h_1$ and receives $h_2$. By individual rationality, $h_2 P_{i_1} h_1$. Since $i_1$ and $i_5$ are of type $t_1$, $i_5$ inherits these preferences and points to all individuals holding $h_2$.

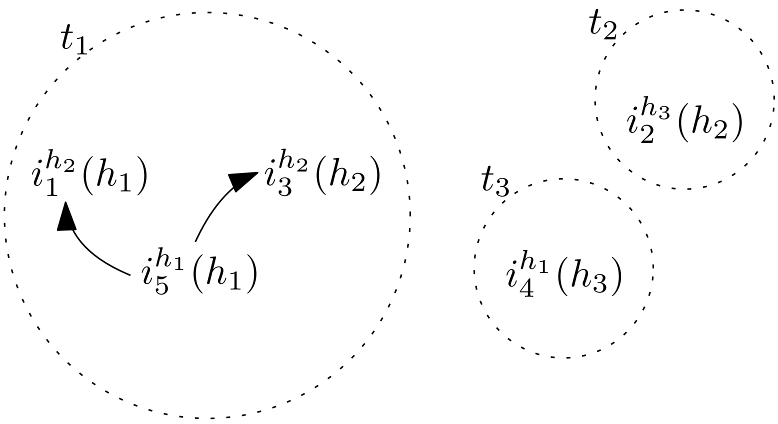


Figure 2: Allocation graph for Example 2

One of the central results of our paper is the following characterization. Theorem 2 establishes that the aggregate allocation graph contains no cycles if and only if there exists a preference profile of types under which the observed allocation is PI. This result generalizes the classical approach of verifying Pareto efficiency and individual rationality in settings where individual preferences are directly observable. To check whether an allocation is PI, one typically constructs an envy graph whose vertices represent individuals. There is a directed edge from individual $i$ to $j$ whenever $i$ prefers $j$'s object to their own. If the envy graph contains a cycle, then each individual in the cycle can obtain the object of the next individual along the cycle, leaving all of them strictly better off. This contradicts Pareto efficiency.

**Theorem 2.** *An aggregate allocation $x$ is PI-rationalizable if and only if $AG = (AV, AE)$ is acyclic.*

*Proof.* See Appendix A.1. □

**Proof sketch.** If $x$ is PI, then by construction any edge $(i, i') \in AE$ indicates that $i$ prefers the object of $i'$ to their own. The aggregate allocation graph therefore functions as an *envy graph*: a cycle corresponds to a coalition of agents who can cyclically exchange objects so that all strictly benefit, violating Pareto efficiency. Hence PI implies acyclicity.

Conversely, if $AG$ is acyclic, we can construct a preference profile consistent with all edges in $AE$ such that $x$ satisfies IR and no cycle of mutual improvements exists. This profile makes $x$ both Pareto efficient and individually rational, so $x$ is PI-rationalizable.

# 4 The Bipartite Graph

In this section, we present a bipartite graph representation of the aggregate allocation data. A bipartite graph consists of two distinct sets of vertices, with edges only connecting vertices from different sets. Using this representation, we provide an alternative characterization of PI-rationalizability.

For each type $t \in T$, let $\Omega^t$ be the set of objects initially endowed to individuals in $I^t$ and $H^t$ be the set of objects allocated to the same group. We define the bipartite graph $\mathcal{G} = (\mathcal{O}, \mathcal{K}, \mathcal{E}, w)$ as follows. For each $h \in \Omega^t$, create a vertex $o_h^t \in \mathcal{O}$ and for each $h \in H^t$, create a vertex $k_h^t \in \mathcal{K}$.

There are two types of edges, namely *horizontal and vertical.* An edge $e = \{o_h^t, k_h^t\}$ is *horizontal* if $t$ and $h$ are the same on both endpoints. It represents that $h$ is both initially endowed by some individual and is allocated to some individual in $I^t$. It is not necessary that this individual must be the same. Its weight is

$$w(e) = \big|\{i \in I^t : x(i) = h\}\big|.$$

An edge $e = \{o^t_h, k^{t'}_{h'}\}$ is *diagonal* if $t \neq t'$ and $h \to h'$ in the aggregate allocation graph $AG$. It represents that some $i \in I^t$ with endowment $h$ receives $h'$ which is allocated to some $j \in I^{t'}$.

**Definition 4.** *A* cycle *in $\mathcal{G}$ is a sequence of vertices $\gamma = \langle v_0, v_1, \ldots, v_n, v_0\rangle$ such that for $j \in \{0, \ldots, n-1\}$, $\{v_j, v_{j+1}\} \in \mathcal{E}$, and $\{v_n, v_0\} \in \mathcal{E}$.*

**Definition 5.** *An* alternating cycle *is a cycle in which consecutive edges are of different types (horizontal followed by diagonal, or vice versa).*

A bipartite graph is *alternating-cycle-free* if it contains no alternating cycle. Alternating cycles are important because of their direct correspondence to cycles in $AG$.

**Lemma 1.** *If $\mathcal{G}$ contains an alternating cycle, then $AG$ contains a cycle.*

*Proof.* See Appendix A.2. □

Thus, an alternating cycle in $\mathcal{G}$ implies a violation of Pareto efficiency.

Any cycle in $AG$ admits a *minimal version*, a cycle whose vertices alternate between different types. Formally, let $\gamma$ be a cycle in an aggregate allocation graph $AG = (AV, AE)$. A minimal version of $\gamma$, denoted $\gamma^m = \langle i_0, i_1, \ldots, i_m, i_0\rangle$, is a cycle in $AG = (AV, AE)$ such that the set of vertices in $\gamma^m$ is a subset of vertices in $\gamma$ and for every $j = 0, \ldots, m-1$, $(i_j, i_{j+1}) \in AE$, $\tau(i_j) \neq \tau(i_{j+1})$, $(i_m, i_0) \in AE$, and $\tau(i_m) \neq \tau(i_0)$.

**Lemma 2.** *Every cycle $\gamma$ in $AG$ has a minimal version $\gamma^m$.*

*Proof.* See Appendix A.3. □

**Lemma 3.** *If $\mathcal{G}$ is alternating-cycle-free, then $AG$ is acyclic.*

*Proof.* See Appendix A.4. □

From Lemmas 1–3 and Theorem 2 we obtain:

**Theorem 3.** *An aggregate allocation $x$ is PI-rationalizable if and only if $\mathcal{G}$ is alternating-cycle-free.*

*Proof.* See Appendix A.5. □

The property of being alternating-cycle-free defines a *matroid*. A matroid consists of a finite *ground set* $Z$ and a collection $\mathcal{I}$ of *independent sets* satisfying: (i) $\emptyset \in \mathcal{I}$; (ii) if $A \in \mathcal{I}$ and $B \subseteq A$, then $B \in \mathcal{I}$; (iii) if $A, B \in \mathcal{I}$ and $|A| < |B|$, there exists $b \in B \setminus A$ with $A \cup \{b\} \in \mathcal{I}$. The exchange property (iii) ensures that all maximal independent sets have the same size and that certain optimization problems can be solved greedily.

In our setting, the ground set is $\mathcal{E}$ and independence means being alternating-cycle-free.

**Lemma 4.** *Let $\mathcal{G} = (\mathcal{O}, \mathcal{K}, \mathcal{E}, w)$ be as above. Let $\mathcal{I}$ be the collection of subsets $A \subseteq \mathcal{E}$ such that $(\mathcal{O}, \mathcal{K}, A, w)$ is alternating-cycle-free. Then $(\mathcal{E}, \mathcal{I})$ is a matroid.*

*Proof.* See Appendix A.6. □

The matroid property implies that a maximum-weight alternating-cycle-free subgraph can be found by adding edges greedily in order of weight.

**Algorithm 2** Greedy construction of a maximum-weight alternating-cycle-free subgraph

**Require:** Bipartite graph $\mathcal{G} = (\mathcal{O}, \mathcal{K}, \mathcal{E}, w)$; partition $\mathcal{E} = \mathcal{H} \dot{\cup} \mathcal{D}$ (horizontal/diagonal); list $\mathcal{H}^{\triangleleft}$ of $\mathcal{H}$ sorted by nonincreasing $w$ (ties broken deterministically).

**Ensure:** Set $S \subseteq \mathcal{E}$ that is alternating-cycle-free, contains all $\mathcal{D}$, and has maximum total weight among such sets.

```
S ← D, W ← 0
for e ∈ H^◁ do
    if S ∪ {e} is alternating-cycle-free then
        S ← S ∪ {e}
        W ← W + w(e)
    end if
end for
return S, W
```

# 5 Conclusion

We have developed a revealed-preference framework for testing Pareto efficiency and individual rationality (PI) in aggregate object allocations without observing individual preferences. Using a type-based preferences structure, we derived necessary and sufficient conditions for PI-rationalizability and expressed them in graph theoretic terms.

Two equivalent characterizations emerge. An allocation is PI-rationalizable if and only if its allocation graph is acyclic; equivalently, if and only if its associated bipartite graph is free of alternating cycles. The latter formulation reveals a matroid structure on the set of alternating-cycle-free edge sets, allowing violations of PI to be measured via a simple greedy algorithm that identifies the minimal set of individual–object assignments whose removal restores rationalizability.

These results provide the first complete revealed-preference test for PI allocations in matching markets, offering both a theoretical characterization and an implementable procedure for empirical applications.

# A Proofs of the results in the main text

## A.1 Proof of Theorem 2

*Proof.* Let $P$ be the preference profile (which is a complete, transitive and anti-symmetric binary relation over objects) of individuals such that $x$ is a PI allocation according to $P$. We need to show that there is no cycle in $AG$.

**Lemma 5.** *Let $x$ be an IR allocation with $P$. If there exists $e = (i, i') \in AE$ then $x(i')\ P_i\ x(i)$.*

*Proof.* Let $e$ be constructed by Step 1. There exists some $j \in I$ such that $\omega(j) \neq x(j)$. Since $x$ is IR, then $x(j)\ P_j\ \omega(j)$. If $i \in I^t$ such that $\tau(j) = t$ then $x(j)\ P_i\ \omega(j)$. If $x(i) = \omega(j)$ and $x(i') = x(j)$, then $x(i')\ P_i\ x(i)$.

Let $e$ be constructed by Step 2. Before $e$ was constructed in Step 2, all edges in $AG$ are constructed by Step 1. If there is a path from $i$ to $i'$ then by transitivity $x(i')\ P_i\ x(i)$. Hence if there is an edge from $i$ to $i'$ constructed by Step 2 then $x(i')\ P_i\ x(i)$.

Let $e$ be constructed by Step 3. There is no edge from $i$ to $i'$ before Step 3. There is an edge $(i, j) \in AE$ constructed before Step 3 and $x(i') = x(j)$. Hence, $x(j)\ P_i\ x(i)$. Since $x(i') = x(j)$, $x(i')\ P_i\ x(i)$. □

Suppose for a contradiction that there is a cycle in $AG$. Since $x$ is IR, by Lemma 5 all the steps in the edge construction indicate that if there is an edge $e = (i, i') \in AE$ then $x(i')\ P_i\ x(i)$. Hence, a cycle in $AG$ implies that there is a set of individuals who can improve their situations by exchanging their objects. This contradicts with PE.

We assume that the graph $AG$ contains no cycles. Our goal is to show that there exists a preference profile $P^{\star}$ such that the allocation $x$ is PI with $P^{\star}$.

To achieve this, we construct the preference profile for individuals. Let $P_t^{\star}$ denote the preference profile of individuals $i \in I^t$, and let $H^t$ represent the set of objects allocated to individuals of type $t$, i.e.,

$$H^t = \{h \in H : x(i) = h, i \in I^t\}.$$

To construct the preference profile $P^*$ we need to define two binary relationships in the following way: For every $t \in T$, let $P_t$ be a binary relationship defined for the objects $h, h' \in H$ in the following way:

$$h \; P_t \; h' \text{ if and only if } \text{ there is an individual } i \in I^t \text{ such that } \begin{cases} h = x(i) \\ h' = \omega(i) \end{cases}$$

**Lemma 6.** *$P_t$ is acyclic.*

*Proof.* Suppose for a contradiction that $P_t$ is cyclic. $\exists$ a sequence $h_1, \ldots, h_m$ in $H$ such that $h_1 \; P_t \; h_2 \; P_t \ldots h_m$ and $h_m \; P_t \; h_1$. So, for $j \in \{1, \ldots, m-1\}$,

$$h_j \; P_t \; h_{j+1} \implies \exists i_j \in I^t \text{ such that } h_j = x(i_j) \text{ and } h_{j+1} = \omega(i_j)$$

and

$$h_m \; P_t \; h_1 \implies \exists i_m \in I^t \text{ such that } h_m = x(i_m) \text{ and } h_1 = \omega(i_m)$$

Hence, $(i_{j+1}, i_j) \in AE$ for $j \in \{1, \ldots, m-1\}$. Since $\exists i_m \in I^t$ such that $h_m = x(i_m)$ and $h_1 = \omega(i_m)$, $(i_1, i_m) \in AE$ because $h_1 = x(i_1)$. Hence, $\langle (i_m, i_{m-1}), \ldots, (i_2, i_1), (i_1, i_m) \rangle$ is a cycle in $AG$. This is a contradiction. □

By Szpilrajn's theorem[6] there exists a complete, asymmetric and transitive relation $P_t^+$ such that $h \; P_t \; h'$ implies $h \; P_t^+ \; h'$ for all $h, h' \in H$ (i.e. $P_t \subseteq P_t^+$).

Let $\succ$ be a binary relationship defined for the objects $h, h' \in H$ in the following way:

$$h \succ h' \text{ if and only if } h' \to h$$

[6] See Szpilrajn (1930)

**Lemma 7.** $\succ$ *is acyclic.*

*Proof.* Suppose for a contradiction that $\succ$ is cyclic. $\exists$ a sequence $h_1, \ldots, h_m$ in $H$ such that $h_1 \succ h_2 \succ \ldots h_m$ and $h_m \succ h_1$. So, for $j \in \{1, \ldots, m-1\}$,

$$h_j \succ h_{j+1} \implies \exists e = (i_{j+1}, i_j) \in AE \text{ such that } x(i_j) = h_j, x(i_{j+1}) = h_{j+1}$$

and

$$h_m \succ h_1 \implies \exists e = (i_1, i_m) \in AE \text{ such that } x(i_m) = h_m, x(i_1) = h_1.$$

Hence, $\langle (i_m, i_{m-1}), \ldots, (i_2, i_1), (i_1, i_m) \rangle$ is a cycle in $AG$. This is a contradiction. □

By Szpilrajn's theorem there exists a complete, asymmetric and transitive relation $\succ^+$ such that $h \succ h'$ implies $h \succ^+ h'$ for all $h, h' \in H$ (i.e. $\succ \subseteq \succ^+$).

Now we can construct the preference profile for every type $t \in T$ in the following way: Every object $h \in H^t$ is strictly preferred to any object $h' \notin H^t$, i.e.,

$$h \; P_t^* \; h' \quad \text{for all } h \in H^t, h' \notin H^t.$$

For every type $t \in T$ and objects $h, h' \in H^t$, $h \; P_t^* \; h'$ if and only if $h \succ^+ h'$. For every type $t \in T$ and objects $h, h' \in H \setminus H^t$, $h \; P_t^* \; h'$ if and only if $h \; P_t^+ \; h'$.

We need to show that $P_t^*$, for every $t \in T$, is complete, asymmetric and transitive.

1. The relation $P_t^*$ is complete:
   - If $h, h' \in H^t$, this follows from completeness of $\succ^+$.
   - If $h \in H^t$ and $h' \notin H^t$ then $h \; P_t^* \; h'$.
   - If $h, h' \notin H^t$, this follows from completeness of $P_t^+$.
2. The relation $P_t^*$ is asymmetric:
   - If $h, h' \in H^t$, this follows from asymmetry of $\succ^+$.
   - If $h \in H^t$ and $h' \notin H^t$ then $h \; P_t^* \; h'$ and not $h' \; P_t^* \; h$.
   - If $h, h' \notin H^t$, this follows from asymmetry of $P_t^+$.
3. The relation $P_t^*$ is transitive. Let $x \; P_t^* \; y \; P_t^* \; z$. We need to show that $x \; P_t^* \; z$. Then:
   - if $z \in H^t$, then also $y \in H^t$ so $x \succ^+ y \succ^+ z$ so $x \succ^+ z$ (by transitivity of $\succ^+$) and therefore, by definition, $x \; P_t^* \; z$.

- if $z \notin H^t$ then either $y \in H^t$ or $y \notin H^t$. If $y \in H^t$, then $x \in H^t$. Hence $x\ P_t^*\ z$. If $y \notin H^t$, then either $x \in H^t$ or $x \notin H^t$. If $x \in H^t$, $x\ P_t^*\ z$. If $x \notin H^t$, $x, y, z \notin H^t$. By transitivity of $P_t^+$, $x\ P_t^+\ z$. So, $x\ P_t^*\ z$.

First we need to show that $x$ satisfies IR with $P^*$. Suppose for a contradiction that $x$ is not IR. There is an individual $i \in I^t$ such that $\omega(i)\ P_t^*\ x(i)$. This implies that $\omega(i) \in H^t$, $\exists j \in I^t$ such that $x(j) = \omega(i)$. Hence, there must exist $(i, j) \in AE$. Since $\exists i, j \in I^t$ such that $\omega(i) \neq x(i)$, $x(j) = \omega(i)$, then $(j, i) \in AE$. So, $\langle (i, j), (j, i) \rangle$ is a cycle in AG. This is a contradiction.

Now, we need to show that $x$ is PE with the preference profile $P^*$. Let $G^* = (V^*, E^*)$ be the envy graph of $x$ with the preference profile $P^*$ where the set of vertices $V^*$ is composed of the individuals and from each $v \in V^*$ there is an edge to $v' \in V^*$ if and only if $x(v')\ P_v^*\ x(v)$.

**Lemma 8.** *Given an IR allocation $x$ with $P^\star$, the aggregate allocation graph $AG = (AV, AE)$ and envy graph $G^* = (V^*, E^*)$ of $x$ with the preference profile $P^*$, we have the following: $AE \subseteq E^*$.*

*Proof.* Suppose for a contradiction that there exists an edge $e = (i, i') \in AE$ such that $e \notin E^*$. Since $x$ is IR with $P^\star$ and $e \in AE$, then by Lemma 5 $x(i')\ P_i^*\ x(i)$. Hence $e$ must be in $E^*$. This is a contradiction. □

By Lemma 8, we know that either $AE = E^*$ or $AE \subset E^*$. If $AE = E^*$ then $x$ is PE with $P^*$.

Assume that $AE \subset E^*$. Suppose, for a contradiction, $x$ is not PE. Hence there exists a cycle $c = \langle e_1, \ldots, e_m \rangle$ in the envy graph $G^* = (V^*, E^*)$ such that for all $j = \{1, \ldots, m-1\}$, $e_j = (i_j, i_{j+1}) \in E^*$ and $e_m = (i_m, i_1) \in E^*$. If each edge in $c$ is also an edge in $AG$ then $c$ is also a cycle in $AG$. This is a contradiction. There must exist some $e_k$ such that $e_k \in E^* \setminus AE$. For all $e_k = (i_k, i_{k+1}) \in E^* \setminus AE$, we know that there exists a type $t \in T$ such that $h_{k+1}\ P_t^*\ h_k$, $x(i_{k+1}) = h_{k+1}$ and $x(i_k) = h_k$. This implies that for all $e_k = (i_k, i_{k+1}) \in E^* \setminus AE$, $x(i_{k+1}) \succ^+ x(i_k)$. Hence for all $j = \{1, \ldots, m-1\}$, $x(i_{j+1}) \succ^+ x(i_j)$, and $x(i_1) \succ^+ x(i_m)$. This means that $\succ^+$ has a cycle which is a contradiction.

□

## A.2 Proof of Lemma 1

*Proof.* We proceed by constructing a cycle in $AG$ based on $c$. We know that

(i) for $j = 0, \ldots, n-1$, $o_{h_j}^{t_j} \in \mathcal{O}$, $k_{h_j}^{t_j} \in \mathcal{K}$, $\{o_{h_j}^{t_j}, k_{h_j}^{t_j}\} \in \mathcal{E}$,

(ii) for $j = 0, \ldots, n-1$, $\{o_{h_j}^{t_j}, k_{h_{j+1}}^{t_{j+1}}\} \in \mathcal{E}$,

(iii) $o_{h_n}^{t_n} \in \mathcal{O}$, $k_{h_n}^{t_n} \in \mathcal{K}$, $\{o_{h_n}^{t_n}, k_{h_n}^{t_n}\} \in \mathcal{E}$, and $\{o_{h_n}^{t_n}, k_{h_0}^{t_0}\} \in \mathcal{E}$.

By (i), for $j = 0, \ldots, n$, $\exists i_j \in I^{t_j}$ such that $x(i_j) = h_j$. By (ii) and (iii), $h_0 \to h_1 \to \cdots \to h_n \to h_0$. Hence $\gamma = \langle i_0, i_1, \ldots, i_n, i_0 \rangle$ is a cycle in $AG$. □

## A.3 Proof of Lemma 2

*Proof.* Let $\gamma = \langle i_0, i_1, \ldots, i_n, i_0 \rangle$ be a cycle in $AG = (AV, AE)$. We introduce an algorithm below to convert $\gamma = \langle i_0, i_1, \ldots, i_n, i_0 \rangle$ to the minimal version of it, $\gamma^m$.

**Algorithm 3** The Converter

Consider $\gamma = \langle i_0, i_1, i_2, \ldots, i_n, i_0 \rangle$. Set $i^\star = i_0 = i_0^m$, $l = 1$, and $\gamma^m = \langle i_0^m \rangle$.
**for** $k = 1, \ldots, n+1$ **do**,
    **if** $\tau(i_k) \neq \tau(i^\star)$ and $k \neq n+1$ **then**
        $i_l^m \leftarrow i_k$
        add $i_l^m$ to $\gamma^m$
        $i^\star \leftarrow i_k$
        increase $l$ by one
    **else if** $k = n+1$ **then**
        add $i_0^m$ as the last vertex to $\gamma^m$ if $\tau(i_0) \neq \tau(i^\star)$
        remove the first vertex $i_0^m$ and add $i_1^m$ as the last vertex to $\gamma^m$ if $\tau(i_0) = \tau(i^\star)$
    **end if**
**end for**

We need to show that the following hold:

(i) The set of vertices in $\gamma^m$ is a subset of the set of vertices in $\gamma$.

(ii) For every consecutive edges $(i_j^m, i_{j+1}^m)$ in $\gamma^m$, $(i_j^m, i_{j+1}^m) \in AE$ and $\tau(i_j) \neq \tau(i_{j+1})$.

(iii) $\gamma^m$ is a cycle in $AG = (AV, AE)$.

By construction, all vertices in $\gamma^m$ are selected from the original cycle $\gamma$, so $\gamma^m$'s vertex set is a subset of $\gamma$'s vertex set. Hence $(i)$ is satisfied.

We need to show that for every pair of consecutive vertices $(i_j^m, i_{j+1}^m)$ in $\gamma^m$, the following hold: $(i_j^m, i_{j+1}^m) \in AE$ and $\tau(i_j^m) \neq \tau(i_{j+1}^m)$.

We prove this by induction:

Base Case ($j = 0$):

- Let $i_0^m = i_0$ and $i_1^m = i_k$ be the first two vertices of $\gamma^m$, with $\tau(i_0) \neq \tau(i_k)$ by construction.
- Since both $i_0$ and $i_k$ occur in $\gamma$, and all intermediate individuals from $i_0$ to $i_k$ (in $\gamma$) are of the same type, then by Step 2 of Algorithm 1, the edge $(i_0, i_k) \in AE$.

Inductive Step:

Assume that for all $j < r$, the pairs $(i_j^m, i_{j+1}^m)$ are connected in $AE$ and belong to different types.

Now consider $(i_r^m, i_{r+1}^m)$:

- Again, by construction, $\tau(i_r^m) \neq \tau(i_{r+1}^m)$.
- In $\gamma$, there is a directed path from $i_r^m$ to $i_{r+1}^m$ where all intermediate nodes have the same type $\tau(i_r^m)$.
- Thus, by Step 2 of Algorithm 1, $(i_r^m, i_{r+1}^m) \in AE$.

Now we need to show that $\gamma^m$ is a cycle.

There are two cases depending on whether the last selected vertex $i_k^m$ has the same type as $i_0$:

- Case 1: $\tau(i_k^m) \neq \tau(i_0)$:
  - Then we simply append $i_0$ to close the cycle.
  - There is a path in $\gamma$ from $i_k^m$ to $i_0$ via same-type nodes, so again $(i_k^m, i_0) \in AE$ by Step 2.
- Case 2: $\tau(i_k^m) = \tau(i_0)$:
  - Then we discard $i_0$ and instead append $i_1^m$ to close the cycle.
  - Since $\tau(i_k^m) \neq \tau(i_1^m)$ by construction, and there is a path from $i_k^m$ to $i_1^m$ within $\gamma$ through same-type nodes, Step 2 again ensures $(i_k^m, i_1^m) \in AE$.

Since the first and the last vertices of $\gamma^m$ are the same and there is an edge for any two consecutive pair of vertices, it is a cycle in $AG$.

This concludes the proof.

□

## A.4 Proof of Lemma 3

*Proof.* Suppose for a contradiction that there exists a cycle $c$ in $AG = (AV, AE)$. Let $c^m = \langle i_0^m, \ldots, i_k^m, i_0^m \rangle$ be the minimal version of $c$. We know that there exists a minimal version of $c$ by Lemma 2. For $j = 0, \ldots, k$, $k+1 = 0$, and every edge $(i_j^m, i_{j+1}^m)$ in $c^m$, let $x(i_j^m) = h_j$, $x(i_{j+1}^m) = h_{j+1}$, $\tau(i_j^m) = t_j$, and $\tau(i_{j+1}^m) = h_{j+1}$. So, $k_{h_j}^{t_j}, k_{h_{j+1}}^{t_{j+1}} \in \mathcal{K}$. By the construction of each edge $(i_j^m, i_{j+1}^m)$ (by Algorithm 1), there exists $i' \in I^{t_j}$ such that $\omega(i') = h_j$ and $x(i') = h_{j+1}$. So $o_{h_j}^{t_j} \in \mathcal{O}$. So, for $j = 0, \ldots, k$, $k+1 = 0$, and every edge $(i_j^m, i_{j+1}^m)$ in $c^m$, $\{o_{h_j}^{t_j}, k_{h_j}^{t_j}\} \in \mathcal{E}$. Since $(i_j^m, i_{j+1}^m) \in AE$, $h_j \rightarrow h_{j+1}$. Hence, $\{o_{h_j}^{t_j}, k_{h_{j+1}}^{t_{j+1}}\} \in \mathcal{E}$. So, $\langle o_{h_0}^{t_0}, k_{h_0}^{t_0}, o_{h_0}^{t_0}, k_{h_1}^{t_1}, o_{h_1}^{t_1}, k_{h_1}^{t_1}, o_{h_1}^{t_1}, k_{h_1}^{t_1}, \ldots, o_{h_n}^{t_n}, k_{h_0}^{t_0} \rangle$ such that for $j = 0, \ldots, n$, $x(i_j^m) = h_j$ and $\tau(i_j^m) = t_j$ forms an alternating cycle in $\mathcal{G}$. This is a contradiction.

□

## A.5 Proof of Theorem 3

*Proof.* By Lemma 1, if there is an alternating cycle in $\mathcal{G}$, then there exists at least one cycle in $AG$. By Lemma 3, if there is no alternating cycle in $\mathcal{G}$, then there exists no cycle in $AG$. Hence, by Theorem 2, the result follows. □

## A.6 Proof of Lemma 4

*Proof.* It is straightforward to show that the first and second conditions are satisfied. To see the third condition, let $E$ and $E'$ be two elements of $\mathcal{I}$ such that $|E'| > |E|$, and $(\mathcal{O}, \mathcal{K}, E, w)$ and $(\mathcal{O}, \mathcal{K}, E', w)$ are alternating-cycle-free. We need to show that there exists an element $e \in E' \setminus E$ such that $E \cup \{e\}$ is also independent.

Suppose for a contradiction that for every $e \in E' \setminus E$ there is a distinct alternating cycle in $(\mathcal{O}, \mathcal{K}, E \cup \{e\}, w)$. Let $C_e$ denote such cycle. Since $|E'| > |E|$, $|E \setminus E'| < |E' \setminus E|$. This implies that for each $C_e$ there is at most one $e$ such that $e \in C_e$ and $e \in E \setminus E'$. Since there is no alternating cycle in $(\mathcal{O}, \mathcal{K}, E', w)$, for each $C_e$ there must be at least one $e$ such that $e \in C_e$ and $e \in E \setminus E'$. Hence, each $e \in E \setminus E'$

belongs to a unique $C_e$. This implies that $|E \setminus E'| \geq |\{C_e : e \in E \setminus E'\}| = |E' \setminus E|$. This is a contradiction.

□

## References


Choo, E., Siow, A., 2006. Who marries whom and why. Journal of political Economy 114 (1), 175–201.

Demuynck, T., Salman, U., 2022. On the revealed preference analysis of stable aggregate matchings. Theoretical Economics 17 (4), 1651–1682.

Echenique, F., Lee, S., Shum, M., Yenmez, M. B., 2013. The revealed preference theory of stable and extremal stable matchings. Econometrica 81 (1), 153–171.

Gale, D., Shapley, L. S., 1962. College admissions and the stability of marriage. The American Mathematical Monthly 69 (1), 9–15.

Szpilrajn, E., 1930. Sur l'extension de l'ordre partiel. Fundamenta mathematicae 16 (1), 386–389.

Tai, A., 2022. Revealed preferences of one-sided matching. arXiv preprint arXiv:2210.14388.