Group Formation with Fixed Group Size: Complementarity vs Substitutability

Christian Ahlin*

April 2012

Abstract

One-sided matching under transferable utility is well understood when groups of size two are being formed. In this setting, complementarity or substitutability of types in the group payoff function pins down key features of the equilibrium matching pattern, without it being necessary to know the distribution of types or specifics of the payoff function. This note explores equilibrium matching patterns when groups of fixed size $n \geq 2$ are being formed. Complementarity of types continues to pin down a unique equilibrium grouping. Substitutability of types rules out much less. We show that it rules out groupings in one which one group rank-wise dominates the other. Conversely, we show in a simple setting that any grouping not exhibiting rankwise dominance can be the unique equilibrium. Thus, "no rank-wise dominance" is all that can be said generically about matching under substitutability. In the case of $2n$ agents, "no rank-wise dominance" leaves a unique grouping when $n = 2$ but rules out a vanishingly small fraction of groupings ($\frac{2}{n+1}$) as n increases. In general, the results suggest that substitutability by itself has much less predictive power than complementarity, when group size is bigger than two.

*I thank Ashley Ahlin, Pierre-Andre Chiappori, Vivek Dhand, Patrick Legros, Andrew Newman, Debraj Ray, Katerina Sherstyuk, and Myrna Wooders for valuable input. All errors are mine.

1 Introduction

Models of group formation as frictionless, one-sided matching under transferable utility have been used in various contexts to shed light on economic questions. Examples include skill matching and underdevelopment (Kremer, 1993); skill diversity and trade patterns (Grossman and Maggi, 2000); and micro-credit group formation (Ghatak, 1999, 2000). Some stark results on matching have emerged from this and related literature (see especially Legros and Newman, 2002). If the group payoff function exhibits type complementarity (or supermodularity), the unique matching equilibrium involves perfectly homogeneous matching (or "segregation"): groups are made up of homogeneous types. If the group payoff function exhibits type substitutability (or submodularity), the unique matching equilibrium involves a specific form of heterogeneous ("onion-style" or "median") matching: each agent matches with a type from the complementary percentile, e.g. a 95th percentile type with a 5st percentile type, and so on. Evidently, complementarity and substitutability pin down key aspects of the matching pattern for any distribution of types and payoff function specifics.

However, most of the literature has stayed within the confines of two-person matches – typically for tractability reasons.[1] While valuable insights have been obtained with this simplification, it would be helpful to know how the matching results generalize to groups of size larger than two. This is especially so because size-two groups are often counterfactual – e.g., firms often have more than two workers, and micro-credit groups typically have more than two members.

This paper generalizes these matching results to groups of fixed size $n \geq 2$. The gener-

[1] There are a few significant exceptions for homogeneous group formation. Legros and Newman (2002) state that several of their results on homogeneous matching generalize to larger group size. In the case of n-sided matching, the classic positive assortative matching result for (2-sided) marriage of Becker (1973) has been generalized by Lin (1992) and Sherstyuk (1999).

For substitutability, the lone exception we are aware of is Saint-Paul (2001), who characterizes matching patterns under complementarity, substitutability, and hybrid cases, with fixed-*measure* group size. However, his paper differs from both this paper and the other literature referenced here in that he examines groups with a continuum of members, while here the focus is on finite-member groups. As a result, his framework does not deliver onion-style (median) matching results for any size of groups under substitutability. See conclusion for more discussion.

alization for complementarity is easy to foresee: perfect homogeneity obtains regardless of group-size n. Less straightforward is how heterogeneous, onion-style matching generalizes. If n is even, onion-style matching could occur with $n/2$ agents of one type matching with $n/2$ types from the complementary percentile. But if n is odd, 3 for example, and two agents match with types at percentiles c and $1-c$, respectively, then which type joins them, a type from c, $1-c$, or somewhere else entirely?

We find that complementarity does indeed affect matching in the predictable way for groups of size $n > 2$. For substitutability, however, it is not onion-style (or median) matching that generalizes whether n is odd or even. We show that what substitutability rules out is matching patterns in which any group **rank-wise dominates** another, i.e. has higher types at each rank. We also show by analyzing a particular setting that rank-wise dominance is all that substitutability rules out in general. However, rank-wise dominance can be relatively rare. Consider the simple case of $2n$ agents forming two groups of size n: as group size gets large, a vanishingly small fraction of groupings exhibit rank-wise dominance and can thus be ruled out by substitutability alone. An asymmetry between the predictions of complementarity and substitutability is thus uncovered: complementarity pins down equilibrium groupings for any group size, while substitutability rules out much less for larger groups.

The implication for theoretical work is that to pin down matching patterns under substitutability, typically assumptions will need to be made on the distribution of types and/or the specifics of the group payoff function; while under complementarity, no further assumptions are needed. The implication for empirical work is that, since it is potentially compatible with many different groupings, substitutability per se is likely to be much harder to rule out empirically than complementarity. Structural estimation that uses an explicit payoff function and information on types may be necessary.

The baseline model and results under complementarity are in section 2, while the results on substitutability are in section 3. Section 4 provides some discussion and concluding remarks. All proofs are in the Appendix.

2 Model and Complementarity Results

A set of agents match to form n-person groups to produce, $n \geq 2$. Examples include workers sorting into firms, or into production teams within firms; firms forming alliances; and households forming microcredit groups to obtain joint liability loans, or households forming groups to share risk.

Each agent has a type $p \in \mathcal{P}$, where $\mathcal{P}$ is a bounded subset of $\Re$. Type could capture human capital or ability, firm size or reputation, or risk or return of household income streams or of projects needing funding. Two cases will be considered, 1) a $\mathcal{P}$ with a finite number of elements; and 2) a convex $\mathcal{P}$ where each type in $\mathcal{P}$ has positive density. As in Legros and Newman (2002), where there is a continuum of agents, we assume there is a continuum of agents of each type. On the other hand, if the number of players is finite, we assume the total number of players is divisible by n, so all players can be in groups.

The joint payoff function of n agents with types $(p_1, p_2, ..., p_n)$ is $\Pi(p_1, p_2, ..., p_n)$. Π is assumed twice continuously differentiable and symmetric.[2] Agent utility is assumed transferable, so n agents in a group are able to split their joint payoff in any way.

In this context, a **grouping** is an assignment of each agent to a group, such that each group has exactly n agents and the number or measure of each type of agent across groups is consistent with the number or measure of agents of each type. An equilibrium (core) grouping is a grouping for which individual payoffs exist that a) are feasible, meaning the sum of individual payoffs in a group equals the group payoff; and b) cannot be blocked by any n agents reorganizing into a group so that each achieves a higher payoff.[3]

Let $\mathcal{N} = \{1, 2, \ldots, n\}$ and $\mathcal{N}_{-i} = \mathcal{N} \setminus i$. Let $G = (p_1^G, p_2^G, \ldots, p_n^G) \in \mathcal{P}^n$ denote the set of types of a group of agents; and let $\Pi_G \equiv \Pi(p_1^G, p_2^G, \ldots, p_n^G)$. Define $\underline{\mathbf{p}}^{\mathbf{G}}$ ($\overline{\mathbf{p}}^{\mathbf{G}}$) as the minimum (maximum) type in group G: $\underline{p}^G = \min_{i \in \mathcal{N}} p_i^G$ and $\overline{p}^G = \max_{i \in \mathcal{N}} p_i^G$. Types are said to be

[2]In the context of two-person matching, Legros and Newman (2002) explore more general conditions on the joint production function that do not require differentiability or symmetry. For our results (except Proposition 5), differentiability is not needed, only super/submodularity.

[3]We use the same basic setup as Legros and Newman (2002), who also draw on Kaneko and Wooders (1986); see their papers for more details.

complements (**substitutes**) if the group payoff function exhibits strictly positive (strictly negative) cross-partials – $\partial^2\Pi/\partial p_j \partial p_k$, for all $j, k \in \mathcal{N}$, $j \neq k$ – everywhere on its domain.

Proposition 1. *Assume types are complements. Then, groups are rank-ordered (non-overlapping) in any equilibrium grouping. That is, if L and M are two groups in an equilibrium grouping, then $\overline{p}^L \leq \underline{p}^M$ or $\overline{p}^M \leq \underline{p}^L$.*

This result holds whether there is a continuum or a finite number of agents. In the finite case, it implies that equilibrium group formation is unique and simple: the highest n types in the first group, the next highest n in the next group, and so on.[4] This is the result of supermodularity (implied by complementarity) of the group payoff function, which implies that payoffs can always be raised by taking the maximum n types and the minimum n types from any two groups. Continually applying this fact leaves the rank-ordered grouping as uniquely efficient; and any equilibrium grouping must be efficient.

When there is a continuum of agents, rank-ordering is squeezed to perfect intra-group homogeneity, or "segregation". To prove existence, we first establish the following lemma.

Lemma 1. *If $f : \Re^n \to \Re$ is symmetric and supermodular, then for any $(x_1, x_2, ..., x_n) \in \Re^n$,*

$$f(x_1, x_2, ..., x_n) \leq \frac{\sum_{k=1}^{n} f(x_k, ..., x_k)}{n}.$$

Proposition 2. *Assume types are complements. If there is a continuum of agents with types drawn from a convex set such that every type has positive density, there exists a unique equilibrium grouping. Every group is homogeneous in this equilibrium.*

This result is familiar, in the sense of being no different qualitatively whether $n = 2$ or $n > 2$. It is essentially a Folk Theorem that we include here for completeness, though the

[4] More precisely, in the finite case this is the unique equilibrium *candidate*, in the sense of being the unique *efficient* grouping; but an *equilibrium* grouping may not exist. While supermodularity guarantees existence in the n-sided matching case (Proposition 2 of Sherstyuk, 1999), it is not sufficient here. However, while we have examples of non-existence, we also have existence proofs for some special cases. Thus in general, existence in the finite case depends on specifics of the set of agent types and the payoff function.

existence proof is new and relies explicitly on supermodularity.[5]

The results are also strong. Existence and uniqueness obtain, and homogeneous matching is the rule regardless of group size (n), and independent of the type distribution and details of the production function.

3 Substitutability Results

Grossman and Maggi (2000) show that in the $n = 2$ case when types are substitutes, onion-style matching occurs. Every group has one member above and one member below the median, both equidistant from it in percentile terms. Here we explore which restrictions from the $n = 2$ case generalize to larger group sizes.

Say that group L **rank-wise dominates** group M if L dominates M at every rank, that is, the jth largest type in group L is greater than the jth largest type in group M for all $j \in \mathcal{N}$.[6] Define two groups L and M as **nearly rank-wise identical** if there exists an $i \in \mathcal{N}$ such that the jth largest elements of L and M, respectively, are equal for all $j \in \mathcal{N}_{-i}$. That is, L and M are nearly rank-wise identical if at $n-1$ (or n) ranks they have the same value.[7]

Proposition 3. *Assume types are substitutes. In any equilibrium grouping, one group rank-wise dominates another only if the two groups are nearly rank-wise identical.*

This appears to be the main implication of substitutability of risk-types in higher dimensions: groups are intertwined, in that each dominates the other at some rank. The only possible exception occurs when they are identical, at least at all but one rank. The intuition is essentially the reverse of that for Proposition 1. Submodularity (implied by substitutability) and symmetry imply that any two groups that can be (permuted so as to be) written

[5] Formal equilibrium characterizations of one-sided matching exist for $n = 2$; we have not found explicit characterizations for $n > 2$, though existing results are stated to be easily extendable, e.g. Legros and Newman (2002) (see also Lin, 1992). The results of Kaneko and Wooders (1996) also establish equilibrium existence in this context, at least of an "approximately feasible" core.

[6] This is equivalent to first-order stochastic dominance of one group's distribution of types over the other's.

[7] So $(1, 2, 5)$ is nearly rank-wise identical to $(1, x, 5)$ iff $1 \leq x \leq 5$.

as the element-by-element maximum and minimum, respectively, of two other groups can be rearranged into these other groups so as to raise payoffs. In turn, iff one group rank-wise dominates the other can two groups be (permuted so as to be) written as the maximum and minimum of two other groups. The exception is where the two groups are so similar (nearly rank-wise identical) that the above "operator" leaves them unchanged.

Consider the simplest example of group size $n = 2$ and four agents: $p_1 < p_2 < p_3 < p_4$. There are $\binom{4}{2}/2 = 3$ groupings of these agents. By Proposition 1, complementarity rules out two of the three groupings, leaving only the rank-ordered one: $\{(p_1, p_2), (p_3, p_4)\}$. By Proposition 3, substitutability also rules out two of these groupings, in which one group rank-wise dominates the other:

$$\begin{array}{ccc} \xcancel{\begin{matrix}(p_3,p_4)\\(p_1,p_2)\end{matrix}} & \xcancel{\begin{matrix}(p_2,p_4)\\(p_1,p_3)\end{matrix}} & \begin{matrix}(p_2,p_3)\\(p_1,p_4)\end{matrix} \end{array}.$$

The only possible equilibrium grouping is the onion-style one.

However, consider $n = 3$ and six agents: $p_1 < p_2 < p_3 < p_4 < p_5 < p_6$. There are $\binom{6}{3}/2 = 10$ groupings of these agents. Complementarity rules out nine of the ten groupings, again leaving only the rank-ordered one: $\{(p_1, p_2, p_3), (p_4, p_5, p_6)\}$. However, the no rank-wise dominance condition of Proposition 3 only rules out five of these groupings:

$$\begin{array}{ccccc} \xcancel{\begin{matrix}(p_4,p_5,p_6)\\(p_1,p_2,p_3)\end{matrix}} & \xcancel{\begin{matrix}(p_3,p_5,p_6)\\(p_1,p_2,p_4)\end{matrix}} & \xcancel{\begin{matrix}(p_3,p_4,p_6)\\(p_1,p_2,p_5)\end{matrix}} & \begin{matrix}(p_3,p_4,p_5)\\(p_1,p_2,p_6)\end{matrix} & \xcancel{\begin{matrix}(p_2,p_5,p_6)\\(p_1,p_3,p_4)\end{matrix}} \\ \xcancel{\begin{matrix}(p_2,p_4,p_6)\\(p_1,p_3,p_5)\end{matrix}} & \begin{matrix}(p_2,p_4,p_5)\\(p_1,p_3,p_6)\end{matrix} & \begin{matrix}(p_2,p_3,p_6)\\(p_1,p_4,p_5)\end{matrix} & \begin{matrix}(p_2,p_3,p_5)\\(p_1,p_4,p_6)\end{matrix} & \begin{matrix}(p_2,p_3,p_4)\\(p_1,p_5,p_6)\end{matrix} \end{array}.$$

Evidently, when $n > 2$ there is less predictive content in no rank-wise dominance than in rank-ordering. Further, no rank-wise dominance does not leave only "onion-style" groupings: e.g. $\{(p_1, p_4, p_5), (p_2, p_3, p_6)\}$.

The following corollaries show that the exceptional case of nearly rank-wise identical groups, which may rank-wise dominate each other in equilibrium, does not arise in some standard contexts.

Corollary 1. *Assume types are substitutes and that there is a continuum of agents with types drawn from a convex set such that every type has positive density. In any equilibrium grouping, if two groups are picked at random, one group rank-wise dominates the other with probability zero.*

Corollary 2. *Assume types are substitutes, that there are a finite number of agents, and that no two agents have the same type. No group rank-wise dominates another in any equilibrium grouping.*

Ruling out rank-wise dominance also sheds light on the overall matching pattern, and in particular, what aspect of the "median" matching of the 2-person case (Legros and Newman, 2002, Grossman and Maggi, 2000) generalizes. It turns out that in the $n > 2$ setting also, all groups match around a common type:

Corollary 3. *Assume types are substitutes. In any equilibrium grouping, there exists a type* $\tilde{p} \in \Re$*, such that for any group* G *in this grouping,* $\underline{p}^G \leq \tilde{p} \leq \overline{p}^G$*.*

However, the type(s) around which all groups match need not include the median type when $n > 2$.[8]

Two follow-up questions arise. First, we have shown the no rank-wise dominance condition is necessary for any grouping to be an equilibrium; is it sufficient? Second, how much predictive power is there in the no rank-wise dominance condition?

3.1 Predictive power of no rank-wise dominance

We address the second question first: how many groupings does the no rank-wise dominance condition (henceforth, "**no-RWD**") rule out as potential equilibria?

[8] For example, let $n = 3$ and $\mathcal{P} = \{p_l, p_m, p_h\}$, where $p_l < p_h$ and $p_m = p_h \cdot 2/3 + p_l \cdot 1/3$. Assume eight agents of type p_h, three of type p_m, and four of type p_l. Then there exists a unique equilibrium grouping under substitutability and a group payoff function of the form used in Section 3.2. It has four groups of the form (p_l, p_h, p_h) and the other (p_m, p_m, p_m); this last group is entirely below the median type, p_h.

Fix an $n \geq 2$. Consider the simple case where group size is n and there are $2n$ uniquely-typed agents. Thus, $\mathcal{P} = \{p_1, p_2, ..., p_{2n}\}$, with $p_j < p_{j+1}$ for $j \in \{1, 2, ..., 2n-1\}$. A grouping is a partition of $\mathcal{P}$ into two size-n subsets; given $\mathcal{P}$, there are $\binom{2n}{n}/2$ groupings.

Recall that complementarity implies rank-ordering, which rules out all but one grouping regardless of n. Recall also that in the $n = 2$ substitutability case (illustrated above), RWD occurs in two of the three possible groupings, and thus no-RWD also leaves only one potential equilibrium grouping. However, in the $n = 3$ substitutability case, RWD occurs only in five of ten groupings (illustrated above), so that no-RWD rules out exactly half of the groupings. As n gets larger, the no-RWD condition becomes even less restrictive:

Proposition 4. *Consider all groupings of $2n$ uniquely-typed agents into two groups of size n. The fraction of such groupings in which one group rank-wise dominates the other is $\frac{2}{n+1}$.*

That is, the larger is n, the smaller the fraction of groupings the no-RWD condition rules out as potential equilibria. In the limit, it rules out a vanishing fraction of groupings. This does not mean there are multiple equilibrium groupings, just that *by itself* the no-RWD condition does not go very far in ruling out groupings as potential equilibria.

If no-RWD is all that is generically implied by substitutability – a proposition we support in the next section – then this result uncovers a stark contrast between substitutability and complementarity: complementarity rules out all but one grouping regardless of n, while substitutability rules out a vanishingly small fraction of groupings as n gets large. Only in moving away from $n = 2$ does the asymmetry in the predictive power of complementarity and substitutability emerge.

3.2 Is rank-wise dominance all that substitutability rules out?

Are there groupings that do not exhibit rank-wise dominance but can be ruled out by substitutability? Here we show that rank-wise dominance ("RWD") is **all** that substitutability rules out in one setting. Hence there is nothing beyond no-RWD that is generically implied by substitutability.

We continue with the assumption of $2n$ unique types of agents, with agents forming groups of size n. We assume an equal number (integer) or measure (continuum) of each type of agent. We further restrict attention to the class of payoff functions that can be written

$$\Pi(p_1, ..., p_n) = \sum_{j \in \mathcal{N}} g(p_j) + h\left(\sum_{j \in \mathcal{N}} p_j\right) \tag{1}$$

for any differentiable function g and twice-differentiable function h. The key feature is that the interaction between types in the payoff function comes through a function of the sum of types.[9] Π exhibits substitutability iff $h''(\cdot) < 0$ and complementarity iff $h''(\cdot) > 0$.

In the context of the $n = 3$ example illustrated above, our goal is to show that any of the five groupings satisfying no-RWD can be the unique equilibrium grouping depending on the exact values of the p_i's. To make this claim, we use the following notation to differentiate the grouping patterns from the actual values of the p_i's. Let $\mathcal{T} = \{1, 2, ..., 2n\}$. Let $\mathcal{G}$ be a **group-pattern** if $\mathcal{G}$ is an n-element subset of $\mathcal{T}$. Let $\mathcal{G}$ and $\mathcal{G}'$ together be called a **grouping-pattern** if $\mathcal{G}$ and $\mathcal{G}'$ are disjoint group-patterns of $\mathcal{T}$. When $\mathcal{P} = \{p_1, p_2, ..., p_{2n}\}$, with $p_j < p_{j+1}$ for $j \in \{1, 2, ..., 2n - 1\}$, there is a natural mapping from the set of all group-patterns of $\mathcal{T}$ to size-n groups of agents with types from $\mathcal{P}$, which maps (reverse) rank into type ($i \to p_i$). We will say a group G taken from $\mathcal{P}$ **corresponds to** a group-pattern $\mathcal{G}$ in $\mathcal{T}$ if this natural mapping maps $\mathcal{G}$ into G.[10]

Proposition 5. *Assume types are substitutes and that group payoffs can be written as in equation 1. Fix any grouping-pattern $\mathcal{G}$, $\mathcal{G}'$ of $\mathcal{T} = \{1, 2, ..., 2n\}$ in which neither group-pattern rank-wise dominates the other. Then there exists a $\mathcal{P} = \{p_1, p_2, ..., p_{2n}\}$, with $p_j < p_{j+1}$ for $j \in \{1, 2, ..., 2n - 1\}$, such that the unique equilibrium grouping of an equal number or measure of agents of each type in $\mathcal{P}$ involves half of the groups corresponding to $\mathcal{G}$ and half corresponding to $\mathcal{G}'$.*

[9]This is essentially the class of payoff functions that Saint-Paul (2001) analyzes.

[10]For example, group-pattern $(1, 4, 6)$ maps to group (p_1, p_4, p_6), so group (p_1, p_4, p_6) "corresponds to" group-pattern $(1, 4, 6)$.

For example, consider the $n = 3$ case above generalized to a finite number or a continuum of each type. Proposition 5 guarantees that any of the five groupings not ruled out due to RWD can form the basis for the unique equilibrium grouping – in that half of the groups are identical to one of the two groups, and half to the other – for a group payoff function as above and proper choice of the set of types, $(p_1, p_2, ..., p_6)$.

More generally, under substitutability, a sum-based group payoff function, and $2n$ types of equal representation forming groups of size n, any no-RWD grouping-pattern can be the basis for the unique equilibrium grouping. Coupled with Proposition 4, this implies that a fraction $\frac{n-1}{n+1}$ of the partitions of the $2n$ types can form the basis for the unique equilibrium. That is, substitutability by itself says no more than no-RWD (Proposition 5), which itself does not say much (Proposition 4); when types are substitutes, equilibrium matching depends on the specifics of $\mathcal{P}$ and/or the specifics of the payoff function.

4 Discussion and Conclusion

We have characterized equilibrium matching patterns with finite group size greater than two. In the substitutability case, it is not "onion-style" or "median" matching that generalizes, but no rank-wise dominance. As group size grows, complementarity of types continues to pin down a unique equilibrium grouping, while substitutability seems to rule out much less, and in one setting rules out a vanishing fraction of groupings. Thus, the focus on group size of two costs significant generality in the case of substitutability, and masks an asymmetry between substitutability vs. complementarity in predictive power for group formation.

It is worth comparing these results to earlier work by Saint-Paul (2001). He examines group formation where each group has a fixed-*measure continuum* of agents, under complementarity, substitutability, and hybrid cases. Given that groups are so "large", a simple solution in the substitutability case is to put all types in each group, so that all groups are identical and mimic the overall distribution of types – trivially satisfying "no rankwise dominance". The key distinction between his work and ours (as well as other contributions we follow, including Grossman and Maggi, 2000, and Legros and Newman, 2002) is that we look at finite-size groups. This includes the central case where the number of types is greater than group size, and hence groups must differ. Our results contribute to understanding how matching patterns generalize in this setting.

We have mainly characterized equilibrium groupings assuming an equilibrium exists. (Equivalently, we have characterized efficient groupings.) In the continuum cases for both complementarity and substitutability, existence is guaranteed by others' results (Kaneko and Wooders, 1996, at least under "approximate feasibility") and, for complementarity, by our Proposition 2. However, in the finite case we have examples showing that an equilibrium need not exist, at least under complementarity. On the other hand, we also can prove existence in either finite case under additional assumptions. For example, in the substitutability setting considered in Proposition 5, Sherstyuk's (1999) results are used to guarantee existence. Thus equilibrium existence is certainly possible, though not inevitable, in the finite case.

Though we consider group size to be fixed, it may often be a choice variable. The results here provide building blocks for understanding matching patterns, for any given n. With further assumptions, e.g. on the payoff function, the optimal n could potentially be traced out. Examples of endogenizing n are in Kremer (1993), Saint-Paul (2001), and Ahlin (2011).

Finally, future work may find plausible assumptions in the substitutability case that help pin down matching patterns. In some cases, the problem can be simplified by assuming group size no less than the number of types (as in Saint-Paul, 2001). In cases where this is not a desirable assumption, there may exist other plausible assumptions that help pin down the matching pattern enough to aid applied theoretical work. At any rate, because substitutability appears to rule out very little by itself, empirical work will likely need to proceed structurally if it aims to identify the nature of interactions in group formation.

References

[1] Christian Ahlin. Group size matters in microlending under adverse selection or moral hazard. Working Paper, September 2011.

[2] Gary S. Becker. A theory of marriage: Part I. *Journal of Political Economy*, 81(4):813–846, July 1973.

[3] Thomas A. Dowling and Douglas R. Shier. Special sequences. In Kenneth H. Rosen, John G. Michaels, Jonathan L. Gross, Jerrold W. Grossman, and Douglas R. Shier, editors, *Handbook of Discrete and Combinatorial Mathematics*. CRC Press, 2000.

[4] Maitreesh Ghatak. Group lending, local information and peer selection. *Journal of Development Economics*, 60(1):27–50, October 1999.

[5] Maitreesh Ghatak. Screening by the company you keep: Joint liability lending and the peer selection effect. *Economic Journal*, 110(465):601–631, July 2000.

[6] Gene M Grossman and Giovanni Maggi. Diversity and trade. *American Economic Review*, 90(5):1255–1275, December 2000.

[7] Mamoru Kaneko and Myrna Holtz Wooders. The core of a game with a continuum of players and finite coalitions: The model and some results. *Mathematical Social Sciences*, 12(2):105–137, October 1986.

[8] Mamoru Kaneko and Myrna Holtz Wooders. The nonemptiness of the *f*-Core of a game without side payments. *International Journal of Game Theory*, 25(2):245–258, 1996.

[9] Michael Kremer. The O-Ring theory of economic development. *Quarterly Journal of Economics*, 108(3):551–575, August 1993.

[10] Patrick Legros and Andrew F. Newman. Monotone matching in perfect and imperfect worlds. *Review of Economic Studies*, 69(4):925–942, October 2002.

[11] Y. Joseph Lin. The Becker-Brock efficient matching theorem with a supermodular technology defined on an n-lattice. *Mathematical Social Sciences*, 24:105–109, 1992.

[12] Gilles Saint-Paul. On the distribution of income and worker assignment under intrafirm spillovers, with an application to ideas and networks. *Journal of Political Economy*, 109(1):1–37, February 2001.

[13] Katerina Sherstyuk. Multisided matching games with complementarities. *International Journal of Game Theory*, 28(4):489–509, November 1999.

Appendix

Proof of Proposition 1. Since Π is twice continuously differentiable and all cross-partials are strictly positive, it follows that Π is strictly supermodular.[11]

Pick any two groups in an equilibrium grouping; denote them L and M. Suppose L and M are not rank-ordered. This implies that $\overline{p}^L > \underline{p}^M$ and $\overline{p}^M > \underline{p}^L$. Let L' and M' be (re-)orderings of L and M satisfying $p_1^{L'} \leq p_2^{L'} \leq \ldots \leq p_n^{L'}$ and $p_1^{M'} \geq p_2^{M'} \geq \ldots \geq p_n^{M'}$. The fact that L and M are not rank-ordered implies that $p_1^{L'} < p_1^{M'}$ and that $p_n^{M'} < p_n^{L'}$; hence $L' \ngeq M'$ and $M' \ngeq L'$. Finally, let $L'' = L' \wedge M'$ and $M'' = L' \vee M'$. Then

$$\Pi_{L''} + \Pi_{M''} > \Pi_{L'} + \Pi_{M'} = \Pi_L + \Pi_M \ ,$$

where the inequality follows from strict supermodularity of Π and the equality from symmetry of Π. Since L'' and M'' represent an alternative, feasible grouping of the $2n$ agents that produces higher total payoffs, this contradicts L and M being equilibrium groups – at least one of the groups L'' and M'' earns more in total for its members than they earned in the original grouping, so all agents in L'' or M'' can be made better off by defecting from the L and M grouping. Hence, the hypothesis is wrong; any two groups in an equilibrium grouping are rank-ordered.

Proof of Lemma 1. The claim is clearly true when $n = 2$, for

$$2f(x_1, x_2) = f(x_1, x_2) + f(x_2, x_1) \leq f(x_1, x_1) + f(x_2, x_2),$$

the equality by symmetry and the inequality by supermodularity of f.

Now suppose the claim holds for any symmetric and supermodular $f : \Re^n \to \Re$ and any $(x_1, x_2, ..., x_n) \in \Re^n$. By induction, it remains to show that if $g : \Re^{n+1} \to \Re$ is symmetric and supermodular, then for any $(x_1, x_2, ..., x_{n+1}) \in \Re^{n+1}$,

$$(n+1)\ g(x_1, x_2, ..., x_{n+1}) \leq \sum_{k=1}^{n+1} g(x_k, ..., x_k) \tag{2}$$

Fix $(X_1, X_2, ..., X_{n+1}) \in \Re^{n+1}$ and a symmetric and supermodular function $g : \Re^{n+1} \to \Re$.

First, note that $\tilde{G} : \Re^n \to \Re$ defined as $\tilde{G}(x_1, x_2, ..., x_n) \equiv g(x_1, x_2, ..., x_n, X_{n+1})$ satisfies symmetry and supermodularity, since g does. The hypothesis then gives that

$$n\ \tilde{G}(X_1, X_2, ..., X_n) \leq \sum_{k=1}^{n} \tilde{G}(X_k, ..., X_k).$$

[11] A function $f : D \subseteq \Re^n \to \Re$ is supermodular if for any $x, y \in D$, $f(x \wedge y) + f(x \vee y) \geq f(x) + f(y)$, where $x \wedge y$ and $x \vee y$ denote the component-wise minimum and maximum of x and y, respectively. It is *strictly* supermodular if the inequality is strict for any $x, y \in D$ such that $x \ngeq y$ and $y \ngeq x$. The function f is (strictly) submodular if $-f$ is (strictly) supermodular.

Replacing $\tilde{G}$ with g and multiplying by $(n+1)/n$ gives

$$(n+1)\ g(X_1, X_2, ..., X_{n+1}) \leq \frac{1}{n}\sum_{k=1}^{n}(n+1)\ g(X_k, ..., X_k, X_{n+1}).$$

Next, fix a $k \in \{1, ..., n\}$ and note that there are $(n+1)$ terms in the right-hand side sum of the form $g(X_k, ..., X_k, X_{n+1})$. By symmetry of g, we can permute the arguments of these $(n+1)$ terms so that X_{n+1} appears in a different position in each term (1st, 2nd, ..., $(n+1)$st), without changing the sum's value. Finally, we can iteratively apply the supermodularity condition to pairs of these $(n+1)$ terms, always choosing pairs in which X_{n+1} is an argument, to (weakly) increase the sum. After n iterations there is one term of the form $g(X_{n+1}, ..., X_{n+1})$ and n of the form $g(X_k, ..., X_k)$. This holds for every $k \in \{1, ..., n\}$, so the above right-hand side term satisfies

$$\leq \frac{1}{n}\sum_{k=1}^{n}\left[n\ g(X_k, ..., X_k) + g(X_{n+1}, ..., X_{n+1})\right] = \sum_{k=1}^{n+1} g(X_k, ..., X_k).$$

Since $(X_1, X_2, ..., X_{n+1})$ was arbitrary, this establishes the hypothesis for $(n+1)$. ■

Proof of Proposition 2. First, we show that any grouping that is not homogeneous is not an equilibrium. This is simply a corollary of Proposition 1. If one equilibrium group, G say, is not homogeneous, then $\underline{p}^G < \overline{p}^G$. Since $\mathcal{P}$ is convex and each type has positive measure, there must be another equilibrium group, G' say, with an agent of type $p \in (\underline{p}^G, \overline{p}^G)$. But then G and G' are not rank-ordered, contradicting Proposition 1.

Next, we show that the homogeneous grouping, with agents sharing group payoffs equally, is an equilibrium. This follows directly from Lemma 1 (given that Π is supermodular, as argued in the proof of Proposition 1), since for any $(p_1, p_2, ..., p_n)$,

$$\Pi(p_1, p_2, ..., p_n) \leq \frac{\sum_{k=1}^{n}\Pi(p_k, ..., p_k)}{n}.$$

This guarantees that the payoff from any potentially blocking group (the left-hand side) is no greater than the sum of the agents' equilibrium payoffs (the right-hand side).

Proof of Proposition 3. Pick any two groups in an equilibrium grouping. Suppose that one rank-wise dominates the other though the two groups are not nearly rank-wise identical; denote the dominant (dominated) group as M (L). Let L' and M' be (re-)orderings of the types in L and M, respectively, satisfying $p_1^{L'} \leq p_2^{L'} \leq \ldots \leq p_n^{L'}$ and $p_1^{M'} \leq p_2^{M'} \leq \ldots \leq p_n^{M'}$. Since M rank-wise dominates L, $p_i^{L'} \leq p_i^{M'}$ for all $i \in \mathcal{N}$; further, since L and M are not nearly rank-wise identical, at least two of these inequalities are strict, say for $j, k \in \mathcal{N}$. Let L'' and M'' be groups formed from L' and M' by interchanging $p_j^{L'}$ and $p_j^{M'}$: $L'' = (p_1^{L'}, \ldots, p_{j-1}^{L'}, p_j^{M'}, p_{j+1}^{L'}, \ldots, p_n^{L'})$ and $M'' = (p_1^{M'}, \ldots, p_{j-1}^{M'}, p_j^{L'}, p_{j+1}^{M'}, \ldots, p_n^{M'})$. Note that rank-wise dominance implies $L' = L'' \wedge M''$ and $M' = L'' \vee M''$. Also, $M'' \ngeq L''$ (since

$p_j^{L'} < p_j^{M'}$) and $L'' \not\geq M''$ (since $p_k^{L'} < p_k^{M'}$). Thus

$$\Pi_{L''} + \Pi_{M''} > \Pi_{L'} + \Pi_{M'} = \Pi_L + \Pi_M,$$

where the inequality follows from strict submodularity of Π_G (which follows from twice continuous differentiability of Π and the strictly negative cross-partials of Π) and the equality from symmetry of Π. Since L'' and M'' represent an alternative, feasible grouping of the $2n$ agents that produces higher total payoffs, this contradicts L and M being equilibrium groups – at least one of the groups L'' and M'' earns more in total for its members than they earned in the original grouping, so all agents in L'' or M'' can be made better off by defecting from the L and M grouping. Hence, the hypothesis is wrong; no group rank-wise dominates another in an equilibrium grouping, provided the two groups are not nearly rank-wise identical.

Proof of Corollary 1. Fix any group L. Since there are no mass points in the distribution of types, the maximum measure of groups that can be formed that are nearly rank-wise identical to L is zero. The result then follows from proposition 3.

Proof of Corollary 2. This follows from proposition 3, since no two groups can be nearly rank-wise identical when all agents have distinct types.

Proof of Corollary 3. Let $\mathcal{G}$ be the set of groups in an equilibrium grouping, $\underline{p} = \sup_{G\in\mathcal{G}} \underline{p}^G$, and $\overline{p} = \inf_{G\in\mathcal{G}} \overline{p}^G$. If $\overline{p} < \underline{p}$, then there exist groups L and M in $\mathcal{G}$ such that $\overline{p}^L < \underline{p}^M$. But then M rank-wise dominates L without being nearly rank-wise identical, contradicting proposition 3. So, $\underline{p} \leq \overline{p}$ and for any $\tilde{p} \in [\underline{p}, \overline{p}]$, each equilibrium group G has $\underline{p}^G \leq \underline{p} \leq \tilde{p}$ and $\overline{p}^G \geq \overline{p} \geq \tilde{p}$.

Proof of Proposition 4. For group size n and population size $2n$, the total number of groupings is $\binom{2n}{n}/2$. The division by 2 is because group labels are irrelevant and hence $\binom{2n}{n}$ counts each grouping twice. It suffices to show that the total number of groupings in which one group rank-wise dominates the other is $\binom{2n}{n}/(n+1)$.

Without loss of generality, consider $p_1 < p_2 < ... < p_{2n}$. Any grouping of these $2n$ agents can be expressed uniquely in a $2\times n$ matrix as follows: each group is placed on a single row in increasing order, with the group containing p_1 in the first row. (Without the normalization that p_1 goes in the first row, each grouping has two such matrix expressions.) Clearly, any such grouping exhibits rank-wise dominance of one group over the other iff the matrix is monotone increasing going down each column. Thus, the number of rank-wise dominance groupings is equal to the number of ways to construct a $2 \times n$ matrix of $2n$ ordered numbers that is monotonically increasing within each row and column. This is the nth Catalan number: $\binom{2n}{n}/(n+1)$. (See Dowling and Shier, 2000, pp. 145-147, especially Example 11.)

Proof of Proposition 5. This paragraph previews the proof, which is constructive. For an arbitrary no-RWD grouping-pattern of $\mathcal{T}$, $\mathcal{G}$ and $\mathcal{G}'$, we aim to construct a set of types $\mathcal{P} = \{p_1, p_2, ..., p_{2n}\}$ such that $\mathcal{P}$'s groups corresponding to $\mathcal{G}$ and $\mathcal{G}'$ have the same sum of types: $\sum_{j\in\mathcal{G}} p_j = \sum_{j\in\mathcal{G}'} p_j$. Given the sum-based group payoff function (equation 1) and its strict concavity (implied by substitutability), having all groups with the same sum of types maximizes the sum of group payoffs, i.e. achieves an efficient grouping. Hence, having half the groups correspond to $\mathcal{G}$ and the other half correspond to $\mathcal{G}'$ ensures that all agents are matching into groups that have the same sum of types; thus this grouping maximizes efficiency. That this grouping is the unique efficiency maximizer, and hence the unique equilibrium grouping candidate, we guarantee by using square roots of primes in constructing $\mathcal{P}$. This ensures that no other combination of n types from $\mathcal{P}$ can produce the same sum as the two groups corresponding to $\mathcal{G}$ and $\mathcal{G}'$ – hence all other groupings do not achieve equal sums of types for all groups, fall short of efficiency, and thus cannot be equilibrium groupings. Finally, existence is guaranteed using Sherstyuk's (1999) results.

Fix a no-RWD grouping-pattern of $\mathcal{T}$, $\mathcal{G}$ and $\mathcal{G}'$. Our first step is to construct a $\mathcal{P}$ whose groups corresponding to $\mathcal{G}$ and $\mathcal{G}'$ have an equal sum of types in each group. First, let m_j denote the square root of the jth prime number: $m_1 = \sqrt{2}$, $m_2 = \sqrt{3}$, $m_3 = \sqrt{5}$, and so on. The critical property here of the m_j's is incommensurability among themselves, i.e. one cannot produce any m_j by rational-coefficient linear combinations of other m_j's. Let

$$\Sigma_{\mathcal{G}} \equiv \sum_{j \in \mathcal{G}} m_i \ , \qquad \Sigma_{\mathcal{G}'} \equiv \sum_{j \in \mathcal{G}'} m_i \ , \qquad D \equiv \Sigma_{\mathcal{G}} - \Sigma_{\mathcal{G}'} \ .$$

Without loss, let $\Sigma_{\mathcal{G}} \geq \Sigma_{\mathcal{G}'}$, which can be accomplished by switching the labels $\mathcal{G}$ and $\mathcal{G}'$ if need be. Thus $D \geq 0$. Define a function $rank(G, k)$ that gives the kth ranked element (the kth smallest element) of a finite set G; e.g., $rank(\{1, 3, 5, 7\}, 4) = 7$. Now, let k^* be the maximum rank at which $\mathcal{G}'$ dominates $\mathcal{G}$: $k^* = max\{k : rank(\mathcal{G}', k) > rank(\mathcal{G}, k)\}$. By no-RWD, such a k^* exists. Finally, let $j^* = rank(\mathcal{G}', k^*)$.

We now define $\mathcal{P}$: $\mathcal{P} = \{p_1, p_2, ..., p_{2n}\}$, where

$$p_j = \begin{cases} m_j & \text{if } j < j^* \\ m_j + D & \text{if } j \geq j^* \end{cases} \ . \tag{3}$$

Since $D \geq 0$, it is clear that $p_j < p_{j+1}$, for $j \in \{1, 2, ..., 2n - 1\}$. Now,

$$\sum_{j \in \mathcal{G}'} p_j = \sum_{j \in \mathcal{G}'} m_j + (n - k^* + 1)D \ ;$$

the last term is because all types of rank k^* and higher in the group corresponding to $\mathcal{G}'$ get D added (see equation 3). Similarly,

$$\sum_{j \in \mathcal{G}} p_j = \sum_{j \in \mathcal{G}} m_j + (n - k^*)D \ .$$

The last term is because a) if $k^* = n$, in which case $j^* = 2n$, no types in the group corresponding to $\mathcal{G}$, call it G, get D added; and b) if $k^* < n$, all types of rank k^*+1 and higher in G get D added: we know this because $rank(\mathcal{G}, k^*) < rank(\mathcal{G}', k^*) < rank(\mathcal{G}', k^* + 1) < rank(\mathcal{G}, k^*+1)$, the first inequality because $\mathcal{G}'$ dominates $\mathcal{G}$ at k^* by construction; the second by monotonicity of the rank function; and the third because $\mathcal{G}$ dominates $\mathcal{G}'$ at all ranks higher than k^* by construction. Using the last two equations,

$$\sum_{j\in\mathcal{G}} p_j - \sum_{j\in\mathcal{G}'} p_j = \sum_{j\in\mathcal{G}} m_j - \sum_{j\in\mathcal{G}'} m_j - D = 0 \ ;$$

the last equality is by the definition of D. Thus, as desired, we have constructed a $\mathcal{P}$ such that its two groups corresponding to $\mathcal{G}$ and $\mathcal{G}'$, respectively, have the same sum of types: $\sum_{j\in\mathcal{G}} p_j = \sum_{j\in\mathcal{G}'} p_j$.

Our second step is to show no other group from $\mathcal{P}$ achieves the same sum of types as the two groups corresponding to $\mathcal{G}$ and $\mathcal{G}'$. Using the above expressions and the definition of D, and letting $z \equiv n - k^*$, one can show that both groups achieve the following sum:

$$\sum_{j\in\mathcal{G}} p_j = \sum_{j\in\mathcal{G}'} p_j = (z+1)\sum_{j\in\mathcal{G}} m_j - z\sum_{j\in\mathcal{G}'} m_j \equiv \Sigma \ , \tag{4}$$

say. Now take an arbitrary group Q, which we will define by a function $\nu(j) : \{1, 2, ..., 2n\} \to \mathbb{Z}_+$ which gives the number of each type $j \in \{p_1, p_2, ..., p_{2n}\}$ in the group Q. Since group size is n, the function must also satisfy $\sum_{j=1}^{2n} \nu(j) = n$. Let z' be the number of group members in Q with D added, i.e. $z' = \sum_{j=j^*}^{2n} \nu(j)$. Of course, $z' \in \{0, 1, 2, ..., n\}$. Then the sum of types in group Q, call it Σ_Q, is

$$\Sigma_Q = \sum_{j=1}^{2n} \nu(j) m_j + z' \sum_{j\in\mathcal{G}} m_j - z' \sum_{j\in\mathcal{G}'} m_j \ .$$

Given incommensurability of the m_j's, $\Sigma_Q = \Sigma$ iff the coefficient on each m_j in Σ_Q is identical to the coefficient on the same m_j in Σ. That is, examining equation 4, it is clear that the coefficient on m_j for $j \in \mathcal{G}$ must equal $(z+1)$ and the coefficient on m_j for $j \in \mathcal{G}'$ must equal $-z$. Thus,

$$\Sigma_Q = \Sigma \qquad \Longleftrightarrow \qquad \nu(j) = \begin{cases} z + 1 - z' & j \in \mathcal{G} \\ z' - z & j \in \mathcal{G}' \end{cases} \ .$$

Clearly $z' = z$ or $z' = z+1$ is needed to keep $\nu(j)$ non-negative. It is also clear that if $z' = z$, Q is the group corresponding to $\mathcal{G}$; and if $z' = z + 1$, Q is the group corresponding to $\mathcal{G}'$. Thus, we have shown that the only groups that can be assembled from agents with types in $\mathcal{P}$ and that have sum of types equal to Σ are the ones corresponding to $\mathcal{G}$ and $\mathcal{G}'$.

Our third and final step is to show that the grouping with half the groups corresponding to $\mathcal{G}$ and the other half corresponding to $\mathcal{G}'$ is the unique efficient grouping, given $\mathcal{P}$. Note that since a) $\sum_{i\in\mathcal{T}} p_i = 2\Sigma$, b) groups have n members, and c) there is an equal number/measure of each type, then the average group must have sum of types equal to Σ. Given substitutability, which implies concavity, and the sum-based group payoff function 1, the sum of group payoffs is maximal in any grouping where all groups have sums of types equal to the average sum of types. Thus the proposed grouping is an efficient grouping, since all groups have sum of types equal to Σ. Now consider any other grouping. It must involve a strictly positive number/measure of groups that do not correspond to $\mathcal{G}$ or $\mathcal{G}'$. (Note that any grouping can have at most half of its groups corresponding to $\mathcal{G}$ since at half, all the relevant types are used up; and the same for $\mathcal{G}'$). But the previous paragraph establishes that these groups have sums of types not equal to Σ, and in fact bounded away by some strictly positive quantity since the set of all n-person groups that can be formed from $\mathcal{P}$ is finite. Thus, any other grouping has a positive number/measure of groups with sum of types not equal to the average sum of types, and given strict concavity of the sum-based group payoff function, produces a strictly lower sum of group payoffs than a grouping where all groups have the average sum of types. Thus the proposed grouping is the unique efficient grouping.

As is well known, in this context any equilibrium grouping must be an efficient one – if not, there would exist n agents who could achieve higher payoffs by forming a new group, one that belongs to an efficient grouping. Thus, if there is an equilibrium, it is the proposed grouping and is unique. It remains to demonstrate that the core is non-empty. In the continuum case, existence is guaranteed by the results of Kaneko and Wooders (1986, 1996; at least under approximate feasibility). Consider the finite case with ng agents, say, and thus g groups of n agents each in the uniquely efficient grouping; call it $\{\tilde{G}^r\}_{r=1}^{g}$. Straightforward modification of Lemma 1 of Sherstyuk (1999) gives that the core is non-empty if

$$\sum_{r=1}^{mg} \Pi(G^r) \leq m \sum_{r=1}^{g} \Pi(\tilde{G}^r)$$

for any positive integer m and any collection of mg (not necessarily distinct) size-n groups $\{G^r\}_{r=1}^{mg}$ such that each agent belongs to exactly m of these groups. Note that the right-hand side essentially sums the payoffs of mg groups all with the same sum of types, from an m-replica of the set of players. The left-hand side also sums the payoff of mg groups formed from an m-replica of the set of players; but it can do no better than the right-hand side, since the average sum of types across groups is fixed on both sides, and given concavity of Π, payoffs are maximized when each group has the same sum of types.