Chiappori, Pierre-André; Reny, Philip J.

Article
Matching to share risk

Theoretical Economics

Provided in Cooperation with:
The Econometric Society

Suggested Citation: Chiappori, Pierre-André; Reny, Philip J. (2016) : Matching to share risk, Theoretical Economics, ISSN 1555-7561, Wiley, Hoboken, NJ, Vol. 11, Iss. 1, pp. 227-251, http://dx.doi.org/10.3982/TE1914

This Version is available at:
http://hdl.handle.net/10419/150277

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

https://creativecommons.org/licenses/by-nc/3.0/
Matching to share risk

PIERRE-ANDRÉ CHIAPPORI
Department of Economics, Columbia University

PHILIP J. RENY
Department of Economics, University of Chicago

We consider a matching model in which individuals belonging to two populations ("males" and "females") can match to share their exogenous income risk. Within each population, individuals can be ranked by risk aversion in the Arrow–Pratt sense. The model permits nontransferable utility, a context in which few general results have previously been derived. We show that in this framework a stable matching always exists, it is generically unique, and it is negatively assortative: for any two matched couples, the more risk averse male is matched with the less risk averse female.

Keywords. Negatively assortative matching, risk sharing, stable match.

JEL classification. C78, D00.

1. Introduction

The analysis of matching patterns when individuals can transfer utility between them, initiated by Becker (1973) and Shapley and Shubik (1971), has recently attracted renewed attention. In general, these works rely on a transferable utility (TU) assumption, which posits that under a convenient representation of utilities, the Pareto frontier generated by any possible match is a straight line with slope $-1$. The TU framework has several, interesting features. Stability, in that context, is equivalent to aggregate surplus maximization. Not only can existence and generic uniqueness results be readily derived from this fact, but the comparative statics tend to be significantly easier, and numerical simulations are especially simple to perform, since looking for a stable match boils down to solving a linear programming model, a task for which fast algorithms exist.

The TU assumption, however, has its drawbacks, which have also been pointed out. In a static context, an implication of the TU assumption is that groups behave as single
decision makers—a claim that has been repeatedly challenged, both from a theoretical and an empirical perspective. Moreover, some specific problems cannot in general be easily cast into a TU framework. A particularly interesting example is risk sharing. While the ability to share otherwise uninsurable risk has long been considered a major motivation for household or group formation, the corresponding models typically fail to satisfy the restrictions implied by TU. Specifically, Schulhofer-Wohl (2006) shows that a necessary condition for a risk-sharing problem to be compatible with TU is that individual risk aversions belong to the identical shape harmonic absolute risk-aversion (ISHARA) class (as defined by Mazzocco 2004a). In particular, for individual preferences to exhibit constant relative risk aversion (CRRA) and to also be in the ISHARA class, it must be the case that individuals have identical risk aversions—quite a restrictive condition indeed.

The general case in which utility can be transferred between individuals, but not necessarily at a constant “exchange rate” as implied by TU, is therefore extremely interesting but quite complex. Recently, Legros and Newman (2007) have provided a general investigation of such models. In particular, these authors introduce two conditions, generalized increasing differences (GID) and generalized decreasing differences (GDD), which they show are sufficient for positive and negative assortative matching, respectively. Moreover, if either condition holds strictly, say GID, then it is immediate that every stable match must be positive assortative. The main difficulty when facing any particular problem, including the present one, is to determine whether any one of these conditions holds.

In an illustrative example, Legros and Newman consider a risk-sharing problem in which individuals with different risk aversions (ordered by the Arrow–Pratt measure) match in pairs to share some exogenous, uninsurable risk. In the special case in which each couple’s joint income can be either high or low (i.e., there are just two possible states of the world), Legros and Newman establish that any stable match must be negative assortative, i.e., among matched individuals, the $k$th most risk averse male is matched with the $k$th least risk averse female.6

That any stable matching in this risk-sharing context must be negative assortative with respect to risk tolerance is intuitively appealing, and one might hope that this property holds far more generally and, in particular, for general types of income risks involving an arbitrary number of states of the world and for general risk averse preferences. The purpose of the present paper is to show that this intuition is correct.

---

3See Browning et al. (2014) for a recent overview.
5ISHARA requires that (i) individual utilities be of the harmonic absolute risk-aversion form (i.e., the index of absolute risk aversion must be a harmonic function of income), and (ii) the corresponding income coefficient must be identical across individuals.
6Legros and Newman (2007) show that the same result holds with an arbitrary number of states of the world under the additional assumption that each individual $i$’s utility function takes the form

$$u_i(x) = \log(1 + a_i + x).$$

But note that because these utilities belong to the ISHARA family, the resulting risk-sharing problem is in fact a TU problem for which submodularity of the surplus function is well known to be sufficient for a negative assortative stable match.
Specifically, we consider a setting with two populations (“males” and “females”). Within each population, individual incomes are drawn from some joint distribution; we assume only that each joint distribution is exchangeable (so that males have ex ante identical incomes from any female’s viewpoint and vice versa), and that male and female incomes are independent. Any male and any female can form a group (“household”), pool their income, and share the corresponding aggregate risk efficiently. Last, we assume that individuals within each population are expected utility maximizers and can be ranked according to their risk aversion. That is, for any two males, one is more risk averse than the other in the sense of Arrow–Pratt, and similarly for any two females. No further assumptions, with the exception of an Inada-type condition at the origin, are made on the shape of individual preferences. A result due to Kaneko (1982) ensures that a stable matching always exists.7

Our main result (Theorem 2) establishes that any stable matching is negative assortative among matched couples, i.e., for any \( k \), the \( k \)th most risk averse matched male is matched with the \( k \)th least risk averse matched female. In general there may be multiple stable matches. However, there are two senses in which uniqueness obtains. First, a complementary form of uniqueness always holds: across all stable matches, either an individual’s mate is unique or his/her payoff is unique (Theorem 4). Second, “generically” in our domain of utility functions, there is a unique stable match (Theorem 5).

Interestingly, if populations are not of equal sizes, so that some individuals must remain unmatched, the identity of the unmatched individuals depends on the exact distribution of preferences. In particular, we provide robust examples in which the unmatched individuals are the most risk averse, the least risk averse, or of intermediate risk aversion.

The proof of our main result relies on establishing a strict single-crossing property (Lemma 1). Once this property is established, the negative assortativeness of any stable match becomes immediate (Theorem 3).

Finally, a variation of the problem is studied in the last section. There, we consider a one-sided or “roommate” version of the problem, whereby individuals belonging to the same population match (pairwise) to share risk. Roommate matching problems are in general more complex than bipartite ones; for instance, stable matchings can easily fail to exist.8 Still, we show that when the total number of individuals is even, a stable matching exists, is unique, and is negative assortative.

The present analysis may provide a basis for other applications.9 Consider, for example, a labor market composed of high-skill and low-skill individuals who must work in teams, of two say, to be productive. Each team that forms must agree, ex ante, to an incentive contract that specifies how the stochastic profits generated from their private efforts are to be divided between them. High-skill individuals may be attracted to low-skill individuals and vice versa not only for risk-sharing reasons, but also because specialization makes it possible to avoid counterproductive effort. As in our risk-sharing

---

7 We thank an anonymous referee for this reference.
8 See Chiappori et al. (2014a).
9 We are grateful to an anonymous referee for this suggestion.
scenario, a stable match in this labor market setting involves an endogenous choice of teams as well as an endogenous contract choice for each team.

Beyond its theoretical appeal, our negative assortativeness conclusion has important empirical consequences. Most empirical studies of efficient risk sharing within social groups refer to an underlying, theoretical framework characterized by two features: (i) households are taken to be the basic decision units and are each characterized by some von Neumann–Morgenstern (VNM) utility (generally of the CRRA type), and (ii) risk aversion is identical across households.10 These approaches have been criticized on the grounds that, empirically, risk aversion appears to be very heterogeneously distributed across individuals.11 Therefore, if the determinants of household formation are not directly linked to risk sharing, not only should different households have different risk attitudes, but the mere notion of a VNM utility defined at the household level is basically flawed, since CRRA preferences with different levels of relative risk aversions do not satisfy Mazzocco’s ISHARA conditions for aggregation and cannot therefore be represented by a single VNM utility.12 Our results reinforce this criticism by showing that if, alternatively, matching is related to individual attitudes to risk, the situation is even worse: negative assortativity implies that differences in risk aversion should on average be larger within households (or risk-sharing groups) than across them. This suggests that a large fraction of the empirical literature on risk sharing should be considered with some caution.13

2. The model

Consider a one-to-one matching model in which a finite set \{1, \ldots, M\} of males match with a finite set \{1, \ldots, F\} of females. Males and females are strictly risk averse, and \( u_i : [0, \infty) \rightarrow \mathbb{R} \) denotes the von Neumann–Morgenstern utility function of income of male \( i \), and \( w_j : [0, \infty) \rightarrow \mathbb{R} \) denotes that of female \( j \). For every male \( i \), \( u_i(\cdot) \) is bounded and continuous on its domain, \( u_i'(x) > 0 \), \( u_i''(x) < 0 \) for all \( x \in (0, \infty) \), \( u_i(0) = 0 \), and \( \lim_{x \downarrow 0} u_i'(x) = +\infty \). Female utility functions satisfy the same conditions.

We assume that individuals in each population can be ranked according to their absolute risk aversion in the Arrow–Pratt sense. That is, for any two males, one of them is more risk averse than the other, and similarly for any two females. We may therefore index each set of individuals in order of increasing risk aversion, so that male \( i \) is (strictly)
more risk averse than male \( j \) (i.e., \(-u''_i(x)/u'_i(x) > -u''_j(x)/u'_j(x) \forall x > 0\)) if and only if \( i > j \), and similarly for females.

The statement that individual \( i \) is more risk averse than \( j \) only means that \( i \)'s Arrow–Pratt measure of risk aversion is strictly greater than \( j \)'s when their income realizations are identical. Because incomes are random and the incomes received through risk sharing will be endogenous, there is no reason to expect, and we do not assume, that \( i \)'s Arrow–Pratt measure of risk aversion is greater than \( j \)'s at their respective ex post incomes when these are distinct.

Each individual is endowed with an exogenous nonnegative and nondegenerate random income denoted by \( \tilde{x}_i \) for male \( i \) and \( \tilde{z}_j \) for female \( j \).\(^{14}\) The \( M \) random variables \( \tilde{x}_1, \ldots, \tilde{x}_M \) are assumed to be exchangeable and independent of the \( F \) exchangeable random variables \( \tilde{z}_1, \ldots, \tilde{z}_F \).\(^{15}\) Note that incomes within each population may be correlated and the marginal income distributions of any two members of the same population are identical.

Let \( \tilde{y} = \tilde{x}_i + \tilde{z}_j \) be the nondegenerate random total income of any male and female. Exchangeability guarantees that the distribution of \( \tilde{y} \) is independent of \( i \) and \( j \). Let \( Y \subseteq [0, \infty) \) denote the support of \( \tilde{y} \).

Individuals who choose to remain single receive their random income. However, if male \( i \) and female \( j \) choose to match and share risk, they can enter into a binding agreement, ex ante, prior to the realization of their incomes, specifying how their income will be shared. The “stability” condition introduced below implies that this binding agreement will be ex ante Pareto efficient. Consequently, as shown by Borch (1962), and given our assumptions, the allocation of income between male \( i \) and female \( j \) must satisfy the mutuality principle: each individual’s income share depends only on the couple’s total income \( \tilde{y} \). In particular, a couple’s efficient risk-sharing agreement, or sharing rule, can be denoted by a function, \( x : Y \rightarrow [0, \infty) \), of total income, where \( x(y) \) denotes the amount of total income \( y \) given to the male and \( y - x(y) \geq 0 \) is the amount given to the female. Under this sharing rule, male \( i \)'s expected utility is \( E_{u_i}(x(\tilde{y})) \) and female \( j \)'s expected utility is \( E_{u_j}(\tilde{y} - x(\tilde{y})) \).

Because individuals can choose to remain single, we focus on sharing rules that are individually rational in the sense that both the male’s and the female’s ex ante payoffs are at least as large under the agreed upon sharing rule as their respective ex ante payoffs when single.

We wish to determine which pairs of men and women will choose to share risk and, for those that do, we also wish to determine the sharing rules they adopt. To accomplish this we employ a standard stability criterion.

An ordered pair \( (i, j) \) is called a couple if \( i \) is a male and \( j \) is a female. A match is a subset of couples, \( C \), such that for any male \( i \) there is at most one female \( j \) with \( (i, j) \in C \), and for any female \( j \) there is at most one male \( i \) with \( (i, j) \in C \). Any individual appearing

\(^{14}\)We adopt the convention that tildes denote random variables and the absence of a tilde denotes a realization of that random variable.

\(^{15}\)Collections of random variables are exchangeable if their joint cumulative distribution function (c.d.f.) is a symmetric function. Note that income distributions may be discrete or continuous.
in some couple in \( C \) is said to be \textit{matched}, while any individual not appearing in any couple in \( C \) is said to be \textit{unmatched} or \textit{single}.

A match \( C \) is \textit{stable} if for each couple \((i, j) \in C\) there is an individually rational sharing rule \( x_{ij}(\cdot) \) such that the resulting vector of sharing rules, \( S = (x_{ij}(\cdot))_{(i, j) \in C} \), has the following property. For any sharing rule \( x(\cdot) \), there does not exist a male \( i \) and a female \( j \) who each strictly prefer \( x(\cdot) \) to his/her sharing rule from \( S \) if matched or to his/her income distribution if unmatched. In this case, we say that the stable match \( C \) is \textit{supported} by \( S \) and we call \((C, S)\) a \textit{stable outcome}.

Thus, a stable match refers to the set of couples that form, while a stable outcome includes also the sharing rules that matched couples employ. Consequently, there can be a unique stable match and yet multiple stable outcomes when the stable match is supported by more than one vector of sharing rules.

\textbf{Remark 1.} It is inefficient for a man and a woman to each be unmatched. Indeed, one way they could share income is to mimic being single (i.e., male \( i \) keeps \( \tilde{x}_i \) and female \( j \) keeps \( \tilde{z}_j \)). Such a sharing rule will however never be efficient, since efficiency requires male and female income shares to rise and fall together (by the mutuality principle and Wilson 1968, Theorem 5), while \( \tilde{x}_i \) and \( \tilde{z}_j \) are nondegenerate and independent. Consequently, there exist sharing rules that are preferred by both individuals over being single.

It follows that at any stable match there can be single males (if and only if \( M > F \)) or single females (if and only if \( M < F \)), but not both. In particular, stability implies that all individuals are matched if \( M = F \).

A modest amount of additional notation will be useful in what follows. First, let \( r_j = Ew_j(\tilde{z}_j) \) denote female \( j \)'s reservation utility, and let \( \bar{v}_j = Ew_j(\tilde{y}) \) denote female \( j \)'s expected utility when she is matched with a male and receives all of the joint income \( \tilde{y} \) (recall that the joint income distribution is independent of the match by exchangeability). Note that because male and female incomes are nonnegative and nondegenerate, \( 0 < r_j < \bar{v}_j \). Second, let \( U_{ij}(v_j) \) denote male \( i \)'s maximum expected utility when he is matched with female \( j \), where the maximum is taken over all possible sharing rules that ensure female \( j \) an expected payoff of at least \( v_j \in [0, \bar{v}_j] \).

Let us first record that a stable outcome, and hence also a stable match, always exists.

\textbf{Theorem 1.} \textit{There is at least one stable outcome, and so there also is at least one stable match.}

\textbf{Proof.} The nontransferable utility game that is defined by considering the payoffs that any individual can attain by staying single and that any couple can attain with some sharing rule is a central assignment game in the sense of Kaneko (1982). Hence, by Kaneko’s Theorem 1, a stable outcome exists. \( \Box \)

\textsuperscript{16}The analyses of Borch (1962) and Wilson (1968), together with the assumptions made here, ensure that \( U_{ij}(\cdot) \) is well defined and continuous on \([0, \bar{v}_j]\). In particular, continuity follows by the logic of Berge’s theorem of the maximum because sharing rule solutions here are always uniformly bounded nondecreasing functions, and sequences of such functions always have pointwise convergent subsequences by Helly’s theorem.
We now consider the most interesting part of the problem, namely the characterization of the stable matches.

3. Stability

3.1 The main result

Our main result is the following.

**Theorem 2.** Any stable match has the following features:

- Either all of the men are matched or all of the women are matched.
- The matching is negative assortative in the sense that among matched individuals, the $k$th most risk averse male is matched with the $k$th least risk averse female for all $k$.

In particular, if there are equal numbers of males and females, the unique stable match is negative assortative and all individuals are matched.

Before considering the most general case involving any number of males and females, let us consider first the case in which there are two males and two females. Theorem 2 then reduces to the following more fundamental result.

**Theorem 3.** When there are precisely two males and two females, the negative assortative match, namely that in which the most risk averse male (female) is matched with the least risk averse female (male), is the unique stable match.

The proof of Theorem 3, and also ultimately of Theorem 2, relies on the following key lemma, which establishes a strict single-crossing property. Recall from Section 2 that $\bar{v}_j$ is female $j$'s expected payoff from matching with any male and receiving the entire joint income. Recall also that $U_{ij}(v)$ denotes male $i$'s maximum expected payoff when he is matched with female $j$, where the maximum is taken over all possible sharing rules that ensure female $j$ an expected payoff of at least $v \in [0, \bar{v}_j]$. Finally, recall that male (female) 2 is more risk averse than male (female) 1.

**Lemma 1.** For any real numbers $v_1$ and $v_2$ satisfying $0 \leq v_1 < \bar{v}_1$ and $0 < v_2 \leq \bar{v}_2$, we have

$$U_{11}(v_1) \geq U_{12}(v_2) \implies U_{21}(v_1) > U_{22}(v_2).$$

Lemma 1, whose proof is given in the Appendix, says the following. If, given the utilities that two females must receive, a male weakly prefers matching with the less risk averse female, then any strictly more risk averse male will strictly prefer matching with this less risk averse female. Intuitively, the more risk averse male prefers more insurance, any amount of which is cheaper to obtain from the less risk averse female since she is willing to bear more income risk than the more risk averse female.
Given the compelling intuition, the proof of Lemma 1 is perhaps more subtle than one might expect. Indeed, in a typical proof of negative assortative matching, one simply observes that higher types are willing to pay more to match with a low type than lower types are willing to pay. The argument here is complicated by the fact that payment comes in the form of an endogenous sharing rule, and that the opportunity cost of that payment depends in a complex way on the concavity of a person's utility. This point can be understood in reference to the transferable utility benchmark. Transferable utility implies that, for a well chosen cardinalization of individual expected utilities, the "exchange rate" of the spouses' utilities is constant and equal to $-1$. In our nontransferable utility framework, on the contrary, the exchange rate is not constant, and is endogenous to the choice of the sharing rule.

The strict single-crossing property in the statement of Lemma 1 is closely related to, but stronger than, Legros and Newman's (2007) generalized decreasing differences condition. In particular, their condition replaces the strict inequality in the displayed equation of Lemma 1 with a weak inequality. The strict inequality is important for two reasons. First, it implies that every stable match must be negative assortative, whereas a weak inequality implies only that every stable match is payoff equivalent to one that is negative assortative (Proposition 1, part (ii) in Legros and Newman 2007). Second, the strict inequality is crucial for establishing that, generically in our domain of utility functions, there is a unique stable match (Theorem 5 below).

Given Lemma 1, let us now prove Theorem 3.

**Proof of Theorem 3.** We first argue that $\{(1,1),(2,2)\}$ is not stable. Suppose that it is. Then there are individually rational and efficient sharing rules associated with this match giving each female $j$ utility $v_j$, say. Individual rationality ensures that $0 < v_j < \tilde{v}_j$ for each female $j$, and efficiency implies that male $i$'s payoff must be $U_{ii}(v_i)$. By stability, male 1 must be unable to strictly improve his payoff by matching with female 2 and ensuring her a payoff of at least $v_2$; otherwise he could also make female 2 strictly better off by giving her slightly more income. Hence, $U_{11}(v_1) \geq U_{12}(v_2)$. But then, according to Lemma 1, $U_{21}(v_1) > U_{22}(v_2)$. That is, male 2 can strictly improve his payoff by matching with female 1 and choosing a sharing rule that leaves female 1 at least indifferent. But then male 2 can make both himself and female 1 strictly better off by giving female 1 slightly more income. We conclude that $\{(1,1),(2,2)\}$ is not stable. Consequently, by Theorem 1 and Remark 1, $\{(1,2),(2,1)\}$ is the unique stable match. $\square$

The proof of Theorem 2 is now straightforward.

**Proof of Theorem 2.** The first part of the theorem follows from Remark 1. Consider next the third part of the theorem in which there are equal numbers of males and females. By Theorem 1, there exists at least one stable match, and by the first part of the theorem, every individual is matched. It therefore suffices to show that any match different from the negative assortative one cannot be stable. For any nonnegative assortative match, there must be two males, $i$ and $i'$, and two females, $j$ and $j'$, such that $i$ is more risk averse than $i'$ and $j$ is more risk averse than $j'$ such that $i$ is matched with $j$ and
By Theorem 3, the submatch \((i, j), (i', j')\) is not stable and so the overall match is not stable either. Finally, the second part of the theorem follows from the third.

3.2 Uniqueness

When the number of males equals the number of females, Theorem 2 implies that there is a unique stable match, namely the negative assortative one. Theorem 2 also implies that, for a given set of matched individuals, the stable match is unique. The difficulty however, comes from the fact that the set of matched individuals need not be unique. To see why, consider the following simple example.

Example 1. There are three individuals, one male and two females with independent and identically distributed (i.i.d.) income distributions. One of the females is nearly risk neutral while the other is very risk averse. If the male's utility function is identical to the nearly risk neutral female's, then, in the unique stable match, he will—for optimal risk-sharing purposes—match with the very risk averse female. However, if the male's utility function is identical to the nearly risk neutral female's, then he will match with the nearly risk neutral female. By continuity, there is a convex combination of the two females' utility functions such that if the male's utility function were equal to that convex combination, then there would be two stable matches. Each of the two females is single in one of these stable matches and all three individuals are indifferent between the two stable matches.

Simple as it may be, this example provides two interesting insights. First, if the stable match is not unique, then the payoffs must be. More precisely, if there is more than one stable outcome, then any individual who does not have the same mate in all stable outcomes (e.g., all three individuals in our example) must be indifferent among all stable outcomes. Second, nonuniqueness of the stable match is not "generic." Indeed, in the example there is exactly one convex combination of the two female utility functions that leads to multiple stable matches.

The two theorems to follow show that these insights are generally valid. The first theorem says that across all stable outcomes, either an individual's mate is unique or the individual's payoff is unique. The second theorem says that for a "generic" set of utility functions in our domain, there is a unique stable match. The proofs of both theorems are given in the Appendix.

**Theorem 4.** Each individual has either a unique mate or a unique payoff across all stable outcomes. That is, if an individual does not have the same mate in every stable outcome, then he/she receives the same payoff in every stable outcome.

To get some intuition for the proof, suppose there are two females and one male, and that some individual has a different mate in two stable outcomes. Then there is more than one stable match. By Remark 1, the male has a mate in every stable match. Consequently, there are precisely two stable matches, with one of the two females matched...
with the male in each of them. Hence, no individual has the same mate in every stable outcome. We will argue that all three individuals must be indifferent among all of the stable outcomes.

When female \( j = 1, 2 \) is unmatched, she receives her reservation utility \( r_j \). Let \( v_j \) denote her utility in some stable outcome in which she is matched to the male \((i = 1)\). Then

\[
r_1 \leq v_1,
\]

by individual rationality for female 1 when she is matched to the male.

But then we must also have

\[
U_{11}(v_1) \leq U_{12}(v_2),
\]

since otherwise the match between the male and female 2 when her payoff is \( v_2 \) would not be stable. Indeed, if (1) holds and (2) fails, the male would strictly prefer matching with female 1 and giving her a payoff slightly above \( v_1 \geq r_1 \) (which female 1 would strictly prefer to being single) than matching with female 2 and giving her a payoff of \( v_2 \). Hence, (2) holds.

But then we must also have

\[
v_2 \leq r_2,
\]

since otherwise the match between the male and female 1 when her payoff is \( v_1 \) would not be stable. Indeed, if (2) holds and (3) fails, female 2 would strictly prefer to match with the male and accept a payoff slightly less than \( v_2 \) (which, by (2), the male would strictly prefer to being matched with female 1 and giving her \( v_1 \)) than being single. Hence, (3) holds.

Reversing the roles of females 1 and 2 in the argument above, we conclude that each of the inequalities in (1)–(3) must be an equality. Therefore, all three individuals must be indifferent between these two stable outcomes.

Since the match defined by any other stable outcome must be distinct from one of the two matches considered above, we may repeat the argument and conclude that all three individuals must be indifferent between this third stable outcome and one of the two—and hence both, by transitivity—stable outcomes considered above. We conclude that all three individuals are indifferent among all stable outcomes, as promised.

Our next result says that generically in the space of utility functions, there is a unique stable match. In preparation for the formal statement, let \( \mathcal{U} \) denote the set of vectors of utility functions \((u_1, \ldots, u_M)\) such that (i) each \( u_i : [0, \infty) \to \mathbb{R} \) is bounded and continuous and satisfies \( u_i(0) = 0, \lim_{x \to 0^+} u'_i(x) = +\infty, \) and \( u'_i(x) > 0, u''_i(x) < 0 \) \( \forall x \in (0, \infty) \), and (ii) \( -u''_i(x)/u'_i(x) > -u''_j(x)/u'_j(x) \) \( \forall x \in (0, \infty), \forall i < M \).

Hence, \( \mathcal{U} \) is the domain of vectors of utility functions for the males that we have been considering throughout. Similarly, let \( \mathcal{W} \) denote the domain of vectors of utility functions for the females that we have been considering throughout.
For any real-valued functions $f_1, \ldots, f_M, g_1, \ldots, g_F$ on $[0, \infty)$, define the norm of $(f, g) = (f_1, \ldots, f_M, g_1, \ldots, g_F)$ by

$$\|(f, g)\| = \sup_{i=1, \ldots, M} \sup_{x \geq 0} |f_i(x)| + \sup_{j=1, \ldots, F} \sup_{x \geq 0} |g_j(x)|.$$ 

Then $(U \times W, \| \cdot \|)$ is a metric space.\(^{17}\)

**Theorem 5.** There is an open and dense subset $G$ of the metric space $(U \times W, \| \cdot \|)$ such that for every vector of utility functions $(u_1, \ldots, u_M, w_1, \ldots, w_F)$ in $G$, there is a unique stable match.

To get a feel for why this result holds, let us once again suppose that there are two females and one male (with equal numbers of males and females there is a unique stable match—the negative assortative one). As we argued just above, for any two distinct stable matches, (1)–(3) must hold. In particular,

$$U_{11}(r_1) = U_{12}(r_2).$$

This says that the male must be indifferent between matching with either female when the females must receive their reservation utilities.

Recall that we have indexed the individuals so that female 1 is less risk averse than female 2. If the male were to become even slightly more risk averse, then he should strictly prefer to obtain additional insurance that can be more affordably purchased from the less risk averse female 1. That is, a perturbation of 1’s utility function from $u_1$ to $\hat{u}_1$ that makes him more risk averse should (and does, in fact, by Lemma 1) result in

$$\hat{U}_{11}(r_1) > \hat{U}_{12}(r_2).$$

But then (1)–(3), which we have seen are necessary for the existence of multiple stable matches, can never hold. We conclude that if there are multiple stable matches, then an arbitrarily small perturbation of the male’s utility function leads to a unique stable match.

Alternatively, if we began in a situation in which $U_{11}(r_1) \neq U_{12}(r_2)$, then there would be a unique stable match. Moreover, given our metric, the inequality, and therefore also uniqueness, would hold for an open set of utility functions for the males and females.

Altogether, this explains why uniqueness holds on an open and dense set of utility functions in our domain.

**Remark 2.** Theorems 4 and 5 are valid in any matching environment in which the strict single-crossing property (or the positive assortative analogue) that is displayed in Lemma 1 holds. Thus, these results extend beyond our efficient risk-sharing setting.

### 3.3 Who are the singles?

An ambiguity remains, however, regarding the identity of those who are left single. For example, if there are more males than females, so that some males must be single, is it

\(^{17}\)The distance between $(u, w)$ and $(\hat{u}, \hat{w})$ in $U \times W$ is $\|(u - \hat{u}, w - \hat{w})\|$.
the most risk averse males? The least risk averse? Or could the single males be located somewhere within the distribution of risk aversion? We now proceed to show, through a simple example, that any of these situations is possible.

**An example** The example we consider has the following features:

- Both $\tilde{x}_i$ and $\tilde{z}_j$ are uniformly distributed over $[1/2, 2]$.
- Individual preferences are CARA.\(^{18}\) In particular, for a match between a male and a female with respective Arrow–Pratt risk-aversion (RA) indices of $\mu$ and $\phi$, an efficient sharing rule, $x(\cdot)$, is given by

$$x(\tilde{y}) = \frac{\phi \tilde{y}}{\mu + \phi} + k$$

$$\tilde{y} - x(\tilde{y}) = \frac{\mu \tilde{y}}{\mu + \phi} - k,$$

where $k$ is an arbitrary constant.

We first compute the maximum level of expected utility that a female with RA $\phi$ can obtain when matched with a male with RA $\mu$. This corresponds to an efficient sharing rule in which the constant $k$ is such that the male is indifferent between being matched with the female and being single. This implies that $k$ must satisfy

$$e^{-\mu k} = \frac{\mu \phi^2 (e^{-\mu} - e^{-2\mu})}{(\mu + \phi)^2 ((e^{-2\mu} - e^{-2\mu}/(\mu + \phi)))^2}.$$

Letting $W(\phi, \mu)$ denote the corresponding utility of the female, it follows that

$$W(\phi, \mu) = -\left(\frac{\mu}{(e^{-\mu} - e^{-2\mu})}\right)^{\phi/\mu} \left(\frac{\mu + \phi}{\mu \phi}\right)^2 \left((e^{-2\mu}/(\mu + \phi)) - e^{-2\mu}/(\mu + \phi)\right)^{1+\phi/\mu}.$$

In particular, $W(7, \mu)$ is not monotonic in $\mu$, as illustrated in Figure 1, which plots $W(\phi = 7, \mu) \times 10^6$ as a function of $\mu$.

Consider now a situation in which three males (with respective RA parameters $\mu_1$, $\mu_2$, and $\mu_3$) are to be matched with two females with identical RA parameter $\phi$. We shall make use of the following lemma.

**Lemma 2.** In any stable match of the example, male $i$ remains single only if

$$W(\phi, \mu_i) = \min_{i'} W(\phi, \mu_{i'}).$$

---

\(^{18}\)Technically, CARA functions do not satisfy our assumption that the derivative at zero should be infinite. However, in our example individual incomes under stable sharing rules are always larger than 0.5 for the specific parameters we consider. Hence one can replace the CARA form $-\exp(-\sigma x)$ by the function

$$u(x) = \begin{cases} 
-\exp(-\sigma x) & \text{if } x \leq 0.5 \\
-\exp(-\sigma x) & \text{if } x > 0.5,
\end{cases}$$

without changing the conclusions. The redefined function satisfies our conditions.
Proof. Fix any stable match. By Remark 1, precisely one male, say male 1, is single. Assume, by way of contradiction, that \( W(\phi, \mu_i) < W(\phi, \mu_1) \) for some matched male \( i \). Male \( i \)'s mate cannot receive more than \( W(\phi, \mu_i) \). Hence, a match with 1 in which \( i \)'s mate receives \( \frac{1}{2}(W(\phi, \mu_i) + W(\phi, \mu_1)) \) would be strictly preferred by her and by male 1. Hence, the match is not stable. \( \square \)

In particular, when \( \phi = 7 \), consider any three values of \( \mu \) on the horizontal axis of Figure 1. The one giving the smallest value to the function determines the male who remains single.\(^{19}\) Clearly then, all negative assortative matches are possible. Referring to Figure 1, we see that for \( \mu_1 = 1 \), \( \mu_2 = 2 \), and \( \mu_3 = 3 \), the most risk averse male 3 is single; for \( \mu_1 = 4 \), \( \mu_2 = 5 \), \( \mu_3 = 6 \), the least risk averse male 1 is single; finally, for \( \mu_1 = 1 \), \( \mu_2 = 4 \), and \( \mu_3 = 7 \), the intermediate male 2 is single.

The intuition is as follows. When selecting a mate, a female must take into account two conflicting considerations. On the one hand, she may prefer a less risk averse partner because he is willing to accept a larger share of the income risk. On the other hand, she may prefer a more risk averse partner because he is willing to accept a smaller amount of joint income as long as it is not too uncertain. Depending on the context, one aspect may dominate (in which case the partners at the other extreme are single) or they may compensate each other (then intermediate potential partners are less attractive than the two extremes). The examples above bear this out.\(^{20}\)

4. Extension: One-sided markets

We have so far assumed the matching market is two-sided, where individuals on one side can match only with individuals on the other side. However, it may be even more natural to suppose that any individual can match with any other, i.e., that the market is one-sided.

Let us then consider a one-sided market, but maintain our assumption that risk-sharing groups can be of size at most 2. Hence, we are in fact now considering what is known as the roommate problem. The typical roommate problem cannot be guaranteed

\(^{19}\)To apply Theorems 1 and 2, simply perturb one of the female's \( \phi \) slightly.

\(^{20}\)A natural conjecture stemming from this intuition, but which we have been unable to prove, is that the set of unmatched individuals is an interval.
to possess a stable match. However, we can obtain an existence result here on the basis of our previous analysis.

So, consider our model above, but without the females, and allow any two males to match and share income. Assume further that the incomes of all of the males are jointly independent. We have the following result.

**Theorem 6.** Consider a one-sided one-to-one matching risk-sharing model. If all incomes are i.i.d., if there are an even number, n, of individuals, and if all individuals can be ranked according to their Arrow–Pratt measure of risk aversion, then a unique stable match exists. Moreover, the stable match is negative assortative in the sense that the kth most risk averse individual is matched with kth least risk averse individual for every k.

**Proof.** Create a fictitious two-sided market by adding, for each male k, a female k who is identical to male k in terms of preferences and income distribution. From our previous analysis we know that a stable match exists in this two-sided market and that it is negative assortative. Hence, male i is matched with female j such that \(i + j = n + 1\). But then, because n is even, this same matching pattern together with the sharing rules is feasible for the males in the one-sided market. Moreover, it is also stable in the one-sided market since any blocking opportunity, say between male i and male j in the one-sided market, would imply the existence of an analogous blocking opportunity in the fictitious two-sided market between male i and female j.

---

5. **Concluding remarks**

The theoretical nature of our investigation has led us to consider a very simple model. Nonetheless, even our simple framework has interesting consequences. In particular, risk sharing leads to negative assortative matching. Relatively risk averse individuals are eager to match with less risk averse partners, who can provide the coverage they need at low cost; conversely, relatively risk neutral individuals exploit their comparative advantage by matching with the risk averse, who are willing to give up a large risk premium in exchange for coverage. To the extent that risk sharing may play a role in marital decisions, one would expect intrahousehold differences in risk aversion to be large—a conclusion that fits empirical evidence pretty well.

Our findings appear to be especially relevant to the empirical literature on risk sharing, as discussed in the Introduction, although a complete analysis of the impact of our results on this literature, while surely important, is beyond the scope of the present note. For instance, our results provide support for the “individualistic” approach of Mazzocco (2008), who shows that Euler equations can be estimated at the individual level, even for couples, using labor supply behavior. At the very least, our results suggest that

---

Note that we continue to assume that all males have the same income distribution.

For instance, Mazzocco (2004a), using Health and Retirement Study (HRS) data, shows that even when individuals are gathered into four wide (and potentially heterogenous) classes of risk aversion, half of married men are found to belong to a different category than their spouse. A different investigation, using the Consumer Expenditure Survey, leads to the same conclusion (Mazzocco 2004b).

Mazzocco finds that traditional Euler equations, estimated at the household level, are rejected for couples but not for singles. Moreover, his “individualistic” generalization, which independently analyzes individuals within the couple, is not rejected.
risk-aversion heterogeneity, being endogenous, should be taken seriously, which may lead to a more extensive use of long panels.\textsuperscript{24}

In practice of course, matching is not exclusively based on risk aversion; other characteristics also play a key role. An interesting extension would permit the individuals to choose their income distributions prior to matching. There could then be a trade-off between choosing a distribution that fits one’s risk preferences and making oneself attractive to others for the purposes of risk sharing. This direction is investigated by Wang (2013) in a specific model in which preferences are CRRA and shocks are normally distributed. In that case, clean conditions for assortative matching can be derived. Of particular interest are the welfare implications of her work; for instance, a policy that reduces aggregate risk may, through its impact on matching, worsen the situation of some individuals and increase inequality—a point already stressed by Schulhofer-Wohl (2008) in a different context. All in all, this body of work strongly suggests that when it comes to risk sharing, issues related to endogenous group formation should be taken very seriously.

\textbf{Appendix}

The following well known result will be helpful in proving Lemma 1.

\textbf{Theorem A.1.} If a differentiable real-valued function defined on an interval is nonnegative at \(x_0\) and its derivative is positive whenever the function vanishes, the function is positive at all \(x > x_0\) and can be zero at no more than one point.\textsuperscript{25}

\textbf{Remark.} The interval, \(I\) say, need not be open and \(x_0\) need not be an interior point of \(I\) when the derivative at \(x \in I\) is defined by \(f'(x) = \lim_{x' \to x, x' \in I} (f(x') - f(x))/(x' - x)\).

\textbf{Proof of Lemma 1.} Fix \(v_1\) and \(v_2\) as in the statement of the lemma. If \(v_1 = 0\), then we are done because \(U_{21}(v_1) = Eu_{22}(\tilde{y}) > U_{22}(v_2)\), where the second inequality follows because \(v_2 > 0\) requires the male to strictly share the joint income with the female with positive probability. Hence, we may assume that both \(v_1\) and \(v_2\) are strictly positive. We may similarly assume that \(v_1 < \tilde{v}_1\) and \(v_2 < \tilde{v}_2\).

Let \(x_{ij}: Y \to [0, \infty)\) denote the sharing rule employed by male \(i\) and female \(j\) that maximizes male \(i\)'s utility subject to female \(j\) receiving at least utility \(v_j\). By definition, male \(i\) then receives utility \(U_{ij}(v_j) = Eu_i(x_{ij}(\tilde{y}))\).

We first extend \(x_{ij}(\cdot)\) to all of \([0, \infty)\). As shown in Wilson (1968), there are Pareto weights, \(\lambda_i, \lambda_j > 0\) (strict positivity follows because \(0 < v_j < \tilde{v}_j\)), such that \(x_{ij}(\cdot)\) solves

\[
\max_{x: Y \to [0, \infty)} \lambda_i Eu_i(x(\tilde{y})) + \lambda_j Eu_j(\tilde{y} - x(\tilde{y}))
\]

\textsuperscript{24}See Chiappori et al. (2014b).

\textsuperscript{25}A proof of the first part is as follows. If \(N = \{x > x_0 : f(x) \leq 0\}\) is nonempty, it contains a smallest member, \(\bar{x} > x_0\); otherwise \(f(x_0) = 0\) and \(f'(x_0) \leq 0\), a contradiction. Consequently, \(f(\bar{x}) = 0\) and \(f\) assumes a minimum at \(\bar{x}\) on the interval \([x_0, \bar{x}]\), implying that \(f'(\bar{x}) \leq 0\), a contradiction. Hence, \(N\) is empty. The second part follows immediately from the first.
subject to $0 \leq x(y) \leq y$ for all $y \in Y$. Hence, $x_{ij}(0) = 0$ and because $u_j'(0) = w_j'(0) = +\infty$, $x_{ij}(y)$ is the unique solution to

$$
\lambda_i u_j'(x) = \lambda_j u_j'(y - x)
$$

(A.1)

for almost every positive $y \in Y$. Clearly, we can use (A.1) to uniquely extend $x_{ij}(-)$ to all of $[0, \infty)$. Moreover, $x_{ij}'(y)$ exists for all $y > 0$ by the implicit function theorem. Consider now the difference $x_{11}(y) - x_{12}(y)$ as a function of $y \in [0, \infty)$. A first claim is the following:

**Claim 1.** There exists $\bar{y} \geq 0$, possibly infinite, such that $x_{12}(y) - x_{11}(y)$ is negative for all $y$ in $(0, \bar{y})$ and is positive for all $y$ in $(\bar{y}, \infty)$.

To prove Claim 1, suppose that $x_{12}(\bar{y}) = x_{11}(\bar{y}) = \bar{x}$ for some $\bar{y} > 0$. Because $v_2 > 0$ and $u_j'(0) = w_j'(0) = +\infty$, we know that $0 < \bar{x} < \bar{y}$. Consequently, (A.1) implies that

$$
x_{ij}'(\bar{y}) = \frac{\phi_j(\bar{y} - \bar{x})}{\mu_i(\bar{x}) + \phi_j(\bar{y} - \bar{x})} \in (0, 1)
$$

for $j = 1, 2$,

where $\phi_j(z) = -w_j''(z)/w_j'(z)$ and $u_j(x) = -u_j''(x)/u_j'(x)$ are the female and male Arrow–Pratt measures of risk aversion. Since $\phi_2(\bar{y} - \bar{x}) > \phi_1(\bar{y} - \bar{x})$ by assumption, we must have $x_{12}'(\bar{y}) > x_{11}'(\bar{y})$. Hence, whenever $x_{12}(y) - x_{11}(y)$ vanishes on $(0, \infty)$, its derivative is positive. Claim 1 now follows from Theorem A.1.

By Claim 1, the function $\Delta(y) = u_1(x_{11}(y)) - u_1(x_{12}(y))$ is positive for $0 < y < \bar{y}$ and negative for $y > \bar{y}$. Let $[0, \bar{u}_1)$ denote the range of $u_1(\cdot)$. Then defining $f : [0, \bar{u}_1) \rightarrow \mathbb{R}$ so that $f(u_1(x)) = u_2(x)$ for all $x > 0$, the fact that male 2 is strictly more risk averse than male 1 implies that $f(\cdot)$ is strictly increasing and strictly concave. Because $x_{11}(y)$ is strictly increasing in $y$ and $u_1(\cdot)$ is strictly increasing, the function $g(y) = f[u_1(x_{11}(y))]$ is positive and strictly decreasing in $y$. Hence, if $\bar{y}$ has c.d.f. $H(\cdot)$,

$$
E[g(\bar{y})\Delta(\bar{y})] = \int_0^{\bar{y}} g(y)\Delta(y) \, dH(y) + \int_{\bar{y}}^{\infty} g(y)\Delta(y) \, dH(y)
$$

$$
> \int_0^{\bar{y}} g(\bar{y})\Delta(y) \, dH(y) + \int_{\bar{y}}^{\infty} g(\bar{y})\Delta(y) \, dH(y)
$$

$$
= g(\bar{y})E[u_1(x_{11}(\bar{y})) - u_1(x_{12}(\bar{y}))]
$$

$$
= g(\bar{y})[U_{11}(v_1) - U_{12}(v_2)]
$$

$$
\geq 0,
$$

where the final inequality follows by hypothesis. Hence, substituting the definition of $\Delta(\cdot)$ into the left-hand side above yields

$$
E[g(\bar{y})(u_1(x_{11}(\bar{y})) - u_1(x_{12}(\bar{y})))] > 0.
$$

(A.2)

The concavity of $f(\cdot)$ implies that

$$
f[u_1(x_{11}(y))] - f[u_1(x_{12}(y))] \geq f'[u_1(x_{11}(y))]u_1(x_{11}(y)) - u_1(x_{12}(y))
$$

(A.3)
Since \( u_2 = f \circ u_1 \), we may combine (A.2) and (A.3) to conclude that
\[
E[u_2(x_{11}(\tilde{y})) - u_2(x_{12}(\tilde{y}))] > 0.
\] (A.4)

For \( j = 1 \) and 2, define the value functions for \( 0 \leq \lambda \leq 1 \),
\[
\pi_j(\lambda) = \max_{x,y - [0,\infty)} E[(1 - \lambda)u_1(x(\tilde{y})) + \lambda u_2(x(\tilde{y}))] \quad \text{s.t.} \quad E[w_j(\tilde{y} - x(\tilde{y}))] \geq v_j,
\]
and subject to \( 0 \leq x(y) \leq y \) for almost every \( y \in Y \). A key property of the these value functions is the following claim.

**Claim 2.** If \( \tilde{\lambda} \in [0, 1] \) is such that \( \pi_1(\tilde{\lambda}) = \pi_2(\tilde{\lambda}) \), then \( \pi'_1(\tilde{\lambda}) > \pi'_2(\tilde{\lambda}) \).

To prove **Claim 2**, assume first that \( \tilde{\lambda} = 0 \) and \( \pi_1(0) = \pi_2(0) \). Then \( \pi_j(0) = E[u_1(x_{1j}(\tilde{y}))] \) and, by the envelope theorem (see Milgrom and Segal 2002, Theorem 3),
\[
\pi'_j(0) = E[u_2(x_{1j}(\tilde{y}))] - E[u_1(x_{1j}(\tilde{y}))].
\]
Consequently,
\[
\pi'_1(0) - \pi'_2(0) = (E[u_2(x_{11}(\tilde{y}))] - E[u_1(x_{11}(\tilde{y}))]) - (E[u_2(x_{12}(\tilde{y}))] - E[u_1(x_{12}(\tilde{y}))])
\]
\[
= E[u_2(x_{11}(\tilde{y}))] - E[u_2(x_{12}(\tilde{y}))]
\]
\[
> 0,
\]
where the second equality follows because \( \pi_1(0) = \pi_2(0) \) and the inequality follows from (A.4).

Hence, if \( \pi_1(0) = \pi_2(0) \), then \( \pi'_1(0) > \pi'_2(0) \). Of course, this conclusion holds for all utility functions satisfying our hypotheses. Consequently, if instead \( \tilde{\lambda} > 0 \), and we define \( \tilde{u}_1 = (1 - \tilde{\lambda})u_1 + \tilde{\lambda}u_2 \) and we define for \( j = 1, 2 \) the value function \( \tilde{\pi}_j(\lambda) \) as before, but replacing the utility function \( u_1 \) with \( \tilde{u}_1 \), then, because \( u_2 \) is a concavification of \( u_1 \) (e.g., consider the Arrow–Pratt measures), we may similarly conclude that if \( \tilde{\pi}_1(0) = \tilde{\pi}_2(0) \), then \( \tilde{\pi}'_1(0) > \tilde{\pi}'_2(0) \). But this is equivalent to the statement that if \( \pi_1(\tilde{\lambda}) = \pi_2(\tilde{\lambda}) \), then \( \pi'_1(\tilde{\lambda}) > \pi'_2(\tilde{\lambda}) \). This establishes **Claim 2**.

In view of **Theorem A.1**, a direct consequence of **Claim 2** is that if \( \pi_1(0) \geq \pi_2(0) \) implies \( \pi_1(1) > \pi_2(1) \). Since \( \pi_j(0) = U_{1j}(v_j) \) and \( \pi_j(1) = U_{2j}(v_j) \), this completes the proof of **Lemma 1**.

**Proof of Theorem 4.** If \( F = M \), then all individuals are matched negative assortatively by **Theorem 2** and so there is a unique stable match, which implies that every individual has the same mate across all stable outcomes. So assume without loss of generality that \( M < F \), and suppose that there are two distinct stable matches, \( C \) and \( \hat{C} \), say.

By **Remark 1**, no male is single in either match. We claim that at least one female, \( j_1 \) say, must be single in one match, \( C \) say, and have a mate in the other match, \( \hat{C} \) say. Otherwise, the set of matched individuals is the same in both matches. But then, because

\footnote{Milgrom and Segal's (2002) Theorem 3 applies because, under our assumptions, the solution to the optimization problem defining \( \pi_j(\lambda) \) is unique and, as \( \lambda \to \lambda_0 \), the corresponding sequence of such solutions (being a sequence of nondecreasing functions) has a pointwise convergent subsequence whose limit is a solution to the problem for \( \lambda_0 \). See **footnote 16**.}
each match, being stable, must be negative assortative by Theorem 2, both matches would be identical, a contradiction.

Hence, for some \( N \geq 2 \) there must be \( N \) distinct females \( j_1, j_2, \ldots, j_N \) and \( N - 1 \) distinct males \( i_1, i_2, \ldots, i_{N-1} \) such that (i) \( j_1 \) is single in \( C \), (ii) for \( n \) odd, \( j_n \) is matched to \( i_n \) in \( \hat{C} \), and \( i_n \) is matched to \( j_{n+1} \) in \( C \), and (iii) \( j_N \) is single in \( \hat{C} \). (The sequence can always be chosen to end with a female who is unmatched in \( \hat{C} \) by simply continuing until such a female is reached, which finiteness guarantees will eventually occur. This is why \( N \) is at least 2.) These males and females need not exhaust the sets of males and females or even the sets of males and females who switch mates between the two matches. See Figure 2.

For any female \( j \), let \( v_j \) denote her utility in some stable outcome with match \( C \) and let \( \hat{v}_j \) denote her utility in some stable outcome with match \( \hat{C} \). Then by stability, if male \( i \) is matched to female \( j \) in match \( C \), male \( i \)'s utility must be \( U_{i,j}(v_j) \) and similarly for match \( \hat{C} \). In addition, stability implies that the following inequalities must hold, starting with female \( j_1 \):

\[
\begin{align*}
(I.1) & \quad r_{j_1} \leq \hat{v}_{j_1}, \quad \text{else } j_1 \text{ blocks } \hat{C} \\
(I.2) & \quad U_{i_1j_1}(\hat{v}_{j_1}) \leq U_{i_1j_2}(v_{j_2}), \quad \text{else, given (I.1), } (i_1, j_1) \text{ blocks } C \\
(I.3) & \quad v_{j_2} \leq \hat{v}_{j_2}, \quad \text{else, given (I.2), } (i_1, j_2) \text{ blocks } \hat{C} \\
(I.4) & \quad U_{i_2j_2}(\hat{v}_{j_2}) \leq U_{i_2j_3}(v_{j_3}), \quad \text{else, given (I.3), } (i_2, j_2) \text{ blocks } C \\
& \vdots \\
(I.2N - 2) & \quad U_{i_{N-1}j_{N-1}}(\hat{v}_{j_{N-1}}) \leq U_{i_{N-1}j_N}(v_{j_N}), \\
& \quad \text{else, given (I.2N - 3), } (i_{N-1}, j_{N-1}) \text{ blocks } C \\
(I.2N - 1) & \quad v_{j_N} \leq r_{j_N}, \quad \text{else, given (I.2N - 2), } (i_{N-1}, j_N) \text{ blocks } \hat{C}.
\end{align*}
\]

For example, (I.1) must hold because \( j_1 \) can obtain utility \( r_{j_1} \) on her own. For inequality (I.2), suppose that it fails. Then \( U_{i_1j_1}(\hat{v}_{j_1}) > U_{i_1j_2}(v_{j_2}) \) so that male \( i_1 \) strictly prefers
matching with \( j_1 \) and giving her utility \( \hat{v}_{j_1} \) (which is feasible for them since this is what she receives when matched with \( i_1 \) in \( \hat{C} \)) than the outcome from match \( C \). So if male \( i_1 \) matches with \( j_1 \) and gives her just slightly more income than the sharing rule from match \( \hat{C} \), then \( i_1 \)'s utility would still be strictly above \( \hat{U}_{i_1j_1}(v_{j_1}) \) and \( j_1 \)'s utility, given inequality (I.1), would be strictly above \( r_{j_1} \). Hence, \((i_1, j_1)\) would block \( C \), contradicting the stability of \( C \) and proving (I.2). The remaining inequalities follow analogously.

If we apply the same logic but start instead with female \( j_N \) and work backward toward female \( j_1 \), all of the inequalities are reversed. We conclude that all of the inequalities (I.1)–(I.2N − 1) are in fact equalities. Hence, all of these particular individuals who switch mates are indifferent between the two matches.

What about the payoffs of other individuals who switch mates between the two matches \( C \) and \( \hat{C} \)? Since in the two matches, the mates of females \( j_1, \ldots, j_N \) are either themselves (if single) or one of the males \( i_1, \ldots, i_{N-1} \) and vice versa, we may remove all of these males and females and repeat the argument above for the remaining individuals if any of those remaining individuals switch mates. Hence, any individual who switches mates between \( C \) and \( \hat{C} \) is indifferent between any two stable outcomes with those matches.

Since any stable match \( C' \) must be distinct from either \( C \) or \( \hat{C} \), so let us say \( C \), we may repeat the above argument using \( C \) and \( C' \), and conclude that any individual who switches mates between \( C \) and \( C' \) must be indifferent between any two stable outcomes with those two matches and, hence, by transitivity, indifferent among all stable outcomes with matches that are either \( C \), \( \hat{C} \), or \( C' \). We conclude that if an individual does not have the same mate in every stable match, then that individual has the same payoff in every stable outcome.

\[ \square \]

**Proof of Theorem 5.**

**Preliminaries.** In Section 2, we defined \( U_{ij}(v_j) \) to be male \( i \)'s maximum expected utility when he is matched with female \( j \), where the maximum is taken over all possible sharing rules that ensure female \( j \) an expected utility of at least \( v_j \in [0, \hat{v}_j] \). It will be helpful to make explicit the dependence of \( U_{ij} \) on the utility functions \( u_i \) and \( w_j \) of male \( i \) and female \( j \), so we now write \( U_{ij}(v_j, u_i, w_j) \).

It is straightforward to establish that for the given metric on \( \mathcal{U} \times \mathcal{W} \), each value function \( U_{ij}(v_j, u_i, w_j) \) is continuous, i.e., if \( v^k \to_k v_j \in [0, \infty) \) and \((u^k, w^k) \to_k (u, w) \in \mathcal{U} \times \mathcal{W} \), then \( U_{ij}(v^k_j, u^k_i, w^k_j) \to_k U_{ij}(v_j, u_i, w_j) \).

Next, we formalize the set of solutions to the equations that we encountered in the proof of Theorem 4. For any \((u, w) \in \mathcal{U} \times \mathcal{W} \), any \( N \geq 2 \), any \( N - 1 \) distinct males \( i_1, \ldots, i_{N-1} \), and any \( N \) distinct females \( j_1, \ldots, j_N \), let \( \mathcal{V}(i_1, \ldots, i_{N-1}, j_1, \ldots, j_N, u, w) \) denote the set of vectors \((v_1, \ldots, v_N, \hat{v}_1, \ldots, \hat{v}_N) \in (\times_{j=1}^{\hat{F}}[0, \hat{v}_j])^2 \) that solve the following system of \( 2N - 1 \) equations, where \( \hat{v}_j = E[w_j(\hat{y})] \) is the maximum utility that female \( j \) can

\[ \text{The proof follows the logic of Berge's theorem of the maximum. See footnote 16.} \]
obtain in any match and \( r_j = E[w_j(\tilde{z}_j)] \) is her reservation utility when her utility function is \( w_j \):

\[
\begin{align*}
E.1 & \quad r_{j_1} = \hat{v}_{j_1} \\
E.2 & \quad U_{i_1j_1}(\hat{v}_{j_1}, u_{i_1}, w_{j_1}) = U_{i_1j_2}(v_{j_1}, u_{i_1}, w_{j_2}) \\
E.3 & \quad v_{j_2} = \hat{v}_{j_2} \\
E.4 & \quad U_{i_2j_2}(\hat{v}_{j_2}, u_{i_2}, w_{j_2}) = U_{i_2j_3}(v_{j_2}, u_{i_2}, w_{j_3}) \\
E.5 & \quad v_{j_3} = \hat{v}_{j_3} \\
\vdots \\
E.2N-2 & \quad U_{i_{N-1}j_{N-1}}(\hat{v}_{j_{N-1}}, u_{i_{N-1}}, w_{j_{N-1}}) = U_{i_{N-1}j_N}(v_{j_{N-1}}, u_{i_{N-1}}, w_{j_N}) \\
E.2N-1 & \quad v_{j_N} = r_{j_N}.
\end{align*}
\]

Observe that some of the variables \( v_j \) and \( \hat{v}_j \) will not appear in any equation if \( j \) is not in \( \{j_1, \ldots, j_N\} \).

Call a vector, \( \sigma \), of male and female indices feasible if for some \( N \geq 2 \), \( \sigma = (i_1, \ldots, i_{N-1}, j_1, \ldots, j_N) \), where \( i_1, \ldots, i_{N-1} \in \{1, \ldots, M\} \) are distinct indices for a subset of males and \( j_1, \ldots, j_N \in \{1, \ldots, F\} \) are distinct indices for a subset of females. Let \( \Sigma \) denote the finite set of all feasible \( \sigma \). Hence, for any \( (u, w) \in \mathcal{U} \times \mathcal{W} \), each \( \sigma \in \Sigma \) determines a set of equations of the form (E.1)–(E.2N–1) with solution set \( V(\sigma, u, w) \).

With these preliminaries out of the way, suppose first that \( F = M \). Then all individuals are matched negative assortatively by Theorem 2 and so there is a unique stable match for any \( (u, w) \in \mathcal{U} \times \mathcal{W} \). Thus letting \( G = \mathcal{U} \times \mathcal{W} \) suffices and the proof is complete.

Of the two remaining possibilities, \( F > M \) and \( M > F \), we need consider only one of them since their proofs are analogous.

So suppose that \( F > M \) and that the vector of utility functions is \( (u, w) \in \mathcal{U} \times \mathcal{W} \). By the proof of Theorem 4, there can be more than one stable match only if some system of equations of the form (E.1)–(E.2N–1) has a solution, indeed a solution in which all female utility values \( v_j \) and \( \hat{v}_j \) are nonnegative (since utility is always nonnegative) and no greater than \( \bar{v}_j = E[w_j(\tilde{y})] \), the maximum utility that female \( j \) can receive with utility function \( w_j \). Consequently, there is a unique stable match if \( V(\sigma, u, w) \) is empty for every \( \sigma \in \Sigma \).

Let

\[
G = \{(u, w) \in \mathcal{U} \times \mathcal{W} : V(\sigma, u, w) = \emptyset \forall \sigma \in \Sigma \}.
\]

By the conclusion of the previous paragraph, it suffices to show that \( G \) is open and dense. We break the remainder of the proof into two parts.

---

28Since the \( U_{ij} \) functions are strictly decreasing in their first argument, it is not difficult to see that the system of equations (E.1)–(E.2N–1) has at most one solution. Hence, \( V(\sigma, u, w) \) is either a singleton or the empty set.
Part I: G is open. Since \( G = \bigcap_{\sigma \in \Sigma} \{(u,w) \in \mathcal{U} \times \mathcal{W} : V(\sigma, u, w) = \emptyset\} \) and because \( \Sigma \) is finite, it suffices to show that for any \( \sigma \in \Sigma \), the set

\[
G_\sigma = \{(u,w) \in \mathcal{U} \times \mathcal{W} : V(\sigma, u, w) = \emptyset\}
\]

is open.

Fix any \( \sigma \in \Sigma \). To show that \( G_\sigma \) is open, we will show that its complement \( G_\sigma^c = (\mathcal{U} \times \mathcal{W}) \setminus G_\sigma \) is closed. So suppose that \((u^k, w^k)\) is a sequence of points in \( G_\sigma^c \), converging to \((u,w) \in \mathcal{U} \times \mathcal{W} \). We must show that \((u,w) \in G_\sigma^c \).

Consider any \( k \) and suppose that the vector of utility functions is \((u^k, w^k)\). Then male \( i \)'s value function when he is matched with female \( j \) is \( U_{ij}(\cdot, u^k_i, w^k_j) \).

For any female \( j \), let \( r_j^k = E[w_j^k(\hat{z}_j)] \) denote her reservation utility and let \( \bar{v}_j^k = E[w_j^k(\bar{y})] \) denote the maximum utility that she can obtain in any match. Because \((u^k, w^k) \rightarrow (u, w)\), we have \( r_j^k \rightarrow r_j = E[w_j(z)] \) and \( \bar{v}_j^k \rightarrow \bar{v}_j = E[w_j(y)] \) for each female \( j \).

Since \((u^k, w^k) \in G_\sigma^c \), the solution set \( V(\sigma, u^k, w^k) \) is nonempty and so we may suppose that \((v_1^k, v_2^k, \ldots, v_F^k, \hat{v}_1^k, \ldots, \hat{v}_F^k)\) is a member. Since each \( \bar{v}_j^k \) and \( \hat{v}_j^k \) are in \([0, \bar{v}_j^k]\) and each \( \bar{v}_j^k \) converges to \( \bar{v}_j \), we may assume without loss of generality, by taking a subsequence if necessary, that each \( v_j^k \) converges to some \( v_j \in [0, \bar{v}_j] \) and that each \( \hat{v}_j^k \) converges to some \( \hat{v}_j \in [0, \bar{v}_j] \).

Since \((u^k_i, v^k_j) \rightarrow (u_i, w_j)\) for every male \( i \) and female \( j \), the continuity of \( U_{ij}(\cdot, \cdot, \cdot) \) implies that \( U_{ij}(v^k_j, u^k_i, w^k_j) \) converges to \( U_{ij}(v_j, u_i, w_j) \) and \( U_{ij}(\hat{v}^k_j, u^k_i, w^k_j) \) converges to \( U_{ij}(\hat{v}_j, u_i, w_j) \) for every male \( i \) and every female \( j \). Therefore, because the continuity of each \( U_{ij} \) clearly implies that \( V(\sigma, u, w) \) is upper hemi-continuous in \((u, w)\), we may conclude that \((v_1, \ldots, v_F, \hat{v}_1, \ldots, \hat{v}_F) \in V(\sigma, u, w)\), proving that \((u,w) \in G_\sigma^c \), as desired. Hence, \( G_\sigma^c \) is closed and so \( G_\sigma \) is open. But then, as already observed, \( G \) too is open. It remains only to establish that \( G \) is a dense subset of \( \mathcal{U} \times \mathcal{W} \).

Part II: \( G \) is dense. For any \( u \in \mathcal{U} \), for any \( \lambda \in (0,1) \), and for any male \( i \), let \( u^\lambda_i = (1-\lambda)u_i + \lambda u_{i+1} \), where \( u_{M+1}(x) = \sqrt{\mu_{M}(x)} \) \( \forall x \geq 0 \). For any \( i < M \), recall that we have indexed the males so that male \( i + 1 \) with utility function \( u_{i+1} \) is strictly more risk averse than male \( i \) with utility function \( u_i \). If \( i = M \), then \( i \) is the most risk averse male and \( u^\lambda_i = (1-\lambda)u_M + \lambda \sqrt{\mu_M} \). So for every \( i \), \( u^\lambda_i \) is strictly more risk averse than \( u_i \) (simply compare the Arrow–Pratt measures) and \((u^\lambda_i, u_{-i}, w) \in \mathcal{U} \times \mathcal{W} \) since the strict ordering of the males via risk aversion does not change when \( u_i \) is replaced with \( u^\lambda_i \) for \( \lambda \in (0,1) \).

We first show that for every \((u, w) \in \mathcal{U} \times \mathcal{W} \) and for every \( \sigma \in \Sigma \), there is a male \( i \) such that

\[
V(\sigma, (u^\lambda_i, u_{-i}), w) = \emptyset \quad \text{for all } \lambda \in (0,1) \text{ sufficiently small.} \tag{A.5}
\]

Consider any \((u, w) \in \mathcal{U} \times \mathcal{W} \) and any \( \sigma = (i_1, \ldots, i_{N-1}, j_1, \ldots, j_N) \in \Sigma \). There are two cases.

Case I: \( V(\sigma, u, w) = \emptyset \). If \( V(\sigma, u, w) = \emptyset \), then \( (A.5) \) follows from our result from Part I that \( G_\sigma \) is open and because, for any male \( i \), \((u^\lambda_i, u_{-i}, w) \rightarrow (u, w) \) as \( \lambda \rightarrow 0 \).

Case II: \( V(\sigma, u, w) \neq \emptyset \). If \( V(\sigma, u, w) \neq \emptyset \), then it contains some \((v_1, \ldots, v_F, \hat{v}_1, \ldots, \hat{v}_F) \), say. Hence, \((v_1, \ldots, v_F, \hat{v}_1, \ldots, \hat{v}_F) \) solves \((E.1)-(E.2N-1)\).
Since $U_{ij}(\cdot, u_i, w_j)$ is strictly decreasing, starting from $(E.2N - 1)$ and working backward until $(E.3)$, we see that all of the values $v_{jN}, \hat{v}_{j_{N-1}}, \ldots, \hat{v}_{j_2}, v_{j_2}$ are uniquely determined. In particular, the value of $v_{j_2}$ is uniquely determined independently of $(E.2)$ and so independently of the utility function $u_i$ of male $i_1$ that appears in $(E.2)$. But since both $(E.1)$ and $(E.2)$ hold, we must have

$$U_{ij_1}(r_{j_1}, u_{i_1}, w_{j_1}) = U_{ij_2}(v_{j_2}, u_{i_1}, w_{j_2}). \tag{A.6}$$

Also, $0 < r_{j_1} < \hat{v}_{j_1} = E[w_{j_1}(\hat{y})]$ (see the paragraph following Remark 1) and $0 < v_{j_2} < \hat{v}_{j_2} = E[w_{j_2}(\hat{y})]$.\footnote{To see the latter, observe that because each $U_{ij}(\cdot)$ is strictly decreasing, $U_{ij_1}(\hat{v}_{j_1}) = 0 = U_{ij_2}(\hat{v}_{j_2}) < U_{ij_1}(r_{j_1}) < U_{ij_1}(0) = Ew_i(\hat{y}) = U_{ij_2}(0)$. So (A.6) implies $U_{ij_1}(\hat{v}_{j_1}) < U_{ij_1}(v_{j_2}) < U_{ij_2}(0)$, which gives $0 < v_{j_2} < \hat{v}_{j_2}$.} Therefore, since $u^\lambda_{i_1}$ is more risk averse than $u_i$, Lemma 1 and (A.6) imply that for all $\lambda \in (0, 1)$,

$$U_{ij_1}(r_{j_1}, u^\lambda_{i_1}, w_{j_1}) \neq U_{ij_2}(v_{j_2}, u^\lambda_{i_1}, w_{j_2}). \tag{A.7}$$

Since perturbing $u_{i_1}$ does not affect any of $(E.3)$–$(E.2N - 1)$, any solution to the perturbed set of equations has the same value of $v_{j_2}$. But then (A.7) implies that for every $\lambda \in (0, 1)$, the perturbed system has no solution. That is,

$$V(\sigma, (u^\lambda_{i_1}, u_{-i_1}), w) = \emptyset \quad \text{for all } \lambda \in (0, 1).$$

This completes the proof of (A.5) and we are now ready to show that $G$ is dense.

Let $G^c = (\mathcal{U} \times \mathcal{W}) \setminus G$ denote the complement of $G$ and let $\epsilon > 0$ be given. Suppose $(u, w) \in G^c$. We must show that some element of $G$ is within distance $\epsilon$ of $(u, w)$.

Since $\Sigma$ is a finite set, we may write it as $\Sigma = \{\sigma^1, \ldots, \sigma^K\}$. Let $\lambda$ be an upper bound for all the male utility functions $u_i, \ldots, u_M$ and also for $\sqrt{u_M}$, and choose $\bar{\lambda} \in (0, 1)$ such that $\bar{\lambda} < \epsilon/(2LK)$. We will use (A.5) repeatedly to successively perturb $u$ for each element of $\Sigma$.

Define the mapping $T: \mathcal{U} \times (0, 1) \times \{1, \ldots, M\} \to \mathcal{U}$ by

$$T(u, \lambda, i) = (u^\lambda_i, u_{-i}).$$

Starting with $\sigma^1$, by (A.5) there is a male $i(1)$ and $\lambda_1 \in (0, \bar{\lambda})$ such that

$$V(\sigma^1, u^1, w) = \emptyset,$$

where $u^1 = T(u, \lambda_1, i(1))$. We next define $u^2, \ldots, u^K$ inductively.

Given $k \in \{2, \ldots, K\}$ and $u^{k-1} \in \mathcal{U}$ satisfying

$$V(\sigma^K, u^{k-1}, w) = \emptyset \quad \forall n = 1, \ldots, k - 1, \tag{A.8}$$

define $u^k \in \mathcal{U}$ as follows.

By (A.5), there is a male $i(k)$ such that

$$V(\sigma^k, T(u^{k-1}, \lambda, i(k)), w) = \emptyset \quad \forall \lambda \in (0, 1) \text{ sufficiently small.} \tag{A.9}$$
By Part I above, $G_{\sigma^n}$ is open for every $n = 1, \ldots, k - 1$. Consequently, because $T(u^{k-1}, \lambda, \iota(k)) \to u^{k-1}$ as $\lambda \to 0$, (A.8) implies that

$$V(\sigma^n, T(u^{k-1}, \lambda, \iota(k)), w) = \emptyset \quad \forall n = 1, \ldots, k - 1, \forall \lambda \in (0, 1)$$ sufficiently small. (A.10)

Together, (A.9) and (A.10), imply that there is $\lambda_k \in (0, \bar{\lambda})$ such that,

$$V(\sigma^n, T(u^{k-1}, \lambda_k, \iota(k)), w) = \emptyset \quad \forall n = 1, \ldots, k.$$

Define $u^k = T(u^{k-1}, \lambda_k, \iota(k))$. We therefore have

$$V(\sigma^n, u^k, w) = \emptyset \quad \forall n = 1, \ldots, k,$$ (A.11)

which completes the inductive definition.

Setting $k = K$ in (A.11), we see that $u^K \in U$ satisfies

$$V(\sigma, u^K, w) = \emptyset \quad \forall \sigma \in \Sigma,$$

which implies that $(u^K, w) \in G$. It remains only to show that $(u^K, w)$ is within $\varepsilon$ of $(u, w)$.

Let $u^0 = u$. For each $k \in \{1, \ldots, K\}$, each utility function that is a coordinate of $u^k$ is bounded by $L$ because each such coordinate function is a convex combination of the coordinates of $u$, each of which (as well as $u_{M+1}$) is bounded by $L$. Also, for any male $i$, either $u^k_i = u^{k-1}_i$ or $u^k_i = (1 - \lambda)u^{k-1}_i + \lambda u^1_{i+1}$ for some $\lambda \in (0, \bar{\lambda})$. Consequently,

$$\sup_{i, x \geq 0} |u^k_i(x) - u^{k-1}_i(x)| \leq 2\bar{\lambda}L < \varepsilon / K,$$

where the strict inequality follows from the definition of $\bar{\lambda}$. Hence,

$$\sup_{i, x \geq 0} |u^K_i(x) - u_i(x)| \leq \sum_{k=1}^{K} \sup_{i, x \geq 0} |u^k_i(x) - u^{k-1}_i(x)| < \varepsilon,
$$

which implies that

$$\| (u^K, w) - (u, w) \| = \sup_{i, x \geq 0} |u^K_i(x) - u_i(x)| < \varepsilon,$$

which completes the proof. \qed

References


