Matching with Multilateral Contracts*

Marzena Rostek[†] and Nathan Yoder[‡]

April 16, 2022

Abstract

In many environments, agents form agreements which are multilateral and/or have externalities. We show that stable outcomes exist in these environments when the *irrelevance of rejected contracts* condition survives aggregation, either across all agents or within two *implicit* sides of the market for whom contracts are substitutes. In settings where agents are strategically sophisticated, in the sense that they make correct conjectures about how other agents will choose from each set of contracts, we show this is ensured by a mild criterion on those conjectures. When each agent is strategically sophisticated about the behavior of all other agents, stable outcomes always exist: No conditions on preferences or market structure are necessary. Our characterization of these outcomes allows the application of matching theory to new settings, such as legislative bargaining or free trade agreement formation.

Keywords: Multilateral matching, externalities, matching with contracts, stability

* The authors are grateful to numerous colleagues for their helpful comments. This material is based upon work supported by the National Science Foundation under Grant No. SES-1357758.

† University of Wisconsin-Madison, Department of Economics; E-mail: `mrostek@ssc.wisc.edu`.

‡ University of Georgia, Terry College of Business, Department of Economics; E-mail: `nathan.yoder@uga.edu`.

1 Introduction

Agreements among more than two agents are common. Legislators bargaining over which bills they will support, firms starting joint ventures, American professional sports teams seeking multi-team trades, or nations bargaining over international treaties each seek to form such *multilateral* agreements. Moreover, *externalities* are pervasive in both multilateral and bilateral agreements. Agreements to form a cartel, include a doctor in an insurance network, form a military alliance, or lower tariffs often affect the behavior of other producers, insurers, or countries. Using matching theory to provide predictive results has been seen as challenging in settings where agreements have either of these features. In this paper, we consider these settings in a model based on the matching with contracts framework introduced by Hatfield and Milgrom (2005), and show how to overcome these challenges to find outcomes which are *stable*, or robust to both individual deviations to remove agreements and joint deviations to form new ones. We do so by making two key observations.

First, generating predictive results is challenging in settings with multilateral contracts and those with externalities for the same reason: both features lead to interdependence in choice behavior that causes choice to be poorly behaved under aggregation. This interdependence does not appear in most of the existing literature, because of the restrictions it places on the market's structure: the number of agents who can sign a contract (bilateral contracting), the agents they can sign them with (two-sidedness or acyclicity), and the set of agents those contracts can affect (no externalities). We do not restrict the structure of the market in these ways.

To see why interdependence can cause problems for aggregation, recall that to guarantee the existence of stable outcomes in two-sided matching markets, individuals' choice functions must satisfy *irrelevance of rejected contracts (IRC)* (Aygün and Sönmez, 2013): removing contracts which an agent rejects does not change her choices. In such models, the aggregate choice of each side of the market is separable in the choices of individuals, and so irrelevance of rejected contracts is equivalent at the individual and aggregate levels. But when agents are influenced by agreements made by others on the same side of the market (as with externalities) or contract with them directly (e.g., with multilateral agreements), this separability need not hold. As a consequence, the irrelevance of rejected contracts condition may fail to survive aggregation among agents on the same side.

We show that this failure of IRC to aggregate is the key challenge for matching-theoretic tools created by agreements that are multilateral or have externalities. In particular, we provide two existence results showing that when IRC survives aggregation, stable outcomes exist in these environments.

Theorem 1 shows that the Gale-Shapley (1962) algorithm used to identify stable outcomes in canonical two-sided matching markets (e.g., Hatfield and Milgrom (2005)) can be generalized to characterize outcomes in settings with two *implicit* sides whose aggregate choice functions satisfy IRC, even when contracts may involve or affect multiple agents on the same side. Like other results that rely on the Gale-Shapley algorithm, Theorem 1 requires substitutability among contracts; we extend this condition to accommodate externalities.[1]

When the aggregate choice of the entire market satisfies irrelevance of rejected contracts, however, Theorem 2 shows that a stable outcome not only exists, but is unique. In fact, this is true even when IRC only holds locally at the set of all contracts. In contrast to Theorem 1, this result does not rely on the techniques of Gale and Shapley (1962), and does not require substitutability.[2] In particular, the same contracts can be complementary for some agents and substitutable for others.

Our second key observation concerns the reason why IRC may not survive aggregation with interdependence in choice behavior. In particular, the failure of IRC to aggregate among a group of agents arises from group members' incorrect assumptions about which contracts the others will choose to sign.

Consider, for instance, a group consisting of two agents, each of whose agreement is required for contracts x and y. One agent's most preferred outcome is x alone, and the other's is y alone; both agents prefer either contract to autarky. Suppose, as is standard in the matching literature, that the agents' choices from a set of available contracts are simply their most preferred subset of those contracts. When making choices this way, both agents implicitly take as given that if they choose a contract, it will go into effect — or equivalently, that the other agent (who must also agree to the contract) will choose it as well. Then both x and y will be rejected by one of them, and hence by the group as a whole, from $\{x, y\}$. But when the rejected contract y is unavailable, both agents, and thus the group, will choose x. This causes the group's aggregate choice to violate IRC, precisely because each agent incorrectly assumed that each available contract would be chosen from $\{x, y\}$ by the other agent. If both of them had instead assumed that only x, or only y, would be chosen by the other agent, they would each choose that contract, and no IRC violation would occur.

But suppose that the agents in a group are strategically sophisticated about one another's behavior, in the sense that their choice functions are derived from utility maximization given *conjectures* about the contracts that other agents in the group will choose to sign, and those conjectures are correct. With such strategic sophistication, Theorem 4 shows that the irrelevance of rejected contracts condition on the group's aggregate choice is equivalent to a mild

[1]Note that we extend substitutability to settings with externalities in a different way than Pycia and Yenmez (2021) — see Section 3.1.

[2]Interestingly, its outcome can be thought of as resulting from a *one-sided* Gale-Shapley algorithm. This algorithm converges within two rounds even when it is not monotone — and hence *even when contracts are not substitutes*.

robustness criterion on the agents' conjectures. Specifically, this criterion captures the idea that no agent should believe that any other agent in the group would change their choices when contracts that *every* agent in the group rejected become unavailable; i.e., contracts rejected by each agent in the group should be irrelevant to an agent's beliefs.[3] Proposition 1 shows that this is equivalent to the irrelevance of rejected contracts condition *as applied to agents' conjectures rather than their choice functions.* In other words, IRC on the aggregate choice of a group of strategically sophisticated agents is equivalent to a refinement on those agents' *beliefs*; this refinement amounts to the epistemic irrelevance of contracts that *all* agents reject; and mathematically, it is equivalent to applying IRC to each agent's beliefs rather than their choice functions.

The endogenous profiles of a group's choice functions and beliefs that we consider, which we call *strategically consistent assessments*, are not standard in the matching literature. However, if the group is one side of a two-sided market without externalities, conjectures about the behavior of other agents on the same side are not relevant: the choice functions in such assessments are always derived from independent optimization problems, as in the literature, and so the assessment is unique. But if the group contains agents from more than one side, or if contracts are multilateral or have externalities, agents' choices will depend on their beliefs. Hence, rather than independent optimization by each agent, choice functions and beliefs are determined by individual optimization that is jointly consistent among all agents in the group.

When agents' choice functions are part of a strategically consistent assessment for the entire market — i.e., when each agent is strategically sophisticated about the behavior of every other agent — we can apply Theorems 2 and 4 to characterize stable outcomes. In fact, Theorem 5 shows that this is possible in *any* setting where agents make agreements: No restrictions on preferences (e.g., complementarity or (full) substitutability), market structure (e.g., acyclic trading networks), or who can be involved (e.g., bilateral contracts) or affected (e.g., no externalities) by the contract are necessary.

Theorem 5's characterization is in terms of the choice functions that agents would have if they always took all available contracts as given (as is standard in the literature) — what we call *myopically consistent* choice functions. Specifically, Theorem 5 shows that an outcome is stable for some strategically consistent assessment with regular beliefs if and only if it would be individually rational with myopically consistent choice functions.

On their own, the predictions of Theorem 5 are fairly weak. However, Theorem 6 shows that they can be significantly strengthened — in particular, all but the *maximal* outcomes predicted by Theorem 5 can be eliminated — with a refinement on agents' beliefs based on

[3] To see what this means in practice, consider the example above: If neither agent chose any contracts from $\{x, y\}$, neither agent should believe that the other would choose x when it is the only contract available. Hence, both agents' choices from $\{x\}$ would be empty, eliminating the IRC violation that we show arises in the agents' aggregate choice when they are not strategically sophisticated.

forward induction. This refinement, which we call *weak forward induction (WFI)*, requires that when a proposal to add contracts[4] is credible, in the sense that no agent can benefit from rejecting any of the proposed contracts, either directly (by getting a higher payoff) or indirectly (by leading to new opportunities for further deviations), each agent should believe that the others will go along with it.

Formally, Theorem 6 shows that outcomes are stable for some assessment where beliefs are both reasonable (in the sense that they satisfy IRC) and robust to forward induction (in the sense that they satisfy WFI) if and only if they are maximal among outcomes that would be individually rational if the agents were given myopically consistent choice functions. In settings where the matching literature offers predictive results — i.e., where stable outcomes exist when choice is myopically consistent — this never overturns those predictions: Corollary 2 shows that if an outcome is stable when choice is myopically consistent, it is stable for some strategically consistent assessment with beliefs satisfying IRC and WFI. And with complementarity, strategic sophistication among all agents leaves a matching model's predictions completely unchanged: The unique maximal outcome that is individually rational when agents have myopically consistent choice is the only one that can be stable, whether all agents are strategically sophisticated (as in Theorem 6) or none are (as in Rostek and Yoder (2020)).

While Theorem 6 identifies the set of outcomes that are stable for *some* strategically consistent assessment for all agents with beliefs satisfying WFI and IRC, under any *particular* assessment, Theorems 2 and 4 show that the stable outcome is unique. When these outcomes differ — that is, when Theorem 6 identifies multiple outcomes — we can think of each assessment as describing a different way that bargaining power might endogenously accrue to different agents.[5]

In Theorem 7, we consider a canonical setting — two-sided markets — where contracts have externalities, and agents are strategically sophisticated about the behavior of other agents on the same side (rather than that of each other agent). There, we provide conditions on agents' preferences ensuring the existence of strategically consistent assessments for both sides that give rise to aggregate choice functions satisfying the prerequisites of Theorem 1. In particular, Theorem 7 shows that when agents' myopically consistent choice functions satisfy the *standard substitutability* and *monotone externalities* conditions introduced by Pycia and

[4] That is, a block which does not result in the removal of existing contracts.

[5] For instance, in the two-agent, two-contract example given earlier, there are two strategically consistent assessments with beliefs satisfying WFI and IRC: one where both agents choose x when both contracts are available (and thus $\{x\}$ is stable), and one where both agents choose y (and thus $\{y\}$ is stable). We can think of the first assessment as assigning greater bargaining power to the agent that prefers x, and the second as assigning greater bargaining power to the agent that prefers y. Similarly, we show in Example 3 that in the canonical three-agent *roommate problem*, three different outcomes are stable for some assessment with beliefs satisfying WFI and IRC, each of which can be thought of as a different distribution of bargaining power among the agents.

Yenmez (2021), each side of the market has a strategically consistent assessment with choice functions satisfying substitutability and beliefs satisfying IRC. Then as a consequence of Theorems 1 and 4, these conditions — which Pycia and Yenmez (2021) show guarantee the existence of stable outcomes in two-sided markets with externalities when choice is myopically consistent — also ensure that stable outcomes exist with strategic sophistication within each side of the market.

Related Literature

Hatfield and Kominers (2015) initiated the study of multilateral agreements in a matching with contracts setting. They examine settings with continuously divisible contracts and transferable utility, and use the concavity of agents' valuations to show that competitive equilibria (and thus, as they show, stable outcomes) exist. We work with a finite set of contracts instead, and take an approach based on agents' strategic sophistication and the survival of IRC under aggregation that it implies.

On the other hand, Bando and Hirai (2021) consider a setting with finitely many multilateral contracts, as we do in our model. Unlike this paper or Hatfield and Kominers (2015), they work with conditions on the market's structure that they show ensure a stable outcome exists, regardless of agents' preferences.

Pycia (2012) considers the formation of disjoint coalitions — a type of multilateral agreement. He shows that when the agents' preferences are *pairwise aligned*, a stable coalition structure exists. Our results add to his by accommodating coalition formation settings where agents are strategically sophisticated, even without alignment among their preferences. In addition, our framework allows us to consider settings where agents can be part of multiple coalitions.

Our work also contributes to the literature on matching with externalities. One strand of this literature explores externalities in settings of applied interest, such as labor market matching with couples (e.g., Kojima et al. (2013)). Some authors, such as Bando (2012) and Fisher and Hafalir (2016), consider general environments in matching markets with two explicit sides.

Of these, the paper whose setting is closest to ours is Pycia and Yenmez (2021), who also use a matching with contracts framework. Though our work extends beyond the canonical two-sided setting they consider, it also contributes to the existing literature there by showing that stable outcomes always exist (and characterizing them) when each agent is strategically sophisticated about each other agent's behavior. Moreover, we extend the results of Pycia and Yenmez (2021) by showing that the conditions they introduce ensure the existence of stable outcomes when agents are strategically sophisticated about the behavior of agents on the same side.

Within the literature on two-sided matching markets with externalities, papers like Sasaki and Toda (1996) and Hafalir (2008) take an approach that is related to our notion of assessments that are strategically consistent. In the models they consider, an agent determines what to take as given about other agents' matchings through the use of an *estimation function* — an important precursor to the concept of a *conjecture* used in this paper. Unlike a conjecture about other agents' choices given a proposed set of contracts, estimation functions give a *set* of outcomes that an agent thinks are plausible, given the identity of the individual she is matched to. The agent then evaluates potential partners by taking as given the *least preferred* outcome that is plausible according to her estimation function.

While Sasaki and Toda (1996) take these estimation functions as a primitive of the model, Hafalir (2008) allows them to be endogenously determined according to a consistency condition. This condition pins down estimations differently from the way strategic consistency pins down beliefs: in his model, an agent's estimation function treats matchings as plausible if they are stable when the agent and her partner are removed from the market, whereas with strategic consistency, an agent's conjecture correctly matches the choices made by other agents from the available set of contracts.

2 Model

2.1 Setting[6]

There is a finite set I of agents and a finite set X of contracts they can sign with one another. Each contract $x \in X$ requires the agreement of a set of agents $N(x) \subseteq I$ to enter into force. For sets of contracts $X' \subseteq X$, we write $N(X') \equiv \bigcup_{x \in X'} N(x)$. We assume that each contract names at least two agents: For all x, $|N(x)| \geq 2$. We say contract x is *multilateral* if $|N(x)| > 2$ and *bilateral* if $|N(x)| = 2$. For each agent $i \in I$, denote the set of contracts requiring i's agreement as $X_i \equiv \{x | i \in N(x)\}$. In keeping with the literature, we say that X_i is the set of contracts that *name* i. Similarly, let $X_J \equiv \bigcup_{i \in J} X_i$, let $X_{-i} \equiv X \setminus X_i$, and for sets of contracts $Y \subseteq X$, write $Y_i \equiv Y \cap X_i$ and $Y_{-i} \equiv Y \cap X_{-i}$.

Each agent i has a *choice function* $C_i : 2^{X_i} \times 2^{X_{-i}} \to 2^{X_i}$. $C_i(Y_i|Y_{-i})$ gives the set of contracts that agent i chooses from the set of available contracts Y_i, given the presence of contracts in Y_{-i}. Likewise, define each agent's *rejection function*, $R_i(X'|Y) \equiv X' \setminus C_i(X'|Y)$. Though we take choice functions as primitive, they may arise from the maximization of payoff functions $u_i : 2^X \to \mathbb{R}$. We give a detailed discussion of this in Section 4.

[6]We consider a similar setting to Rostek and Yoder (2020); hence, we import much of the description from our other paper.

Several assumptions often present in the literature are absent in our setting. Specifically, we do not assume a certain market structure (such as two-sidedness or acyclicity) or that only two agents can be part of any agreement. Our first example explains what such restrictions would look like in the setting we study.

Example 1 (Nested Environments). Our environment allows us to consider other market structures that have been studied in the literature. Here, we map several of these to the primitives of our model.

Network formation. $X = \{(i,j)|i,j \in I, i \neq j\}$; $X_i = \{(i,j)|j \in I, i \neq j\}$.

One-to-one matching. $X = \{(i,j)|i \in I_1, j \in I_2\}$; $X_i = \{(i,j)|j \in I_2\}$, $i \in I_1$; $X_i = \{(i,j)|j \in I_1\}$, $i \in I_2$; $|C_i(Y)| \leq 1$ for all $i \in I$ and $Y \subseteq X_i$.

Many-to-one matching with contracts. For all $x \in X$, $N(x) = \{i,j\}$ for some $i \in I_1$ and $j \in I_2$; for $i \in I_1$, $|C_i(Y)| \leq 1$ for each $Y \subseteq X_i$.

Many-to-many matching with contracts. For all $x \in X$, $N(x) = \{i,j\}$ for some $i \in I_1$ and $j \in I_2$.

Matching with contracts on networks. For all $x \in X$, $N(x) = \{b(x), s(x)\}$ for some $b(x), s(x) \in I$.

(Disjoint) Coalition formation. $X \subseteq \{J \subseteq I|\ |J| > 1\}$; $N(J) = J$ for all $J \subseteq I$; $|C_i(Y)| \leq 1$ for all $i \in I$ and $Y \subseteq X_i$.

Overlapping coalition (hypergraph) formation. $X \subseteq \{J \subseteq I|\ |J| > 1\}$; $N(J) = J$ for all $J \subseteq I$.

The interactions between pairs of agents specified by a multilateral contract cannot be represented as independent bilateral contracts in a matching on networks model. We illustrate this fact using an example from professional sports.

Example 2 (Multilateral vs. Bilateral Agreements). On July 31, 2014, Major League Baseball's Detroit Tigers, Tampa Bay Rays, and Seattle Mariners traded a total of five players as part of a single multilateral agreement. Detroit sent two players to Tampa Bay; Tampa Bay sent one player to Detroit; Detroit sent one player to Seattle; and Seattle sent one player to Tampa Bay.[7]

If we tried to model this transaction as three independent bilateral contracts, instead of a single multilateral contract, we would generally fail to predict that it would take place. The reason is that a contract representing the bilateral interaction between the Seattle Mariners and Detroit Tigers would not be individually rational for the Tigers: they would send Seattle a valuable player without receiving anything in return. Likewise, a contract representing the bilateral interaction between the Mariners and Rays would not be individually rational for the

[7]Source: http://mlb.mlb.com/mlb/transactions/index.jsp#month=7&year=2014.

Mariners. These interactions are only possible because they were conducted as part of a single multilateral agreement. To accurately model environments with multilateral agreements, matching models must explicitly allow for the interdependence those agreements create.

2.2 Stability

Our solution concept is the usual matching-theoretic definition of *stability*, generalized to our setting with multilateral contracts and externalities. [8,9]

Definition (Stability). A set of contracts $Y \subseteq X$ is *stable* if it is

1. *Individually rational*: $Y_i = C_i(Y_i|Y_{-i})$ for all $i \in N$.

2. *Unblocked*: There does not exist $Z \subseteq (X \setminus Y)$ such that for all $i \in N(Z)$, $Z_i \subseteq C_i((Z \cup Y)_i|(Z \cup Y)_{-i})$.

We next show that we can give an equivalent characterization of stability in terms of the *aggregate choice function* $C : 2^X \to 2^X$ defined by $C(Y) \equiv \bigcap_{i \in I}(C_i(Y_i|Y_{-i}) \cup Y_{-i})$. In contrast to aggregate choice functions of either side of a two-sided market in, e.g., Hatfield and Milgrom (2005), C is the aggregate choice *of the entire market as a whole.* Because of this — and because we consider matching settings which need not have two sides, and which may include contracts that are multilateral or have externalities — this aggregate choice function cannot be defined as the direct sum of individual choice functions. Instead, we define it as an intersection, so that it discards information about contracts which are chosen by some agents, but rejected by others. Lemma 1 shows that this is precisely the information about individual choice that is not relevant for stability.

Lemma 1 (Stability in Aggregate). *Y is stable if and only if* 1. $C(Y) = Y$ *and* 2. $Y' \nsubseteq C(Y' \cup Y)$ *for all* $Y' \nsubseteq Y$.

Part 1 of Lemma 1 restates individual rationality in terms of the aggregate choice function. Part 2 is the additional restriction that an individually rational outcome must satisfy if it is unblocked.

[8]In particular, our solution concept coincides with those of Gale and Shapley (1962) (one-to-one matching), Hatfield and Milgrom (2005) (many-to-one matching with contracts), and Hatfield and Kominers (2012) (matching on networks) in the settings they consider.

[9]This solution concept differs slightly from the one introduced by Pycia and Yenmez (2021) in two-sided matching markets with externalities. In their concept, agents in a blocking coalition do not anticipate any changes to the set of contracts signed by other agents, even the other members of the blocking coalition. Our stability concept instead assumes that agents in a blocking coalition are aware of the contracts added by the agents they negotiate with, but (in keeping with the literature) may disagree about which contracts are to be deleted. In the presence of substitutability, this means that our concept allows a smaller set of blocks than theirs does.

Lemma 1 can be interpreted as a formulation of stability in terms of market clearing. Seen this way, it says that stable outcomes are precisely those where there is no excess demand (there is no $Y' \nsubseteq Y$ such that $C(Y \cup Y') \supseteq Y'$) and no excess supply (there is no $Y' \subseteq Y$ such that $C(Y) = Y'$) *in aggregate*.[10]

This result relies on similar arguments to those used in Rostek and Yoder (2020) to show that stability can be characterized using aggregate choice. Unlike the other paper's result, however, Lemma 1 does not require complementarity; in fact, it does not require any restrictions on choice behavior. This makes its implications slightly different: Without complementarity, we cannot restrict attention in part 2 of Lemma 1 to $Y' \supset Y$.

Lemma 1 also allows us to characterize stable outcomes using the Blair (1988) revealed preference order $\succeq_C$ associated with the aggregate choice function C. Recall that for any choice function (without externalities) $\tilde{C}$, we write $Y \succeq_{\tilde{C}} Z$ whenever $\tilde{C}(Y \cup Z) = Y$. Then we have the following corollary.

Corollary 1 (Stability and Aggregate Revealed Preference). *Y is stable if and only if there is no $Y' \subseteq X$ such that $Y' \succ_C Y$.*

3 Stable Outcomes: Existence and Characterization

As we discuss in the introduction, models in the matching literature typically place restrictions on the market structure. Many papers consider environments where the set of agents is separable into two sides, such that agents can only sign contracts with others on the opposite side. Others capture markets that can be described as acyclic trading networks. Moreover, the literature generally assumes that contracts are bilateral and do not have externalities. As we show in Rostek and Yoder (2020), each of these assumptions is unnecessary for stable outcomes to exist when contracts are complements.

These assumptions can also be dispensed with in more general settings. Instead, we show that stable outcomes exist whenever the *irrelevance of rejected contracts (IRC)* condition (Aygün and Sönmez, 2013) survives aggregation, either among agents on each of two *implicit* sides (Theorem 1), or among all agents (Theorem 2).

It is well known that in two-sided markets, *individual* choice must satisfy IRC for stable outcomes to exist generically (Aygün and Sönmez, 2013). That is, making rejected contracts unavailable must not change the set of contracts that each agent chooses. Formally, an individual agent's choice function C_i satisfies IRC if for any $Y, Z \subseteq X_i$ and $X' \subseteq X_{-i}$,

[10] Authors such as Hatfield and Kominers (2012) and Hatfield et al. (2013) have interpreted stability in terms of a lack of excess demand and supply at an *individual* level. But Lemma 1 and its counterpart in Rostek and Yoder (2020) show that considering *aggregate* choice is without loss.

$C_i(Z|X') \subseteq Y \subseteq Z$ implies $C_i(Y|X') = C_i(Z|X')$.[11] Holding the set of contracts available to other agents constant, IRC is equivalent to the weak axiom of revealed preference: see Alva (2018) Theorem 1.

In canonical two-sided settings without externalities (e.g., Hatfield and Milgrom (2005)), individual agents' choice functions satisfy IRC if and only if the aggregate choice of each side of the market does: Since agents on the same side of the market are not named or affected by each other's contracts, their aggregate choice is just the direct sum of their individual choice functions (Hatfield and Milgrom, 2005). With multilateral contracts or externalities, however, there is no longer an equivalence between IRC at the individual and aggregate levels, since the aggregate choice of a group cannot be defined as a direct sum.

Instead, we show that the relevant notion of a group's aggregate choice is an *intersection* of the contracts that no agents in the group reject, just as we did for the aggregate choice of the entire market C. With a group's aggregate choice defined in this manner, its aggregate rejection function is the union of the individual rejection functions of its members (Lemma 8). This property is crucial to ensure that outcomes which are immune to deviations based on the aggregate choice of a group — such as each of two implicit sides (Lemma 3) or the entire market (Lemma 1) — are also immune to deviations based on individual choice.

Formally, for each nonempty $J \subseteq I$, define the *aggregate choice function for group* J, $C_J : 2^X \to 2^X$, such that $C_J(Y) \equiv \bigcap_{i \in J}(C_i(Y_i|Y_{-i}) \cup Y_{-i})$; by definition, $C_I = C$. Accordingly, define the *rejection function for group* J, $R_J : 2^X \to 2^X$, such that $R_J(Y) \equiv Y \setminus C_J(Y)$. Applied to the aggregate choice of a group J, irrelevance of rejected contracts requires that for all $Y, Z \subseteq X$, whenever $C_J(Y) \subseteq Z$ and $C_J(Z) \subseteq Y$, then $C_J(Y) = C_J(Z)$.

3.1 Orderly Aggregation on Two Implicit Sides

In settings characterized by two implicit sides whose aggregate choice functions satisfy IRC, Theorem 1 shows that the Gale-Shapley algorithm functions (i.e., finds every stable outcome) even in settings with multilateral contracts and externalities. These implicit sides may or may not also be explicit, in the sense that agents on one side do not contract with agents on the other. They need not be disjoint: they must form a covering of the set of agents, but not necessarily a partition. However, use of the Gale-Shapley algorithm requires their choice functions to satisfy the familiar *substitutability* property.

Formally, we say that C_J satisfies substitutability if R_J is monotone. Like IRC, this is an aggregate version of an individual property: We say that an agent's choice function C_i satisfies substitutability if their rejection function R_i is monotone in both arguments, i.e., for

[11] We extend Aygün and Sönmez' (2013) definition to allow for externalities by holding contracts among other agents constant, just as Pycia and Yenmez (2021) do.

all $Y \subseteq Z \subseteq X$ and $i \in I$, $R_i(Y_i|Y_{-i}) \subseteq R_i(Z_i|Z_{-i})$. In words, substitutability means that whenever an agent or group rejects a contract from some choice set, they will continue to reject that contract when more contracts are available or when they expect other agents to sign more contracts. This extends the substitutes condition to encompass externalities in a different way than Pycia and Yenmez (2021): Ours implies that contracts one agent might sign are substitutable for contracts signed by other agents, while theirs implies that one agent's contracts are substitutable for *better outcomes* for other agents.

Unlike IRC, substitutability in aggregate is implied by substitutability at the individual level.

Lemma 2 (Substitutes and Aggregation). *If $\{C_i\}_{i\in J}$ satisfy substitutability, then so does C_J.*

When the aggregate choice functions of two implicit sides satisfy both substitutability and IRC, stability can be characterized in terms of a system of equations. In particular, Lemma 3 shows that stable outcomes are the intersection of two choice sets, one for each implicit side. Each of these choice sets contains all contracts except those rejected from the other choice set by the opposite implicit side. This characterization extends Hatfield and Milgrom (2005) Theorem 2. Because agents can sign contracts with others on the same side (e.g., when contracts are multilateral), and can be affected by contracts between other agents (i.e., when contracts have externalities), IRC is required at the aggregate rather than individual level.

Lemma 3 (Stability as a Fixed Point). *Let $J \cup K = I$. If C_J and C_K both satisfy substitutability and IRC, then X' is stable if and only if $X' = Y \cap Z$ for some solution (Y, Z) to the system of equations*

$$Y = X \setminus R_J(Z), \qquad\qquad Z = X \setminus R_K(Y). \tag{1}$$

The "if" part of Lemma 3 only requires the weak axiom *locally*, rather than for all pairs of choice sets. We say C_J satisfies *irrelevance of rejected contracts at $Y \subseteq X$ (IRC at Y)* if for all $Z \subseteq Y$, whenever $C_J(Y) \subseteq Z$, then $C_J(Y) = C_J(Z)$. IRC at Y specifies that contracts rejected from Y are irrelevant to choices from subsets of Y.

Lemma 4 (Stability as a Fixed Point with Local IRC). *Let $J \cup K = I$. If (Y, Z) solves the system of equations (1), C_J satisfies IRC at Z, C_K satisfies IRC at Y, and both C_J and C_K satisfy substitutability, then $X' = Y \cap Z$ is stable.*

Lemmas 3 and 4 allow us to define a monotone operator F whose fixed points correspond to stable outcomes: Given a pair of implicit sides $J, K \subseteq I$ with $J \cup K = I$, define F :

$2^X \times 2^X \to 2^X \times 2^X$ by

$$F_1(Y, Z) \equiv X \setminus R_J(Z), \qquad\qquad F_2(Y, Z) \equiv X \setminus R_K(F_1(Y, Z)).$$

When the aggregate choice functions for group J and group K both satisfy substitutability, F is monotone in the partial order $\succeq_F \equiv (\supseteq, \subseteq)$. Lemma 3 implies that when, in addition, the aggregate choice functions of J and K each satisfy the weak axiom, the set $\{Y \cap Z \mid F(Y, Z) = Y, Z\}$ is exactly the set of stable outcomes. More generally, Lemma 4 shows that for each fixed point (Y, Z) of F, $Y \cap Z$ is stable as long as C_J satisfies IRC at Z and C_K satisfies IRC at Y, even if C_J and C_K do not satisfy IRC everywhere.

As in Hatfield and Milgrom (2005), iterated applications of F correspond to rounds of a generalized Gale-Shapley algorithm in which agents in one group make offers and agents in the other conditionally accept them. Tarski's theorem ensures that this algorithm converges to a fixed point of F (parts i, iv, and v of Theorem 1), and thus — so long as the aggregate choice of each implicit side satisfies IRC there — a stable outcome (parts ii and iii). It also ensures that these fixed points form a lattice (part i).

Theorem 1 (Stability with Aggregate IRC on Two Implicit Sides). *Let $J \cup K = N$ and suppose both C_J and C_K satisfy substitutability.*

i. The set of fixed points of F is a nonempty lattice with the order $\succeq_F \equiv (\supseteq, \subseteq)$.

ii. If (Y, Z) is a fixed point of F, C_J satisfies IRC at Z, and C_K satisfies IRC at Y, then $Y \cap Z$ is stable.

iii. If both C_J and C_K satisfy IRC, the set of stable outcomes is $\{Y \cap Z \mid F(Y, Z) = Y, Z\}$.

iv. Starting from $(X, \emptyset)$, the generalized Gale-Shapley algorithm converges to the $\succeq_F$-largest fixed point $(\overline{Y}, \overline{Z})$ of F.

v. Starting from $(\emptyset, X)$, the generalized Gale-Shapley algorithm converges to the $\succeq_F$-smallest fixed point $(\underline{Y}, \underline{Z})$ of F.

In Theorem 1, irrelevance of rejected contracts ensures that contracts rejected over the course of the Gale-Shapley algorithm are unable to block the algorithm's outcome. Without it, an implicit side might reject a pair of contracts when both are available, but choose one if the other were absent. This could allow the first contract to form a blocking set once the algorithm terminates.[12]

[12] The Gale-Shapley algorithm in Theorem 1 may remind readers of tâtonnement processes in general equilibrium. (We thank Alex Teytelboym for encouraging us to think about this connection.) When commodities are gross substitutes, the aggregate excess demand for them satisfies the weak axiom of revealed preference

The local version of irrelevance of rejected contracts suffices for this purpose. This is because the final choice set for each implicit side is made up of precisely those contracts which it has been offered (or which have not been rejected) by the other implicit side over the course of the algorithm. But if C_J and C_K only satisfy IRC at these sets, there may be other stable outcomes not found by the Gale-Shapley algorithm: The "only if" part of Lemma 3 relies on IRC holding everywhere. In contrast, when C_J and C_K each satisfy IRC everywhere (as in standard two-sided matching models), Theorem 1 shows that the set of stable outcomes found using the Gale-Shapley algorithm is exhaustive.

3.2 Orderly Aggregation Among All Agents

In settings where it applies to the aggregate choice of the entire market (rather than two implicit sides), irrelevance of rejected contracts allows us to dispense not only with requirements on the market's structure, but also with the requirement of substitutability. In particular, when IRC holds locally at the set of all contracts, we show that a stable outcome exists and is unique.

Theorem 2 (Stability with Aggregate IRC Among All Agents). *If the aggregate choice function, C, satisfies irrelevance of rejected contracts at the set of all contracts, X, then $C(X)$ is the unique stable set of contracts.*

Theorem 2's stable outcome can be interpreted as the result of a *one-sided* version of our generalized Gale-Shapley algorithm. Choose $J = I$ and $K = \emptyset$, letting $C_\emptyset(Y) = Y$ for all $Y \subseteq X$ since an empty group of agents cannot reject any contracts. Then $F_2(Y, Z) = X$ for all Y, Z — and from any starting point, the algorithm converges to $(C(X), X)$ within two rounds. This convergence does not rely on monotonicity, and hence does not require a substitutability condition. Like in Theorem 1, we need to invoke IRC, this time for $C_I = C$, to ensure that the algorithm's outcome is stable. And as in part (ii) of Theorem 1, we need only do so at the final choice set, which will always be X. But when all agents are on one implicit side, this is *all* we need: unlike in Lemma 3, no substitutability condition is necessary.

We show in Rostek and Yoder (2020) that a unique stable outcome exists when contracts are complementary. It bears emphasizing that complementarity is different than the condition we use here. In particular, C can fail to satisfy IRC at X when contracts are complementary,

at market-clearing prices; consequently, tâtonnement processes converge to those prices (Arrow et al., 1959). In a matching model, substitutability is not sufficient for aggregate demand for contracts C (or C_J, or C_K) to obey irrelevance of rejected contracts — or equivalently, the weak axiom — even at a specific set Y. It is, however, sufficient for the Gale-Shapley algorithm to converge, since it causes the algorithm to be monotone. In the absence of substitutability, the discussion following Theorem 2 shows that we can also get the algorithm to converge when the aggregate choice of the entire market satisfies irrelevance of rejected contracts at the set X of all contracts — the set from which the unique stable outcome is chosen.

and conversely, C can satisfy IRC at X when contracts are not complementary. Naturally, when both conditions hold, the results give the same outcome even though their proofs rely on different properties of aggregate choice. Complementarity is about monotonicity, whereas IRC is about orderly aggregate behavior.

3.3 Aggregate IRC as a Necessary Condition

In general, preservation of the irrelevance of rejected contracts condition under aggregation — either among all agents, or among the agents on each of two implicit sides — is not a necessary condition for the generic existence of stable outcomes. However, when the market can be decomposed into two implicit "sides", there is a related condition that is necessary in a maximal domain sense. We say that C_J has *no 3-cycles* if there exist no distinct contracts $\{y(1), y(2), y(3)\}$ such that for each m, $y(m)$ is chosen by J when both $y(m)$ and $y(m-1)$ are available, and $y(m-1)$ is not; i.e., $y(m) = C_J(\{y(m), y(m-1 \mod 3)\})$.[13]

When C_J satisfies irrelevance of rejected contracts, it has no 3-cycles.[14] This relationship allows us to give a partial converse to Theorem 1 in the vein of Abeledo and Isaak (1991). Theorem 3 shows that when the market has two implicit sides, if the aggregate behavior of one implicit side does not obey at least the minimal no 3-cycles axiom of rational choice, then stable outcomes will not generically exist within a rich class of choice behavior for the other implicit side.

Theorem 3 (Stable Outcomes and 3-Cycles). *Suppose that each contract names an agent in K, i.e., $X_K = X$, and let $J = N \setminus K$. If the aggregate choice of group J, C_J, has a 3-cycle, then there exist choice functions C_k for the agents in K such that C_K satisfies IRC and substitutability but no stable outcome exists.*

While other authors have noted that the presence of a 3-cycle among members of a group poses problems for stability in other settings, the point made by Theorems 1 and 3 is different. It is not the presence of a 3-cycle that rules out the existence of stable outcomes. Rather, it is the presence of a 3-cycle *that cannot be eliminated by identifying two implicit sides of the market whose aggregate choice satisfies irrelevance of rejected contracts.*

Theorem 3 tells us that if we want to guarantee the existence of stable outcomes, the requirement that IRC survives aggregation — either on two implicit sides, or among all agents

[13] This property is similar in principle to the no 3-cycles condition used by Pycia (2012), but rules out a 3-element loop of contracts in a coalition's choice function rather than a triple of coalitions which cause the preferences of a triple of agents to form a loop.

[14] No 3-cycles is a weaker version of the *strong axiom of revealed preference (SARP)*, which requires that there is no loop of choice sets $\{Y(m)\}_{m=1}^{M}$ such that $C_J(Y(m)) \subseteq Y(m+1 \mod M)$. When contracts are substitutes, Alva (2018) shows that IRC and SARP are equivalent, and hence either implies the absence of 3-cycles.

— cannot be weakened in isolation (at least not by much). Instead, we must replace it with different conditions on the environment. For instance, we show in Rostek and Yoder (2020) that in settings where contracts are complementary instead of substitutable, no "sides", whether implicit or explicit, are necessary for stable outcomes to exist. Alternatively, Bando and Hirai (2021) show that certain conditions on the structure of the market ensure existence of a stable outcome, even in the absence of substitutability or complementarity.

4 Foundations for Orderly Aggregation

At first glance, the conditions on aggregate choice which Theorems 1 and 2 rely on appear quite strong. In particular, suppose that each agent's choice function C_i is derived from maximization of her payoffs $u_i : 2^X \to \mathbb{R}$ — which we assume do not have indifferences[15] — in the manner that is common in the literature:

$$C_i(Y_i|Y_{-i}) = \arg\max_{S \subseteq Y_i} u_i(S \cup Y_{-i}) \text{ for each } Y \subseteq X. \tag{2}$$

As Aygün and Sönmez (2013) note, such choice functions will always satisfy irrelevance of rejected contracts. But in general, there is no reason to expect the IRC condition to survive aggregation among a group of agents: for instance, a pair of agents might disagree about which of two contracts naming both of them they want to sign, in which case neither is chosen by the pair in aggregate; but if one of those contracts is removed, they might both choose the one which remains.

The choice function described by (2) can be thought of as *myopic*, in the sense that each agent behaves as if all contracts which are available will actually be signed by the other agents. But suppose instead that the members of a group are strategically sophisticated, in the sense that each makes conjectures about the contracts that others will choose, and those conjectures are correct. In this section, we show that in contrast to myopic choice functions, the resulting *strategically consistent* choice functions aggregate across agents in an orderly manner in settings with externalities or multilateral contracts. Theorem 4 shows that in environments with strategcally sophisticated agents, the IRC condition on aggregate choice amounts to a natural refinement on agents' conjectures ensuring that they are consistent across different sets of proposed contracts. Moreover, this refinement also ensures that IRC is satisfied at the level of individual agents.

[15] Formally, we assume that payoff functions have no indifferences, conditional on the set of contracts that do not name the agent: $u_i(Y \cup X') \neq u_i(Z \cup X')$ for each $Y, Z \subseteq X_i$ and $X' \subseteq X_{-i}$.

4.1 Strategic Consistency

We begin by formalizing our key notions of strategic and myopic choice behavior.

Definition (Strategic and Myopic Consistency). Given a profile of payoff functions over outcomes $\{u_i : 2^X \to \mathbb{R}\}_{i \in J}$,

- The choice functions $\{C_i\}_{i \in J}$ and conjectures $\{\mu_i : 2^X \to 2^X\}$ form a *strategically consistent* assessment if for each $i \in J$,

  1. C_i is rational given μ_i: For each $Y \subseteq X$, $C_i(Y_i|Y_{-i}) = \arg\max_{S \subseteq \mu_i(Y)_i} u_i(S \cup \mu_i(Y)_{-i})$, and
  2. μ_i is correct given $\{C_j\}_{j \in J}$: For each $Y \subseteq X$, $\mu_i(Y) = C_{J \setminus \{i\}}(Y)$.

  When the profile of choice functions $\{C_i\}_{i \in J}$ is part of a strategically consistent assessment, we say those choice functions are strategically consistent.

- The choice function C_i is *myopically consistent* if it is defined according to (2).[16]

Strategic consistency is motivated by assumptions about agents' epistemic sophistication: when faced with a set of contracts that has been proposed to a specific group they belong to, such as their side of the market (in two-sided markets with externalities), their implicit "side" (in markets with two implicit "sides"), or the market as a whole (in more general settings), they make correct conjectures about which contracts the other agents will choose. In the classical two-sided matching literature, agents do not need to make such conjectures: Externalities are absent, so the behavior of others on the same side of the market is not relevant to an agent's choice. Hence, in such settings, myopic consistency pins down agents' choice behavior in exactly the same way as strategic consistency among agents on the same side.[17] More generally, choice functions that are myopically consistent take as argument the set of contracts that other agents are assumed to have *signed*, while strategically consistent choice functions take as argument a set of contracts that has been *proposed* to a group.

While (as noted earlier) the myopically consistent choice functions of individual agents will always satisfy irrelevance of rejected contracts, it is possible for strategically consistent choice functions to fail IRC. Consider a strategically consistent assessment $\{C_i, \mu_i\}_{i \in I}$ in which each agent accurately believes that when Y is available, the other agents will collectively choose $Z \subseteq Y$: we have $C_i(Y_i|Y_{-i}) = Z_i$ for each i. Suppose we remove the rejected contracts that name some agent j from the proposal, leaving $Z_j \cup Y_{-j}$. Then strategic consistency allows for

[16] Note that this definition means that given a profile of payoff functions (without indifferences — see footnote 15), the profile of myopically consistent choice functions is unique.

[17] That is, in a two-sided market without externalities, the choice functions of each side are part of a strategically consistent assessment if, and only if, they are myopically consistent.

— to pick an extreme example — each agent to correctly believe that *every other agent rejects all contracts* from that set: $\mu_i(Z_j \cup Y_{-j}) = \emptyset$. Given these beliefs, each agent's only rational choice from $Z_j \cup Y_{-j}$ is the empty set of contracts; and with those choices, their beliefs at $Z_j \cup Y_{-j}$ are correct. Hence, we have $C_j(Y_j|Y_{-j}) = Z_j$, but $C_j(Z_j|Y_{-j}) = \emptyset$, and C_j does not satisfy IRC, despite the fact that it is part of a strategically consistent assessment.

The possibility of such "sunspot" behavior motivates a refinement on agents' beliefs. In particular, we propose that an agent i should not believe that the other members of their group would change their behavior because of the removal of contracts which are not chosen by any of them. That is, for each Y, Z with $Y \setminus X_J = Z \setminus X_J$, and each $i \in J$,

$$Y \supseteq Z \supseteq C_j(Y_j|Y_{-j}) \text{ for all } j \in J \Rightarrow \mu_i(Z) \cap X_{J\setminus\{i\}} = \mu_i(Y) \cap X_{J\setminus\{i\}}.$$

We call this criterion on strategically consistent assessments *irrelevance of unanimously rejected contracts (IURC)*. Proposition 1 shows that in fact, with strategic consistency, IURC amounts to a condition on conjectures alone. In particular, the "unanimous" quantifier is superfluous: IURC is equivalent to irrelevance of rejected contracts, as applied to the agents' conjectures; i.e., $\mu_i(Z) \subseteq Y \subseteq Z \Rightarrow \mu_i(Y) = \mu_i(Z)$. That is, beliefs are independent of contracts rejected by *every agent* in group J (as in IURC) exactly when they are independent of contracts rejected by *some other agent* in the group (which is what the IRC condition captures).

Proposition 1 (Irrelevance of Rejected Contracts in Choices and Beliefs). *Suppose $\{C_i, \mu_i\}_{i\in J}$ is strategically consistent for some profile of payoff functions.*

i. $\{C_i, \mu_i\}_{i\in J}$ satisfies IURC if and only if $\{\mu_i\}_{i\in J}$ each satisfy IRC.

ii. If $\{\mu_i\}_{i\in J}$ each satisfy IRC, then $\{C_i\}_{i\in J}$ each satisfy IRC.

Even though IURC involves the choices of every agent in the group, strategic consistency ensures that each individual agent's beliefs are consistent with those choices; consequently, refining agents' conjectures using IURC is equivalent to refining them using the IRC condition. When agents' choice functions are derived from payoff maximization given these beliefs, it follows that they will also satisfy IRC. In fact, this refinement on agents' conjectures not only ensures that individual choice satisfies IRC, it is *equivalent* to the irrelevance of rejected contracts condition on the group's aggregate choice.

Theorem 4 (Irrelevance of Rejected Contracts in Beliefs and Aggregate Choice). *Suppose $\{C_i, \mu_i\}_{i\in J}$ is strategically consistent for some profile of payoff functions. $\{\mu_i\}_{i\in J}$ each satisfy IRC if and only if C_J satisfies IRC.*

Theorem 4 allows us to apply our existence theorems in settings with strategically sophisticated agents. First, in a (explicitly or implicitly) two-sided market, suppose that agents are only strategically sophisticated regarding the behavior of other agents on the same side. This is natural, for instance, when agents participate in a mechanism based on a generalized Gale-Shapley algorithm. In such settings, Theorem 1 characterizes the set of stable outcomes that correspond to each strategically consistent assessment with reasonable beliefs. Alternatively, in the general setting, suppose that each agent is strategically sophisticated about the behavior of each of the others. Then Theorem 2 shows that each strategically consistent assessment with reasonable beliefs uniquely pins down a stable outcome.

The arguments for Proposition 1 and Theorem 4 rely on three implications of strategic consistency which merit independent attention. The first is that if a group of agents' choice functions are part of a strategically consistent assessment, then the group's aggregate choice is separable in individual choice, even if it is not one side of a two-sided market; i.e., the set of contracts chosen by any agent is equal to the set of contracts which name that agent and are chosen in aggregate by the group. Intuitively, if an agent rejects a contract, it cannot be part of the group's aggregate choice, while if an agent chooses a contract, it cannot have been rejected by another agent in the group, since then the agent would have correctly believed it to be unavailable. This leads to our second observation: the aggregate choice of a group can be written as the union of its members' choice functions and the set of contracts not involving the group. That is, even when the group does not represent one side of a two-sided market, its aggregate choice is a direct sum of individual choices, and so aggregation does not discard information about contracts chosen by any of the agents. Finally, each agent's conjecture is endogenously pinned down by the available contracts between him and agents outside the group, plus the aggregate choice of the whole group.

Lemma 5 (Structure of Strategically Consistent Assessments). *Suppose $\{C_i, \mu_i\}_{i \in J}$ is strategically consistent for some profile of payoff functions.*

i. C_J is separable in individual choice: For each $i \in J$ and $Y \subseteq X$, $C_i(Y_i|Y_{-i}) = C_J(Y) \cap X_i$.

ii. C_J can be written as a union: For each $Y \subseteq X$, $C_J(Y) = (Y \setminus X_J) \cup \bigcup_{i \in J} C_i(Y_i|Y_{-i})$.

iii. Agents' conjectures have a common component, and this common component is aggregate choice: For each $i \in J$ and $Y \subseteq X$, $\mu_i(Y) = C_J(Y) \cup (Y_i \setminus X_{J \setminus \{i\}})$.

The intuition for (iii) is as follows. Recall that an agent i's correct conjecture $\mu_i(Y)$ is equal to the set of contracts $C_{J \setminus \{i\}}(Y)$ which other agents in the group do not reject. Contracts that do not involve other agents in the group cannot be rejected by them (and thus $C_{J \setminus \{i\}}(Y) \supseteq Y_i \setminus X_{J \setminus \{i\}}$); contracts chosen by the other agents in the group must not be rejected by agent

i, since those other agents' beliefs are correct (and thus $C_J(Y) \cap X_{J\setminus\{i\}} = C_{J\setminus\{i\}}(Y) \cap X_{J\setminus\{i\}}$); and contracts that do not name any agent in J cannot be rejected by any of them (and thus $C_J(Y) \setminus X_J = Y \setminus X_J = C_{J\setminus\{i\}}(Y) \setminus X_J$).

4.2 Discussion

Strategically consistent assessments are equilibrium objects, in the sense that each agent in the group makes optimal choices, given the choices of the others. This means that, unlike myopically consistent choice functions, they are not necessarily unique. Instead, for any given profile of payoff functions, there may be multiple strategically consistent assessments, and each may make different predictions about the set of stable outcomes. Unlike the set of stable outcomes in a classical matching model, however, these predictions can be refined using criteria on agents' beliefs. In particular, in Section 5.1, we use a mild forward induction criterion to narrow down the set of outcomes that are stable when each agent is strategically sophisticated about the behavior of the others, similarly to the way that the intuitive criterion of Cho and Kreps (1987) refines the predictions of signaling games.

Moreover, because strategically consistent assessments are equilibrium objects, their existence may not be immediate for a given profile of payoff functions. In Section 5, we provide existence results for two classes of environments. First, Section 5.2 considers two-sided markets with externalities where agents are strategically sophisticated about the behavior of others on the same side. Second, Section 5.1 considers settings where agents are strategically sophisticated about the behavior of all other agents.

5 Foundations for Strategic Consistency

Together with Theorems 1 and 2, Section 4 shows that strategic consistency combined with a refinement on beliefs ensures that stable outcomes exist and form a lattice (with strategic consistency on two implicit "sides" and substitutability) or are unique (with strategic consistency at the level of the entire market). Here, we discuss what the features of agents' underlying preferences imply about the existence and properties of these strategically consistent assessments. We can do so by focusing attention on the agents' myopically consistent choice functions, since these completely characterize the way their payoffs influence any strategically consistent assessment.

Lemma 6 (Relationship Between Myopic and Strategic Consistency). *Given a profile of payoff functions $\{u_i : 2^X \to \mathbb{R}\}_{i\in J}$, suppose that $\{\hat{C}_i\}_{i\in J}$ are myopically consistent. Then the assessment $\{C_i, \mu_i\}_{i\in J}$ is strategically consistent if and only if for each $i \in J$,*

1. $C_i(Y_i|Y_{-i}) = \hat{C}_i(\mu_i(Y)_i|\mu_i(Y)_{-i})$ *for each* $Y \subset X$*; and*

2. μ_i *is correct given* $\{C_j\}_{j\in J}$.

We begin by showing in Section 5.1 that there *always* exists a strategically consistent assessment for the entire market with beliefs satisfying IRC. No conditions on the market's structure or complementarity or substitutability among contracts are required. Hence, taking agents' preferences (or their myopically consistent choice functions) as given, we can use Theorems 2 and 4 to find stable outcomes in *any* matching setting where agents are strategically sophisticated. In particular, stable outcomes exist regardless of whether multilateral contracts or externalities are present, and without regard to subsititutability or complementarity among contracts.

We also characterize the set of outcomes which are stable for *some* strategically consistent assessment with beliefs satisfying IRC, given a profile of agents' payoff functions. Since this set may include many outcomes, we then introduce a further refinement on agents' conjectures based on forward induction, and characterize the outcomes which are stable for some strategically consistent assessment with beliefs that satisfy it.

In Section 5.2, we turn our attention to two-sided markets with externalities. There, we give conditions ensuring that each side has a strategically consistent assessment comprised of choice functions satisfying substitutability and beliefs satisfying IRC. Consequently, when agents' myopically consistent choice functions satisfy these conditions, Theorems 1 and 7 ensure that when agents are able to correctly anticipate the behavior of other agents on the same side of the market — perhaps because they face similar incentives — stable outcomes exist and form a lattice, even if agents are not strategically sophisticated about the behavior of others on the opposite side.

5.1 Foundations for Strategic Consistency in General Settings

It is always possible for each agent to simultaneously make conjectures about the behavior of each of the others which are both correct and reasonable (i.e., satisfy IRC). That is, there always exists a strategically consistent assessment for the set of all agents with beliefs satisfying IRC. In fact, we can give a complete characterization of the set of outcomes that can arise from such strategically consistent behavior, and show that it is nonempty. Formally, we say that $Y \subseteq X$ is *stable for* $\{C_i, \mu_i\}_{i\in I}$ if it is stable in an environment with choice functions given by $\{C_i\}_{i\in I}$.

Theorem 5 (Stability with Strategic Consistency for All Agents). *Given a profile of payoff functions* $\{u_i : 2^X \to \mathbb{R}\}_{i\in I}$*, suppose that* $\{\hat{C}_i\}_{i\in I}$ *are myopically consistent.*

1. *If Y is stable for some strategically consistent assessment $\{C_i, \mu_i\}_{i \in I}$, it is myopically individually rational: $\hat{C}_i(Y_i|Y_{-i}) = Y_i$ for each $i \in I$.*

2. *If Y is myopically individually rational, i.e., $\hat{C}_i(Y_i|Y_{-i}) = Y_i$ for each $i \in I$, then Y is stable for some strategically consistent assessment $\{C_i, \mu_i\}_{i \in I}$ such that $\{\mu_i\}_{i \in I}$ satisfy IRC.*

Since $\hat{C}_i(\emptyset|\emptyset) = \emptyset$ for any myopically consistent choice functions $\{\hat{C}_i\}_{i \in I}$, it follows that given any profile of payoff functions, $\emptyset$ is stable for some strategically consistent assessment with beliefs satisfying IRC. This is because an assessment where each agent believes no contracts will be chosen by the other agents from any set of available contracts, and thus never chooses any contracts himself, is always strategically consistent, regardless of the underlying payoff functions. However, if other outcomes exist which are individually rational for agents' myopic choice functions, this assessment might seem somewhat unnatural. In particular, it is not robust to forward induction, in the following sense.

Definition (Weak Forward Induction). Given a profile of payoff functions $\{u_i\}_{i \in I}$ and myopically consistent choice functions $\{\hat{C}_i\}_{i \in I}$, we say that beliefs μ_i satisfy *weak forward induction (WFI)* if $\mu_i(Y) = Y$ whenever Y is myopically individually rational, i.e., $\hat{C}_i(Y_j|Y_{-j}) = Y_j$ for each $j \in I$.

To understand this criterion, suppose that given payoff functions $\{u_i\}_{i \in I}$, the assessment $\{C_i, \mu_i\}_{i \in I}$ is strategically consistent and $\{\mu_i\}_{i \in I}$ satisfy IRC. Consider a proposed set of contracts Y such that each agent prefers to sign all of the contracts in Y that name him, provided that every other agent does the same. This proposal is a credible block of any smaller set of contracts: Agents cannot benefit directly from rejecting any of the additional contracts that have been proposed, since $u_i(Y) > u_i(S \cup Y_{-i})$ for each $S \subsetneq Y$ and $i \in I$. Furthermore, they cannot benefit indirectly by creating blocking opportunities that did not exist before Y was proposed: By Theorem 4, $C(Y \cup Z) = Z \Rightarrow C(Y' \cup Z) = Z$ for each $Z \subseteq X$ and $Y' \subseteq Y$. Thus, a forward induction argument should lead the agents to infer that none of the contracts in Y will be rejected — and hence that Y will block each smaller set of contracts, including $\emptyset$.

In general, weak forward induction requires each agent to believe that the members of a blocking coalition will go along with credible proposals to add contracts. We call it *weak* forward induction because it does not require agents to similarly believe in credible proposals to *change* the set of contracts, i.e., add some contracts while deleting others. Evaluating the credibility of the latter set of blocking proposals is more difficult for the agents: In addition to determining whether agents in the blocking coalition can benefit from rejecting newly proposed contracts, they must also determine whether they might benefit from keeping their existing contracts. Moreover, unlike with proposals to add contracts, the absence of an indirect benefit from the creation of new blocking opportunities does not follow from the absence of a direct benefit. While we could consider a stronger forward induction refinement requiring agents to evaluate the credibility of such proposals, doing so would add limited predictive power.[18]

Theorem 6 characterizes the set of outcomes that can arise from strategically consistent behavior when beliefs are both reasonable (in the sense that they satisfy IRC) and robust to forward induction (in the sense that they satisfy WFI).

Theorem 6 (Stability with Weak Forward Induction). *Given a profile of payoff functions $\{u_i : 2^X \to \mathbb{R}\}_{i \in I}$, suppose that $\{\hat{C}_i\}_{i \in I}$ are myopically consistent. Y is stable for some strategically consistent assessment $\{C_i, \mu_i\}_{i \in I}$ with beliefs $\{\mu_i\}_{i \in I}$ satisfying IRC and WFI if and only if Y is a maximal myopically individually rational outcome: (1) $\hat{C}_i(Y_i|Y_{-i}) = Y_i$ for each $i \in I$ and (2) there is no $Y' \supset Y$ such that $\hat{C}_i(Y'_i|Y'_{-i}) = Y'_i$ for each $i \in I$.*

The predictions of Theorem 6 coincide with those of Rostek and Yoder (2020) Theorem 1 in the presence of complementarity between contracts (i.e., monotonicity in myopically consistent choice functions). Such complementarity ensures that the set of myopically individually rational outcomes is a lattice, and so has a unique maximal element. This outcome is then uniquely stable, whether agents behave myopically (as in Rostek and Yoder (2020)) or are strategically sophisticated about the behavior of each of the other agents, and have reasonable beliefs that are robust to forward induction (as in Theorem 6).

The reason that strategic sophistication does not alter predictions about stable outcomes in the presence of complementarity is that — as we point out in Rostek and Yoder (2020) — with complementarity, if choice is myopically consistent, no agent will choose to delete existing contracts from an outcome Y as part of a successful block Z. Instead, the combination of new and old contracts $Y \cup Z$ must be myopically individually rational. Hence, taking the entire set of available contracts $Y \cup Z$ as given when they evaluate the block myopically turns out to be correct given their behavior — and is thus consistent with strategic sophistication. In fact, strategically sophisticated agents must do so if their beliefs satisfy WFI, since $Y \cup Z$ is credible, in the sense of being myopically individually rational.

However, when agents have myopically consistent choice functions that do not satisfy complementarity, Theorem 6 does not necessarily produce a unique prediction.

Example 3 (Roommate Problem). Consider the classical roommate problem: Three friends must come to an agreement about which two of them will rent an apartment to-

[18]In particular, suppose that among the maximal myopically individually rational outcomes, there is no element that dominates each of the others in the Blair (1988) order associated with the aggregate myopically consistent choice function, and conversely, no element which is dominated by each of the others in that order. Then for any maximal myopically individually rational outcome Y, we can construct a strategically consistent assessment which satisfies this stronger refinement, and for which Y is the unique stable outcome.

gether. Hence $I = \{1, 2, 3\}$, $X = \{x_{12}, x_{23}, x_{31}\}$, and $N(x_{ij}) = \{i, j\}$ for each $i, j \in I$. The three friends have payoff functions with the following properties. Agents are indifferent about roommate agreements that do not involve them: $u_i(Y) = u_i(Y \cup \{x_{jk}\})$ for each $i \in I$ and each $j \neq k \neq i$. Agents never choose to be part of two roommate agreements: $u_i(\{x_{ij}, x_{ik}\}) < u_i(\emptyset)$ for each $i \in I$ and each $j \neq k \neq i$. Agents prefer having a roommate to being unmatched: $u_i(\emptyset) < u_i(\{x_{ij}\})$ for each $i \in I$ and $j \neq i$. Their preferences over roommates form a cycle: $u_1(x_{12}) > u_1(x_{31})$, $u_2(x_{23}) > u_1(x_{12})$, $u_3(x_{31}) > u_1(x_{23})$.

When choice is myopically consistent given these payoff functions, no stable outcome exists. Any outcome with two or three agreements fails to be individually rational, since some agent will be named by multiple agreements, and will reject one of them. Any outcome with one agreement is blocked, since $\{x_{12}\}$ is blocked by $\{x_{23}\}$, $\{x_{23}\}$ is blocked by $\{x_{31}\}$, and $\{x_{31}\}$ is blocked by $\{x_{12}\}$. And the outcome with no agreements is blocked by any single agreement.

This seems at odds with the observation that in the real world, people find roommates. But this is precisely the prediction of Theorem 6: Each of $\{x_{12}\}$, $\{x_{23}\}$, and $\{x_{31}\}$ are myopically individually rational, and there are no larger myopically individually rational outcomes. Hence, they are each stable for some strategically consistent assessment with beliefs that are reasonable and robust to forward induction.

Some discussion of the assessments for which these outcomes are stable is in order. Since the environment is symmetric, we focus on those for which $\{x_{12}\}$ is stable. There are two of these, described below:

$$\{C_i, \mu_i\}_{i \in I} : \begin{array}{ll} \mu_i(Y) = \{x_{12}\}, \text{ if } Y \ni x_{12}; & \mu_i(Y) = \{x_{31}\}, \text{ if } Y \ni x_{31} \text{ and } x_{12} \notin Y; \\ \mu_i(\{x_{23}\}) = x_{23}; & C_i(Y_i | Y_{-i}) = \mu_i(Y)_i, \text{ for all } Y \subseteq X. \end{array}$$

$$\{C'_i, \mu'_i\}_{i \in I} : \begin{array}{ll} \mu'_i(Y) = \{x_{12}\}, \text{ if } Y \ni x_{12}; & \mu'_i(Y) = \{x_{23}\}, \text{ if } Y \ni x_{23} \text{ and } x_{12} \notin Y; \\ \mu'_i(\{x_{31}\}) = x_{31}; & C'_i(Y_i | Y_{-i}) = \mu'_i(Y)_i, \text{ for all } Y \subseteq X. \end{array}$$

In both assessments, agent 2 chooses not to participate in a block of $\{x_{12}\}$ by $\{x_{23}\}$. For $\{C_i, \mu_i\}_{i \in I}$, the reasoning is straightforward: Although agent 2 would prefer having agent 3 as a roommate rather than agent 1, he knows that going along with this block would allow agents 1 and 3 to successfully propose a further block to $\{x_{13}\}$, leaving him unmatched. Hence, even though agents 2 and 3 both prefer $\{x_{23}\}$ to $\{x_{12}\}$, proposing a block of the latter outcome by the former is not credible under $\{C_i, \mu_i\}_{i \in I}$.

This "farsighted" inference does not apply under $\{C'_i, \mu'_i\}_{i \in I}$: In that assessment, a block of $\{x_{12}\}$ by $\{x_{23}\}$ does not create any new blocking opportunities, and agents 2 and 3 would choose to go along with it if they either believed the other agent would or if they behaved myopically. Hence, $\{C'_i, \mu'_i\}_{i \in I}$ would fail a stronger forward induction criterion that required agents to believe in credible proposals to change the set of contracts. But $\{C_i, \mu_i\}_{i \in I}$ would pass such a criterion, and so using it to eliminate $\{C'_i, \mu'_i\}_{i\in I}$ would not change the set of outcomes that are stable for some assessment.

More generally, Theorem 6 allows us to compare the collection of outcomes that are stable for some strategically consistent assessment with beliefs satisfying WFI and IRC to the collection of outcomes that are stable when choice is myopically consistent. The former collection is a superset of the latter in settings where no outcomes are stable for myopically consistent choice functions (as in the roommate problem) or where myopically consistent choice functions satisfy complementarity (in which case they coincide).

Corollary 2 shows that this relationship holds more generally. This shows the impact of strategic sophistication among all agents in environments where myopically consistent choice functions are known to produce stable outcomes, such as two-sided markets with substitutability and no externalities (e.g., Hatfield and Milgrom (2005)) or bilateral trading networks with full substitutability and no externalities (e.g., Hatfield and Kominers (2012); Hatfield et al. (2013)). In such settings, the assumption that each agent is strategically sophisticated about the behavior of the others — and that their beliefs satisfy the WFI and IRC refinements — can only expand the set of predictions that are consistent with the primitives of the model, i.e., payoff functions or (as in Lemma 6) myopically consistent choice functions. Conversely, the assumption of any *specific* strategically consistent assessment satisfying IRC and WFI can either narrow the set of predictions offered by the model with myopically consistent choice to a single element (by Theorem 2) or change its prediction to a different outcome.

Corollary 2 (Stability with Myopic Consistency vs. Strategic Consistency and WFI). *Given a profile of payoff functions $\{u_i : 2^X \to \mathbb{R}\}_{i\in I}$, if Y is stable when choice functions are myopically consistent, then it is stable for some strategically consistent assessment $\{C_i, \mu_i\}_{i\in I}$ such that $\{\mu_i\}_{i\in I}$ satisfies WFI and IRC.*

Intuitively, if an outcome Y is stable when choice is myopically consistent, it must be myopically individually rational, and there cannot be a larger myopically individually rational outcome Y' (since then when choice is myopically consistent, $Y' \setminus Y$ would block Y). Thus, Theorem 6 implies that it is stable for some strategically consistent assessment with beliefs satisfying IRC and WFI.

5.2 Two-Sided Markets with Externalities

When it applies only to one side of a two-sided market with externalities — such as those studied by Pycia and Yenmez (2021) — strategic consistency requires less strategic sophistication than it does in more general settings. In particular, agents do not need to make inference about which available contracts others on the same side will choose to sign *with them*. Instead, they only need to infer which contracts the others will choose to sign *with agents on the other side*.

Formally, we say the market is *two-sided* if there exists a partition $\{J, K\}$ of I such that for each $x \in X$, $N(x) = \{j, k\}$ for some $j \in J$ and $k \in K$. In a two-sided market, the set of all contracts X is partitioned by the sets of contracts $\{X_i\}_{i\in J}$ that name each agent on a side J. This fact leads to the following lemma:

Lemma 7. *If the market is two-sided with sides J and K, then for each $Y \subseteq X$, $L \subseteq J$, and $M \subseteq K$, $C_L(Y) = \left(\bigcup_{i\in L} C_i(Y_i|Y_{-i})\right) \cup \left(\bigcup_{i\in J\setminus L} Y_i\right)$ and $C_M(Y) = \left(\bigcup_{i\in M} C_i(Y_i|Y_{-i})\right) \cup \left(\bigcup_{i\in K\setminus M} Y_i\right)$.*

In a two-sided market with externalities, it may not be possible for agents on the same side of the market to simultaneously make accurate predictions about each other's behavior. That is, there may not be a strategically consistent assessment for each side of the market. Here, we give conditions on preferences — in the form of conditions on the myopically consistent choice functions they give rise to — which are sufficient for the existence of strategically consistent assessments on both sides of the market such that choice functions satisfy substitutability, and beliefs satisfy IRC. These sufficient conditions are precisely those introduced to this setting by Pycia and Yenmez (2021) to ensure that stable outcomes exist when agents behave myopically. This clarifies the connection between our results for two-sided markets with externalities and theirs. It also provides a microfoundation for the irrelevance of rejected contracts condition on the aggregate choice of each side.

In order to present this result, we repeat their conditions here. $\hat{C}_i$ satisfies *standard substitutability* if its rejection function $\hat{R}_i$ is monotone in its first argument: For all $Y, Z \subseteq X_i$ and $X' \subseteq X_{-i}$, $Y \subseteq Z \Rightarrow \hat{R}_i(Y|X') \subseteq \hat{R}_i(Z|X')$. For a side J, the preorder $\succeq_J$ on 2^X is *consistent* with $\{\hat{C}_i\}_{i\in J}$ if $\bigcup_{i\in J} \hat{C}_i(Y'_i|Z'_{-i}) \succeq \bigcup_{i\in J} \hat{C}_i(Y_i|Z_{-i})$ for each $Y' \supseteq Y$ and $Z' \succeq Z$. Finally, $\{\hat{C}_i\}_{i\in J}$ satisfy *monotone externalities* if there is a consistent preorder $\succeq_J$ such that for all $Y \subseteq X$ and $Z' \succeq Z$, $\hat{R}_i(Y_i|Z'_{-i}) \supseteq \hat{R}_i(Y_i|Z_{-i})$.

Theorem 7 (Monotone Externalities as a Foundation For Strategic Consistency). *Suppose that the market is two-sided with sides J and K. Given payoff functions $\{u_i\}_{i\in J}$, if the myopically consistent choice functions $\{\hat{C}_i\}_{i\in J}$ satisfy irrelevance of rejected contracts, standard substitutability, and monotone externalities, then a strategically consistent assessment $\{C_i, \mu_i\}_{i\in J}$ exists such that each μ_i satisfies IRC and each C_i satisfies subsitutability.*

Theorem 7 shows that in a two-sided matching setting where agents' myopically consistent choice functions satisfy Pycia and Yenmez' (2021) conditions, it is possible for agents on the same side of the market to simultaneously make accurate predictions about each other's behavior. Hence, Theorems 1 and 4 imply that whenever they are strategically sophisticated enough to do so, stable outcomes exist. This complements the main result of Pycia and Yenmez (2021), who show that these conditions ensure the existence of stable outcomes when agents behave myopically.

Moreover, Theorem 1 (i) implies that under these conditions, when the agents' correct conjectures about others' behavior — and thus the strategically consistent assessment — are held fixed, the set of stable outcomes forms a lattice.[19] This contrasts with settings where agents behave myopically: there, Pycia and Yenmez (2021) show that a lattice characterization of the set of stable outcomes is unavailable.

6 Applications

Here, we discuss two applications of matching theory that Theorem 6 allows us to consider.

6.1 Legislative Bargaining

A rich political economy literature considers settings where legislators bargain multilaterally over which of several policies to enact. Following Baron and Ferejohn (1989), this literature generally takes a noncooperative approach, modeling a negotiation as a dynamic game where legislators take turns proposing a division of surplus which is then subjected to a majoritarian vote. As Ali et al. (2019) show, the outcome of this bargaining protocol is sensitive to its extensive form: giving the legislators a small amount of information about the identity of future proposers can dramatically change the division of surplus predicted by the Baron and Ferejohn (1989) model.

When we represent each possible agreement by a majority of legislators to pass a bill as a contract, Theorem 6 allows us to use matching theory to model legislative bargaining in a way which does not rely on the specifics of the bargaining process. By doing so, we can also accommodate settings where policies might not represent divisions of surplus, and where multiple proposals can pass simultaneously. Theorem 6 then tells us that when the legislators are strategically sophisticated, and have beliefs that are both reasonable and satisfy forward induction, the legislature will pass a maximal set of bills that command the support of a majority, given the voting behavior of the other legislators. Moreover, Theorem 6 pins down the way that legislators' beliefs determine which of these outcomes will obtain.

[19] However, we should not expect that, given a profile of payoff functions, the set of outcomes that are stable for *some* strategically consistent assessment will have a lattice characterization: If 2^X is not a lattice in the consistent preorder $\succeq_J$, then the set of fixed points of $G_Y(Z) \equiv \bigcup_{i \in J} \hat{C}_i(Y_i|Z_{-i})$ — and hence the set of choice function profiles generated by the externalities in myopic choice — does not form a lattice.

6.2 Trade Agreement Formation

The question of when and why nations form free trade agreements with one another has attracted considerable interest in the international trade literature. While some papers explicitly model the bargaining process, many (e.g., Saggi et al. (2017)) adopt a stability concept of the same general nature as ours. In particular, a set of agreements is typically required to be *individually rational* and *robust to coalitional deviations*, for various definitions of these terms.

Because the collection of possible free trade agreements grows geometrically in the number of countries — and hence the collection of possible coalitional deviations grows even faster — manually checking for the presence of these deviations can be intractable in models with more than a few countries. This can make some of the insights provided by these models inaccessible: For instance, Saggi et al. (2017) show that expanding their model from three countries to four changes its qualitative predictions, but do not consider settings with more than four countries.

Myopically individually rational outcomes, on the other hand, are simple to compute in these models, even with many countries. Hence, Theorem 6 makes it computationally feasible to consider the full implications of these models of FTA formation under the assumption that nations are strategically sophisticated.

References

Abeledo, H. and G. Isaak (1991): "A Characterization of Graphs That Ensure the Existence of Stable Matchings," *Mathematical Social Sciences*, 22, 93–96.

Ali, S. N., B. D. Bernheim, and X. Fan (2019): "Predictability and Power in Legislative Bargaining," *The Review of Economic Studies*, 86, 500–525.

Alva, S. (2018): "WARP and Combinatorial Choice," *Journal of Economic Theory*, 173, 320–333.

Arrow, K. J., H. D. Block, and L. Hurwicz (1959): "On the Stability of the Competitive Equilibrium, II," *Econometrica: Journal of the Econometric Society*, 82–109.

Aygün, O. and T. Sönmez (2013): "Matching With Contracts: Comment," *American Economic Review*, 103, 2050–51.

Bando, K. (2012): "Many-to-One Matching Markets With Externalities Among Firms," *Journal of Mathematical Economics*, 48, 14–20.

Bando, K. and T. Hirai (2021): "Stability and Venture Structures in Multilateral Matching," *Journal of Economic Theory*, 105292.

Baron, D. P. and J. A. Ferejohn (1989): "Bargaining in Legislatures," *American Political Science Review*, 83, 1181–1206.

Blair, C. (1988): "The Lattice Structure of the Set of Stable Matchings with Multiple Partners," *Mathematics of Operations Research*, 13, 619–628.

Cho, I.-K. and D. M. Kreps (1987): "Signaling Games and Stable Equilibria," *The Quarterly Journal of Economics*, 102, 179–221.

Fisher, J. C. and I. E. Hafalir (2016): "Matching with Aggregate Externalities," *Mathematical Social Sciences*, 81, 1–7.

Gale, D. and L. S. Shapley (1962): "College Admissions and the Stability of Marriage," *The American Mathematical Monthly*, 69, 9–15.

Hafalir, I. E. (2008): "Stability of Marriage with Externalities," *International Journal of Game Theory*, 37, 353–369.

Hatfield, J. W. and S. D. Kominers (2012): "Matching in Networks with Bilateral Contracts," *American Economic Journal: Microeconomics*, 4, 176–208.

——— (2015): "Multilateral Matching," *Journal of Economic Theory*, 156, 175–206.

Hatfield, J. W., S. D. Kominers, A. Nichifor, M. Ostrovsky, and A. Westkamp (2013): "Stability and Competitive Equilibrium in Trading Networks," *Journal of Political Economy*, 121, 966–1005.

Hatfield, J. W. and P. R. Milgrom (2005): "Matching With Contracts," *American Economic Review*, 913–935.

Kojima, F., P. A. Pathak, and A. E. Roth (2013): "Matching with Couples: Stability and Incentives in Large Markets," *The Quarterly Journal of Economics*, 128, 1585–1632.

Pycia, M. (2012): "Stability and Preference Alignment in Matching and Coalition Formation," *Econometrica*, 80, 323–362.

Pycia, M. and M. B. Yenmez (2021): "Matching With Externalities," Working paper.

Rostek, M. and N. Yoder (2020): "Matching with Complementary Contracts," *Econometrica*, 88, 1793–1827.

Saggi, K., W. F. Wong, and H. M. Yildiz (2017): "Should the WTO Require Free Trade Agreements to Eliminate Internal Tariffs?" *Journal of International Economics*, forthcoming.

Sasaki, H. and M. Toda (1996): "Two-Sided Matching Problems with Externalities," *Journal of Economic Theory*, 70, 93–108.

Proofs

Proof of Lemma 1 (Stability in Aggregate)

($\Rightarrow$) Suppose Y is stable. We begin by proving condition 1: Since Y is individually rational, $Y_i = C_i(Y_i|Y_{-i})$ for all $i \in N$. Then $Y = C_i(Y_i|Y_{-i}) \cup Y_{-i}$ for all $i \in N$. From the definition of C, $Y = C(Y)$.

Now suppose that condition 2 fails, and there exists $Y' \nsubseteq Y$ such that $Y' \subseteq C(Y' \cup Y)$. From the definition of C, $Y' \subseteq C(Y' \cup Y) \subseteq C_i((Y' \cup Y)_i|(Y' \cup Y)_{-i}) \cup (Y' \cup Y)_{-i}$ for all $i \in N$; then $(Y' \setminus Y)_i \subseteq Y'_i \subseteq C_i((Y' \cup Y)_i|(Y' \cup Y)_{-i})$ for all $i \in N$, and $Y' \setminus Y$ blocks Y, a contradiction.

($\Leftarrow$) Suppose that conditions 1 and 2 hold. By definition of C, for each $i \in N$, condition 1 implies $Y = C(Y) \subseteq C_i(Y_i|Y_{-i}) \cup Y_{-i} \subseteq Y$, hence $Y_i \subseteq C_i(Y_i|Y_{-i}) \subseteq Y_i$, and Y is individually rational.

Z is unblocked: For any $Z \subseteq X \setminus Y$, condition 2 implies $Z \neq Z \cap C(Z \cup Y) = Z \cap \bigcap_{i \in N} (C_i((Z \cup Y)_i|(Z \cup Y)_{-i}) \cup (Z \cup Y)_{-i})$. Then $Z \cap (C_i((Z \cup Y)_i|(Z \cup Y)_{-i}) \cup (Z \cup Y)_{-i}) \neq Z$ for some $i \in N$; hence $Z_i \cap C_i((Z \cup Y)_i|(Z \cup Y)_{-i}) \cup (Z \cup Y)_{-i}) \neq Z_i$, and thus $Z_i \nsubseteq C_i((Z \cup Y)_i|(Z \cup Y)_{-i})$. Since this requires $Z_i \neq \emptyset$, we can conclude that Z does not block Y. $\square$

Proof of Corollary 1 (Stability and Aggregate Revealed Preference) (If) Suppose Y is undominated in $\succeq_C$. Then for any $Y' \subset Y$, we have $C(Y) = C(Y \cup Y') \neq Y'$. Then we must have $C(Y) = Y$. So if Y is not stable, by Lemma 1, we must have $Y' \subseteq C(Y' \cup Y)$ for some $Y' \nsubseteq Y$. Then $C(Y' \cup Y) \cup Y = Y' \cup Y$ and $C(Y' \cup Y) \neq Y$. Hence $C(C(Y' \cup Y) \cup Y) = C(Y' \cup Y) \neq Y$, and $C(Y' \cup Y) \succ_C Y$, a contradiction.

(Only if) Suppose Y is stable. By Lemma 1 (1), for any $Y' \subset Y$, $C(Y \cup Y') = C(Y) = Y$, and so $Y' \not\succ_C Y$. By Lemma 1 (2), for any $Y' \nsubseteq Y$, $C(Y \cup Y') \nsupseteq Y$; hence, $C(Y \cup Y') \neq Y'$ and $Y' \not\succ_C Y$. $\square$

Lemma 8. *For each $Y \subseteq X$ and $J \subseteq I$, $R_J(Y) = \bigcup_{i \in J} R_i(Y_i|Y_{-i})$.*

Proof. From De Morgan's laws,

$$R_J(Y) \equiv Y \setminus C_J(Y) = Y \setminus \left(\bigcap_{i \in J} (C_i(Y_i|Y_{-i}) \cup Y_{-i}) \right) = \bigcup_{i \in J} Y \setminus (C_i(Y_i|Y_{-i}) \cup Y_{-i}) = \bigcup_{i \in J} R_i(Y_i|Y_{-i}).$$

$\square$

Proof of Lemma 2 (Substitutes and Aggregation) By Lemma 8, $R_J(Y) = \bigcup_{i \in J} R_i(Y_i|Y_{-i})$. Then since the R_i are nondecreasing in Y, so is R_J. $\square$

Proof of Lemma 4 (Stability as a Fixed Point with Local IRC) Suppose (Y, Z) solves (1). Then $Y \cap Z = Z \setminus R_J(Z) = C_J(Z)$. Similarly, $Y \cap Z = Y \setminus R_K(Y) = C_K(Y)$. Since C_J satisfies IRC at Z, $Y \cap Z = C_J(Z) = C_J(Y \cap Z)$; since C_K satisfies IRC at Y, $Y \cap Z = C_K(Y) = C_K(Y \cap Z)$. Since $J \cup K = N$, $C(Y \cap Z) = C_J(Y \cap Z) \cap C_K(Y \cap Z) = Y \cap Z$.

For any $Y' \nsubseteq Y \cap Z$, $C_J((Y' \cap Z) \cup (Y \cap Z)) = Y \cap Z$, since C_J satisfies IRC at Z. Since contracts are substitutes for J, $R_J(Y' \cup (Y \cap Z)) \supseteq R_J((Y' \cap Z) \cup (Y \cap Z)) = (Y' \setminus Y) \cap Z$. Then we have

$$\begin{aligned}
C_J(Y' \cup (Y \cap Z)) &= (Y' \cup (Y \cap Z)) \setminus R_J(Y' \cup (Y \cap Z)), \\
&\subseteq (Y' \cup (Y \cap Z)) \setminus ((Y' \setminus Y) \cap Z), \\
&= ((Y' \setminus Z) \cup ((Y' \cup Y) \cap Z)) \setminus ((Y' \setminus Y) \cap Z), \\
&= (Y' \setminus Z) \cup (Y \cap Z) = (Y' \setminus Z) \cup C_J(Z), \\
&\subseteq (X \setminus Z) \cup C_J(Z) = X \setminus R_J(Z) = Y.
\end{aligned}$$

Likewise, $C_K(Y' \cup (Y \cap Z)) \subseteq Z$. So $C(Y' \cup (Y \cap Z)) = C_J(Y' \cup (Y \cap Z)) \cap C_K(Y' \cup (Y \cap Z)) \subseteq Y \cap Z$, and so $Y' \nsubseteq C(Y' \cup (Y \cap Z))$. Then by Lemma 1, $Y \cap Z$ is stable. $\square$

Proof of Lemma 3 (Stability as a Fixed Point) ((Y, Z) solves (1)$\Rightarrow$ $Y \cap Z$ is stable) Follows from Lemma 4.

(X' is stable $\Rightarrow X' = Y \cap Z$ for some (Y, Z) satisfying (1)) Suppose X' is stable, and let

$$Y = \bigcup_{x \in X} C_J(X' \cup \{x\}), \qquad Z = (X \setminus Y) \cup X'.$$

$X' = Y \cap Z$: By Lemma 1, $X' = C(X') = C_K(X') \cap C_J(X') \subseteq C_J(X') \subseteq X'$, so $C_J(X') = X'$. Then $X' \subseteq Y$, implying $Y \cap Z = X'$.

(Y, Z) solve (1): For all $x \in Y \setminus X'$, $x \in C_J(X' \cup \{x\})$. Since X' is stable, Lemma 1 requires $x \notin C_K(X' \cup \{x\}) \Leftrightarrow x \in R_K(X' \cup \{x\})$. Then since contracts are substitutes for K and $X' \subseteq Y$, $x \in R_K(Y)$. Hence $Y \setminus X' \subseteq R_K(Y)$, implying $C_K(Y) \subseteq X'$. By Lemma 1, since

X' is stable, $X' = C(X') = C_K(X') \cap C_J(X') \subseteq C_K(X') \subseteq X'$, so $C_K(X') = X'$. Then since C_K satisfies IRC and $X' \subseteq Y$, $C_K(Y) = X'$. So $Z = X \setminus R_K(Y)$.

For all $x \in Z \setminus X'$, $x \notin Y$, so $x \notin C_J(X' \cup \{x\})$. Then $x \in R_J(X' \cup \{x\}) \subseteq R_J(Z)$ since contracts are substitutes for J. Hence $Z \setminus X' \subseteq R_J(Z)$, implying $C_J(Z) \subseteq X'$. As noted earlier, $C_J(X') = X'$. Then since C_J satisfies IRC and $X' \subseteq Z$, $C_J(Z) = X'$. So $Y = X \setminus R_J(Z)$. □

Proof of Theorem 1 (Stability with Substitutes and Two Implicit "Sides")

1. Since R_J and R_K are monotone, F is monotone in $\succeq_F$. Part 1 follows from Tarski's fixed point theorem.

2. Since $F(Y, Z) = (Y, Z)$, $Y = F_1(Y, Z) = X \setminus R_J(Z)$ and $Z = F_2(Y, Z) = X \setminus R_K(F_1(Y, Z)) = X \setminus R_K(Y)$. Then since C_J satisfies IRC at Z and C_K satisfies IRC at Y, by Lemma 4, $Y \cap Z$ is stable.

3. If $F(Y, Z) = (Y, Z)$, then $Y = F_1(Y, Z) = X \setminus R_J(Z)$ and $Z = F_2(Y, Z) = X \setminus R_K(F_1(Y, Z)) = X \setminus R_K(Y)$, so by Lemma 3, $Y \cap Z$ is stable. Conversely, if X' is stable, then by Lemma 3, $X' = Y \cap Z$ such that $Y = X \setminus R_J(Z)$ and $Z = X \setminus R_K(Y)$. Then $Y = F_1(Y, Z)$; consequently, $Z = X \setminus R_K(F_1(Y, Z)) = F_2(Y, Z)$ as well.

4. For all $(Y, Z) \in 2^X \times 2^X$, $(X, \emptyset) \succeq_F (Y, Z)$. Then $(X, \emptyset) \succeq_F F(X, \emptyset)$. Since F is monotone in $\succeq_F$, so is $F^n = F \circ \cdots_n \circ F$ for all n. Hence $F^n(X, \emptyset) \succeq_F F^{n+1}(X, \emptyset)$. Let $m = |2^X \times 2^X|$; it follows that $F^m(X, \emptyset) = F^{m+1}(X, \emptyset)$ and so it is a fixed point of F. For any other fixed point (Y, Z) of F, $(X, \emptyset) \succeq_F (Y, Z) \Rightarrow F^m(X, \emptyset) \succeq_F F^m(Y, Z) = (Y, Z)$.

5. Follows symmetrically from part 4. □

Proof of Theorem 2 (Stability: Existence and Uniqueness with Local IRC) $C(X)$ is stable: From IRC at X, we have $C(C(X)) = C(X)$ and for all $X' \nsubseteq C(X)$, $C(C(X) \cup X') = C(X) \nsupseteq X'$. So by Lemma 1, $C(X)$ is stable.

No $X' \neq C(X)$ is stable: If $C(X) \nsubseteq X'$, then from IRC at X, $C(C(X) \cup X') = C(X)$, so by Lemma 1 (2), X' is not stable. If $C(X) \subset X'$, then from IRC at X, $C(X') = C(X) \subset X'$, so by Lemma 1 (1), X' is not stable. □

Proof of Theorem 3 (Stable Outcomes and 3-Cycles) Let $\{y(1), y(2), y(3)\}$ denote the 3-cycle in C_J. For any $Z \subseteq X$, define $m(Z) = 2$ if $y(2) \in Z$, and $m(Z) = 3$ otherwise. For each $k \in K$, let $C_k(Z_k|Z_{-k}) = Z_k \cap \{y(1), y(m(Z))\}$. Then since $Z_K = Z$, $C_K(Z) = Z \cap \{y(1), y(m(Z))\}$.

Contracts are substitutes for K: For $Z \subset Z'$, we must have $m(Z) \geq m(Z')$. Now $R_K(Z) = Z \setminus \{y(1), y(m(Z))\} \subseteq Z' \setminus \{y(1), y(m(Z))\} = R_K(Z')$ if $m(Z) = m(Z')$. If $m(Z) > m(Z')$, then $m(Z) = 3$ and $m(Z') = 2$, so $y(2) \in Z' \setminus Z$ and $R_K(Z) = Z \setminus \{y(1), y(3)\} = Z \setminus \{y(1), y(2), y(3)\} \subseteq Z' \setminus \{y(1), y(2)\} = R_K(Z')$.

If $X^* \nsubseteq Y \equiv \{y(1), y(2), y(3)\}$, then $C_K(X^*) \subseteq X^* \cap Y \neq X^*$. Thus, X^* cannot be individually rational (and hence is not stable).

If $X^* = Y$ then $C_K(X^*) = \{y(1), y(3)\} \neq X^*$. Thus, X^* cannot be individually rational (and hence is not stable).

If $X^* = y(m)$, then $y(m+1 \mod 3)$ blocks X^*, since $y(m+1 \mod 3) \in C_K(\{y(m), y(m+1 \mod 3)\})$ and $y(m+1 \mod 3) = C_J(\{y(m), y(m+1 \mod 3)\})$; hence, X^* is not stable.

If $X^* = \{y(m), y(m-1 \mod 3)\}$, then X^* is not individually rational since $C_J(X^*) = y(m)$; hence, X^* is not stable.

Since the above list of possibilities for X^* is exhaustive, we conclude that no stable outcome exists. $\square$

Lemma 9. *For any $J \subseteq I$ and $Y \subseteq X$, $C_J(Y) \setminus X_J = Y \setminus X_J$.*

Proof. If $x \in Y \setminus X_J$, then $x \in Y_{-j}$ for each $j \in J$, and so $x \in C_J(Y)$. Hence $(Y \setminus X_J) \subseteq C_J(Y)$. Then $(Y \setminus X_J) = C_J(Y) \cap (Y \setminus X_J) = C_J(Y) \setminus X_J$, as desired. $\square$

Proof of Lemma 5 (Structure of Strategically Consistent Assessments)

(i): Since C_i is rational given μ_i, $C_i(Y_i|Y_{-i}) \subseteq \mu_i(Y)$. Since μ_i is correct given $\{C_j\}_{j \in J}$, $\mu_i(Y) = C_{J \setminus \{i\}}(Y)$. Then $C_i(Y_i|Y_{-i}) \subseteq C_{J \setminus \{i\}}(Y)$. Then

$$C_i(Y_i|Y_{-i}) = C_i(Y_i|Y_{-i}) \cap C_{J \setminus \{i\}}(Y) = ((C_i(Y_i|Y_{-i}) \cup Y_{-i}) \cap X_i) \cap C_{J \setminus \{i\}}(Y) = X_i \cap C_J(Y),$$

as desired.

(ii): By (i), $C_J(Y) \cap X_J = \bigcup_{i \in J}(C_J(Y) \cap X_i) = \bigcup_{i \in J} C_i(Y_i|Y_{-i})$. By Lemma 9, $C_J(Y) \setminus X_J = Y \setminus X_J$. The statement follows.

(iii): By (i), $C_j(Y_j|Y_{-j}) = C_J(Y) \cap X_i \subseteq C_J(Y) \subseteq C_i(Y_i|Y_{-i}) \cup Y_{-i}$ for each $j \in J \setminus \{i\}$. Now if $x \in Y_{J \setminus \{i\}}$, then for some $\ell \in J \setminus \{i\}$, we have $x \in Y_\ell$. And if $x \in C_{J \setminus \{i\}}(Y)$, $x \in C_\ell(Y_\ell|Y_{-\ell}) \cup Y_{-\ell}$ for each $\ell \in J \setminus \{i\}$. Thus, if $x \in C_{J \setminus \{i\}}(Y) \cap Y_{J \setminus \{i\}}$, then for some $\ell \in J \setminus \{i\}$, we have $x \in C_\ell(Y_\ell|Y_{-\ell}) \subseteq C_i(Y_i|Y_{-i}) \cup Y_{-i}$. It follows that $C_{J \setminus \{i\}}(Y) \cap Y_{J \setminus \{i\}} \subseteq C_i(Y_i|Y_{-i}) \cup Y_{-i}$. Then since $C_J(Y) = C_{J \setminus \{i\}}(Y) \cap (C_i(Y_i|Y_{-i}) \cup Y_{-i})$, we have $C_{J \setminus \{i\}}(Y) \cap Y_{J \setminus \{i\}} = C_J(Y) \cap Y_{J \setminus \{i\}}$.

By Lemma 9, $C_J(Y) \setminus X_J = Y \setminus X_J$ and $C_{J \setminus \{i\}}(Y) \setminus X_{J \setminus \{i\}} = Y \setminus X_{J \setminus \{i\}}$. Since $Y \setminus X_{J \setminus \{i\}} \supseteq (Y \setminus X_J)$, it follows that $C_{J \setminus \{i\}}(Y) \cap (Y \setminus X_J) = Y \setminus X_J = C_J(Y) \setminus X_J$. Likewise, since $Y \setminus X_{J \setminus \{i\}} \supseteq Y_i \setminus X_{J \setminus \{i\}}$, $C_{J \setminus \{i\}}(Y) \cap (Y_i \setminus X_{J \setminus \{i\}}) = Y_i \setminus X_{J \setminus \{i\}}$.

Then since $Y = Y_J \cup (Y \setminus X_J) = ((Y_i \setminus X_{J\setminus\{i\}}) \cup Y_{J\setminus\{i\}} \cup (Y \setminus Y_J))$, we have

$$\begin{aligned} C_{J\setminus\{i\}}(Y) &= ((Y_i \setminus X_{J\setminus\{i\}}) \cap C_{J\setminus\{i\}}(Y)) \cup ((Y_{J\setminus\{i\}} \cup (Y \setminus X_J)) \cap C_{J\setminus\{i\}}(Y)) \\ &= (Y_i \setminus X_{J\setminus\{i\}}) \cup ((Y_{J\setminus\{i\}} \cup (Y \setminus X_J)) \cap C_J(Y)) \\ &= (Y_i \setminus X_{J\setminus\{i\}}) \cup C_J(Y), \end{aligned}$$

as desired. □

Proof of Proposition 1

(i): (IRC⇒(**??**)) If J is a singleton, then (**??**) is trivially satisfied. Then suppose that $|J| \geq 2$, and that for some Y, Z with $Y \setminus X_J = Z \setminus X_J$, $Y \supseteq Z \supseteq C_j(Y_j|Y_{-j})$ for all $j \in J$. By Lemma 5 (ii), $C_J(Y) = (Y \setminus X_J) \cup \bigcup_{j\in J} C_j(Y_j|Y_{-j}) \subseteq Z$.

Choose $j \neq i$ and consider $\hat{Z} = Z \cup (Y_i \setminus Y_{J\setminus\{j\}})$. Then by Lemma 5 (iii), we have $Y \supseteq \hat{Z} \supseteq C_J(Y) \cup (Y_i \setminus Y_{J\setminus\{j\}}) = \mu_j(Y)$. Then since μ_j satisfies IRC, $\mu_j(Y) = \mu_j(\hat{Z})$. Then by Lemma 5 (iii), $C_J(\hat{Z}) \cap X_{J\setminus\{j\}} = C_J(Y) \cap X_{J\setminus\{j\}}$. Then by Lemma 5 (i), $C_\ell(\hat{Z}_\ell|\hat{Z}_{-\ell}) = C_J(\hat{Z}) \cap X_\ell = C_J(Y) \cap X_\ell = C_\ell(Y_\ell|Y_{-\ell})$ for each $\ell \in J \setminus \{j\}$. Moreover, since C_i is rational given μ_j, $C_j(Y_j|Y_{-j}) = C_j(\hat{Z}_j|\hat{Z}_{-j})$. Then by Lemma 5 (ii), since $\hat{Z} \setminus X_J = Z \setminus X_J$, $C_J(Y) = (Y \setminus X_J) \cup \bigcup_{\ell\in J} C_\ell(Y_\ell|Y_{-\ell}) = (\hat{Z} \setminus X_J) \cup \bigcup_{\ell\in J} C_\ell(\hat{Z}_\ell|\hat{Z}_{-\ell}) = C_J(\hat{Z})$. Then by Lemma 5 (iii), $\mu_i(\hat{Z}) \cap X_{J\setminus\{i\}} = \mu_i(Y) \cap X_{J\setminus\{i\}}$.

Now since $C_J(Y) \subseteq Z$ and $\hat{Z}_i = Z_i$ by definition, we have $\hat{Z} \supseteq Z \supseteq C_J(Y) \cup (\hat{Z}_i \setminus \hat{Z}_{J\setminus\{i\}})$. Since $C_J(Y) = C_J(\hat{Z})$, we have $\hat{Z} \supseteq Z \supseteq C_J(\hat{Z}) \cup (\hat{Z}_i \setminus \hat{Z}_{J\setminus\{i\}}) = \mu_i(\hat{Z})$ by Lemma 5 (iii). Then since μ_i satisfies IRC, $\mu_i(\hat{Z}) = \mu_i(Z)$. It follows that $\mu_i(Z) \cap X_{J\setminus\{i\}} = \mu_i(\hat{Z}) \cap X_{J\setminus\{i\}} = \mu_i(Y) \cap X_{J\setminus\{i\}}$, as desired.

((**??**)⇒IRC) Suppose $Y \supseteq Z \supseteq \mu_i(Y)$. Then by Lemma 5 (iii), $C_J(Y) \cup (Y_i \setminus X_{J\setminus\{i\}}) \subseteq Z$. It follows that $Y_i \setminus X_{J\setminus\{i\}} = Z_i \setminus X_{J\setminus\{i\}}$ (since $Z \subseteq Y$) and $C_j(Y_j|Y_{-j}) = C_J(Y) \cap X_j \subseteq Z$ for each $j \in J$ (by Lemma 5 (i)). Then from (**??**), $C_J(Y) \cap X_{J\setminus\{i\}} = \mu_i(Y) \cap X_{J\setminus\{i\}} = \mu_i(Z) \cap X_{J\setminus\{i\}} = C_J(Z) \cap X_{J\setminus\{i\}}$. From Lemma 9, $C_J(Y) \setminus X_J = Y \setminus X_J$; then since $C_J(Y) \subseteq Z \subseteq Y$, $Y \setminus X_J = Z \setminus X_J$. Then by Lemma 9, $C_J(Y) \setminus X_J = C_J(Z) \setminus X_J$.

Then by Lemma 5 (iii),

$$\begin{aligned} \mu_i(Y) = C_J(Y) \cup (Y_i \setminus X_{J\setminus\{i\}}) &= (C_J(Y) \setminus X_J) \cup (C_J(Y) \cap X_{J\setminus\{i\}}) \cup (Y_i \setminus X_{J\setminus\{i\}}) \\ &= (C_J(Z) \setminus X_J) \cup (C_J(Z) \cap X_{J\setminus\{i\}}) \cup (Z_i \setminus X_{J\setminus\{i\}}) = \mu_i(Z), \end{aligned}$$

as desired.

(ii): Suppose that for $Y, Z \subseteq X_i$ and $X' \subseteq X_{-i}$, $C_i(Y|X') \subseteq Z \subseteq Y$. From Lemma 5 (i), $C_J(Y \cup X') \cap X_i = C_i(Y|X') \subseteq Z$. By definition, $C_J(Y \cup X') \cap X_{-i} \subset (Y \cup X') \cap X_{-i} = X'$. Then $C_J(Y \cup X') \subseteq Z \cup X'$.

Choose $j \in J \setminus \{i\}$. By Lemma 5 (iii), $\mu_j(Y \cup X') = C_J(Y \cup X') \cup ((Y \cup X') \cap (X_j \setminus X_{J\setminus\{j\}}))$. Since $Y \subseteq X_i$, $Y \setminus X_{J\setminus\{j\}} = \emptyset$. Then $\mu_j(Y \cup X') = C_J(Y \cup X') \cup (X' \cap (X_j \setminus X_{J\setminus\{j\}})) \subseteq Z \cup X' \subseteq Y \cup X'$. Since μ_j satisfies IRC, $\mu_j(Y \cup X') = \mu_j(Z \cup X')$.

Then by Lemma 5 (iii), $C_J(Y \cup X') \cap X_{J\setminus\{j\}} = \mu_j(Y \cup X') \cap X_{J\setminus\{j\}} = \mu_j(Z \cup X') \cap X_{J\setminus\{j\}} = C_J(Z \cup X') \cap X_{J\setminus\{j\}}$. Since $i \in J \setminus \{j\}$, by Lemma 5 (i), $C_i(Y|X') = C_J(Y \cup X') \cap X_i = C_J(Z \cup X') \cap X_i = C_i(Z|X')$, as desired. □

Proof of Theorem 4 (Irrelevance of Rejected Contracts in Beliefs and Aggregate Choice) ($\{\mu_i\}_{i\in J}$ satisfy IRC $\Rightarrow$ C_J satisfies IRC) If J is a singleton, the claim follows immediately from Proposition 1 (ii). Suppose $|J| \geq 2$, and that $C_J(Y) \subseteq Z \subseteq Y$. Then $C_J(Y) \setminus X_J \subseteq Z \setminus X_J \subseteq Y \setminus X_J$; since $C_J(Y) \setminus X_J = Y \setminus X_J$ by Lemma 9, $Y \setminus X_J = Z \setminus X_J$.

By Lemma 5 (i), $C_j(Y_j|Y_{-j}) = C_J(Y) \cap X_j \subseteq X \subseteq Y$ for each $j \in J$. By Proposition 1 (i), $\{C_i, \mu_i\}_{i\in J}$ satisfies (**??**). Then for each $i \in J$, $\mu_i(Y) \cap X_{J\setminus\{i\}} = \mu_i(Z) \cap X_{J\setminus\{i\}}$. It follows from (iii) that for each $i \in J$, $C_J(Y) \cap X_{J\setminus\{i\}} = \mu_i(Y) \cap X_{J\setminus\{i\}} = \mu_i(Z) \cap X_{J\setminus\{i\}} = C_J(Z) \cap X_{J\setminus\{i\}}$. Choose distinct $i, j \in J$. Then we have

$$C_J(Y) = (C_J(Y) \cap X_{J\setminus\{i\}}) \cup (C_J(Y) \cap X_{J\setminus\{j\}}) = (C_J(Z) \cap X_{J\setminus\{i\}}) \cup (C_J(Z) \cap X_{J\setminus\{j\}}) = C_J(Z),$$

as desired.

(C_J satisfies IRC $\Rightarrow$ $\{\mu_i\}_{i\in J}$ satisfy IRC) Suppose $\mu_i(Z) \subseteq Y \subseteq Z$. Then by Lemma 5 (iii), $Z_i \setminus X_{J\setminus\{i\}} \subseteq \mu_i(Z) \subseteq Y$; hence, $Z_i \setminus X_{J\setminus\{i\}} = Y_i \setminus X_{J\setminus\{i\}}$. Moreover, by Lemma 5 (iii), $C_J(Z) \subseteq \mu_i(Z) \subseteq Y \subseteq Z$; then since C_J satisfies IRC, $C_J(Z) = C_J(Y)$. Then by Lemma 5 (iii), $\mu_i(Z) = C_J(Z) \cup (Z_i \setminus X_{J\setminus\{i\}}) = C_J(Y) \cup (Y_i \setminus X_{J\setminus\{i\}}) = \mu_i(Y)$, as desired. □

Proof of Lemma 6 (Relationship Between Myopic and Strategic Consistency) Since correct beliefs are part of the definition of strategic consistency, the proof amounts to showing that item 1 is equivalent to rationality of C_i given μ_i. Since $\hat{C}_i$ is myopically consistent, by (2) we have $\hat{C}_i(\mu_i(Y)_i|\mu_i(Y)_{-i}) = \arg\max_{S\subseteq\mu_i(Y)_i} u_i(S \cup \mu_i(Y)_{-i})$. It follows immediately that $C_i(Y_i|Y_{-i}) = \hat{C}_i(\mu_i(Y)_i|\mu_i(Y)_{-i})$ for each $Y \subset X \Leftrightarrow C_i$ is rational given μ_i. □

Proof of Lemma 7 Since the market is two-sided, for each $i \in J$ and $y \in X_i$, $y \in X_{-j}$ for any $j \in J$ such that $j \neq i$. Then for any $Y \subseteq X$ and $i \in L$,

$$C_L(Y) \cap X_i = (C_i(Y_i|Y_{-i}) \cup Y_{-i}) \cap \bigcap_{j\in L, j\neq i} (C_j(Y_j|Y_{-j}) \cup Y_{-j}) \cap X_i = C_i(Y_i|Y_{-i}).$$

And for any $i \in J \setminus L$,

$$C_L(Y) \cap X_i = \bigcap_{j\in L} (C_j(Y_j|Y_{-j}) \cup Y_{-j}) \cap X_i = \bigcap_{j\in L} Y_{-j} \cap X_i = Y_i.$$

Moreover, since the market is two-sided, $\bigcup_{i\in J} X_i = X$. Then

$$C_L(Y) = \bigcup_{i\in J} C_L(Y)\cap X_i = \left(\bigcup_{i\in L} C_i(Y_i|Y_{-i})\right) \cup \left(\bigcup_{i\in J\setminus L} Y_i\right),$$

as desired. The statement for C_K follows identically.

Lemma 10. *For any $Y \subseteq X$ and $i \in I$, $\bigcup_{j\neq i} Y_j = Y$ and $\bigcap_{j\neq i} Y_{-j} = \emptyset$.*

Proof. For each $x \in X$, we have $|N(x)| \geq 2$ by assumption, and so $j \in N(x)$ for some $j \neq i$. Then $\bigcup_{j\neq i} Y_j = Y$. Hence $\bigcap_{j\neq i} Y_{-j} = \bigcap_{j\neq i} Y \setminus Y_j = Y \setminus \left(\bigcup_{j\neq i} Y_j\right) = \emptyset$. □

Proof of Theorem 5 (Stability with Strategic Consistency for All Agents) (1): Suppose Y is stable for the strategically consistent assessment $\{C_i, \mu_i\}_{i\in I}$. By Lemma 1, $C(Y) = Y$. Then by Lemma 5 (iii), $\mu_i(Y) = C(Y) \cup (Y_i \setminus X_{I\setminus\{i\}}) = Y$ for each $i \in I$. The claim then follows from Lemma 6.

(2): Define $\{C_i, \mu_i\}_{i\in I}$ as follows: For each $i \in I$,

$$\mu_i(Z) = \begin{cases} Y, & Z \supseteq Y, \\ \emptyset, & Z \not\supseteq Y; \end{cases} \qquad\qquad C_i(Z_i|Z_{-i}) = \begin{cases} Y_i, & Z \supseteq Y, \\ \emptyset, & Z \not\supseteq Y. \end{cases}$$

For each $i \in I$ and $Z \subseteq X$, if $Z \not\supseteq Y$, we have $C_{I\setminus\{i\}}(Z) = \bigcap_{j\neq i}(\emptyset \cup Z_{-j}) = \emptyset = \mu_i(Z)$ by Lemma 10. Alternatively, if $Z \supseteq Y$, we have $C_{I\setminus\{i\}}(Z) = \bigcap_{j\neq i}(Y_j \cup Z_{-j}) = Y \cup \left(\bigcap_{j\neq i}(Z\setminus Y)_j\right)$. By Lemma 10, $\bigcap_{j\neq i}(Z \setminus Y)_j = \emptyset$, and so $C_{I\setminus\{i\}}(Z) = Y = \mu_i(Z)$. It follows that $\{\mu_i\}_{i\in I}$ are correct given $\{C_i\}_{i\in I}$.

Since $\{\hat{C}_i\}_{i\in I}$ are myopically consistent, $\hat{C}_i(\emptyset|\emptyset) = \emptyset$ for each $i \in I$. Then since $\hat{C}_i(Y_i|Y_{-i}) = Y_i$ for each $i \in I$, we have $\hat{C}_i(\mu_i(Z)_i|\mu_i(Z)_{-i}) = C_i(Z_i|Z_{-i})$ for each $i \in I$. Then by Lemma 6, $\{C_i, \mu_i\}_{i\in I}$ is strategically consistent.

Moreover, $\{\mu_i\}_{i\in I}$ satisfy IRC: Suppose $\mu_i(Z) \subseteq Z' \subseteq Z$ for some $i \in I$ and $Z', Z \subseteq X$. If $Z' \supseteq Y$, then $Z \supseteq Y$ as well, and we have $\mu_i(Z') = Y = \mu_i(Z)$. Alternatively, if $Z' \not\supseteq Y$, then $Y \neq \mu_i(Z)$, and so $Z \not\supseteq Y$ as well; hence $\mu_i(Z') = \emptyset = \mu_i(Z)$.

Then by Theorem 4, C satisfies IRC. Then since $C(X) = Y$, by Theorem 2, Y is stable for $\{C_i, \mu_i\}_{i\in I}$. □

Proof of Theorem 6 (Stability with Weak Forward Induction) (Only if) Suppose that Y is stable for some strategically consistent assessment $\{C_i, \mu_i\}_{i\in I}$ with beliefs $\{\mu_i\}_{i\in I}$ satisfying IRC and WFI. (1) then follows from Theorem 5. For (2), suppose that there exists $Y' \supset Y$ such that $\hat{C}_i(Y'_i|Y'_{-i}) = Y'_i$ for each $i \in I$. Since Y is stable, by Lemma 1, $C(Y') \neq Y'$. Since $\{\mu_i\}_{i\in I}$ satisfy WFI, $\mu_i(Y') = Y'$ for each $i \in I$. Then by Lemma 6, for each $i \in I$,

$C_i(Y_i'|Y_{-i}') = \hat{C}_i(\mu_i(Y')_i|\mu_i(Y')_{-i}) = \hat{C}_i(Y_i'|Y_{-i}') = Y_i'$, and hence $C_i(Y_i|Y_{-i}) \cup Y_{-i} = Y$. Then $C(Y') = Y'$, a contradiction.

(If) Let $\mathcal{M} = \{Z \subseteq X | \hat{C}_i(Z_i|Z_{-i}) = Z_i \text{ for each } i \in I\}$, and label its elements according to the sequence $\{Y^n\}_{n=1}^{|\mathcal{M}|}$, constructed inductively as follows: For the initial element, choose $Y^1 = Y$. Then, given elements $\{Y^n\}_{n=1}^m$, choose $Y^{m+1} \in \mathcal{M} \setminus \{Y^n\}_{n=1}^m$ such that there is no $Y' \in \mathcal{M} \setminus \{Y^n\}_{n=1}^m$ with $Y' \supset Y^{m+1}$.

Since $\{\hat{C}_i\}_{i \in I}$ are myopically consistent, we have $\hat{C}_i(\emptyset|\emptyset) = \emptyset$ for each $i \in I$, and so $\emptyset \in \mathcal{M}$. Then for each $Z \subseteq X$, 2^Z contains at least one element of $\mathcal{M}$. Hence, for each $Z \subseteq X$, we can define $n^*(Z)$ as the earliest element of the sequence $\{Y^n\}_{n=1}^{|\mathcal{M}|}$ that is a subset of Z: $n^*(Z) \equiv \min\{n | Y^n \in 2^Z\}$. Then define $\{C_i, \mu_i\}_{i \in I}$ as follows:

$$\mu_i(Z) = Y^{n^*(Z)}, \qquad\qquad C_i(Z_i|Z_{-i}) = Y_i^{n^*(Z)}.$$

$\{\mu_i\}_{i \in I}$ are correct given $\{C_i\}_{i \in I}$: For each $Z \subseteq X$ and $i \in I$, we have $C_{I \setminus \{i\}}(Z) = \bigcap_{j \neq i}(Y_j^{n^*(Z)} \cup Z_{-j}) = Y^{n^*(Z)} \cup \left(\bigcap_{j \neq i}(Z \setminus Y^{n^*(Z)})_j\right)$. By Lemma 10, $\bigcap_{j \neq i}(Z \setminus Y^{n^*(Z)})_j = \emptyset$; hence $C_{I \setminus \{i\}}(Z) = Y^{n^*(Z)} = \mu_i(Z)$, as desired.

$\{C_i, \mu_i\}_{i \in I}$ is strategically consistent: For each $Z \subseteq X$ and $i \in I$, $\mu_i(Z) = Y^{n^*(Z)} \in \mathcal{M}$, and so $\hat{C}_i(\mu_i(Z)_i|\mu_i(Z)_{-i}) = Y_i^{n^*(Z)} = C_i(Z_i|Z_{-i})$. Since $\{\mu_i\}_{i \in I}$ are correct given $\{C_i\}_{i \in I}$, strategic consistency follows from Lemma 6.

$\{\mu_i\}_{i \in I}$ satisfy IRC: Suppose $\mu_i(Z) \subseteq Z' \subseteq Z$ for some $i \in I$ and $Z', Z \subseteq X$. Since $Z' \subseteq Z$, we have $2^{Z'} \subseteq 2^Z$, so $\{n | Y^n \in 2^{Z'}\} \subseteq \{n | Y^n \in 2^Z\}$, and hence $n^*(Z) \leq n^*(Z')$. Since $\mu_i(Z) = Y^{n^*(Z)} \subseteq Z'$, we have $n^*(Z) \in \{n | Y^n \in 2^{Z'}\}$, and hence $n^*(Z') \leq n^*(Z)$. Then $\mu_i(Z') = Y^{n^*(Z')} = Y^{n^*(Z)} = \mu_i(Z)$, as desired.

$\{\mu_i\}_{i \in I}$ satisfy WFI: By construction, for any $Y^n, Y^m \in \mathcal{M}$, if $Y^n \subset Y^m$, then $n > m$. Hence, for any $Z \in \mathcal{M}$, we have $Y^{n^*(Z)} = Z$, and so $\mu_i(Z) = Z$ for each $i \in I$. By construction, $Z \in \mathcal{M}$ whenever $\hat{C}_i(Z_i|Z_{-i}) = Z_i$ for each $i \in I$; it follows that $\{\mu_i\}_{i \in I}$ satisfy WFI. $\square$

Proof of Corollary 2 (Stability with Myopic Consistency vs. Strategic Consistency and WFI) Let $\{\hat{C}_i\}_{i \in I}$ be myopically consistent, and suppose Y is stable for $\{\hat{C}_i\}_{i \in I}$. Then by definition, it is myopically individually rational: $\hat{C}_i(Y_i|Y_{-i}) = Y_i$ for each $i \in I$. Moreover, there is no $Y' \supset Y$ such that $\hat{C}_i(Y_i'|Y_{-i}') = Y_i'$ for each $i \in I$: Suppose not, and there exists such a Y'. Then for all $i \in N(Y' \setminus Y)$, $Y_i' \setminus Y_i \subseteq Y_i' \subseteq \hat{C}_i(Y_i'|Y_{-i}')$, a contradiction since Y is stable (and therefore unblocked) for $\{\hat{C}_i\}_{i \in I}$. It follows from Theorem 6 that Y is stable for some strategically consistent assessment $\{C_i, \mu_i\}_{i \in I}$ such that $\{\mu_i\}_{i \in I}$ satisfies WFI and IRC.

Proof of Theorem 7 (Monotone Externalities as a Foundation For Strategic Consistency)

Claim 1. $G_Y(Z) \equiv \bigcup_{i\in J} \hat{C}_i(Y_i|Z_{-i})$ **has a** $\succeq_J$**-minimal fixed point** Y^***.** For any $Y \subseteq X$, define the sequence $\{\hat{Y}^t\}$ recursively as follows:

$$\hat{Y}^0 = \emptyset, \qquad \hat{Y}^{t+1} = \bigcup_{i\in J} \hat{C}_i(Y_i|\hat{Y}^t_{-i})$$

Let $\succeq_J$ be the consistent preorder for which $\{\hat{C}_i\}_{i\in J}$ satisfy monotone externalities. We show that $\{\hat{Y}^t\}$ is $\succeq$-increasing. For $t = 0$, since $\succeq_J$ is consistent with $\{\hat{C}_i\}_{i\in J}$, we have

$$\hat{Y}^1 = \bigcup_{i\in J} \hat{C}_i(Y_i|\emptyset) \succeq_J \bigcup_{i\in J} \hat{C}_i(\emptyset|\emptyset) = \emptyset = \hat{Y}^0.$$

Now for $t \geq 1$, suppose $\hat{Y}^t \succeq_J \hat{Y}^{t-1}$. Then since $\succeq_J$ is consistent with $\{\hat{C}_i\}_{i\in J}$,

$$\hat{Y}^{t+1} = \bigcup_{i\in J} \hat{C}_i(Y_i|\hat{Y}^t_{-i}) \succeq_J \bigcup_{i\in J} \hat{C}_i(Y_i|\hat{Y}^{t-1}_{-i}) = \hat{Y}^t.$$

It follows by induction that $\hat{Y}^{t+1} \succeq_J \hat{Y}^t$ for each n. Let $T = |2^X| = 2^{|X|}$. Since $\succeq_J$ is a preorder, it is transitive. Then we must have $\hat{Y}^T \simeq \hat{Y}^t$ for each $t > T$. By monotone externalities, we have $R_i(Y_i|\hat{Y}^T_{-i}) \subseteq R_i(Y_i|\hat{Y}^{T+1}_{-i})$ and $R_i(Y_i|\hat{Y}^T_{-i}) \supseteq R_i(Y_i|\hat{Y}^{T+1}_{-i})$ for each $i \in J$. Then $\hat{C}_i(Y_i|\hat{Y}^T_{-i}) = \hat{C}_i(Y_i|\hat{Y}^{T+1}_{-i})$ for each $i \in J$. Then $\hat{Y}^{T+1} = \bigcup_{i\in J} \hat{C}_i(Y_i|\hat{Y}^T_{-i}) = \bigcup_{i\in J} \hat{C}_i(Y_i|\hat{Y}^{T+1}_{-i}) = \hat{Y}^{T+2}$. Thus $Y^* \equiv \hat{Y}^{T+1}$ is a fixed point of G_Y.

Now suppose there is some other fixed point $Y' \neq Y^*$ of G_Y such that $Y' \preceq_J Y^*$. Since $\succeq_J$ is consistent with $\{\hat{C}_i\}_{i\in J}$,

$$Y' = \bigcup_{i\in J} \hat{C}_i(Y_i|Y'_{-i}) \succeq_J \bigcup_{i\in J} \hat{C}_i(\emptyset|Y'_{-i}) = \emptyset = \hat{Y}^0.$$

Now for $t \geq 1$, suppose $Y' \succeq_J \hat{Y}^{t-1}$. Then since $\succeq_J$ is consistent with $\{\hat{C}_i\}_{i\in J}$,

$$Y' = \bigcup_{i\in J} \hat{C}_i(Y_i|Y'_{-i}) \succeq_J \bigcup_{i\in J} \hat{C}_i(Y_i|\hat{Y}^{t-1}_{-i}) = \hat{Y}^t.$$

It follows by induction that $Y' \succeq_J \hat{Y}^{T+1} = Y^*$. Then since $\{\hat{C}_i\}_{i\in J}$ satisfy monotone externalities, we have $R_i(Y_i|Y'_{-i}) \subseteq R_i(Y_i|Y^*_{-i})$ and $R_i(Y_i|Y'_{-i}) \subseteq R_i(Y_i|Y^*_{-i})$ for each $i \in J$. Then $\hat{C}_i(Y_i|Y'_{-i}) = \hat{C}_i(Y_i|Y^*_{-i})$ for each $i \in J$. Then $Y' = \bigcup_{i\in J} \hat{C}_i(Y_i|Y'_{-i}) = \bigcup_{i\in J} \hat{C}_i(Y_i|Y^*_{-i}) = Y^*$, a contradiction.

Claim 2. $Y^* \preceq_J Z^*$ **for each** $Y \subseteq Z$**.** By definition, $\hat{Z}^0 = \hat{Y}^0 = \emptyset$. Now for $t \geq 1$, suppose

$\hat{Z}^{t-1} \succeq_J \hat{Y}^{t-1}$. Since $\succeq_J$ is consistent with $\{\hat{C}_i\}_{i\in J}$,

$$\hat{Z}^t = \bigcup_{i\in J} \hat{C}_i(Z_i|\hat{Z}^{t-1}_{-i}) \succeq_J \bigcup_{i\in J} \hat{C}_i(Y_i|\hat{Y}^{t-1}_{-i}) = \hat{Y}^t.$$

It follows by induction that $Y^* = \hat{Y}^{T+1} \preceq_J \hat{Z}^{T+1} = Z^*$.

Claim 3. If $Z^* \subseteq Y \subseteq Z$, then $Y^* \supseteq Z^*$. By Claim 2, we have $Y^* \preceq_J Z^*$. Since $\{\hat{C}_i\}_{i\in J}$ satisfy monotone externalities, for each $i \in J$, $\hat{R}_i(Y_i|Y^*_{-i}) \subseteq \hat{R}_i(Y_i|Z^*_{-i})$, or equivalently, $\hat{C}_i(Y_i|Y^*_{-i}) \supseteq \hat{C}_i(Y_i|Z^*_{-i})$. And since $\{\hat{C}_i\}_{i\in J}$ satisfy irrelevance of rejected contracts, $\hat{C}_i(Y_i|Z^*_{-i}) = \hat{C}_i(Z_i|Z^*_{-i})$ for each $i \in J$. Then

$$Y^* = \bigcup_{i\in J} \hat{C}_i(Y_i|Y^*_{-i}) \supseteq \bigcup_{i\in J} \hat{C}_i(Y_i|Z^*_{-i}) = \bigcup_{i\in J} \hat{C}_i(Z_i|Z^*_{-i}) = Z^*.$$

Claim 4. If $Z^* \subseteq Y \subseteq Z$, then $Y^* = Z^*$. By definition, we have $(Z^*)^* = \bigcup_{i\in J} \hat{C}_i(Z^*_i|(Z^*)^*_{-i}) \subseteq Z^*$. And by Claim 3, we have $(Z^*)^* \supseteq Z^*$. Then $(Z^*)^* = Z^*$.

Now by Claim 2, we have $Z^* = (Z^*)^* \preceq_J Y^* \preceq_J Z^*$. Then since $\{\hat{C}_i\}_{i\in J}$ satisfy monotone externalities, for each $i \in J$, $\hat{R}_i(Y_i|Y^*_{-i}) \subseteq \hat{R}_i(Y_i|Z^*_{-i}) \subseteq \hat{R}_i(Y_i|Y^*_{-i})$, or equivalently, $\hat{C}_i(Y_i|Y^*_{-i}) = \hat{C}_i(Y_i|Z^*_{-i})$. And since $\{\hat{C}_i\}_{i\in J}$ satisfy irrelevance of rejected contracts, $\hat{C}_i(Y_i|Z^*_{-i}) = \hat{C}_i(Z_i|Z^*_{-i})$ for each $i \in J$. Then we have

$$Y^* = \bigcup_{i\in J} \hat{C}_i(Y_i|Y^*_{-i}) = \bigcup_{i\in J} \hat{C}_i(Y_i|Z^*_{-i}) = \bigcup_{i\in J} \hat{C}_i(Z_i|Z^*_{-i}) = Z^*.$$

Specification of $\{C_i, \mu_i\}_{i\in J}$. For each $i \in J$ and $Y \subseteq X$, let

$$C_i(Y_i|Y_{-i}) = Y^*_i, \qquad \mu_i(Y) = Y^* \cup Y_i.$$

Claim 5. For each $Y \subseteq X$ and $i \in J$, $Y^*_i = \hat{C}_i(Y_i|Y^*_{-i})$. Since the market is two-sided, X_i and X_j are disjoint for each $i, j \in J$ with $i \neq j$. Since Y^* is a fixed point of G_Y, $Y^* = \bigcup_{i\in J} \hat{C}_i(Y_i|Y^*_{-i})$. Then $\hat{C}_i(Y_i|Y^*_{-i}) = X_i \cap \left(\bigcup_{i\in J} \hat{C}_i(Y_i|Y^*_{-i})\right) = X_i \cap Y^* = Y^*_i$.

Claim 6. $\{C_i, \mu_i\}_{i\in J}$ is strategically consistent. By Lemma 7, for each $i \in J$,

$$C_{J\setminus\{i\}}(Y) = Y_i \cup \left(\bigcup_{j\in J\setminus\{i\}} C_j(Y_j|Y_{-j})\right) = Y_i \cup Y^*_{J\setminus\{i\}} = Y_i \cup Y^*.$$

Then for each $i \in J$, μ_i is correct given $\{C_j\}_{j\in J}$. By Claim 5, for each $Y \subseteq X$ and $i \in J$,

$$Y^*_i = \hat{C}_i(Y_i|Y^*_{-i}) = \hat{C}_i(\mu_i(Y)_i|\mu_i(Y)_{-i}).$$

Then by Lemma 6, $\{C_i, \mu_i\}_{i \in J}$ is strategically consistent.

Claim 7. **$\{\mu_i\}_{i\in J}$ satisfy IRC.** Suppose $\mu_i(Z) \subseteq Y \subseteq Z$. Since $\mu_i(Z) = Z^* \cup Z_i$, we must have $Z_i \subseteq Y_i \subseteq Z_i$, and so $Z_i = Y_i$. Further, we must have $Z^* \subseteq Y \subseteq Z$, and so by Claim 4, $Y^* = Z^*$. It follows that $\mu_i(Z) = Z^* \cup Z_i = Y^* \cup Y_i = \mu_i(Y)$. The claim follows.

Claim 8. **$\{C_i\}_{i\in J}$ satisfy substitutability.** Suppose $Y \subseteq Z$. From Claim 2, $Y^* \preceq_J Z^*$. Then since $\{\hat{C}_i\}_{i\in J}$ satisfy standard substitutes and monotone externalities,
$\hat{R}_i(Y_i|Y^*_{-i}) \subseteq \hat{R}_i(Z_i|Z^*_{-i})$ for each $i \in J$. Then for each $i \in J$,

$$\hat{C}_i(Y_i|Y^*_{-i}) = Y_i \setminus \hat{R}_i(Y_i|Y^*_{-i}) \supseteq Y_i \setminus \hat{R}_i(Z_i|Z^*_{-i}) = \hat{C}_i(Z_i|Z^*_{-i}) \cap Y_i.$$

Then by Claim 5,

$$C_i(Y_i|Y_{-i}) = Y_i^* = \hat{C}_i(Y_i|Y^*_{-i}) \supseteq \hat{C}_i(Z_i|Z^*_{-i}) \cap Y_i = Z_i^* \cap Y_i = C_i(Z_i|Z_{-i}) \cap Y_i.$$

Hence,

$$R_i(Y_i|Y_{-i}) = Y_i \setminus C_i(Y_i|Y_{-i}) \subseteq Y_i \setminus C_i(Z_i|Z_{-i}) \subseteq Z_i \setminus C_i(Z_i|Z_{-i}) = R_i(Z_i|Z_{-i}),$$

as desired. $\square$