COMBINING SOCIAL CHOICE AND MATCHING THEORY TO UNDERSTAND INSTITUTIONAL STABILITY

ASHUTOSH THAKUR

Stanford GSB
April 7, 2021

Abstract. In many organizations, members need to be assigned to certain positions, whether these are legislators to committees, executives to roles, or workers to teams. I show that these assignment problems lead to novel questions about institutional stability. Will the set of agents being assigned prefer or vote in favor of some alternative allocation over their current allocation thereby lobbying to reform the institution? I explore questions of institutional stability where the choice of the institution (i.e., the matching mechanism) is chosen and agreed upon by the very people who are assigned by the assignment procedure. I endogenize an institution's choice of assignment procedures by generalizing social choice to what I call a social allocation choice problem. I discuss a variety of voting rules (plurality, majority, and unanimity) and their institutional stability counterparts in matching theory (popular matching, majority stability, and pareto efficiency). The novel property of majority stability is introduced and its existence and robustness to correlated preferences and interdependent preferences are analyzed. Chains of envy are necessary to overcome the packing problem that arises in reallocating a majority to a new set of assignments under an alternative allocation. This makes majority stability, in sharp contrast to plurality rule, strikingly robust to correlated preferences.

Email address: *adthakur@stanford.edu*
A special thanks to Olivier Tercieux, Steven Callander, Alvin Roth, Fuhito Kojima, Paulo Somaini, Ignacio Rios, Jonathan Bendor, Dinesh Thakur, Felix Bierbrauer and participants of Stanford Market Design Seminar, Political Economy Theory Lunch and PE Mini-Conference at Stanford GSB, and Princeton Political Economy Seminar.

1. **Introduction**.

Institutions are important and enduring, but they are the result of deliberate choice, not only at their founding but also over time. As North (1990) emphasizes, "institutions are created by human beings. They evolve and are altered by human beings." The design of an institution is thus a collective choice. These choices can change over time leading to a question of stability of institutional choice.

In many organizations, the matching mechanism used to assign members to positions is itself an institutional choice, that is chosen, agreed upon, and influenced by the very members being assigned. For example, political parties must decide how to assign their politicians to legislative committees, the bureaucracy must structure internal labor markets for the assignments of civil servants to various parts of the country, and personnel must be organized across teams and departments in various capacities within a firm.

Who gets what position affects the performance and output of an organization, the efficacy of its operations, the synergies and conflicts generated, and the resulting satisfaction and aspirations of its members. Agents' preferences over the positions may change as their environment changes, e.g., new legislative battles or changes to the jurisdictions of committees can affect a politicians' preferences over his committee assignments. Moreover, the organization's members themselves may change over time as new generations enter and overlap with the older generations who gradually exit. All of these considerations can allow different coalitions to mobilize and lobby for institutional change. Successful institutional designs are those that endure and prove robust to such endogenous pressure for reform.

Institutional stability thus considers the stability of the matching mechanism—the institutional structure—itself. Namely, in what environments do we get a matching allocation such that no decisive coalition can change the mechanism/allocation in favor of an alternative allocation it prefers, potentially at the loss of those not in the coalition. Importantly, institutional stability differs from the canonical notion of stability in matching theory, in that the alternative allocation can endogenously generate *both winners and losers*. This is a fundamental feature of institutional change in real organizations.

This connects matching theory to social choice. Social choice is the study of how groups collectively make choices. While typically applied to the choice of a single outcome, I apply it to the choice of a matching allocation, a vector of assignments where each agent gets a different position within the organization over which they might have varied preferences. Studying this *social allocation choice problem* delivers key insights as to when an organization's internal matching procedures admit institutional stability versus when endogenous coalitional pressure will surmount and result in institutional reform.

In this paper, I apply and extend the tools of matching theory and social choice to study the stability of institutional choice. I characterize the preference environments in which a decisive coalition—be it a plurality, simple majority, or unanimous agreement—can change the matching allocation in favor of an alternative allocation that everyone in the coalition prefers, even if those outside the coalition are left worse off. I show how the stability of the institution varies in the voting rule and in the domain of admissible preferences.

A fundamental insight that emerges from my analysis is the importance of how correlated individuals' preferences are over positions to the existence of institutional stability. Correlated preferences increase competition and generate envy over each others' assignments. The more correlated are preferences, the easier it is to mobilize larger coalitions that can benefit from change to the institutional design. However, I show that the robustness to correlated preferences starkly depends on the organization's underlying voting rule, that defines which coalitions are decisive and can undermine institutional stability.

I compare a variety of voting rules (plurality, majority, and unanimity) and their corresponding notions of institutional stability in matching (popular matching, majority stability, and pareto efficiency). I introduce the novel property of *majority stability*, whereby an allocation is majority stable if there exists no alternative allocation that makes a majority of agents strictly better off. Not only are majority and super-majority rules commonly used, both formally and informally, within organizations, but I also highlight their striking robustness to correlation across individuals' preferences and explain why.

While pareto efficient allocations under unanimity always exist and are easily attained by commonly-used mechanisms like a Serial Dictatorship, they are somewhat of a low bar for institutional stability as they do not insulate the institution against reforms that also result in both winners and losers. On the other hand, I show that complete popular matching under plurality rule, though cleanly characterizable (Abraham et al., 2007), is extremely fragile and not robust to even the slightest correlation in preferences across agents. Its lack of existence empirically, and moreover, the stringent requirements it imposes on mechanism details, such as the restrictions on the seniority ordering of a Serial Dictatorship to attain a popular matching when it does exist, makes popular matching practically incapable of generating institutional stability.

In stark contrast, I show that majority stability under majority (or super-majority) rule exhibits a striking robustness to correlated preferences across agents, making its existence surprisingly robust across many environments and a wide parameter space. The key feature restricting majority overrules is a packing problem: the inability to reassign a majority of people into a minority of new seats. Thus, majority overrules requires a majority to be moved into a new set of seats using *chains of envy*: i's seat given to j, j's seat given to k, etc. On the whole, although the more correlated preferences are the harder it is to attain institutional stability, correlation of preferences does not perfectly pick up the deep correlation that is needed to undermine majority stability. Namely, you need not only enough envy, but envy in the 'right' places causing chains of envy that undermine majority stability.

Relaxing the assumption that agents only care about their own assignment from the vector of assignments to allowing selfish but interdependent preferences, whereby an agent can have any preferences over allocations in which he gets the same assignment, only leads to weakly less majority stability. This unidirectional effect does not hold for popular matching under plurality rule.

Characterizing preferences where majority stable allocations exist is a computationally challenging problem, as changing the status quo allocation, completely changes the strategic environment altogether as the entire mapping of who is envious of what completely changes. However, I provide an integer linear program—which I show is computationally tractable—to check whether a given allocation is majority stable for a given a set of preferences.

Well-behaved environments from social choice such as weakly single-peaked preferences and order restricted preferences, I show have no bite in social allocation choice problems. However, I provide computationally simple algorithms that are extremely successful at detecting existence of majority stable allocations, these algorithms can, albeit very rarely, fail to detect the existence of majority stable allocations given a set of preferences.

Using simulations, I compare the success of attaining majority stable and popular matchings with a Serial Dictatorship and a Boston mechanism. The favoring higher ranks property of the Boston mechanism produces less envy and renders the Boston mechanism more institutionally stable than Serial Dictatorship, albeit assuming truthful preferences.

Finally, I show how increasing correlation in preferences has opposing effects on stability in social choice versus social allocation choice. Increasing correlation across agents' preferences leads to fewer cycles in preference aggregation rule thereby bolstering stability in social choice, but it leads to more chains of envy which undermine institutional stability in social allocation choice. For perfectly and near-perfectly correlated preferences, I establish a Chaos Theorem, akin to McKelvey (1976), where a majority-approved sequence of reassignments can be generated from any allocation to any other allocation.

2. **Related Literature**.

Addressing these new questions of institutional stability, brings together two large but, until now, distinct literatures of social choice and matching theory, extending each and unifying them in a particular important context.

In social choice, a single outcome is chosen for a group of agents, each of whom might have varied individual preferences on what they prefer. When the outcome of choice is a vector of assignments, as in our problem, an intricate, richness of preference structure is placed on the social choice problem. This paper analyzes this important sub-case of social choice: *social allocation choice* over a vectors of assignments, where each agent receives their own assignment. Agents might have varied preferences over their own assignment in this vector, leading to intricate preference restrictions within the broader social choice problem. Agents collectively vote to decide upon whether to continue using the matching mechanism at hand or change the institution, i.e., vote in favor of/against the allocation the matching mechanism produces relative to an alternative mechanism/allocation. Distinguishing between the institutional stability of certain allocations and the matching mechanisms that lead to these allocations is the goal of this paper.

Moreover, this notion of institutional stability is quite different from the canonical notion of stability in matching theory. Stability deals with whether the allocation produced by the mechanism admits a coalition that can reallocate their own assignments within the coalition and all be strictly better off. Such coalitional deviations are local in that they do not affect non-coalition members' assignments. On the other hand, my paper considers institutional stability under which a coalition can push for global deviations in assignments that strictly benefit themselves, potentially at the expense of those not in the coalition. This conflict between coalitions seeking to reorganize and to lobby to change the assignment procedures and allocations in their favor, creating both winners and losers in the process, is at the heart of institutional change within real organizations. We address endogenous coalition formation that is a key feature of the political economy of institutional change, characterizing the environments that make these self-governing bodies susceptible to endogenous reorganization and how the organization's choice of internal conflict resolution rules (e.g., the voting hurdles) affects the stability and evolution of institutions.

A small yet growing literature on popular matching (i.e., institutional stability under plurality rule) and its many variants has been developed in the computer science literature (see review articles by Manlove (2013) and Cseh (2017)). This literature considers the existence, characterization, and complexity of finding a Condorcet winner under plurality rule given one-sided voting (Abraham et al., 2007; Sng and Manlove, 2010; Manlove, 2013; McDermid and Irving, 2011; Kavitha and Nasre, 2009; McCutchen, 2008; Kavitha et al.,

2011) and two-sided voting in bipartite graphs (Huang and Kavitha, 2013; Kavitha, 2014; Cseh et al., 2017; Cseh and Kavita, 2016) and non-bipartite graphs (Chung, 2000; Biro et al., 2010; Huang and Kavitha, 2017). The focus on algorithmic complexity has often rendered it as a normatively appealing, yet abstract theoretical solution concept, that hasn't received much attention from the economics and applied market design literature. In this paper I show that existence of complete popular matching (i.e., plurality rule) imposes very stringent requirements on the preferences of individuals and the design details of the matching mechanisms. Even slight correlation across agents' preferences renders popular matching non-existent and any matching mechanism institutionally unstable. In practice, many organizations are majoritarian: using majority or super-majority rules to agree upon or overturn certain group choices. And in fact, I find that when it comes to majority stability, many common mechanisms used in practice, like Serial Dictatorship, are often institutionally stable and strikingly robust in spite of significant correlation across agents' preferences.

Considering the stability of the institutional choice itself, shares the underlying spirit of the self-stable voting rules and constitutional choice literature (Barbera and Jackson, 2004; Messner and Polborn, 2004; Maggi and Morelli, 2006; Koray, 2000). However, I differ in that rather than choosing a voting rule, agents in my model choose a matching mechanism/allocation. I also reveal the importance of correlation across agents' preferences for the stability of institutional choice.

3. Voting rules and their corresponding notions of institutional stability.

The criteria for institutional stability, depends on the voting rule chosen to decide between allocations/mechanisms. In this section I introduce three different voting rules (plurality rule, majority rule and unanimity rule) and their corresponding institutional stability counterparts (popular matching, majority stability, and pareto efficiency).

3.1. The Set-Up.

Let us consider a set of agents $\mathbb{I}$ and a set of seats $\mathbb{C}$, indexed by i and c and of size N and C respectively. Assume N and C odd for ease of notation. Agents $i \in \mathbb{I}$ have complete and transitive preferences relations $\succeq_i$ over seats $\mathbb{C}$ to which they are assigned, hereby referred to as i's ***assignment***. Seats do not have any preference over agents, hence the market is one-sided. If an agent i is not indifferent between two assignments ($\sim_i$ denotes indifference), then we say he has a strict preference ($\succ_i$). If agent i prefers to remain unmatched rather than be matched to j, i.e., if $i \succ_i c$, then c is said to be unacceptable to i.

An ***allocation***, or ***matching***, is defined as $M : \mathbb{I} \cup \mathbb{C} \to \mathbb{I} \cup \mathbb{C}$ such that $i = M(c)$ if and only if $M(i) = c$ and for all i and c either $M(i)$ is in $\mathbb{C}$ or $M(i) = i$, and either $M(c)$ is in $\mathbb{I}$ or $M(c) = c$. In other words, an agent is matched to a seat only if the seat is matched to him, and everyone is either matched to a counterpart from the other side of the market or is left unmatched (i.e., matched to one's self). Let $\mathbb{M}$ denote the set of all possible matchings.

Agent i's preference over two allocations $M, M' \in \mathbb{M}$ is denoted by $M \succeq_i M'$. Let R^N denote the set of all agents' reflexive and complete binary preference relations over the set of possible allocations $\mathbb{M}$ and let $\mathcal{R}$ be the set of all reflexive and complete binary relations over the set of possible allocations $\mathbb{M}$.

Definition 1. *A* ***social allocation choice*** *defines a social preference relation* $\succeq^S \in \mathcal{R}$ *between any two allocations* $M, M' \in \mathbb{M}$ *by aggregating the individual preferences of all agents given some preference aggregation rule* $f : R^N \to \mathcal{R}$.

Unless otherwise stated, I will maintain the following assumptions throughout this paper for ease of exposition:

Assumption 1. *All possible allocations are feasible and each agent's preferences are complete (i.e., no truncation deeming certain assignments unacceptable).*

Any restriction on the feasibility or acceptability of alternative allocations can only lead to more institutional stability, unless an institutionally stable allocation itself is made infeasible.

Assumption 2. *All preferences are strict.*

Allowing for weak preferences—including many-to-one matching which can be modeled using weak preferences over cloned versions of the same seat—only leads to more institutional stability.

Assumption 3. *Market is balanced with $N = C$.*

Asymmetrically increasing agents leads to weakly less institutional stability since you have increased envy. Asymmetrically increasing seats can lead to either more or less institutional stability depending on whether envy is decreased or increased.

Given Assumptions 1-3, unless otherwise stated, we restrict our attention to the set of possible allocations being the set of complete, one-to-one matchings, whereby every agent is assigned to a distinct seat and no agent/seat is left unassigned.

Assumption 4. *Agent i only cares about his own assignment. Namely, $M \sim_i M'$ if $M(i) \sim_i M'(i)$ and $M \succ_i M'$ if $M(i) \succ_i M'(i)$.*

Hence we assume that an agent's preference over two allocations is solely determined by his preference over his own assignment in each of the two allocations. Namely, agent i is indifferent across all allocations M which give him the same assignment a_i. This simplification is in line with the classical matching theory set up where the preference of an agent i is independent of assignments for all $-i$.

In Section 7, I relax this assumption and show that adding any interdependent preferences—as long as an agent does not vote against his own preference ranking over his own seat assignment—leads to weakly more institutional instability under majority stability. However, this unidirectional result does not hold for pareto efficiency.

3.2. Stability Concepts & Corresponding Voting Rules.

Let us denote by $|M' \succ_i M|$, the number of agents who strictly prefer matching M over matching M'. I define three different notions of institutional stability:

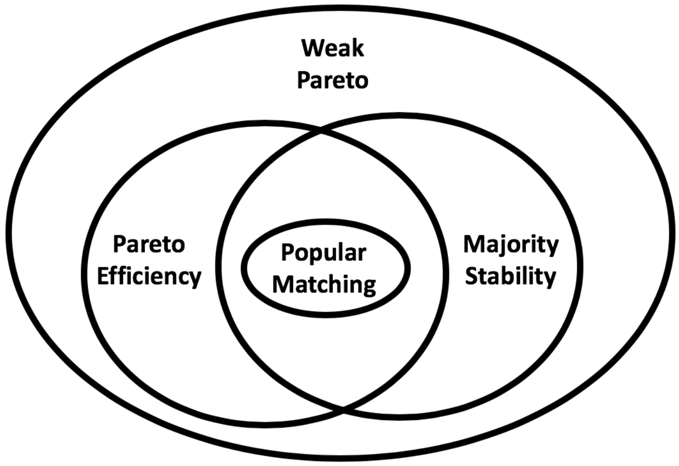
Definition 2. *A matching M is* ***popular*** *if $\nexists M'$ s.t. $|M' \succ_i M| > |M \succ_i M'|$.*

Definition 3. *A matching M is* ***majority stable*** *if $\nexists M'$ s.t. $|M' \succ_i M| \geq \frac{N+1}{2}$.*

Definition 4. *A matching M is* ***pareto efficient*** *if $\nexists M'$ s.t. $|M \succ_i M'| = 0$ and $|M' \succ_i M| > 0$.*

Assuming that all agents who are indifferent between two allocations abstain from voting, popular matching corresponds to the voting rule f being a plurality rule and a matching is popular if it is a Condorcet winner with regards to any alternative matching. Majority

Figure 1. Set inclusions for matchings across various institutional stability notions.



stability corresponds to the voting rule f being a majority rule[1] and pareto efficiency corresponds to a f being an unanimity rule. We also define a matching M to be *weak pareto efficient* if $\nexists M'$ s.t. $|M' \succ_i M| = N$.

As shown in Figure 1, the set of popular matchings are a subset of the set of majority stable matchings and set of pareto efficient matchings, however, the sets of majority stable and pareto efficient matchings are not fully contained in each other. Appendix C provides examples and details.

3.3. **The Relation to Social Choice.**

I highlight how social allocation choice is situated in the general framework of social choice.

Let X be the set of outcomes. Agent i's preference over two outcomes $x, x' \in X$ is denoted by $x \succ_i x'$. Let R^N be the set of all agents' reflexive and complete binary preference relations over the set of outcomes X and let $\mathcal{R}$ be the set of all reflexive and complete binary relations over X.

Definition 5. *A* ***social choice*** *defines a social preference relation* $\succeq^S \in \mathcal{R}$ *between any two outcomes* $x, x' \in X$ *by aggregating the individual preferences of all agents given some preference aggregation rule* $g : R^N \to \mathcal{R}$.

Thus social allocation choice is a sub-case of social choice where the outcome, x, itself is a *vector of assignments*, M assigning each agent a potentially different assignment.[2] This imposes intricate restrictions on agents' preferences over the matchings, when an agent's preference depends on which seat he himself is assigned to from this vector of assignments, and even when we relax this restrictions and allow agents to have complex, interdependent preferences between allocations where they are assigned the same seat (see Section 7). Social allocation choice is an important sub-case of social choice to analyze, as it highlights the endogenous envy of each others' positions within organizations caused by the assignment

[1]More generally, we can characterize any Q-rule by replacing $\frac{N+1}{2}$ by Q in Definition 3

[2]In the case of many-to-1 (e.g., civil servant assignment to states in India (Thakur, 2018a) or many-to-many matching (e.g., as committee assignments in US Senate (Thakur, 2018b) multiple agents can be assigned the same assignment.

procedures and explains which coalitions endogenously mobilize to reform the institution. These forces shape the evolution and longevity of the institutional design.

3.4. **The Relation to Stability.**

I juxtapose institutional stability with the standard notion of stability in matching theory.

Definition 6. *A matching M is stable if there is no group of agents who can reallocate their assigned seats amongst themselves and all be strictly better off.*

Notice that stability is a more 'local' deviation. Namely, unstable allocations admit a coalition whose members, by locally reallocating their own assigned seats amongst themselves, lead to an alternative 'near-by' allocation in which all coalition members are strictly better off. The allocation of all non-coalition agents remains unchanged. The notion of institutional stability is a more 'global' concept, whereby the coalitional deviation can suggest an alternative allocation/mechanism that makes every coalition member strictly better off, potentially at the expense of agents not in the coalition.

3.5. **Questions of Institutional Stability.**

There are three important questions to ask of this model, which I address in each of the three sections that follow.

Firstly, in Section 4 we ask, in which preference environments do institutionally stable allocations exist? This question is concerned with how correlated preferences are, which in turn defines who is envious of whose seats (i.e., competition over seats) and whether coalitions can form endogenously to overturn the proposed allocation in favor of an alternative allocation. Moreover, my analysis will also comment on which mechanisms can reach these stable allocations and what are the constraints on the mechanisms?

Secondly, in Section 5 we ask, for a given set of preferences, is a given allocation institutionally stable? This is a computationally hard problem because for the proposed allocation, we must search over the space of all alternative allocations to check whether sufficiently large coalitions can deviate and strictly improve. I show how to tackle this computational problem using an integer linear program, which simplifies to solving a relaxed linear program that has integer solutions.

Lastly, in Section 6 we ask, which mechanisms are more robust to institutional stability? What aspect of the mechanism explains the difference in institutional stability? This is explored using simulations.

4. **In which preference environments do institutionally stable allocations exist?**

4.1. **Unanimity Rule and Pareto Efficiency.**

Pareto efficiency is easy to guarantee with a simple mechanism like Serial Dictatorship.

With strict preferences, Serial Dictatorship is pareto efficient as the mechanism cannot make the nth senior person in the seniority ordering better off without making some $j \in \{1, ..., n-1\}$ seniorities worse off, otherwise j would have chosen this alternative himself.

If there are weak preferences, then tie-breaking can matter, as seen in Example 1.

Example 1. *Consider two agents $\{1, 2\}$ and two seats $\{A, B\}$. Suppose $A \sim_1 B$ and $A \succ_2 B$ and suppose 1 is more senior to 2. Then Serial Dictatorship yields the allocation $(1-A, 2-B)$ which is pareto dominated by the allocation $(1-B, 2-A)$.*

A Serial Dictatorship can span the entire set of pareto efficient outcomes, depending on the seniority order chosen. However, perhaps unanimity rule is too non-discriminating a voting rule for modeling institutional stability, as a single agent or small group of agents can block massive improvements for very large coalitions. In Example 2, two senior agents 1 and 2 occupy seats A and B and prevent a long chain of improvements amongst all $N-2$ other agents.

Example 2. *Consider N agents in order of seniority 1 more senior than 2, ..., more senior than $N-1$, more senior than N, and seats $\{A, B, ...\}$. Notice how agents 1 and 2 block an entire sequence of potential improvements for all $N-2$ others.*

1	*2*	*3*	*4*	*5*	...
A	***B***	*A*	*A*	*D*	...
B	*A*	*B*	*C*	*B*	...
...	...	***C***	*B*	*A*	...
...	...	...	***D***	*C*	...
...	...	...	...	***E***	...
...	...	...	...	...	...

In economics, we often seek pareto efficiency, because we do not want economic systems to inefficiently leave surplus on the table. However, proofing against unanimously-approved, unambiguous improvements is a somewhat low standard when modeling institutional stability, as in practice, institutional change creates both winners and losers. Considering endogenous coalitions trying to lobby to change the institution in their own favor, potentially at the expense of others, is what we analyze in Section 4.2 under plurality rule and Section 4.3 under majority rule.

4.2. Plurality Rule and Popular Matching.

Building on Abraham et al. (2007)'s characterization of popular matching, I show in this section that the existence of institutionally stable allocations under plurality rule—i.e., complete popular matchings—imposes extremely stringent conditions on both preferences and mechanisms.

Lemma 1. *(Abraham et al. (2007) Lemma 2.4)*
A matching M is popular if and only if

1. *every seat which is somebody's first choice (called a "f-post") is matched in M and*
2. *each agent i is assigned to either his first choice (i.e, $f(i)$) or his most preferred alternative that is not ranked first by any other agent (i.e., $s(i)$) called a "s-post"*

This also characterizes popular matchings that are not complete (i.e., some agents or seats are left unmatched) by introducing the "last resort option" $l(i)$ which is added to the end of each i's preference rank order as their least preferred option. This ensures that the s-post for any agent, $s(i)$, is not empty, and it represents remaining unmatched. A matching where no agent i is assigned to $l(i)$ is called a complete popular matching.

Consider Example 3 (from Abraham et al. (2007)) which illustrates all the concepts in Lemma 1.

Example 3. *Agents $\{1, 2, 3, 4, 5\}$ rank seats $\{A, B, C, D, E\}$. The highlighted seats are the f- and s-posts for each agent. The set of f-posts are $\{A, B, C\}$ and the set of s-posts are*

$\{D, l(3), E\}$. *Hence there are two complete popular matchings (1-A, 2-D, 3-B, 4-E, 5-C) and (1-D, 2-A, 3-B, 4-E, 5-C) and eight not complete popular matchings (1-A, 2-D, 3-l(3), 4-B, 5-C), (1-D, 2-A, 3-l(3), 4-B, 5-C), (1-A, 2-D, 3-l(3), 4-E, 5-C), (1-D, 2-A, 3-l(3), 4-E, 5-C), (1-A, 2-D, 3-l(3), 4-B, 5-C), (1-D, 2-A, 3-l(3), 4-B, 5-C), (1-A, 2-D, 3-l(3), 4-B, 5-E), (1-D, 2-A, 3-l(3), 4-B, 5-E).*

1	*2*	*3*	*4*	*5*
A	***A***	***B***	***B***	***C***
C	***D***	*A*	*C*	***E***
D	*B*	*C*	***E***	*A*
$l(1)$	$l(2)$	***l(3)***	$l(4)$	$l(5)$

Example 3 highlights how popular matchings can have different sizes. Let us restrict our attention to the existence of *complete popular matchings* in a balanced market (i.e., no agent or seat is left unassigned or vacant) and assume all agents have complete strict preferences over all seats in $\mathbb{C}$.

Using Hall's Marriage Theorem,[3] I provide the following corollaries illustrating the stringent restrictions the Lemma 1 characterization imposes on the preference environment and matching mechanism for a complete popular matching to exist and be reached by the matching mechanism.

Corollary 1.1. *A necessary but not sufficient condition for a complete popular matching to exist is that* $\bigcup_{i \in \mathbb{I}} \{f(i), s(i)\} = \mathbb{C}$, *where* $\mathbb{C}$ *is the set of distinct seats to be assigned.*

Example 4. *Agents* $\{1, 2, 3, 4, 5, 6\}$ *rank seats* $\{A, B, C, D, E, F\}$. *Notice that seat F is no one's f- or s-post (highlighted). Hence, no complete popular match exists.*

1	*2*	*3*	*4*	*5*	*6*
A	***A***	***B***	***B***	***C***	***B***
C	***D***	*A*	*C*	***E***	*C*
D	*B*	*C*	***E***	*A*	***E***
F	*F*	***D***	*F*	*F*	*A*
B	*C*	*E*	*A*	*D*	*F*
E	*E*	*F*	*D*	*B*	*D*

Next I define preferences to have n blocks B_n for $n = 1, 2, ...$ as the finest partition of seats such that all seats in the nth block B_n are ranked below all seats in the jth block B_j for all $j < n$ by all $i \in \mathbb{I}$, and are ranked above all seats in the kth block B_k for all $k > n$ by all $i \in \mathbb{I}$.

[3]The graph theoretic formulation of Hall's Marriage Theorem (Hall, 1935) states that for any finite bipartite graph G with bipartite sets P and C ($G := (P + C, E)$). An P-saturating matching is a matching which covers every vertex in P. For a subset W of P, let $N_G(W)$ denote the set of all vertices in C adjacent to some element of W. Hall's Marriage Theorem states that there is an P-saturating matching if and only if for every subset W of P, $|W| \leq |N_G(W)|$. In other words, every subset W of P has sufficiently many adjacent vertices in C. Proofs of the corollaries use the graph G of f- and s-posts imposed by Lemma 1 condition 2), along with Hall's Marriage Theorem.

Corollary 1.2. *No complete popular matching exists if preferences have 3 or more blocks as* $f(i) \in B_1$*,* $s(i) \in \{B_1, B_2\}$ $\forall i \in \mathbb{I}$*, and seats* $c \in B_3, B_4, ...$ *are never* $s(i)$ *for any agent.*

Example 5. *Consider agents* $\{1, 2, 3, 4, 5, 6\}$ *ranking seats* $\{A, B, C, D, E, F\}$*. Notice, that there are three blocks* $B_1 = \{A, B\}$*,* $B_2 = \{C, D\}$*, and* $B_3 = \{E, F\}$*. The highlighted f- and s-posts only consist of seats in* B_1 *and* B_2*, but not those in* B_3*. Hence no complete popular match exists.*

1	*2*	*3*	*4*	*5*	*6*
A	***A***	***B***	***B***	***B***	***A***
B	*B*	*A*	*A*	*A*	*B*
C	***D***	***D***	***C***	***D***	***C***
D	*C*	*C*	*D*	*C*	*D*
F	*E*	*F*	*F*	*E*	*E*
E	*F*	*E*	*E*	*F*	*F*

Corollary 1.3. *If have 2 blocks in preference, then for a complete popular matching to exist, every seat* $c \in B_1$ *must be the first choice of some agent and every seat* $c \in B_2$ *must be the first choice of someone amongst all the seats in* B_2*, else* $\bigcup_{i \in \mathbb{I}} \{f(i), s(i)\} \neq C$

Example 6. *Consider agents* $\{1, 2, 3, 4, 5\}$ *ranking seats* $\{A, B, C, D, E\}$*. There are two blocks* $B_1 = \{A, B\}$ *and* $B_2 = \{C, D, E\}$*, however, seat* E *is no one's third choice, and hence no one's s-post. Hence, there is no complete popular match.*

1	*2*	*3*	*4*	*5*
A	***A***	***B***	***B***	***B***
B	*B*	*A*	*A*	*A*
C	***D***	***D***	***C***	***D***
E	*C*	*C*	*E*	*C*
D	*E*	*E*	*D*	*E*

Thus far, the corollaries have imposed conditions on preferences to guarantee the existence of a complete popular match given any matching mechanism. Now, I consider a particular matching mechanism, the Serial Dictatorship, and illustrate a further restriction complete popular matching imposes on the mechanism's design details.

Corollary 1.4. *If agent(s)* i *when assigned to his s-post* $s(i)$*, would be envious of other agent(s)* j *being assigned to their f-post(s)* $f(j)$*, a complete popular matching is implemented by a Serial Dictatorship only if the seniority ordering is such that the set of agents assigned to such f-posts are more senior to such agents allotted to their s-posts in the popular match.*

Example 7. *Consider agents* $\{1, 2, 3\}$ *ranking seats* $\{A, B, C\}$*. If the seniority is* $1 > 2 > 3$*, the resulting matching* $(1 - A, 2 - B, 3 - C)$ *is not one of two popular matchings—which are* $(1 - A, 2 - C, 3 - B)$ *or* $(1 - C, 2 - A, 3 - B)$*—because agent 3 who must be assigned to his f-post in any complete popular matching was junior to both agents 1 and 2, who would be envious of* B *if assigned to their s-post* C*. Hence, only seniority orderings which have 3 more senior to agent(s) 1 and/or 2 result in a Serial Dictatorship yielding a popular match.*

1	*2*	*3*
A	***A***	***B***
B	*B*	*A*
C	***C***	***C***

Collectively, these Corollaries (also see simulations in Appendix A) establish that existence of complete popular matching is rare; particularly, given that correlation across agents' preferences is empirically common. Correlation across agents' preferences can originate from a general agreement over the most sought-after assignments (e.g., the most powerful, lucrative, and, highly sought after legislative committees in Congressional committee assignment (Thakur, 2018b) or from a shared distaste of bad placements (e.g., the least attractive, distressed, and unequivocally avoided states in civil servant assignments (Thakur, 2018a). Such shared blocks in agents' preferences render complete popular matching hard to attain. Moreover, a perfect lining up of seniority of agents by those getting their f-posts followed by those getting their s-posts under Serial Dictatorship is also hard to get in practice (e.g., is not the case in Republican committee assignment procedures that use Serial Dictatorship-based procedures (Thakur, 2018b). Existence of institutional stability under plurality rule is thus empirically rare and incredibly fragile.

4.3. **Majority Rule and Majority Stability.**
I illustrate the subtleties behind majority stability existence using two examples. Example 8 shows an instance where no popular matching exists, but a majority stable matching exists and can even be implemented by a Serial Dictatorship. Example 9 slightly tweaks preferences, such that correlation across preferences is close but less than in Example 8, yet there exist no majority stable allocations. The correlation coefficient is a generalization of Spearman rank coefficient to N rank order preference lists from Thakur (2018a).[4]

Example 8. *Consider agents* $\{1,2,3,4,5\}$ *ranking seats* $\{A,B,C,D,E\}$ *given a seniority ordering* $1 \succ 2 \succ 3 \succ 4 \succ 5$. *Since preferences have three blocks* $B_1 = A$, $B_2 = B$ *and* $B_3 = \{C,D,E\}$, *no complete popular matching exists. However, Serial Dictatorship in order of seniority produces the underlined allocation* $(1-A, 2-B, 3-C, 4-D, 5-E)$ *which is majority stable. Note: these preferences have a correlation coefficient of 0.84.*

1	*2*	*3*	*4*	*5*
<u>*A*</u>	*A*	*A*	*A*	*A*
B	<u>*B*</u>	*B*	*B*	*B*
C	*C*	<u>*C*</u>	<u>*D*</u>	<u>*E*</u>
D	*D*	*D*	*C*	*C*
E	*E*	*E*	*E*	*D*

The important question is why is this allocation majority stable? All agents' first and second choices are seats A and B respectively, so preferences are quite correlated. However,

[4]When considering the ranking of $i = 1, ..., N$ agents over $c = 1, ..., C$ seats, Thakur (2018a) suggests the following correlation measure

$$\rho = 1 - \frac{\sum_{c=1}^{C}\left(\frac{1}{N-1}\sum_{i=1}^{N}(r_{ic} - \overline{r}_c)^2\right)}{C\left(\frac{(C-1+1)^2-1}{12}\right)}$$

in this very observation lies the key to majority stability. In this example, four agents (2, 3, 4, and 5) are envious of agent 1's assignment to seat A and three agents (3, 4, and 5) are envious of agent 2's assignment to seat B. However, there is a *packing problem* as a majority of $\frac{N+1}{2}$ agents cannot be reassigned to less than a majority ($< \frac{N+1}{2}$) of distinct new seats. In this case, at least 3 agents cannot be assigned to just 2 seats.

Hence, to overrule an allocation by majority rule, $\frac{N+1}{2}$ distinct agents have to be envious of $\frac{N+1}{2}$ distinct seats. And satisfying this condition requires at least one *chain of envy*: a structure where i is envious of j's seat, and j is envious of k's seat, and so on.[5]

The majority stable matching in Example 8 admits 3 chains of envy of length 2: 5-2-1 (i.e., 5 is envious of 2, who is envious of 1), 4-2-1, 3-2-1. However, majority stability is ensured because there is no other envy to overcome the packing problem. As illustrated in Example 9, changing the preferences slightly can generate enough envy and unravel the majority stability.

Example 9. *Consider agents* $\{1,2,3,4,5\}$ *ranking seats* $\{A,B,C,D,E\}$ *given a seniority ordering* $1 \succ 2 \succ 3 \succ 4 \succ 5$. *Notice, the slight difference in agent 5's preference* ($D \succ_5 E$) *compared to that in Example 8* ($E \succ_5 D$). *There is yet another difference relative to Example 8 which had* $C \succ_4 E$, *instead of* $E \succ_4 C$, *that is inconsequential to the majority stability calculations, but used to comment on the correlation difference between Examples 8 and 9 below. This example has the same Serial Dictatorship allocation as Example 8. However, agent* 5 *now being envious of* 4*'s assigned seat* D *leads to institutional instability. Note: these preferences have a correlation coefficient of 0.82.*

1	*2*	*3*	*4*	*5*
$\underline{A}$	A	A	A	A
B	$\underline{B}$	B	B	B
C	C	$\underline{C}$	$\underline{D}$	D
D	D	D	E	$\underline{E}$
E	E	E	C	C

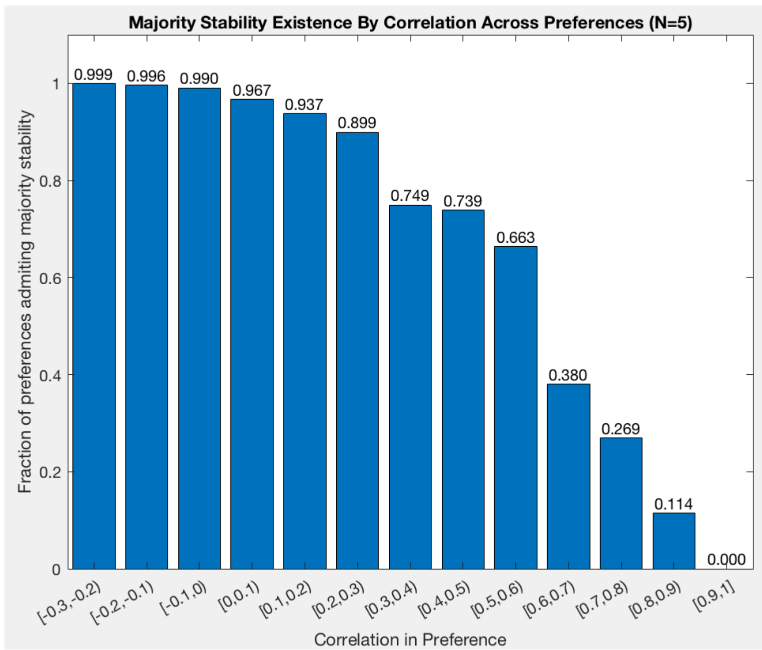
Together, Examples 8 and 9 show that undermining majority stability requires i) 'enough' envy overall to be able to move a majority and make them all better off, ii) the existence of chains of envy to overcome the packing problem, and iii) envy in the 'right' places, for example, multiple chains originating from the same two seat(s) A and B were superfluous as seen in Example 8.

Moreover, these examples show that these features are not perfectly captured by simple correlation across agents' preferences. Example 8 that admits majority stability has a preference correlation of 0.84, which is larger than preference correlation of 0.82 from Example 9 where there are no majority stable allocations. This highlights a more particular structure of preferences, I call *Deep Correlation.* Everyone being envious of the same handful of positions is often not sufficient for institutional instability; there must be sufficient amount of envy, some in the form of chains, and this envy and chains of envy must be in the right places to overcome the packing problem. These intricacies are not picked up by simple measures of correlation in preferences. Our simulation results in Appendix A further illustrate this

[5]The smallest possible chain of envy needed to break majority stability is of length 2, where the remaining $\frac{N-3}{2}$ agents are envious of distinct $\frac{N-3}{2}$ seats remaining.

distinction. Nevertheless, Figure 2 shows how, on the whole, the more correlated preferences are, the harder it is to sustain majority stability.

Figure 2. Majority Stability Existence By Correlation Across Preferences ($N = 5$).
Exhaustively checking whether there exists a majority stable allocation for $N = 5$ for 200,000 simulated set of preferences of differing correlations for 5 agents over 5 seats, shows the striking robustness of majority stability existence even when preferences across agents are considerably correlated.



Importantly however, Figure 2 highlights the striking robustness of majority stability existence even when preferences across agents are considerably correlated. Hence, majority and super-majority rules commonly used to change institutional rules in practice (e.g., changing party or Congressional rules for committee assignment procedures the US Senate (Thakur, 2018b)), might explain why certain assignment procedures are enduring, despite certain correlations across members' preferences, such as an overwhelming preference for the powerful and prestigious Appropriations Committee. Even in the All-India Civil Service, where civil servants systematically rank under-developed, distressed areas with foreign conflict and internal strife at the bottom of their preference lists and correlations are large enough to cause marked imbalances in allocations, correlation coefficients are only around 0.5 (Thakur, 2018a), where around 75% of preferences admit majority stable allocations.

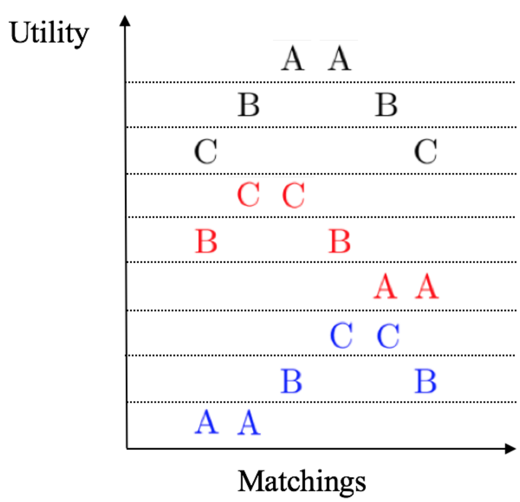
Now that we have an understanding of what makes majority stability more robust to correlated preferences than popular matching, I try to tackle the question: when does a majority stable allocation exist? This turns out to be a hard question. First, I show that common preference restrictions and tractability conditions from social choice—weak single-peaked preferences and order restricted preferences—have no traction in the social allocation choice setting. Second, I show that attempts at constructing candidate majority stable allocations by globally minimizing envy are extremely successful in testing whether there exists a majority stable allocation, but can (very rarely) fail to detect existence.

4.3.1. *Non-existence of Weakly Single-Peaked Preferences in Social Allocation Choice.* Under *weakly single-peaked preferences*, outcomes can be ordered along a left-right scale such that any move away from an agent's most preferred outcome is associated with a weak move down the agent's preference ordering (i.e., allowing flat spots of indifference at the peak of preference order as well as to either side of the peak). For the social allocation choice problem with majority rule, the restriction that an agent's preference over outcomes depends only on their preference over their own individual assignment across the allocations, puts such structure on the set of possible preferences, that there can be no weakly single-peaked preferences for markets of size $N > 3$. Thus, single-peakedness also has no bite in our Social Allocation Choice setting.

Lemma 2. *There is no possible set of weakly single-peaked preferences for a market sized* $N > 3$.

Proof: First consider restricting to any agents $\{1, 2, 3\}$ with preferences over matching allocations, where a matching $\{A, B, C\}$ implies match 1-A, 2-B, and 3-C. Let us maintain our restriction that all agents only care about their own assignment and that preferences are strict. Without loss of generality, consider agent 1's preference to be $A \succ_1 B \succ_1 C$. It is easy to verify that restricting to a market of size 3 (any 3 seats and 3 agents), there is only one possible (ignoring relabeling of seats/agents and reverse order of preferences) ordering of preferences which satisfies weak single-peakedness. This ordering is $\{C, B, A\}, \{B, C, A\}, \{A, C, B\}, \{A, B, C\}, \{B, A, C\}, \{C, A, B\}$, illustrated in Figure 3.

Figure 3. Weakly single-peaked preferences for $N = 3$.
Restricting to a market of size 3 (any 3 seats and 3 agents), there is only one possible (ignoring relabeling of seats/agents and reverse order of preferences) ordering of preferences which satisfies weak single-peakedness. Observe the utility symmetry of 1's preference relative to 2 and 3, and polar asymmetry of 2 and 3's preferences.



1	2	3
A	C	C
B	B	B
C	A	A

Now, consider restricting to any 4 agents $\{1, 2, 3, 4\}$. Agent 1 will be indifferent over all matchings where he gets his first choice, say D. For preferences to be single-peaked for agents 2, 3, and 4 over all allocations where 1 gets D, we know from above, that the ordering must be $\{D, C, B, A\}$, $\{D, B, C, A\}$, $\{D, A, C, B\}$, $\{D, A, B, C\}$, $\{D, B, A, C\}$, $\{D, C, A, B\}$. Now given that Agent 1 has strict preferences, 1 has $D \succ_1 C$. He is indifferent over all allocations where he gets C. For weak single-peakedness to hold, allocations where 1 gets C must either be to the right, to the left, or split across both sides of the ordering above. Either way, there will be an allocation where 2 gets B. And regardless of where this allocation is placed, weak single-peakedness is violated for 2's preference.

It is the intrinsic rivalry and discreteness that the matching problem imposes that prevents the existence of weakly single-peaked preferences, e.g., if I get A, then you are left with B, but then I prefer all the allocations where I get A over those where I get B, and so forth.

4.3.2. *Non-existence of Order Restricted Preferences in Social Allocation Choice.* Under *order-restricted preferences*, agents are ordered along a left-right scale such that for any pair of outcomes, the agents who strictly prefer one outcome all lie to one side of those who strictly prefer the other, and agents in the middle are indifferent. For the Social allocation choice problem with majority rule, the restriction that an agent's preference over outcomes depends only on their preference over their own individual assignment across the allocations, puts such structure on the set of possible preferences, that there can be no order restricted preferences for markets of size $N \geq 3$. Thus, order restriction have no bite in our Social Allocation Choice setting.

Lemma 3. *There is no possible set of order restricted preferences for a market sized $N \geq 3$.*

Proof: It suffices to show that there is no order restriction possible for any subset of 3 seats and 3 agents (i.e., subset of the market of size 3). Consider agents $\{1, 2, 3\}$ with preferences over allocations, where a matching $\{A, B, C\}$ implies match $(1 - A, 2 - B, 3 - C)$. Let us maintain our restriction that all agents only care about their own assignment and that preferences are strict.

So without loss of generality, consider agent 1's preference to be $A \succ_1 B \succ_1 C$. Hence in comparing the two allocations $\{A, B, C\}$ and $\{A, C, B\}$, agent 1 is indifferent. This leaves four cases for the possible preference combinations that agents 2 and 3 can have over seats B and C.

- Case 1: if $B \succ_2 C$ and $B \prec_3 C$.
  Then, the order restriction is 2-1-3 (or reverse) given the comparison of the two allocations above. However, consider the comparison between $\{C, B, A\}$ and $\{B, C, A\}$. Preferences are $\succ_2$, $\prec_1$, and $\sim_3$, which violates the order restriction.
- Case 2: if $B \prec_2 C$ and $B \succ_3 C$.
  Then, the order restriction is 2-1-3 (or reverse) given the comparison of the two allocations above. However, consider the comparison between $\{C, A, B\}$ and $\{B, A, C\}$. Preferences are $\sim_2$, $\prec_1$, and $\succ_3$, which violates the order restriction.
- Case 3: if $B \succ_2 C$ and $B \succ_3 C$.
  Then, the order restriction is either 1-3-2 (or reverse) or 1-2-3 (or reverse) given the comparison of the two allocations above.
  - Case 3a: if order restriction is 1-3-2 (or reverse).
    Consider the comparison between $\{C, A, B\}$ and $\{B, A, C\}$. Preferences are $\prec_1$, $\succ_3$, and $\sim_2$, which violates the order restriction.
  - Case 3b: if order restriction is 1-2-3 (or reverse).

Consider the comparison between $\{C, B, A\}$ and $\{B, C, A\}$. Preferences are $\prec_1$, $\succ_2$, and $\sim_3$, which violates the order restriction.

- Case 4: if $B \prec_2 C$ and $B \prec_3 C$.
  Then, the order restriction is either 1-3-2 (or reverse) or 1-2-3 (or reverse) given the comparison of the two allocations above.
  - Case 4a: if order restriction is 1-3-2 (or reverse).
    Consider the comparison between $\{C, B, A\}$ and $\{B, C, A\}$. Preferences are $\prec_1$, $\sim_3$, and $\prec_2$, which violates the order restriction.
  - Case 4b: if order restriction is 1-2-3 (or reverse).
    Consider the comparison between $\{C, A, B\}$ and $\{B, A, C\}$. Preferences are $\prec_1$, $\sim_2$, and $\prec_3$, which violates the order restriction.

4.3.3. *Constructive Approach to Finding a Majority Stable Allocation: minimizing envy.* In characterizing whether there exists a majority stable allocation for a given set of preferences, since we want to effectively minimize realizable envy, we might try Algorithm 1) minimizing average preference rank or Algorithm 2) greedily minimizing preference ranks allotted[6] in the hopes of finding a majority stable allocation.

Lemma 4. *If preferences admit a majority stable allocation, there exists a pareto efficient and majority stable allocation.*

> **Proof**: If a majority stable allocation is not pareto efficient, then overall envy is weakly decreased by allowing all cycles via Top Trading Cycles because all those who are better off, have weakly less envy; whereas, those who are not any worse off, have the same envy. We thus reach a pareto efficient and majority stable allocation with Top Trading Cycles.

Both Algorithm 1) and Algorithm 2) return pareto efficient allocations, which is useful in light of Lemma 4, and globally minimize envy in different ways. Moreover, greedily minimizing preference ranks allotted also picks up a complete popular matching if one exists.

I find from simulations that Algorithm 1) and Algorithm 2) are extremely successful in picking up a majority stable allocations if any majority stable allocations exist: both algorithms failed to pick up only 1% of random preference environments where majority stable allocations existed for simulations with $N = 7$. Example 10 illustrates how both Algorithms can fall short in identifying majority stable allocations in scenarios where they in fact exist.

Example 10. *Consider agents* $\{1, 2, 3, 4, 5, 6, 7\}$ *ranking seats* $\{A, B, C, D, E, F, G\}$. *A majority stable matching is in bold. Minimum average preference matching is in red, and greedy minimum preference matching is underlined.*

[6]This matching can be calculated by assigning weight N^r to seat ranked rth (where higher rank is better) and then solving the maximal weighted matching using an integer linear program, that is easy to compute. Since $N^r > (N-1)(N^{r-1})$, subject to the overall maximum, the integer linear program will always greedily favor higher ranks. In other words, this matching finds the tie-breaking rules which optimize Boston mechanism's favoring ranks property.

1	*2*	*3*	*4*	*5*	*6*	*7*
F	***G***	***D***	*G*	*F*	*D*	*G*
A	*F*	*A*	***A***	***B***	*A*	*F*
C	*A*	*E*	*D*	*C*	*E*	***E***
G	*C*	*F*	*C*	*D*	*F*	*B*
E	*D*	*C*	*B*	*G*	*B*	*D*
D	*E*	*G*	*E*	*A*	*G*	*C*
B	*B*	*B*	*F*	*E*	***C***	*A*

In Example 10, there exists a majority stable matching $(1-F, 2-G, 3-D, 4-A, 5-B, 6-C, 7-E)$ however, both the minimum average preference matching $(1-F, 2-C, 3-A, 4-G, 5-B, 6-D, 7-E)$ and greedy minimum preference matching $(1-F, 2-C, 3-D, 4-A, 5-B, 6-E, 7-G)$ deliver non-majority stable matchings. Notice how both the Algorithms fail to assign agent 6 his least preferred assignment C, as this increases average rank and doesn't favor higher ranks; however, it turns out that this is needed so that other agents' assignments prevent envy in the necessary places to be able to improve a majority.

5. **Is a given allocation institutionally stable?**

Given a set of preferences, checking whether a matching is majority stable and finding the size of the maximal coalition that can benefit from some alternative matching seems like a computationally hard problem. The number of alternative matchings for a one-to-one matching in a balanced market with N candidates and N seats grows exponentially, $N!$. However, I show how to use an integer linear program to efficiently solve this problem.

The integer linear programming approach takes the graph of envy in matrix form and finds an alternative one-to-one matching with maximal envy resolved. It turns out that the constraint that the alternative matching should be a one-to-one matching, imposes a structure such that the relaxed linear program is guaranteed integer solutions.

Consider the reassignment matrix X, where $X_{ij} = \begin{cases} 1, \text{ if } i \text{ is reassigned to } j\text{'s seat} \\ 0, \text{ otherwise} \end{cases}$.

Note that $X_{ii} = 1$ means that i remains in his own seat.

Consider the envy matrix A, where $A_{ij} = \begin{cases} 1, \text{ if } i \text{ is envious of } j\text{'s seat} \\ 0, \text{ otherwise} \end{cases}$.

Note, i envious of j's seat means that i strictly prefers j's seat to his own assignment.

The goal is thus to find the reassignment X such that the maximal envy is realized subject to the alternative allocation being a one-to-one matching of the balanced market:

$$\begin{aligned} \max_{X} \quad & \textstyle\sum_{diag} A'X \\ s.t. \quad & \textstyle\sum_i X_{ij} = 1 \quad \forall j \\ & \textstyle\sum_j X_{ij} = 1 \quad \forall i \end{aligned} \tag{1}$$

The first set of constraints guarantee that each seat j is assigned to only one agent, while the second set of constraints guarantee that each agent is assigned only one seat. The sum over the diagonal entries of $A'X$ calculates the envy that is realized by the reassignment X

and summing over just the diagonal entries of the matrix ensure no double counting of the envy.

I can now rewrite this problem in vector form:

I stack all the columns of X and of A to form $\tilde{X} = \begin{bmatrix} X_{11} \\ ... \\ X_{N1} \\ X_{12} \\ ... \\ X_{N2} \\ ... \\ ... \\ X_{1N} \\ ... \\ X_{NN} \end{bmatrix}$ and $\tilde{A} = \begin{bmatrix} A_{11} \\ ... \\ A_{N1} \\ A_{12} \\ ... \\ A_{N2} \\ ... \\ ... \\ A_{1N} \\ ... \\ A_{NN} \end{bmatrix}$.

Furthermore, consider the matrices $2N$-by-1 matrix, $b = \begin{bmatrix} 1 \\ 1 \\ ... \\ ... \\ ... \\ 1 \end{bmatrix}$ and the $2N$-by-N^2 matrix,

$$C = \left[\begin{array}{cccccccccccccccc} \multicolumn{4}{c}{\overbrace{\hphantom{1\ 1\ ...\ 1}}^{\text{length N}}} & \multicolumn{4}{c}{\overbrace{\hphantom{0\ 0\ ...\ 0}}^{\text{length N}}} & & & & & \multicolumn{4}{c}{\overbrace{\hphantom{0\ 0\ ...\ 0}}^{\text{length N}}} \\ 1 & 1 & ... & 1 & 0 & 0 & ... & 0 & ... & ... & ... & ... & 0 & 0 & ... & 0 \\ 0 & 0 & ... & 0 & 1 & 1 & ... & 1 & ... & ... & ... & ... & 0 & 0 & ... & 0 \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & 0 & 0 & 0 & ... & 0 & ... & ... & ... & ... & 1 & 1 & ... & 1 \\ \hdashline 1 & 0 & ... & 0 & 1 & 0 & ... & 0 & ... & ... & ... & ... & 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 & 0 & 1 & ... & 0 & ... & ... & ... & ... & 0 & 1 & ... & 0 \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & 1 & 0 & 0 & ... & 1 & ... & ... & ... & ... & 0 & 0 & ... & 1 \\ \multicolumn{4}{c}{\underbrace{\hphantom{0\ 0\ ...\ 1}}_{\text{N-dim identity mtx}}} & \multicolumn{4}{c}{\underbrace{\hphantom{0\ 0\ ...\ 1}}_{\text{N-dim identity mtx}}} & & & & & \multicolumn{4}{c}{\underbrace{\hphantom{0\ 0\ ...\ 1}}_{\text{N-dim identity mtx}}} \end{array}\right]$$

Hence, our integer linear program from Equation (1) can be equivalently expressed as

$$\begin{aligned} & \max_{X} \tilde{A}'\tilde{X} \\ s.t. \quad & C\tilde{X} = b \end{aligned} \tag{2}$$

I can divide the matrix C into submatrices on either side of the dashed line:

$$C = \begin{bmatrix} \overline{C} \\ - \\ \underline{C} \end{bmatrix}$$

, where $\overline{C}x = b_{NxN^2}$ ensures that each seat is allotted to only one agent while $\underline{C}x = b_{NxN^2}$ ensures that each agent is allotted to only one seat.

Note, that matrices $\tilde{A}$, C, and b have all integer entries and C is totally unimodular.[7] Hence, the relaxed linear program will have integer solutions and is computationally easy.

6. Which mechanisms are more robust to institutional stability?

I use simulations to compare a Serial Dictatorship in a given order of seniority with a Boston mechanism with tie-breaks in order of seniority, assuming truthful preference revelation. I make two observations. First, the maximal sized coalition that can be made better off is weakly larger under Serial Dictatorship, than under the Boston mechanism (see Figure 11 left) for most preference environments. Second, Boston mechanism under truthful revelation is more robust for majority stability and popular matching existence (see Figure 11 right). Out of 12,001 simulated preference environments with varying levels of correlation, there were i) 5154 instances where both were majority stable, ii) 4673 instances where both were not majority stable, iii) 2169 instances where the Boston mechanism was majority stable but Serial Dictatorship is not, and iv) only 5 instances where Serial Dictatorship was majority stable but Boston mechanism was not.

It is the favoring higher ranks property—that any seat c that i is envious of is assigned to some j who ranks c at least as high as i—of the Boston mechanism that tends to produce less envy, compared to Serial Dictatorship. Algorithm 2) from Section 4.3.3 globally optimizes on this very property of favoring higher ranks.[8]

I show that there are cases where the Boston mechanism is majority stable and Serial Dictatorship is not (Example 11), and others where Serial Dictatorship is majority stable but Boston mechanism is not (Example 12).

Example 11. *Consider agents* $\{1, 2, 3, 4, 5\}$ *ranking seats* $\{A, B, C, D, E\}$ *given a seniority ordering* $1 \succ 3 \succ 4 \succ 5 \succ 2$. *Boston mechanism with truthful preference and same seniority ordering for tie-breaking is majority stable (1-A, 2-B, 3-C, 4-D, 5-E), while the matching from Serial Dictatorship (1-A, 2-C, 3-B, 4-D, 5-E) is not due to alternative matching (1-C, 2-E, 3-A, 4-D, 5-B) being preferred with a 3:1 vote.*

1	*2*	*3*	*4*	*5*
A	B	A	A	A
B	A	B	B	B
C	D	C	D	E
D	E	D	C	D
E	C	E	E	C

Example 12. *Consider agents* $\{1, 2, 3, 4, 5\}$ *ranking seats* $\{A, B, C, D, E\}$ *given a seniority ordering* $1 \succ 2 \succ 3 \succ 4 \succ 5$. *Serial Dictatorship matching (1-A, 2-B, 3-E, 4-C, 5-D) is majority stable, while Boston mechanism with truthful preference and same seniority ordering for tie-breaking (1-A, 2-C, 3-B, 4-E, 5-D) is not majority stable as alternative matching (1-E, 2-A, 3-D, 4-C, 5-B) is preferred by 3:2 vote.*

[7] Heller and Tompkins (1956) show that the coefficient matrix for a bipartite matching is totally unimodular. And $\tilde{A}$ is an unoriented incidence matrix of a bipartite graph. Equivalently, Hoffman and Gale (1956) characterize more general sufficiency conditions which $\tilde{A}$ satisfies.

[8] Algorithm 2) can be viewed as finding the best tie-breaking criterion for the Boston mechanism such that the favoring higher ranks property is optimized by greedily assigning seats by lowest preference rank.

1	*2*	*3*	*4*	*5*
A	A	B	A	A
B	B	A	B	B
C	C	E	C	D
D	D	D	D	C
E	E	C	E	E

7. Generalizing to allow for Inter-dependent Preferences.

A major simplification I have made in the set up thus far (Assumption 4) is that agent i's preferences over allocations only depend on his own seat assignment, and not on others' assignments (what I will refer to as "*individual preferences*"). In this section I relax this assumption and consider the effect of "*selfish but inter-dependent preferences*:" i first cares about his own assignment lexicographically, but he can have ***any*** preferences over allocations where his assignment is the same. This is a very general class of preferences which includes caring about the well being of one's allies, having enemies who one derives utility from misfortune, caring about one's coalition, caring about one's coalition in an ordered way (caring about the well-being of agent j the most, followed by k,...), and even very idiosyncratic preferences like, as long as I get A, I want j to get C and k to get D; but as long as I get B, I want j to get E and k to get F.[9]

Rather strikingly, I find that introducing selfish but inter-dependent preferences can only lead to weakly more majority instability. This is established by the following two Lemmas.

Lemma 5. *If an allocation is not majority stable with individual preferences, there exist no selfish but inter-dependent preferences that lead to make it majority stable.*

Proof: This is follows from definition of majority stability since an allocation is not majority stable if there exists a majority that strictly prefers an alternative allocation (based on their own selfish preferences over their own allocations). You cannot have selfish but inter-dependent preferences which change this group's voting behavior over these two allocations.

Lemma 6. *If there are individual preferences which generate majority stable allocations, there could be selfish but inter-dependent preferences that lead to non-existence of majority stable allocations.*

Proof: by Example 13.

Example 13. *Consider agents* $\{1, 2, 3, 4, 5\}$ *ranking seats* $\{A, B, C, D, E\}$.

1	*2*	*3*	*4*	*5*
A	A	A	D	E
B	B	B	A	A
C	C	C	B	B
D	D	D	C	C
E	E	E	E	D

[9]The only constraint on selfish but inter-dependent preferences is that agent i never votes against his own interest. For example, i being altruistic—say by preferring to make i's own assignment worse off to benefit a friend's assignment—violates selfish but inter-dependent preferences.

In Example 13, with individual preferences, assigning 1,2, and 3 to seats A, B, and C in any way and matching 4-D and 5-E, are all 6 majority stable allocations.

However, consider if 4's ally was 1 and 5's ally was 2. Then no majority stable allocation exists. Not giving 4 or 5 seats D and E leads to majority instability. Furthermore, matching 1-A leads to overthrow by coalition of 2, 3, and 5; matching 2-A leads to overthrow by 1, 3, and 4; and matching 3-A leads to overthrow by 1, 2, 4, and 5.

Theorem 1. *The set of majority stable allocations with selfish but inter-dependent preferences is a weak subset of the set of majority stable allocations with individual preferences.*

Proof: This is implied by Lemmas 5 and 6 put together.[10]

Notice, that Example 13 also serves as an example of how popular matching (plurality vote) does not have such a unidirectional effect of selfish but inter-dependent voting. Namely, there is no complete popular matching with individual preferences. However, suppose 4 was 1's ally. Then, $(1-A, 2-B, 3-C, 4-D, 5-E)$ is popular. This is because with individual preferences, 2 and 3 would each move up 1 preference rank at the expense of 1. However, with selfish but inter-dependent preferences and 1 having 4 as his ally, the vote for the alternative matching is tied at 2:2. This example thus illustrates how selfish but inter-dependent preferences can only push majority stability towards weakly more institutional instability, while it can lead to more or less institutional stability under popular matching with plurality rule.

For selfish but inter-dependent preferences, I assumed that an agent has strict preferences over his own assignment, but across two allocations where his own assignment is the same, he might have preferences over what others get as well. Namely, I have assumed preferences over one's own assignment are independent of others' assignments. With fully inter-dependent preferences—where your preference over your own assignment depends on what others are assigned[11]—institutional stability—even majority stability—can be both expanded or undermined.

8. **Correlated Preferences in Social Choice vs. Social Allocation Choice.**

Correlation across agents' preferences has opposing effects on stability in social choice as compared to in social allocation choice. Cycles in the preference aggregation rule—whether it be violations of transitivity, acyclicity, or quasi-transitivity—undermine stability in social choice (Austen-Smith and Banks, 2000). The more correlated preferences are across agents, the less likely the preference aggregation rule has cycles. For social allocation choice however, institutional stability is undermined by chains of envy. And the more correlated preferences are across agents, the more likely, preferences form chains of envy leading to institutional instability.

[10]Example 13 had 6 majority stable allocations under individual preferences and 0 majority stable allocations under the selfish but inter-dependent preferences. Consider an example with three agents $\{1, 2, 3\}$ assigned to three seats $\{A, B, C\}$. 1 and 2 have preference $A \succ B \succ C$, and 3 has preference $C \succ B \succ A$. Then 1-A, 2-B, 3-C and 1-B, 2-A, 3-C are two majority stable allocations. But now suppose 3 is in a coalition with 1. Then only 1-A, 2-B, 3-C remains as a majority stable matching.

[11]For example, if agent 2 is assigned C, agent 1 prefers A to B, but if agent 3 is assigned C agent 1 prefers B to A.

This point can be illustrated by two simple examples on the two extreme ends of correlation across agents' preferences.

Example 14. *Correlation of preference across agents is 0. The social choice majority preference relation f lacks transitivity since $A \succ_f B \succ_f C \succ_f A$. However, the underlined allocation is majority stable (moreover, institutionally stable by pareto efficiency and popular matching sense as well) because all agents get their first choice and there is no envy.*

1	*2*	*3*
$\underline{A}$	$\underline{B}$	$\underline{C}$
B	C	A
C	A	B

Example 15. *Correlation of preference across agents is 1. The social choice majority preference relation f is stable and A is the Condorcet winner since $A \succ_f B \succ_f C$ and $A \succ_f C$. However, perfectly correlated preferences render any social allocation choice institutionally unstable under popular matching and majority stability (though any allocation is pareto efficient), allotting C to whoever is assigned A, and then moving the other two agents to 1 higher preference gives both a plurality and majority overrule.*

1	*2*	*3*
$\underline{A}$	A	A
B	$\underline{B}$	B
C	C	$\underline{C}$

Thus, on the whole (recall Examples 8 and 9), increased correlation in preferences across agents tends to undermine institutional stability in social allocation choice while bolstering stability in social choice, as it increases the likelihood of chains of envy while decreasing the likelihood of cycles in the social choice rule.

To further highlight the effect of correlated preferences, I show that when agents' preferences are perfectly correlated, there exists a majority-approved sequence of reassignments from any allocation to any other allocation, akin to McKelvey's Chaos Theorem (McKelvey, 1976).[12] Namely, this result highlights the unrestrained power of the agenda setter.

Theorem 2. *When agents' preferences are perfectly correlated, there exists a majority-approved sequence of reassignments from any allocation to any other allocation, for $N \geq 5$.*

Proof: There are N seats 1,...,N to be assigned to N agents $1, ..., N$. Suppose each agent's preference over seats is identical, say $1 \succ 2 \succ ... \succ N$. Fixing the order of seats[13] to be $1, ..., N$. An allocation hence defines a permutation of agents over this ordered sequence of seats. For example, an allocation denoted by $5, 2, 9,...$ implies that agents $5, 2, 9,...$ were assigned seats 1,2,3,... respectively.

[12]I thank Jonathan Bendor for his engaging discussion and encouragement that lead to this result.

[13]This is key, if you fix order of agents and try permutations of seats on this order, it does not form a group, as movements cannot be composed. Namely, the product of two movements which are individually majority-approved from some initial allocation, cannot necessarily be combined to a majority-approved path (namely, the first movement leads to a new status quo, and hence the second movement might not be majority-approved from this new status quo).

In each step, a feasible movement is defined by the preferences over seats and majority rule (only movements which are approved by a majority are allowed). Because we fixed the order of seats and all agents' preferences are identical, by definition, the composition of feasible movements is also feasible. This is because it only matters which positions are moved, rather than who is in those positions. The set S of permutations defined by the feasible movements is thus closed under composition and thus a subgroup of S_N (the finite symmetric group on N objects) because for any element $s \in S$, it is has a *finite* order k in S_N. Thus s^k and s^{k-1}, which are the identity and inverse of s respectively, are in S because of closure under composition. We want to show that group S is the full group S_N.

An N-cycle is the movement from $1, 2, ..., N$ to $N, 1, 2, ..., N-1$. A $1:2$ transposition is the movement from $1, 2, ..., N-1, N$ to $2, 1, 3, ..., N$. That the set of all possible permutations, S_N, is generated (i.e., a successive composition of these two operators) using these two movements, is a theorem in abstract algebra for finite symmetric groups, stated usually as $(1,2)(1,2,...,N)$ generate the group S_N.

In this setting, an N-cycle is allowed by N-1:1 vote. Moreover, $1:2$ transposition is shown for odd $N \geq 5$ and even $N \geq 6$ with a sequence of majority-approved movements, in Figure 4.[14]

(While N-cycles work for $N \geq 3$ by majority rule, for $N = 3$ and $N = 4$, it is not possible to get any consecutive transposition like $1:2$ or $N:1$ by majority rule).

Corollary 2.1, which is rather the corollary of the proof method, illustrates that Theorem 2 is not just a knife-edge result at the extreme case of perfectly correlated preferences, by exploiting the slack that is available from agendas that are approved by super-majorities under fully correlated preferences.[15] More generally, as the correlation across agents' preferences increases, there tend to be more allocations that can be spanned by majority-approved agendas.

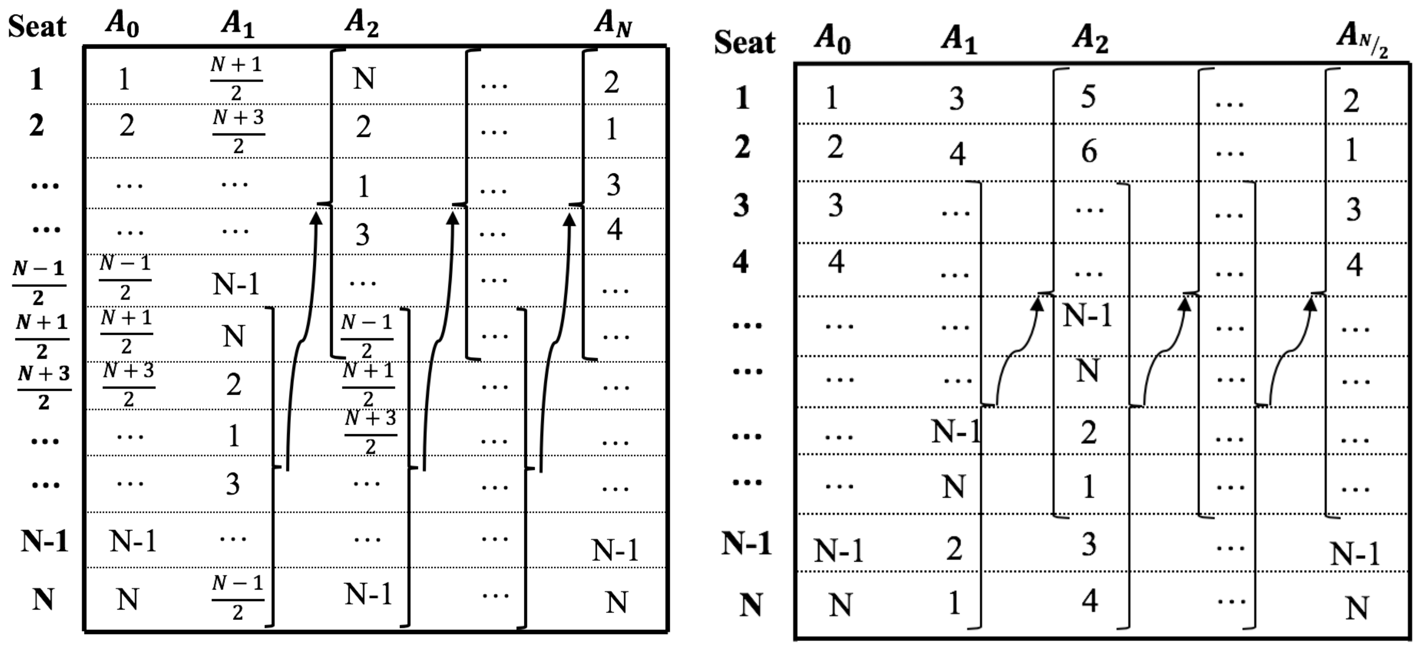
Corollary 2.1. *If* $\left\lceil \frac{N}{2} \right\rceil + 3$ *or more agents have fully correlated preferences, there exists a majority-approved sequence of reassignments from any allocation to any other allocation, for* $N \geq 6$.

Proof: With all N agents having fully correlated preferences over seats $1 \succ 2 \succ ... \succ N$, for N even, Figure 4 (right) gives an path of allocations that is approved in each step by $N-2$ agents. Hence, if preferences of at most $\frac{N}{2} - 3$ agents were arbitrarily changed, each step of this path is still approved by at least $N - 2 - (\frac{N}{2} - 3) = \frac{N}{2} + 1$, which is a majority. Similarly, for N odd, Figure 5 gives a path of allocations that is approved in each step by $N-3$ agents when all N agents have fully correlated preferences. Hence, if preferences of at most $\frac{N+1}{2} - 3$ agents were arbitrarily changed, each step of this path is still approved by at least $N - 3 - (\frac{N+1}{2} - 3) = \frac{N+1}{2}$, which is a majority.

[14] These paths can be generated more efficiently, but for simplicity of the diagrams, I have illustrated paths with N and $\frac{N}{2}$ steps.

[15] Lagrange's theorem also tells us that the number of unique allocations spanned by any initial allocation through majority-approved paths of reassignments must be a factor of $N!$. Note however, that an arbitrary set of preferences and majority rule might not form a group.

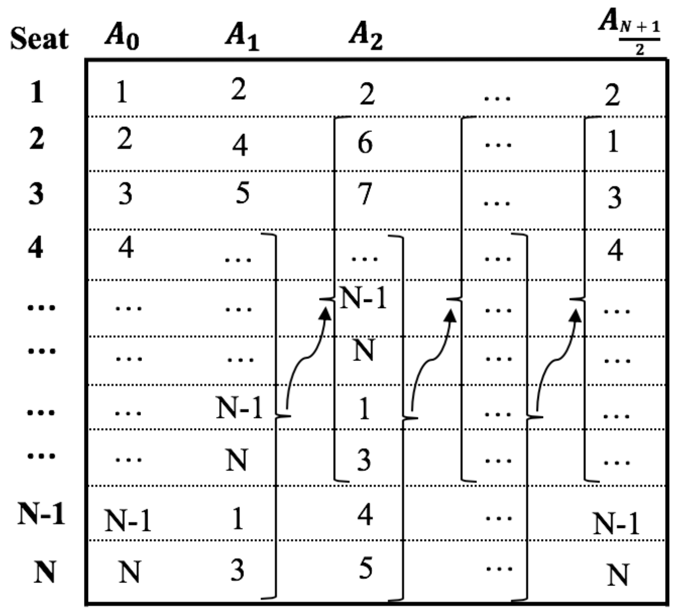
Figure 4. Sequence of reassignments for 1:2 transposition for N odd (left) and N even (right).
Fixing the sequence of seats $1, ..., N$, the figures show the sequences of majority-approved allocations $A_0, A_1, ...$ that lead to a $1 : 2$ transpositions for odd N (left) and even N (right). For odd N, at each step in the sequence from $A_1, ..., A_N$, the bottom $\frac{N+1}{2}$ agents to the top, hence approved by $\frac{N+1}{2}$ agents. For even N, at each step in the sequence from $A_1, ..., A_{\frac{N}{2}}$, the bottom $N-2$ agents are moved to the top, hence approved by $N-2$ agents.



9. **Conclusions.**

In this paper, I have developed a new notion of institutional stability that analyzes the endogenous coalition mobilization and political economy of reorganization and institutional change in self-governing bodies. I juxtapose three voting rules (plurality, majority, and unanimity) which correspond to three notions of institutional stability (popular matching, majority stability, and pareto efficiency), characterize preference environments where institutionally stable allocations exist, and compare the robustness of mechanisms in attaining institutionally stable allocations. Unlike pareto efficiency with unanimity rule which is easy to achieve but too non-discriminating as any agent(s) can block extremely large coalitions from improving at their expense; and unlike popular matching which fails to exist with even little correlation across agents' preferences; I show that majority rule—which is particularly relevant in majoritarian institutions found in political economy such as the US Senate—exhibits a striking robustness to correlation across agents' preferences and is attainable using for commonly used mechanisms such as Serial Dictatorship or Boston mechanism. The key to this robustness lies in the packing problem which arises when using majority rule in

Figure 5. Sequence of reassignments for 1:2 transposition for $N \geq 7$ odd.
Fixing the sequence of seats $1, ..., N$, the figures show the sequences of majority-approved allocations $A_0, A_1, ...$ that lead to a $1:2$ transpositions for $N \geq 7$ odd. At each step in the sequence from $A_1, ..., A_{\frac{N+1}{2}}$, the bottom $N-3$ agents are moved to seats $2, ..., N-2$, while agent 2 remains in seat 1.



social allocation choice settings, necessitating chains of envy for a potential majority overrule. See Appendix B for discussion on properties of mechanism and voting rules that attain institutional stability.

My results indicate why despite certain systematic correlation in preferences, such as Senators consistently ranking that powerful prestige committees in the U.S. Senate at the top of their preference lists (Thakur, 2018b), Indian civil servants consistently ranking underdeveloped, distressed states at the bottom of their preference lists (Thakur, 2018b), or workers preferring teams doing the most exciting, high-profile, and novel work within a firm (Cowgill and Koning, 2018), these institutions might not necessarily undergo institutional reform by changing their matching mechanisms. However, our results also highlight when to expect institutional reform and reorganization. For example a financial crisis that renders new jurisdiction, power, and credit-claiming opportunities to Finance, Banking, and Commerce committees can suddenly align committee preferences across agents. Or if a political party wins a big landslide of multiple elections from a certain geographic region, the newly elected representatives may all have very similar committee preferences of say prestige committees, followed by agriculture, followed by veterans affairs,... In such cases, a sudden increase in correlation across agents' preferences and the resulting chains of envy, can lead to endogenous formation of coalitions seeking to reform the institution.

A complete characterization of the environments in which there exist majority stable allocations, theoretical comparisons across different mechanisms in their robustness to institutional stability, and generalizations of the model to two-sided matching and two-sided voting are all avenues for future work.

References

[1] Abraham, David J. Irving, Robert W., Kavitha, Telikepalli, and Kurt Mehlhorn. "Popular matchings." SIAM Journal on Computing, 37(4) (2007): 1030-1045.

[2] Abraham, David J., and Telikepalli Kavitha. "Voting paths." SIAM Journal on Discrete Mathematics 24.2 (2010): 520-537.

[3] Austen-Smith, David, and Jeffrey S. Banks. Positive political theory. Vol. 2. University of Michigan Press, 1999.

[4] Aziz, Haris, Felix Brandt, and Paul Stursberg. "On popular random assignments." International Symposium on Algorithmic Game Theory. Springer, Berlin, Heidelberg, 2013.

[5] Barbera, Salvador, and Matthew O. Jackson. "Choosing how to choose: Self-stable majority rules and constitutions." The Quarterly Journal of Economics 119.3 (2004): 1011-1048.

[6] Biro, Peter, Robert W. Irving, and David F. Manlove. "Popular matchings in the marriage and roommates problems." International Conference on Algorithms and Complexity. Springer, Berlin, Heidelberg, 2010.

[7] Cowgill, Bo, and Rembrand Koning. "Matching Markets for Googlers. Harvard Business School Case 718-487, March 2018. (Revised August 2018.)

[8] Cseh, Agnes. "Popular matchings." Trends in Computational Social Choice (2017): 105.

[9] Cseh, Agnes, Chien-Chung Huang, and Telikepalli Kavitha. "Popular matchings with two-sided preferences and one-sided ties." SIAM Journal on Discrete Mathematics 31.4 (2017): 2348-2377.

[10] Cseh, Agnes, and Telikepalli Kavitha. "Popular edges and dominant matchings." Mathematical Programming 172.1-2 (2018): 209-229.

[11] Heller, Isidore, and C. B. Tompkins. "An extension of a theorem of Dantzig's." Linear inequalities and related systems 38 (1956): 247-254.

[12] Hoffman, A. J., and D. Gale. "An extension of a theorem of Dantzig's." Linear Inequalities and Related Systems. Princeton UP, 1956. 252-254.

[13] Huang, Chien-Chung, and Telikepalli Kavitha. "Popular matchings in the stable marriage problem." Information and Computation 222 (2013): 180-194.

[14] Huang, Chien-Chung, and Telikepalli Kavitha. "Popularity, mixed matchings, and self-duality." Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms. Society for Industrial and Applied Mathematics, 2017.

[15] Kavitha, Telikepalli. "A size-popularity tradeoff in the stable marriage problem." SIAM Journal on Computing 43.1 (2014): 52-71.

[16] Kavitha, Telikepalli, and Meghana Nasre. "Optimal popular matchings." Discrete Applied Mathematics 157.14 (2009): 3181-3186.

[17] Kavitha, Telikepalli, Julián Mestre, and Meghana Nasre. "Popular mixed matchings." Theoretical Computer Science 412.24 (2011): 2679-2690.

[18] Koray, Semih. "Self-selective social choice functions verify Arrow and Gibbard-Satterthwaite theorems." Econometrica 68.4 (2000): 981-996.

[19] Maggi, Giovanni, and Massimo Morelli. "Self-enforcing voting in international organizations." American Economic Review 96.4 (2006): 1137-1158.

[20] Manlove, David. Popular Matchings, Chapter 7 in Algorithmics of matching under preferences. Vol. 2. World Scientific, 2013.

[21] McCutchen, Richard Matthew. "The least-unpopularity-factor and least-unpopularity-margin criteria for matching problems with one-sided preferences." Latin American Symposium on Theoretical Informatics. Springer, Berlin, Heidelberg, 2008.

[22] McDermid, Eric, and Robert W. Irving. "Popular matchings: structure and algorithms." Journal of combinatorial optimization 22.3 (2011): 339-358.

[23] McKelvey, Richard D. “Intransitivities in multidimensional voting models and some implications for agenda control.” (1976).
[24] Messner, Matthias, and Mattias K. Polborn. “Voting on majority rules.” The Review of Economic Studies 71.1 (2004): 115-132.
[25] North, Douglass. ”Institutions, institutional change and economic performance Cambridge University Press.” New York (1990).
[26] Sng, Colin TS, and David F. Manlove. “Popular matchings in the weighted capacitated house allocation problem.” Journal of Discrete Algorithms 8.2 (2010): 102-116.
[27] Thakur, Ashutosh. “Matching in the Civil Service: A Market Design Approach to Public Administration and Development.” Working Paper (2018)
[28] Thakur, Ashutosh. “Matching Politicians to Committees.” Working Paper (2018)

Appendix A. Simulation Results: existence of Majority Stability and Popular Matching.

A.1. Calibration to the U.S. Senate.

We base our simulation on the application of assigning US Senators to committees (Thakur, 2018b). Institutional stability in this mechanism is undermined by i) increased impatience (e.g., close election or retirement which imply a short time horizon or large discounting of utility from committee assignments in the future), ii) a large freshman class since freshmen have lowest priority and no existing tenant assignment with property rights,[16] and iii) increased correlation across as' preferences, which increases competition over seats and leads to increased likelihood of envy.

We calibrate the vacancies to the 97th Congress (1981-1983), which had the largest proportion of 16/53 freshmen amongst the Republican Senators from the 83rd-113th Congresses. The 97th Congress had 292 seats across the 16 Standing committees. First, I assumed that Republican seats resembled the 53% of seats in each of these committees (in practice, similar to this assumption, seats on a committee for each party resemble the party ratio in the Senate). Next, since the average number of seats held by senator was around 3, but I wanted to simply the matching to be many-to-one (i.e., consider each politician being assigned to only one committee), I divided each of these seats in each committee by 3. Finally, I rounded to get the total number of vacancies in each of the 16 standing committees. The calibration to the 97th Congress with many-to-one matching gives 53 vacancies across 16 standing committees for the 53 Senators shown in Table 1).

To make the simulations more tractable, we make a few simplifications. First, instead of a many-to-many matching (where each committee is assigned multiple agents and each agent is assigned multiple committees), we consider a many-to-one matching where each committee is assigned many agents but each agent is assigned only one committee.[17] Second, we assume maximal impatience, so that all agents care about the static one-shot matching. Lastly, we do not consider existing tenants, i.e., incumbents who have existing committee assignments from previous Congress. Instead, we assume that there is a strict order of seniority, fixed across the senators and compare the Republican mechanism of Serial Dictatorship in order of seniority and the Democrat mechanism of Boston mechanism (assuming truthful preferences and tie-breaking based on seniority.

A.2. Fragility of Popular Matching existence.

First, we note that common mechanisms such as Serial Dictatorship in order of seniority and Boston mechanism (with assumption of truthful revelation) generally fail to guarantee a complete popular matching (see Figure 8).

Second, we consider the necessary but not sufficient for complete popular matching, that the f- and s-posts span the 16 committees (Figure 6). The figures illustrate that even with a slight correlation (i.e., $\rho \approx 0.2$ which is generated by everyone agreeing on their top choice and then having 15 remaining choices randomly generated), the necessary condition begins to be violated. As preferences become increasingly correlated, fewer and fewer of the 16 committees constitute f- and s-posts. Of course, when preferences are perfectly correlated with $\rho = 1$ and everyone has the same preference, only the first and second choice committees are f- and s-posts.

[16]Since a third of the Senators are up for re-election in each Congress, the largest proportion of freshman can of course be 1, but this is empirically extreme.

[17]Of course we assume that agents are indifferent across all the seats within a given committee.

Third, in these simulations, we find that popular matchings exist for some preference environments with intermediate correlation with 2 blocks, however, for fully random preferences, popular matching existence is rare and for higher levels of correlation with 3 or more blocks, popular matchings do not exist (see Figure 7).

A.3. **Existence of Majority Stability.**

The matchings produced by Serial Dictatorship and Boston mechanisms are majority stable for environments with low and intermediate correlations across preferences, however, for highly correlated preferences, the matchings are not majority stable (Figures 9 left and 10 left).

The relationship between correlation across agents' preferences and the maximal sized coalition trying to overturn the matching is positive on the whole but not point-by-point (see Figures 9 right and 10 right).

A.4. **Comparing existence of majority stability under Boston Mechanism vs Serial Dictatorship.**

We consider Boston mechanism with tie-breaks in order of seniority[18] and truthful preference revelation.[19]

Firstly, we see that for almost all preference environments (all but 5/12001 simulations), the maximal sized coalition that can be made better off is weakly larger under Serial Dictatorship than under the Boston mechanism (see Figure 11 right).

Second, Boston mechanism under truthful revelation is more robust for majority stability and popular matching existence (see Figure 11 left). Out of 12,001 simulated preference environments with varying levels of correlation, there were i) 5154 instances where both were majority stable, ii) 4673 instances where both were not majority stable, iii) 2169 instances where Boston mechanism is majority stable but Serial Dictatorship is not, and iv) only 5 instances where Serial Dictatorship was majority stable but Boston mechanism was not.

A.5. **Summary of Simulations Takeaways.**

Thus, the key takeaways from the simulation are that i) popular matchings fail to exist with correlated preferences, ii) Serial Dictatorship and Boston mechanism often fail to produce popular matchings, ii) Majority Stable matchings often exist even with a significant amount of correlation across agents' preferences, iii) as preferences become more and more correlated across agents, f- and s-posts don't span the set of seats and fewer agents can be assigned their f- and s-posts (i.e., smaller the size of the maximal popular matching), iv) as preferences become more and more correlated across agents, maximal coalition size wanting to overturn is larger, and v) Boston mechanism with truthful revelation is more robust to majority stability than Serial Dictatorship.

[18]This is for comparability across Serial Dictatorship in order of seniority and Boston mechanism.

[19]Boston mechanism is not strategyproof, so this assumption is just for illustration purposes.

Table 1. Final list of vacancies by committee for simulation. The first column is the type of committee which is commonly used classification in the political science literature on committees. Second column has the name of the committee, third column has the number of seats on the committee.

Committee Type	Committee Name	# Seats for Simulation
Prestige	Appropriations	6
Prestige	Budget	4
Prestige	Rules	2
Constituency	Agriculture	3
Constituency	Small Business	4
Constituency	Veterans Affairs	2
Constituency	Armed	3
Constituency	Energy	4
Policy	Banking	3
Policy	Commerce	3
Policy	Environment	3
Policy	Finance	4
Policy	Foreign Relations	3
Policy	Govt Affairs	3
Policy	Judiciary	3
Policy	Labor	3
	TOTAL	**53**

Table 2. Generation of Correlated Preferences for Simulations. Varying degrees of correlation of preferences were generated by considering 13 different preference block structures across the 16 committees. Fully random preferences generated preferences in the correlation range of [-.0139, .0404], then fixing block of 1 such that all agents agreed on the top choice and then randomly generating the remaining 15 committees in the next block generated preferences in the correlation range [.1625,.2109], etc. Thus, each row of this table shows the block structure used to generate preferences using randomization within blocks in column 1 and the minimum and maximum correlation ρ this produced in columns 2 and 3. The different preference block schemes are grouped and numbered such that they correspond to close clusters on the simulation figures on the horizontal axis.

	Preference Generation	**Min Correl**	**Max Correl**
1)	16 Fully Random	-.0139	.0404
2)	1;15	.1625	.2109
3)	2;14	.3192	.3547
4)	3;13	.4502	.4748
5)	4:12	.5572	.5809
	12:4	.5585	.5863
6)	8;8	.7488	.7627
7)	2;2;2;2;8	.8684	.8784
	8;2;2;2;2	.8683	.8764
8)	4;4;4;4	.9402	.9429
9)	2;2;2;2;2;2;2;2	.9880	.9887
	1;1;1;1;2;2;2;2;2;2	.9910	.9916
10)	1;1;1;1;1;1;1;1;1;1;1;1;1;1;1;1: Same Pref	1	1

Figure 6. Necessary but not sufficient condition for complete Popular Matching across different Correlations in preferences.
Left: for each simulated set of preferences with a given Correlation on the horizontal axis, the vertical axis plots 1 if f- and s-posts span all 16 committees, 0 otherwise. *Right*: for each simulated set of preferences with a given Correlation on the horizontal axis, the vertical axis plots the number of committees (out of 16 total) which are f- and s-posts.

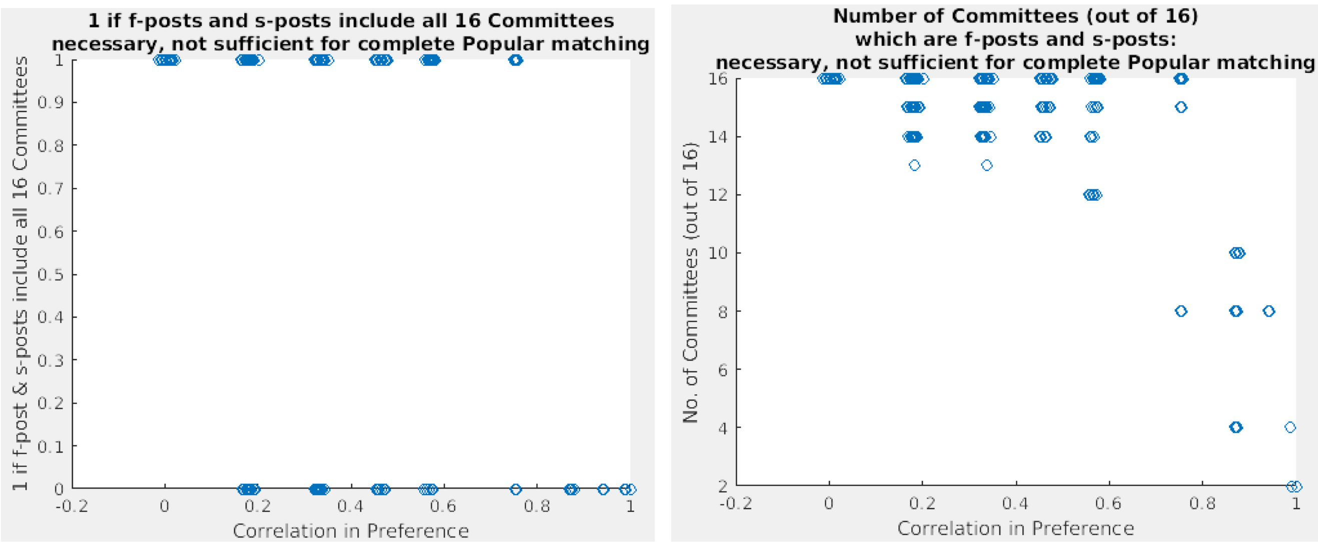


Figure 7. Existence of Popular Matching.
Left: Existence of Popular Matching across different Correlations in preferences. *Right*: Maximal number of candidates (out of 53) able to be assigned their f- and s-posts.

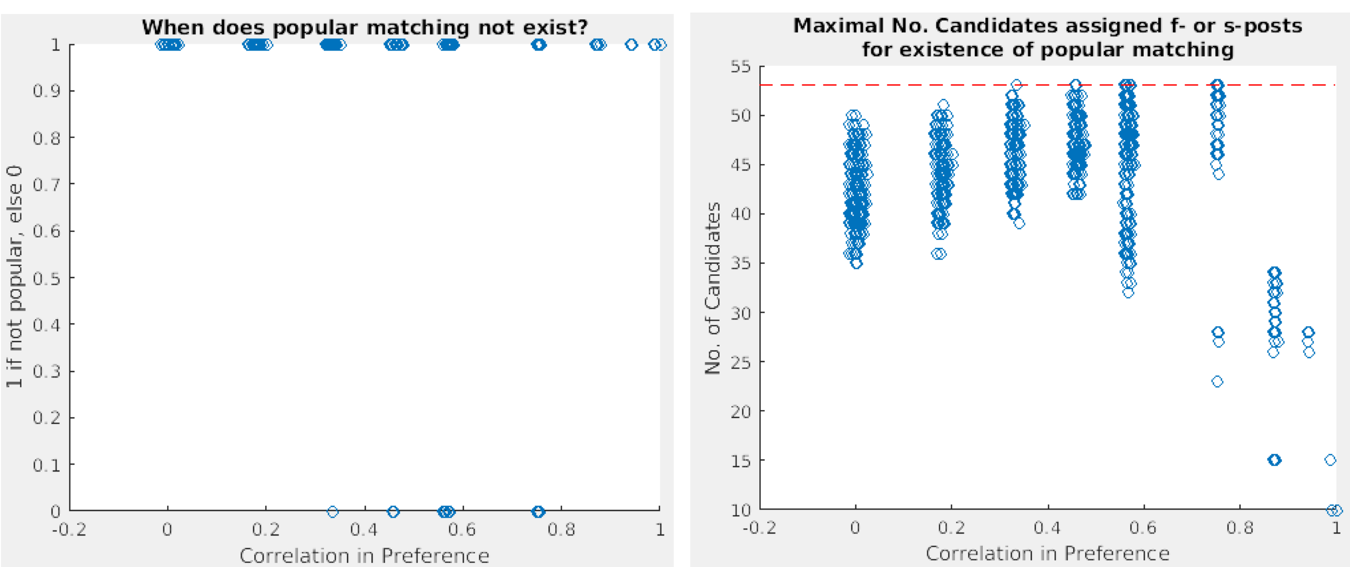


Figure 8. Existence of Popular Matching under Serial Dictatorship and Boston Mechanism.
Under Serial Dictatorship (*Left*) and Boston mechanism assuming truthful revelation and tie-breaking in order of seniority (*Right*), these figures plot whether the matching is popular or not. We see that the matchings produced by these mechanisms are rarely ever complete popular matchings.

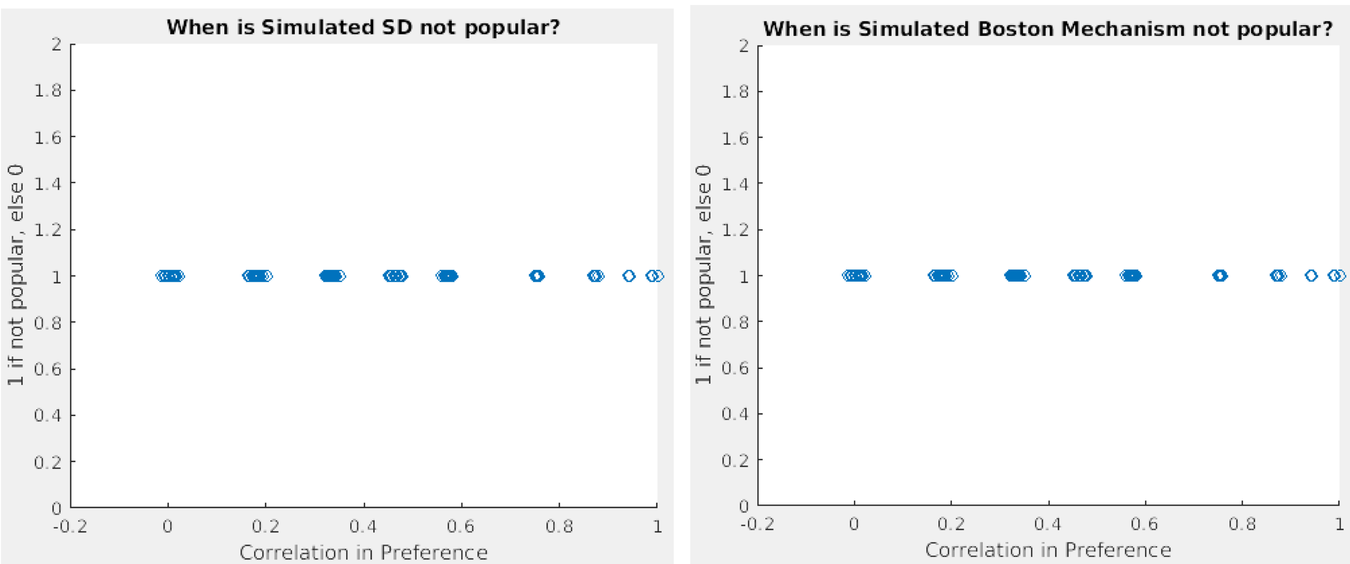


Figure 9. Majority Stability Under Serial Dictatorship.
Simulating the Serial Dictatorship in order of seniority across various preference environments with varying levels of correlation across agents' preferences, *Left* shows when a majority overturns the matching produced by the Serial Dictatorship, while *Right* shows the maximal sized coalition that can improve from each Serial Dictatorship matching.

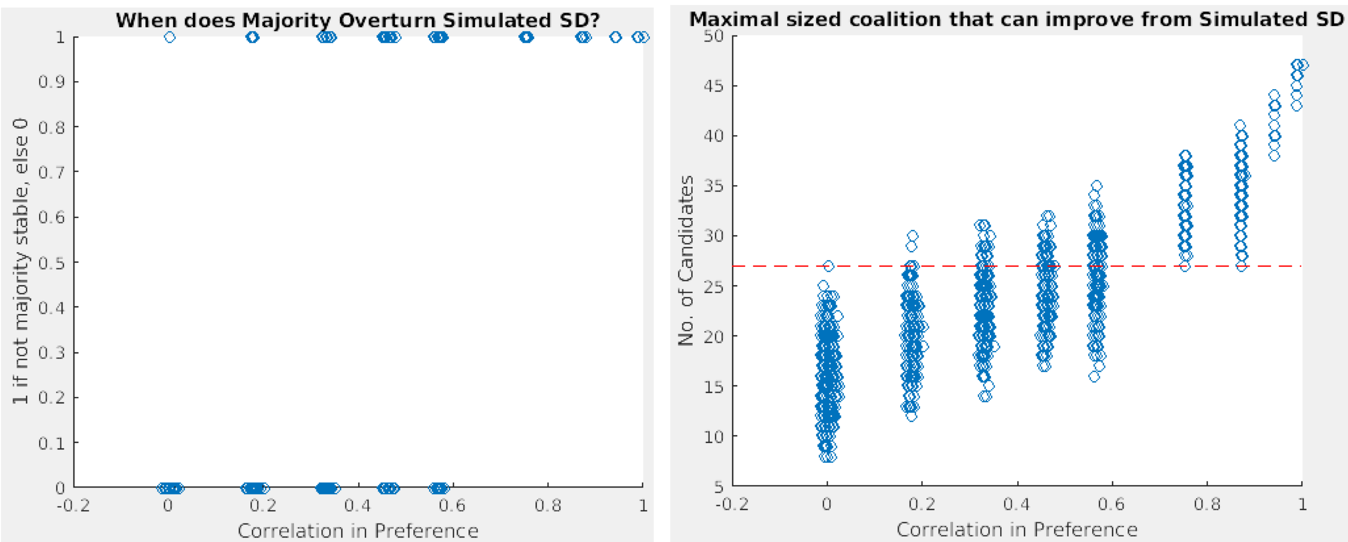


Figure 10. Majority Stability Under Boston Mechanism.
Simulating the Boston mechanism (assuming truthful preference revelation and tie-breaks in order of seniority) across various preference environments with varying levels of correlation across agents' preferences, *Left* shows when a majority overturns the matching produced by the Boston mechanism, while *Right* shows the maximal sized coalition that can improve from each Boston mechanism matching.

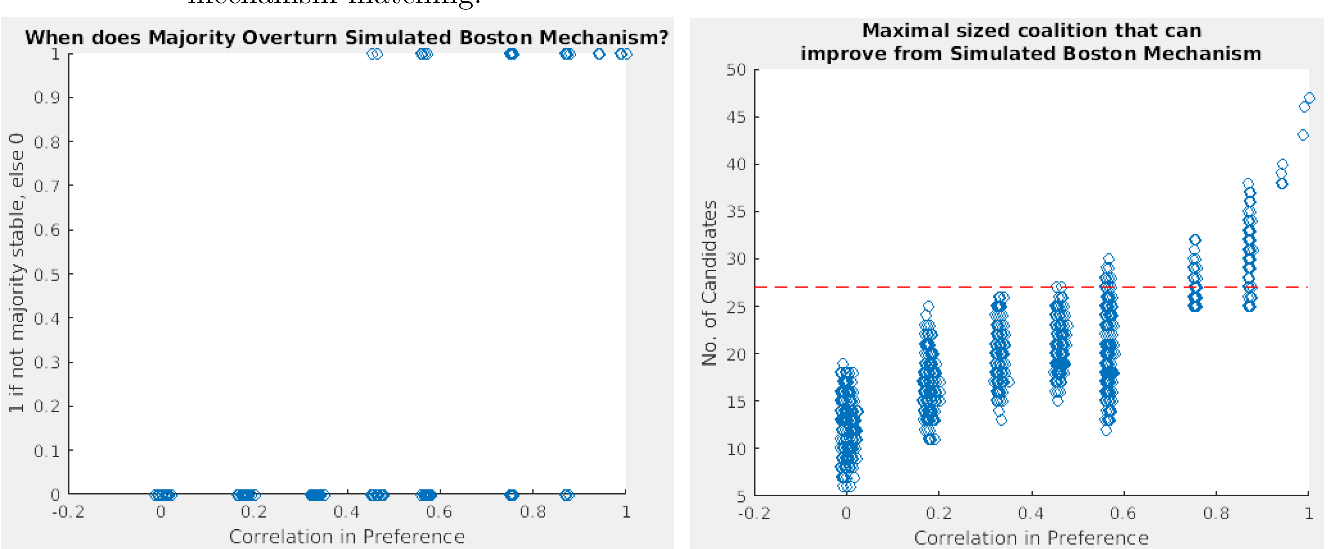
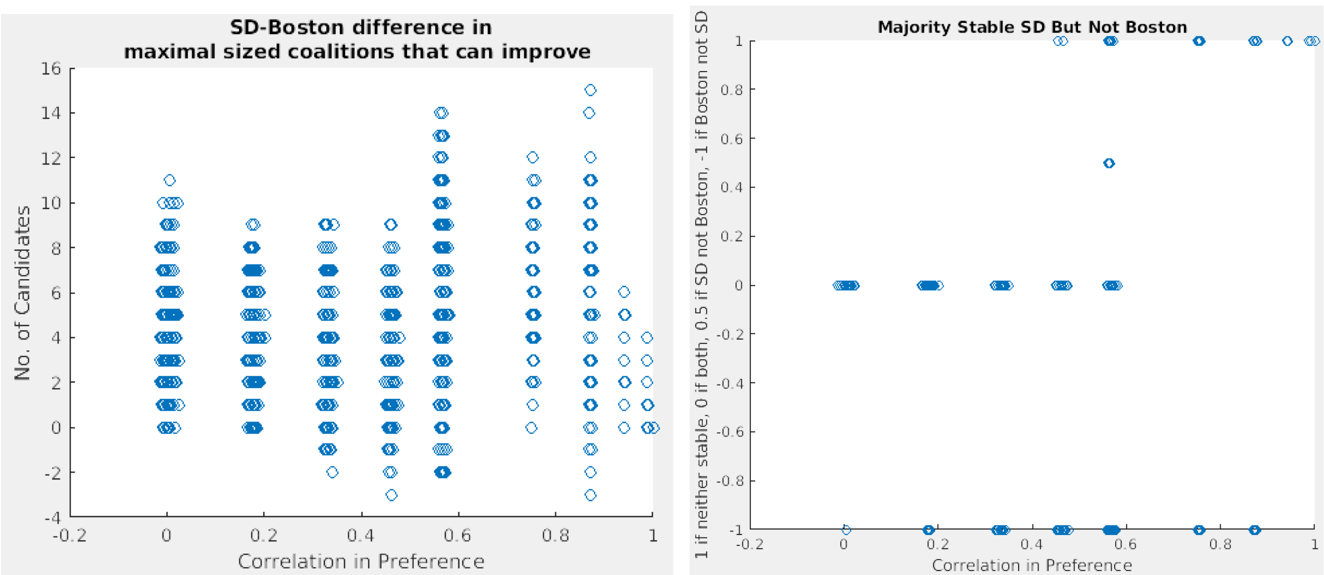


Figure 11. Comparing Majority Stability Under Serial Dictatorship v.s. Boston Mechanism.
Simulating Serial Dictatorship in order of seniority and Boston mechanism (assuming truthful preference revelation and tie-breaks in order of seniority) across various preference environments with varying levels of correlation across agents' preferences, these figures compare the performance of both mechanisms in regards to majority stability. *Left*: shows the difference between the sizes of the maximal sized coalitions which can be strictly improved under an alternative matching, across Serial Dictatorship and Boston mechanism. Right: shows that when matchings are majority stable for both mechanisms (0), under neither mechanism (1), under the Boston mechanism but not Serial Dictatorship (-1), and under Serial Dictatorship but not Boston mechanism (0.5).



Appendix B. **A Discussion on Properties of Voting Rules & Mechanisms Yielding Institutional Stability**.

In this appendix we first analyze the properties of some mechanisms that can yield majority stability and then analyze what is the structure on voting rules that maintain certain normatively nice properties of collective choice in social allocation choice problems.

First, we show that if a preference environment admits majority stable allocations there exists a strategyproof mechanism that delivers at some majority stable allocation. Here we build on Lemma 4:

Corollary 4.1. *If there exists a majority stable allocation, there exists a Serial Dictatorship order which achieves a majority stable allocation.*

> **Proof**: Serial Dictatorship spans all pareto efficient allocations. Hence there exists a pareto efficient allocation that is also majority stable.

Corollary 4.2. *If there exists a majority stable allocation, there always exists a strategyproof mechanism to attain it.*

> **Proof**: Serial Dictatorship in given order is strategyproof.[20]

Corollary 4.4. *If there exists a majority stable allocation, a seniority order of Serial Dictatorship which attains a majority stable allocation has all those getting their first preference seat being senior to those getting their second preference seat, ... (with any seniority order amongst agents being assigned seats of equal preference ranks.).*

> **Proof**: if not, there exists a pareto improving cycle and this contradicts the pareto efficient majority stable allocation.

Next, in the spirit of Arrow's Impossibility Theorem from social choice problems, we discipline the collective choice of the social allocation choice problems to be disciplined by certain properties. It is important to realize that certain results from social choice can often not be imported directly to social allocation choice because under Assumption 4, arbitrary individual preferences over matchings are not feasible.

Theorem 3. *Any preference aggregation rule that is acyclic and weakly Paretian is collegial.*[21]

> **Proof**[22]: Since f is weakly Paretian, there exists a decisive set and hence $\mathbb{L}(f) \neq \emptyset$. Now suppose by contradiction that f is non-collegial. Then for any $i \in \mathbb{I}$, there exists

[20] Serial Dictatorship is also group-strategyproof. It is important to distinguish between institutional stability which is a property of an allocation in a given preference environment and group-strategyproofness which is a property of the mechanism across all preference environments.

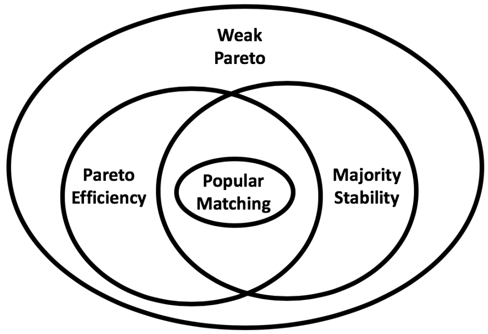
[21] We follow Austen-Smith and Banks (2000) definitions: (1) A preference aggregation rule f is weakly Paretian if, for every profile and for any matchings M_1 and M_2, and all $i \in \mathbb{I}$, $x \succ_i y$ implies $x \succ_f y$. (2) A binary relation $\succeq$ is acyclic on the set of all matchings M if for all $\{M_1, M_2, ..., M_m\} \in M$, $M_1 \succ M_2$, $M_2 \succ M_3$,..., $M_{m-1} \succ M_m$ implies $M_1 \succeq M_m$. (3) A set is decisive for M_1 against M_2 if for every preference profile, $M_1 \succ_i M_2$ for all $i \in \mathbb{I}$ implies $M_1 \succ M_2$. (4) For any preference aggregation rule f, let $\mathbb{L}(f)$ denote the set of decisive coalitions associate with f. (5) A preference aggregation rule f is collegial if and only if $K(\mathbb{L}(f)) = \cap_{L \in \mathbb{L}(f)} L$ is non-empty. The set $K(\mathbb{L}(f))$ is called the collegium.

[22] This proof closely follows Theorem 2.4 from Austen-Smith and Banks (2000), but does not follow immediately since social allocation choice problems with Assumption 4 restrict the domain of preferences.

some $L \in \mathbb{L}(f)$ such that $i \notin L$. Since $\mathbb{L}(f)$ is monotonic (Lemma 2.2 from Austen-Smith and Banks (2000)), for every $i \in \mathbb{I}$, $\mathbb{I} \backslash \{i\} \in \mathbb{L}(f)$. Now consider the preference profile where all N agents have the same rank order preference over seats: $A > B > ....$ Take the set of matchings where i gets his top choice. The size of this set is $(N-1)!$ since, in this set of matchings, the remaining $N-1$ agents can be allotted the $N-1$ remaining seats. For each such permutation, consider all the matchings which are translations over agents of these permutations (e.g., for $N = 3$, a given permutation 1-A,2-B,3-C leads to a set of translation matchings: $M_1 = (1 - A, 2 - B, 3 - C)$, $M_2 = (2 - A, 3 - B, 1 - C)$, and $M_3 = (3 - A, 1 - B, 2 - C)$). Each i has a strict preference over each of these translation matchings. Label these matchings as M_i where i gets everyone's top choice seat, A. Now, order each i's individual preference: i prefers matching where he gets A, over that where he gets B, etc. We now see that for all $j = 2, ..., N$, all agents except j prefer M_{j-1} to M_j, but since $\mathbb{I}\{j\} \in \mathbb{L}(f)$, $M_{j-1} \succ M_j$. Also, all agents except 1 prefer $M_N to M_1$, hence $M_N \succ M_1$. Therefore we have a cycle on $\{M_1, ..., M_N\}$: $M_N \succ M_1 \succ M_2 \succ ... \succ M_N$, which contradicts acyclicity.

Appendix C. Set inclusions of matchings across Institutional Stability notions.

This appendix provides some examples and proofs of the set inclusion from Figure 1 replicated below:



- Claim: All popular matchings are pareto efficient.

  Proof: if allocation M is not pareto efficient, then there exists alternative allocation M' s.t. $|M' \succ_i M| > 0$ and $|M \succ_i M'| = 0$. Thus, $|M' \succ_i M| > |M \succ_i M'| = 0$, thus M is not a popular matching.

- Claim: All popular matchings are majority stable.

  Proof: if allocation M is not majority stable, then there exists alternative allocation M' s.t. $|M' \succ_i M| \geq \frac{N+1}{2}$. Thus, $|M' \succ_i M| < \frac{N+1}{2}$. Hence $|M' \succ_i M| > |M \succ_i M'| = 0$, thus M is not a popular matching.

- A pareto efficient matching that is not popular and no majority stable:

  Example: agents $\{1, 2, 3\}$ ranking seats $\{A, B, C\}$ Consider 1-A, 2-B, 3-C. Since

1	2	3
A	A	B
B	B	A
C	C	C

  Serial Dictatorship with seniority $1 > 2 > 3$ implements it, it is pareto efficient. However, 1-C, 2-A, 3-B is strictly preferred by 3 and 2, hence it is not popular or majority stable.

- A pareto efficient matching that is majority stable, but not popular:

  Example: agents $\{1, 2, 3, 4, 5\}$ ranking seats $\{A, B, C, D, E\}$ given seniority $1 \succ 2 \succ 3 \succ 4 \succ 5$ Consider 1-A, 2-B, 3-C, 4-D, 5-E. Since Serial Dictatorship with

1	2	3	4	5
A	A	A	A	A
B	B	B	B	B
C	C	C	D	E
D	D	D	C	D
E	E	E	E	C

seniority $1 > 2 > 3 > 4 > 5$ implements it, it is pareto efficient. Since envy is only of seats A and B, it is majority stable due to the packing problem. Since there are 3 blocks, there is no complete popular matching.

- A majority stable matching that is not pareto efficient nor popular

  Example: agents $\{1, 2, 3, 4, 5\}$ ranking seats $\{A, B, C, D, E\}$ given seniority $1 \succ 2 \succ 3 \succ 4 \succ 5$ Consider 1-B, 2-A, 3-C, 4-D, 5-E. Only envy is 1 and 2 envious

1	2	3	4	5
A	B	C	D	E
B	A	A	A	A
C	C	B	B	B
D	D	D	C	C
E	E	E	E	D

  of each others' assignments so majority stable, but fails pareto efficiency and popular matching due to 1 and 2 strictly preferring 1-A, 2-B, 3-C, 4-D, 5-E.

- A matching that is majority stable, popular and pareto efficient

  Example: agents $\{1, 2, 3, 4, 5\}$ ranking seats $\{A, B, C, D, E\}$ given seniority $1 \succ 2 \succ 3 \succ 4 \succ 5$ Consider 1-A, 2-B, 3-C, 4-D, 5-E.

1	2	3	4	5
A	B	C	D	E
B	A	A	A	A
C	C	B	B	B
D	D	D	C	C
E	E	E	E	D