



Munich Personal RePEc Archive

Continuous Matching with Single Peaked Preferences

Flanders, Sam

University of North Carolina, Chapel Hill

September 2013

Online at <https://mpra.ub.uni-muenchen.de/53668/>

MPRA Paper No. 53668, posted 16 Feb 2014 01:31 UTC

CONTINUOUS MATCHING WITH SINGLE PEAKED PREFERENCES

SAM FLANDERS

University of North Carolina at Chapel Hill

September 2013

ABSTRACT. This paper derives several simple matching algorithms for special cases of a continuous two sided matching model where agents have single peaked preferences and face no costs or restrictions to matching. These algorithms allow matches to be computed as an explicit function of the parameters of the model, making them convenient for use in theoretical matching market models, such as dating markets. One such theoretical model is solved and analyzed. Finally, several possible extensions and topics for further study are explored, such as adding noise to observed types and allowing misreporting of type.

1. INTRODUCTION

The goal of this paper is to establish tractable stable matching algorithms for a two sided single period matching model with continuous distributions of agents over a single parameter type and nontransferable utility. Further, these agents can freely propose matches to any agent on the opposing side and have strictly single-peaked preferences. By “tractable”, I mean that, for each agent on one side, their set of possible (indifferent) matches can be computed directly from the type of the agent and the parameters of the model (e.g. preferences and distributions of agents) without recourse to iterative, highly computational algorithms. Thus, these algorithms should be convenient for use in theoretical matching market models.

Examples of appropriate models for such an algorithm is are the highly researched dating and marriage markets. While many dating models use horizontal preferences, single peaked preferences allow for the exploration of matching on characteristics like age, cultural background, ethnicity, hobbies, personality type, and life goals, for which there are no universal ordinal rankings and where the distributions may not be identical for both sides. Also, in online dating platforms agents can use search filters or the platform’s matching algorithm to choose compatible type individuals out of the entire distribution at virtually no cost, justifying the use of a model where agents choose from the entire distribution rather than, say, a finite random draw from that distribution.

This paper follows a rich literature on stable matching problems, starting with the seminal paper by Gale and Shapley [5], which introduced an algorithm for deriving stable matches given arbitrary preferences and a finite set of agents. While this algorithm is highly general, it utilizes an iterative process that must be run before matches can be derived. Thus, while it is quite useful for empirical analysis, it does not lend itself to being embedded in a theoretical model. Becker [1] found that positive assortative matching occurs when there is a continuum of types and the utility of a match is increasing in types and nontransferable—that is, when the two matching agents can’t bargain over the apportionment of the utility of the match—and that assortative matching also occurs when utility is transferable and the total utility of a match exhibits increasing differences in the two agents’ types. Unlike Gale-Shapley, this requires no iterative process to pair up agents, so it is suitable for use in theoretical models, but the fairly onerous restriction of vertical preferences—higher types are universally preferred to lower types—limits its application. Legros and Newman [6] extended positive and negative assortative matching results to a class of partially nontransferable utility problems, where there are limitations on the ability of some or all agents to transfer utility to their match. Assuming horizontal preferences where agents want to match to their own type rather than the vertical preferences of Becker, Clark (2003) [2] gives an algorithm for finding stable matchings in a market with a finite set of agents. Clark (2007) [4] then treats the horizontal case with an infinite set of agents, finding a very simple matching result. Clark (2006) [3] also gives a condition guaranteeing a unique stable matching. Klumpp [7] derives a very simple “inside-out” algorithm for horizontal matching with finitely many agents.

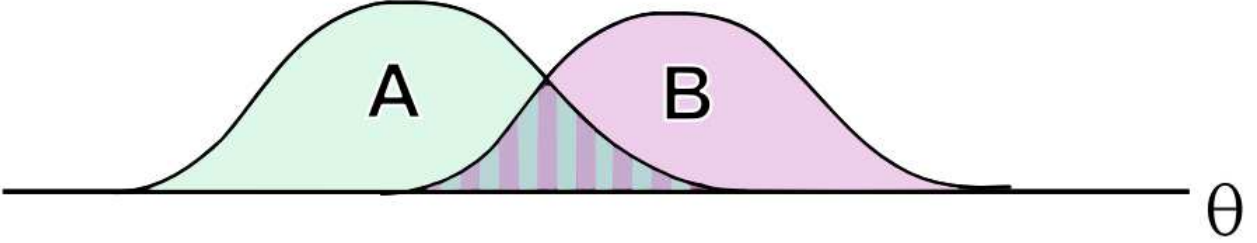


FIGURE 2.1. Initial distributions A and B of agents over type θ . There is an area of overlap in the center where, for any type in the overlap, both distributions have at least as much mass as the overlapped region.

The remainder of this paper is organized as follows: Section 2 demonstrates Clark’s matching algorithm for the simplest case where preferences are homophilic—that is, where the peak preference is the agent’s own type. Section 3 generalizes the model by allow agents to have arbitrary single peaked preferences, and matching algorithms are derived given some additional assumptions. Section 4 relates the single-peaked matching result to the horizontal and vertical preference literature. Section 5 provides interpretation for the results and empirical implications. Section 6.1 sets up a simple dating platform model and section 6.2 solves it using vertical preferences and then solves a generalization of it the single-peaked preferences algorithm. Section 6.3 provides analysis of models and contrasts them. Finally, the concluding section describes directions for further study.

2. MODEL WITH HOMOPHILIC PREFERENCES

2.1. Baseline Model¹. Before we address more general single peaked preferences, it is instructive to review the horizontal preference matching algorithm first derived in Clark 2007 [4]. Consider a two sided matching model with two continuous, integrable distributions A and B over a shared univariate typespace Θ , with A and B representing the masses of agents on each side. Denote an agent i from the mass of agents of type θ on side S $s_{i\theta}$, s_θ if supressing the index is appropriate. Let preferences be strictly single peaked. That is, an agent s with peak preference θ_{peak} for $\theta_1, \theta_2 < \theta_{\text{peak}}$, if $\theta_{\text{peak}} - \theta_1 > \theta_{\text{peak}} - \theta_2$ then $s_1 \prec_s s_2$, and for $\theta_1, \theta_2 > \theta_{\text{peak}}$, if $\theta_1 - \theta_{\text{peak}} > \theta_2 - \theta_{\text{peak}}$ then $s_1 \prec_s s_2$. Suppose \forall agents $s_{i\theta}$, $\theta_{\text{peak}} = \theta$ —that is, agents have homophilic preferences. Suppose further that agents face no search costs or other limitations to matching, i.e. suppose agents optimize over the entire set of agents who are willing to match to them. Note that, while we can normalize either A or B to measure 1 without loss of generality, making both measure 1 is a simplifying assumption, requiring an equal mass of agents on each side. We will proceed for now using this assumption as it simplifies the problem, and relax it later.

- **Assumption 1 (MASS)** : Suppose an equal measure of agents on each side.

This scenario affords an extremely simple solution. First, we match and remove from consideration the area under both curves, if such an area exists.

Lemma 1. *For each type θ , a measure $\mu_\theta = \min(\mu_{\theta A}, \mu_{\theta B})$ of θ agents on side A (B) matches to θ types on side B (A), where $\mu_{\theta S}$ is the mass of agents of type θ on side S .*

Proof. There are at least μ_θ agents of type θ on each side by definition. Since preferences are homophilic, $a_{i\theta}$ strictly prefers $b_{j\theta}$ to any agent b_{ks} , and symmetrically $b_{i\theta}$ prefers $a_{j\theta}$ to any a_{ks} . Then a mass μ_θ of a_θ ’s will strictly prefer to match with any of the measure μ_θ of b_θ ’s, and the b_θ ’s will symmetrically strictly prefer a_θ ’s to any other agents, so they will form stable matches. \square

¹I derived these horizontal results independently, being unaware of Clark’s unpublished paper. However I believe the derivation is a good motivation for the later, novel result, so I have retained this section.

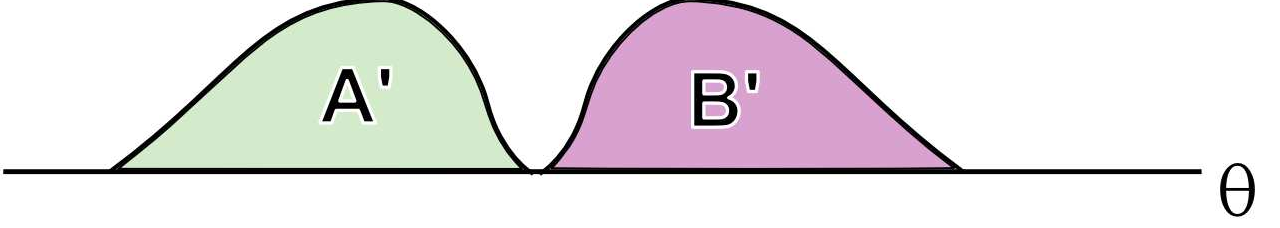


FIGURE 2.2. The remainder distributions A' and B' that are left when the mass of overlapping agents is matched and removed from the market.

Now we can eliminate the stably matched overlap agents from consideration. The remainder distributions can be defined as

$$f_{A'}(\theta) = \max\left\{\frac{f_A(\theta) - f_B(\theta)}{\int_{\Theta} \max\{f_A(\theta) - f_B(\theta), 0\} d\theta}, 0\right\}$$

and

$$f_{B'}(\theta) = \max\left\{\frac{f_B(\theta) - f_A(\theta)}{\int_{\Theta} \max\{f_B(\theta) - f_A(\theta), 0\} d\theta}, 0\right\}$$

where the integral ensures a well-defined probability density function with mass 1. Note that the area under both A and B is the same for both distributions, so the scalar they must be multiplied by is also the same and we don't have any issues of miscounting the measures of agents on each side. Define A' and B' as the distributions with these respective densities.

We now inductively derive a very simple matching algorithm that yields the type s of the match b_{js} for $a_{i\theta}$ (a_{js} for $b_{i\theta}$) as an explicit function depending only on the remainder distributions and $a_{i\theta}$ ($b_{i\theta}$). The intuition here is that we start at the far right of the left remainder distribution (A in this example) and the far left of the right distribution (B in this example), or on other words the innermost points of each distribution, and then iteratively match outward, with the current (innermost remaining) matchers taking the already stably matched interior agents as unavailable. Because agents want the closest match possible, the current matchers on each side strictly prefer the current (innermost remaining) matchers on the other side to anyone else, so they match and the process continues. Note that, because matching is one-to-one, the measure of agents who have been matched on one side must equal the measure of agents who have been matched on the other. Before we complete this proof, we make two additional assumptions.

- **Assumption 2 (SEP):** Suppose the probability density functions have the single crossing property i.e. the probability density functions intersect at only one point.
- **Assumption 3 (OUT):** Suppose that agents prefer any match to no match.

Like MASS, SEP and OUT are not necessary for a tractable answer, but they allow for a very simple baseline result to be derived, against which deviations from these assumptions can later be compared.

Lemma 2. *Without loss of generality, assume A' is to the left of B' . Suppose that all agents in the interval (θ_A, θ_B) have been stably matched and are eliminated from consideration, while no other agents in A' or B' have matched. Then a mass of agents $\min\{f_A(\theta_A), f_B(\theta_B)\}$ of types θ_A and θ_B will match stably.*

Proof. We know $a_{i\theta_A}$ prefers $b_{j\theta_B}$ to any other $b_{k\theta}$ and vice versa, as they are mutually distance-minimal among the set of remaining potential matches, and all agents that have already been matched are closer to their match than they are to any remaining potential match by construction, so they will not prefer to deviate to one of the current matchers. Thus the agents will match stably, as was to be shown. \square

We are now ready to present the algorithm and prove its validity.

Proposition 3. (Homophilic Matching) *Suppose MASS, SEP, and OUT. A measure of agents equal to the measure under both curves and with equal density over θ will match to their own type. For all remaining agents of all types θ , θ_A agents match to agents of type $F_B^{-1}(1 - F_A(\theta_A))$ and θ_B agents match to agents of type $F_A^{-1}(1 - F_B(\theta_B))$.*

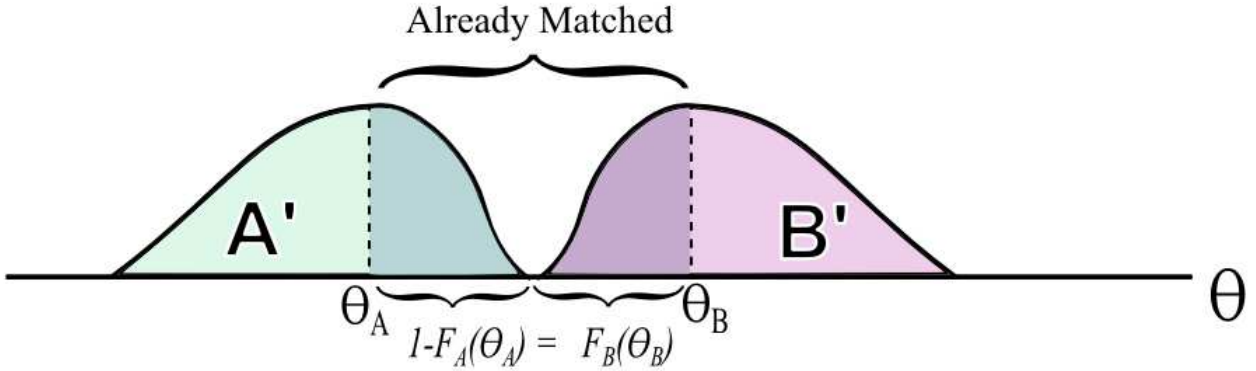


FIGURE 2.3. The stage in the matching process when types θ_A and θ_B are the innermost unmatched types. Note that the darkened areas that have already been matched are of equal mass and θ_A and θ_B are mutually closest to one another among the remaining agents.

Proof. The first portion of Proposition 3 is simply Lemma 1. The second is obtained by inductively applying Lemma 2 starting at the innermost points on the two remainder distributions and moving outwards, and by using the fact that the measures of agents matched on each side, $1 - F_A(\theta_A)$ for A and $F_B(\theta_B)$ for B , must be equal. \square

2.2. Extensions to the Baseline Model. To get the result above, we made three fairly restrictive assumptions. We will now relax them and find the matching outcome in the more general cases. The matching algorithm remains quite simple, although relaxing *SEP* will require a new assumption.

MASS sets the measure of agents on each side of the matching market equal. This is a reasonable assumption in the broader heterosexual dating market, for example, where it is approximately true. However, if one wants to model, say, online dating platforms that attract men and women in disproportionate numbers, or the marriage market in countries that have a significant deficit of men due to war or women due to sex selective abortions, then *MASS* must be abandoned. All this will do is leave the outermost agents of the larger side unmatched, as the supply of agents on the other side will have run out. Specifically, without loss of generality suppose $\mu_B > \mu_A$ and A is to the left of B . Then the rightmost $\mu_B - \mu_A$ B agents—that is, B agents to the right of $\theta_B = F_B^{-1}(\mu_B - \mu_A)$ —will be unmatched, while the rest will match as before.

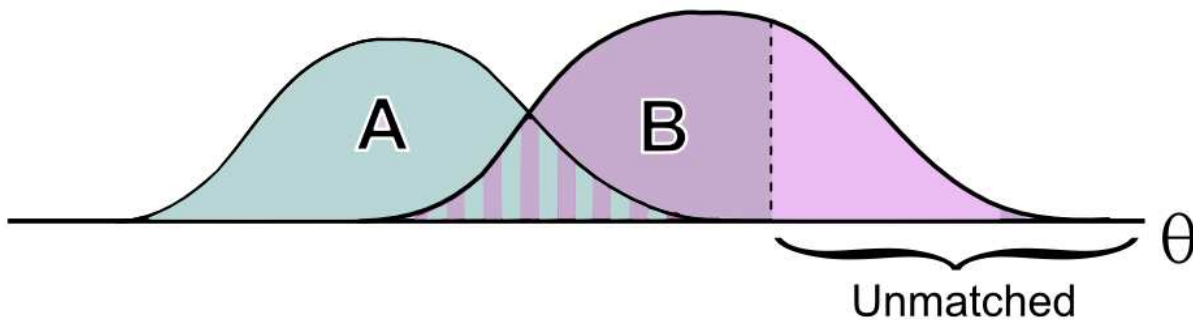


FIGURE 2.4. Matching with unbalanced distributions. Here B has more mass than A , so the furthest (least attractive) B agents do not receive a match unless they are matched in stage one, the overlap matching phase.

SEP ensures that all agents of one distribution are above or below all agents of the other. If this does not hold, we may have a situation like Figure 2.5 where some A agents are above all B agents and some

A agents are below all B agents, and of course much more complicated situations of the same nature could occur. In order to find the matching here, we need to be able to relate preferences for types on the left of the agent to preferences for types to the right of the right of the agent. Assuming we have a utility function or some other means to compare potential matches to the left and right, we can find cutoff agents who are indifferent between their best available match on the right and left. In some cases we can proceed as before from the innermost points on each pair of adjacent “islands” in the two distributions, with indifferent agents determining cutoffs where agents switch from the available match on one side to the match on the other. However, we may run into situations where agents in one island match to agents in a nonadjacent island. This is not the primary focus of this paper, so we will not explore this issue any further.

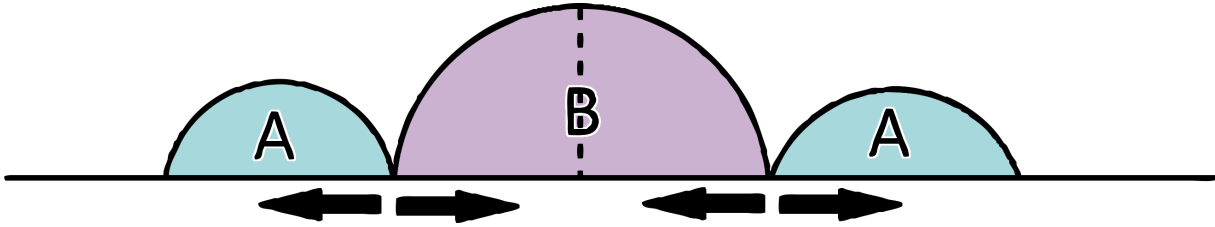


FIGURE 2.5. Matching without the single crossing property—the remainder of A is separated into two “islands”, with the remainder of B in the middle. B agents to the left of the dotted line will match to the left A island, while B agents to the right will match to the right island.

OUT requires that all agents accept whatever the best match available to them is. However the most obvious qualitative characteristic of the matching outcome in this model is that, for the agents of types that are overrepresented relative to that type on the other side, the non-perfect matches quickly deteriorate in quality for fringe agents, as the best remaining match moves further away from them the further to the outside they are. The outermost agents will in fact get their worst possible match, so it seems reasonable that at a certain point agents will prefer no match to a terrible one. The result of dropping *OUT*, assuming that the reservation distance is the same for all types, is simply that matching will terminate once the distance between the innermost remaining agents is equal to the reservation distance, with the rest remaining unmatched.

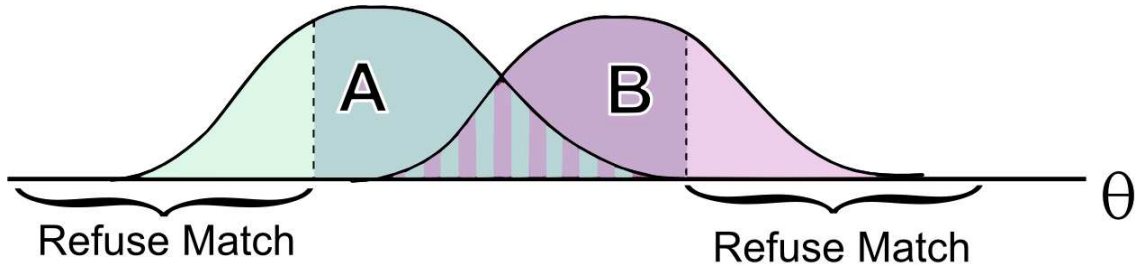


FIGURE 2.6. Matching when agents have the option to refuse. Agents beyond the distance where the cutoff agents are indifferent between matching or staying single do not clear the matching market.

3. GENERALIZATION TO ARBITRARY SINGLE PEAKED PREFERENCES

3.1. Baseline Model. We now allow agent type and agent preference to vary independently, generalizing to arbitrary single peaked preferences. This allows agents to prefer types other than their own. For example,

men may prefer women of a different level of femininity than their own, or may prefer someone of a complementary disposition to make up for their shortcomings. Also, two individuals with the same characteristics may have different preferences over their match’s characteristics, rather than e.g. a man’s height uniquely determining his height preference. Note that, since the “type” of an individual is now a pair of the form (characteristic, peak preference), so if we continue using the term “type” for an agent’s characteristic, we introduce ambiguity. For that reason, an agent’s characteristic (e.g. height, BMI, etc.) will now be referred to as their trait.

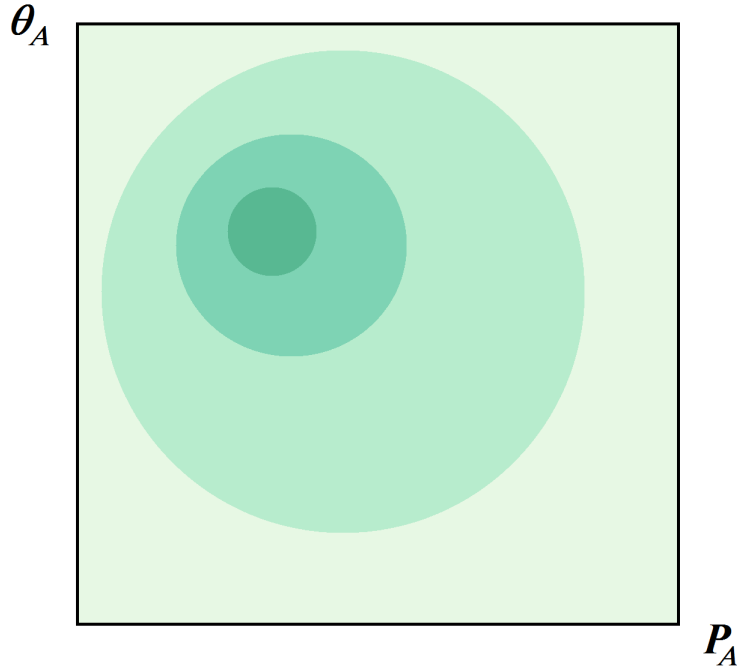


FIGURE 3.1. Contour plot of the trait/preference distribution A. An A agent’s trait runs along the vertical axis and its peak preference over the traits of potential matches runs along the horizontal axis. Darker colors indicate greater mass.

Now that agents have two characteristics, the set of agents on a given side is a bivariate distribution over own trait and peak preference. Denote an agent i with trait θ and peak preference p as $s_{i\theta p}$. Trait in this situation is the sole characteristic over which an agent’s potential matches have preferences, while peak preference determines that agent’s most preferred match. To facilitate easy visualization of the algorithm to be derived, we will overlay the distributions A and B , flipping the axes for B . This will put A agent traits and B agent peak preferences on the vertical axis, and B agent traits and A agent peak preferences on the horizontal axis.

The reason for representing the distributions like this is that A agents evaluate matches based on the distance between their preference and a B agent’s trait, which is now the horizontal distance between $a_{i\theta p}$ and $b_{j\theta p}$ on our graph, and B agents evaluate matches based on the distance between their preference and an A agent’s trait, which is now the vertical distance between $a_{i\theta p}$ and $b_{j\theta p}$, so we can use the graph to easily compare agent preferences over potential matches.

In general, this is a more complicated problem, and no simple formula of great generality will be offered in this paper. The agents under both distributions still match to their preferred traits, but the agents in the remainder distributions are more difficult to deal with. Whether a more general simple solution is possible is a topic for further study. However, under specific assumptions on the distributions, the problem is still very tractable. First, we’ll make the same assumptions as in Section II, with a slight variation to account for the more general environment. Specifically, we keep *MASS* and *OUT* as is and add the following amended assumption:

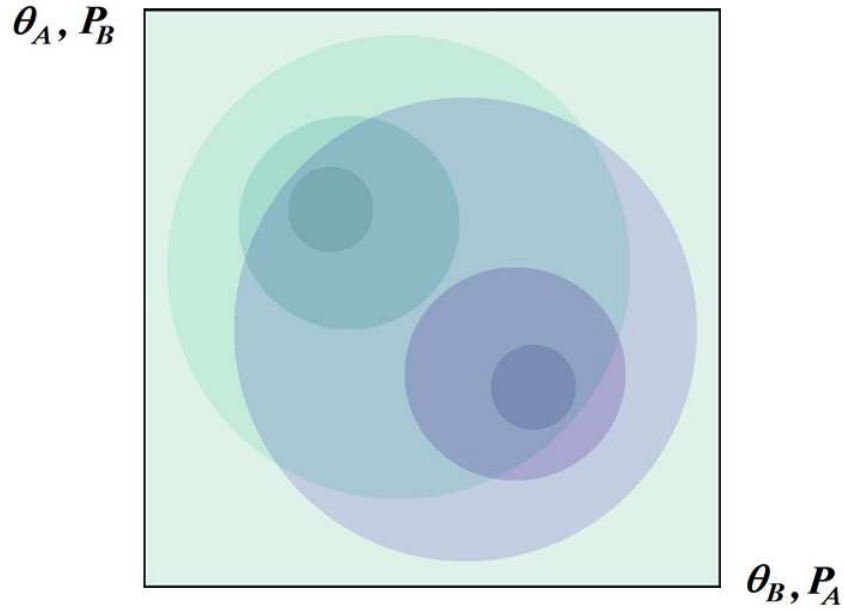


FIGURE 3.2. Overlaid contour plots of distributions A (teal) and B (purple). Darker teal areas indicate more mass in A, and darker purple areas indicate more mass in B.

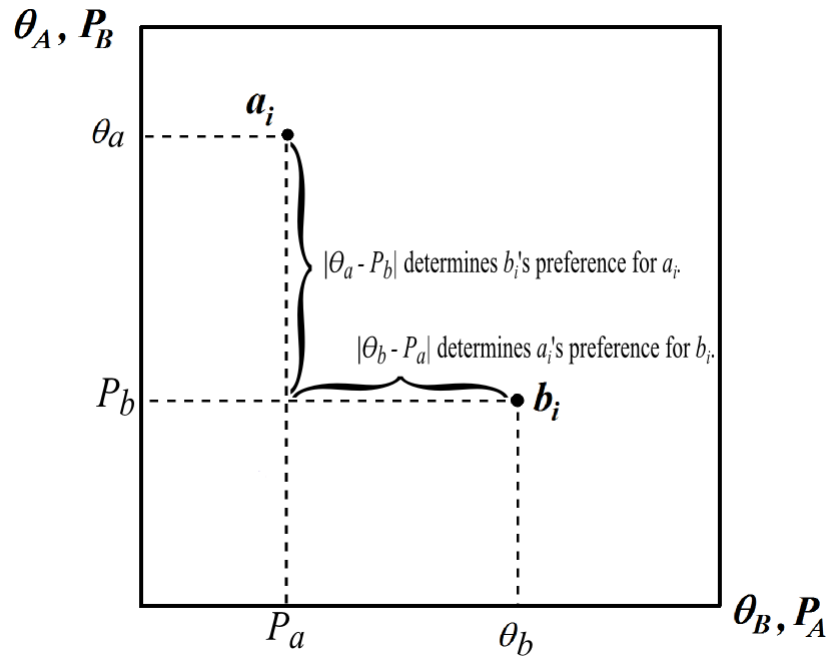


FIGURE 3.3. Comparing matches graphically using the overlay. As shown above, the vertical and horizontal distances between an A and a B agent on this representation tell us the agents' preferences for each other.

- **Assumption 2' (SEP')** : Define $h(x) = y$ all $\forall(x, y) \in s$. Suppose the remainder distributions A' and B' are separated by a single curve s . That is, $\forall x$ and $\forall y' > h(x)$, (x, y') has support only on A for all y' or only on B for all y' , and $\forall y' < h(x)$, the only the opposite distribution has support at (x, y') .

As before, this ensures no complications due to multimodal distributions, varying tail weights, or flat areas between the remainder distributions. Further sufficient but not necessary conditions for a simple solution are as follows:

- **Assumption 4(CURVE1, CURVE2):**
 - h is monotonically increasing (decreasing) in x .
 - the marginal density of A at preference x equals the marginal density of B at preference y for $(x, y) \in s$. That is, $f_{AX}(x) = f_{BY}(y)$.

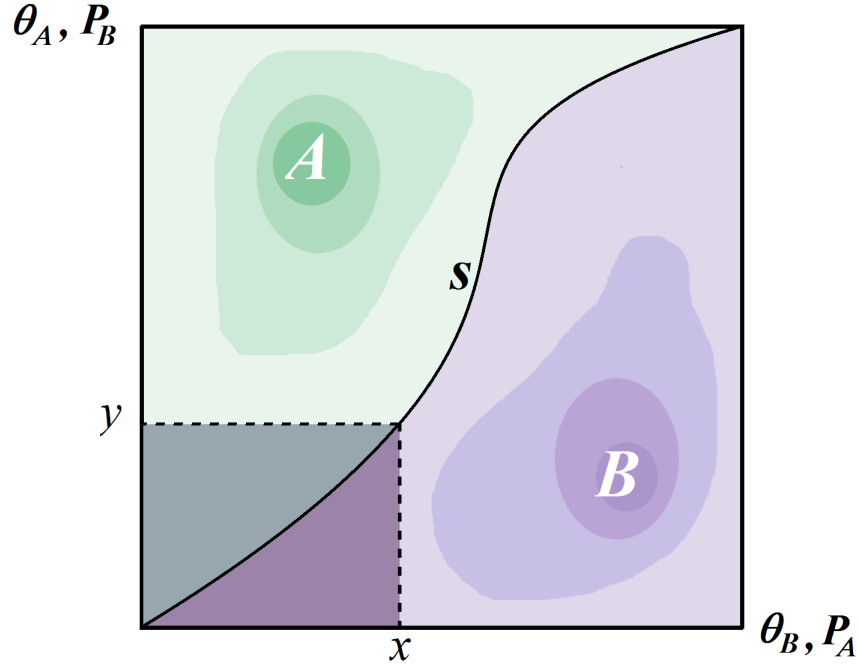


FIGURE 3.4. Matching from the southwest to the northeast on the remainder distributions of A and B . B agents of trait x match to A agents of trait y , regardless of peak preference.

Under these conditions, the matching of the remainder distributions can be solved by matching in an “unzipping” fashion, where at any stage of the matching process the agents in the southwest (southeast) quadrant of the graph below and to the left (right) of some (x, y) on s have all matched (analogous to the interval which has already matched in the homophilic case), while no one else has, and the agents of interest are those on the edges of the quadrant (analogous to the innermost remaining types in the homophilic case). First, we eliminate the overlapping agents, as before.

Lemma 4. *For each point (x, y) , a measure of agents on side $A(B)$ equal to $\mu(x, y) = \min\{f_A(x, y), f_B(y, x)\}$, matches to the types on side $B(A)$ with transposed θ and p .*

Proof. Analogous to Lemma 1, there are at least $\mu_{\theta p}(\theta, p)$ A agents and (p, θ) B agents by definition. Since preferences are homophilic, $a_{i\theta p}$ weakly prefers $b_{jp\theta}$ to any other B agent, and symmetrically $b_{ip\theta}$ prefers $a_{j\theta p}$ to any other A agent. Then a mass μ_{θ} of a_{θ} 's will weakly prefer to match with any of the measure μ_{θ} of b_{θ} 's, and the b_{θ} 's will symmetrically strictly prefer a_{θ} 's to any other agents, so they will form stable matches. \square

Note that in this case the matching outcome described here may not be the only one possible, since agents don't strictly prefer their mirror agent over agents with their ideal trait but a preference for someone other than them. This can be resolved by using lexicographic preferences where preference over θ are as before and, if two potential matches have the same θ , agents prefer matches whose preferences are closer to their own trait, with the rationale that, if someone likes you better, your relationship with them will generally be better. This gives an outcome where agents of a given trait match assortatively in preference (this is an

element of the set of outcomes defined in proposition 6), but complicates the solution in the more generalized case in section 3.2, so we will not make this assumption and simply note that there may be other possible stable matchings.

We now construct the remainder distributions. Define

$$f_{A'}(\theta, p) = \max\left\{\frac{f_A(\theta, p) - f_B(p, \theta)}{\int_p \int_\theta (\max\{f_A(\theta, p) - f_B(p, \theta), 0\}) d\theta dp}, 0\right\}$$

and

$$f_{B'}(\theta, p) = \max\left\{\frac{f_B(\theta, p) - f_A(p, \theta)}{\int_p \int_\theta (\max\{f_B(\theta, p) - f_A(p, \theta), 0\}) d\theta dp}, 0\right\}$$

We are now ready to prove the inductive lemma for this case:

Lemma 5. *Without loss of generality assume A' is to the northwest of B' (separated by a curve as per SEP'). Suppose that all agents with traits and preferences such that $(x, y) < (x^*, y^*)$ have been stably matched and are eliminated from consideration, while no other agents in A' or B' have matched. Then the set of A' agents $\{a_{i\theta_p} : \theta = y^*\}$ will match stably and arbitrarily to the set of B' agents $\{b_{i\theta_p} : \theta = x^*\}$.*

Proof. For each B agent b_{x^*n} , $n \leq y^*$, which is the trait of all A agents on the other edge of the quadrant defined by (x, y) , so $a_{y^*p} \succ_{b_{x^*m}} a_{\theta p}$, \forall unmatched agents $a_{\theta p}$ where $\theta > y^*$. A symmetric argument shows that all a_{y^*p} strictly prefer b_{x^*n} agents to any other unmatched B agents, so the edge agents of both distributions will match to one another. Since $f_{AX}(x) = f_{BY}(y)$, these sets of agents have equal measure, so they exactly and stably match to one another, leaving no remaining b_{x^*n} or a_{y^*p} . \square

Proposition 6. *(Single Peaked Two-sided Matching I) Suppose MASS, SEP' , OUT, and CURVE. A measure of agents equal to the measure under both distributions and with equal density over (x, y) will match to their preferred type, which also finds them optimal. For agents in the remainder distributions A' and B' and for all $(x, y) \in s$, the set of A' agents $\{a_{\theta_p} : \theta = y\}$ will match stably and arbitrarily to the set of B' agents $\{b_{\theta_p} : \theta = x\}$.*

Proof. The first result is simply Lemma 1. Inductively proceeding with Lemma 2 northeast along s , we have that a_{yp} matches with any b_{xn} and vice versa $\forall (x, y) \in s$. Because the marginal densities are equal along this path, the measure of matched agents at any point in the inductive process is identical for both sides, so we don't violate the necessary condition of 1 – 1 matching. \square

It's worth noting that this marginal density assumption is very important. If we did not have $CURVE2$ and tried to proceed as above, we'd have unequal measures of agents being matched at various points in the matching process, a clear contradiction. In fact, what would happen is that each "layer" of A' agents would not completely match out the corresponding layer of B' agents, and the remaining B' agents would match to the next layer of A' agents. We would then no longer be in the the extremely convenient situation where the current matchers are all of one trait and where every current matcher prefers the edge agents on the other side to any other available agent. Similarly, dropping $CURVE1$ would invalidate the procedure, with, for example, agents on one side matching to no one on the other side when the slope of s was negative.

We can also obtain an algorithm for the matching in a one-sided problem with single-peaked preferences from the two-sided algorithm by representing the one-sided problem as a two sided problem. Recall that a one-sided matching problem is one where there is a single set of agents who must be matched to one another in any way that satisfies stability, whereas the two sided problem imposes the additional constraint that agents can only match to individuals on the opposite side.

Corollary 7. *(Single Peaked One-sided Matching) For any distribution $f(p, \theta)$, define $f_A(p, \theta) = f_B(p, \theta) = f(p, \theta)$. Then if f_A and f_B satisfy MASS, SEP' , OUT, and CURVE and $s = (x, x) \forall x \in \mathbb{R}$, Proposition 6 holds. Equivalently, the one-sided matching problem with distribution f has the stable matching given by Proposition 6, where the match of a given agent $a(b)$ is \inf rather than $f_B(f_A)$.*

Proof. The first claim follows directly from Proposition 6, as it is just a special case of the problem considered there. For the second claim, consider the one-sided matching problem with single peaked preferences and distribution f . An agent a with trait θ_a and preference p_a prefers matches b based on \cdot . Similarly, b prefers matches based on $|p_b - \theta_a|$. Then if we overlay $f_A = f$ with an axes-transposed copy of f , f_B , a 's preferences

over f are given by the vertical distance, and b 's preferences over f are given by the horizontal distance. First, we remove agents with perfect matches, so a mass $\min(f(p, \theta), f(\theta, p))$ is matched to its ideal match for each (θ, p) , then we move on to the iterative stage. At every stage of the matching process, the set of unmatched agents on each side in the two-sided problem is equal to the set of unmatched agents in the one-sided problem, where all the matches thus far derived in the two-sided problem are stable in the one-sided problem, then the current matches are optimal among the set of available matches for all agents on both sides, so they are optimal in the one-sided problem. There is one complication here—when agent a is matched to agent b in the two-sided problem, the a is removed from side A and b is removed from side B . However, in the one-sided problem both a and b are on the same side. If some matched agents are not removed from each side, the set of available matches will not correspond to the one-sided problem. However, because the distributions are identical, the remainder distributions are also identical by their definitions, and $x = y \forall (x, y) \in s$, the set of current matches is identical, and agents are indifferent between all possible matches in this set. Since there are infinitely many agents at every point with nonzero support, we can always have half the agents of a given type match to the other half, leaving no agents unmatched and all agents with their preference maximal match among the set of remaining agents.

Note that this yields a very simple matching outcome where as many agents get perfect matches as possibly can and the remainder match to their own trait (positive assortment). \square

3.2. Extensions to the Baseline Model. While the result in 3.1 is extremely simple, the assumptions, especially CURVE2, are unlikely to be even approximately satisfied in a real world application. Having the marginal densities equal at any particular point on s is unlikely, much less at every point. First, then, we will relax this assumption. This significantly complicates the problem, but does not render it insoluble. Without loss of generality, assume that A and B are separated by a monotonically decreasing h , with A above and to the right and B below and to the left. First, find all points $(x_i, y_i) \in s$ such that $1 - F_{B'Y}(y_i) = F_{A'X}(x_i)$ for $i \in \{1, \dots, n\}$ (assume there are finitely many such points). Then, $\forall (x, y) \in s$ where $(x_i, y_i) \leq (x, y) \leq (x_{i+1}, y_{i+1})$, $(x_n, y_n) \leq (x, y)$, or $(x, y) \leq (x_1, y_1)$, $1 - F_{B'Y}(y) \leq F_{A'X}(x)$ or $1 - F_{B'Y}(y) \geq F_{A'X}(x)$. Without loss of generality, suppose $1 - F_{B'Y}(y) \leq F_{A'X}(x)$. Then, with the following amended assumption, we can proceed to a matching solution.

• **Assumption 4' (CURVE2'):**

- When $A'(B')$ has a larger mass matched out, the marginal density of $A'(B')$ at trait $x(y)$ is greater than or equal to the marginal density of $B'(A')$ at trait $y(x)$ for preferences greater than or equal to $y(x)$, where $y(x)$ is such that we have masses $m_a = m_{b1} + m_{b2}$ ($m_b = m_{a1} + m_{a2}$). That is, $f_{A'X}(x) \geq \int_y^\infty f_{B'}(x, p) dp$ ($f_{B'Y}(y) \geq \int_{-\infty}^x f_{A'}(y, p) dp$), .

This assumption ensures that there is never more mass in β_1 in Figure 3.5 than in α , which would invalidate the matching algorithm since, as x moved outward as the matching progressed down and to the right, more mass would be matched in B than in A , even if y didn't decrease at all.

Finally, before we state the proposition, we'll need a definition and two equations.

Definition 8. Define $M_A(x)$ as the set of B agents that an A agent of trait x can stably match to. That is, $M_A(x) = \{b_{j\theta p} : \theta = y\} \cap \{b_{j\theta p} : p = x \wedge \theta \geq y\}$. Similarly, $M_B(y) = \{a_{j\theta p} : \theta = x\} \cap \{a_{j\theta p} : p = y \wedge \theta \leq x\}$.

We also define two equations guaranteeing equal masses of agents have been matched out at each step (this is equivalent to the equal masses condition in CURVE2').

$$(3.1) \quad \int_{-\infty}^{\infty} \int_{-\infty}^x f_{A'}(\theta, p) d\theta dp = \int_{-\infty}^x \int_y^{\infty} f_{B'}(\theta, p) d\theta dp$$

$$(3.2) \quad \int_{-\infty}^{\infty} \int_y^{\infty} f_{B'}(\theta, p) d\theta dp = \int_y^{\infty} \int_{-\infty}^x f_{A'}(\theta, p) d\theta dp$$

Proposition 9. (Single Peaked Two-sided Matching II) Suppose MASS, SEP', OUT, and CURVE'. A measure of agents equal to the measure under both distributions and with equal density over (x, y) will match to their converse type, who also finds them optimal. For agents in the remainder distributions A' and B' and for all $(x, y) \in s$, if $1 - F_{B'Y}(y) \leq F_{A'X}(x)$ A' agents $\{a_{i\theta p} : \theta = x\}$ will match stably and arbitrarily to

elements of the set of B' agents $M_A(x)$ and vice versa, where y satisfies Eqn. 3.1 If $1 - F_{B'Y}(y) > F_{A'X}(x)$ B' agents $\{b_{i\theta_p} : \theta = y\}$ will match stably and arbitrarily to elements of the set of A' agents $M_B(y)$ and vice versa, where x satisfies Eqn. 3.2.

Proof. Then A agents of trait x will match to b agents of trait y and $p \leq x$ or b agents where $h^{-1}(x) \leq p \leq y$ and trait x . To check that this is stable, consider a_1, b_1 and a_2, b_2 matching this way, where $x_2 > x_1$ and consequently $y_2 < y_1$. Without loss of generality, either (1) $1 - F_{B'Y}(y) \leq F_{A'X}(x)$ for both 1 and 2, or (2) $1 - F_{B'Y}(y) \leq F_{A'X}(x)$ for pair 1 and not for pair 2. If (1), $a_1 \succ_{b_1} a_2$ since $p_{b_1} \leq \theta_{a_1} < \theta_{a_2}$. Then we need only consider the potential $a_1 - b_2$ blocking pair. $a_1 \succ_{b_2} a_2$ only if $p_{b_2} < \theta_{a_2}$, else a_2 is b_2 's perfect match. But if so then the trait of b_2 is y_2 . Then $b_1 \succ_{a_1} b_2$ as $y_2 < \theta_{b_1} < p_{a_1}$. If (2), $b_1 \succ_{a_1} b_2$ since $p_{a_1} > \theta_{b_1} > \theta_{b_2}$ and $a_1 \succ_{b_1} a_2$ as $p_{b_1} \geq \theta_{a_1} > \theta_{a_2}$, so there is no blocking pair. Finally, if both pairs come from the same stage in the matching algorithm, either the b 's are indifferent between the a 's or the a 's are indifferent between the b 's, so there is no blocking pair. \square

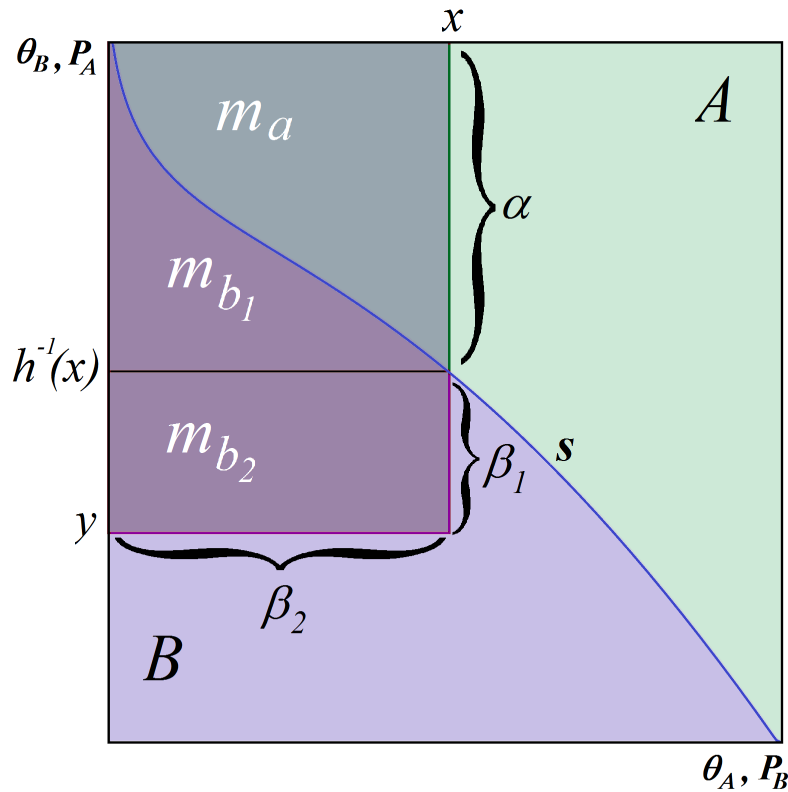


FIGURE 3.5. Matching with unequal marginal distributions. A has more mass in m_a than B has in m_{b_1} , so agents in m_{b_2} have also been matched to equalize the mass on both sides.

4. RELATIONSHIP TO THE LITERATURE

As mentioned previously, there are well known results for two-sided matching with vertical preferences, and Clark 2007[4] gives the results for horizontal preferences shown in section 2. It is obvious from the previous exposition that horizontal preferences are a special case of single-peaked preferences—specifically, they are the case where preference is set equal to trait. Then with these preferences, we should find that

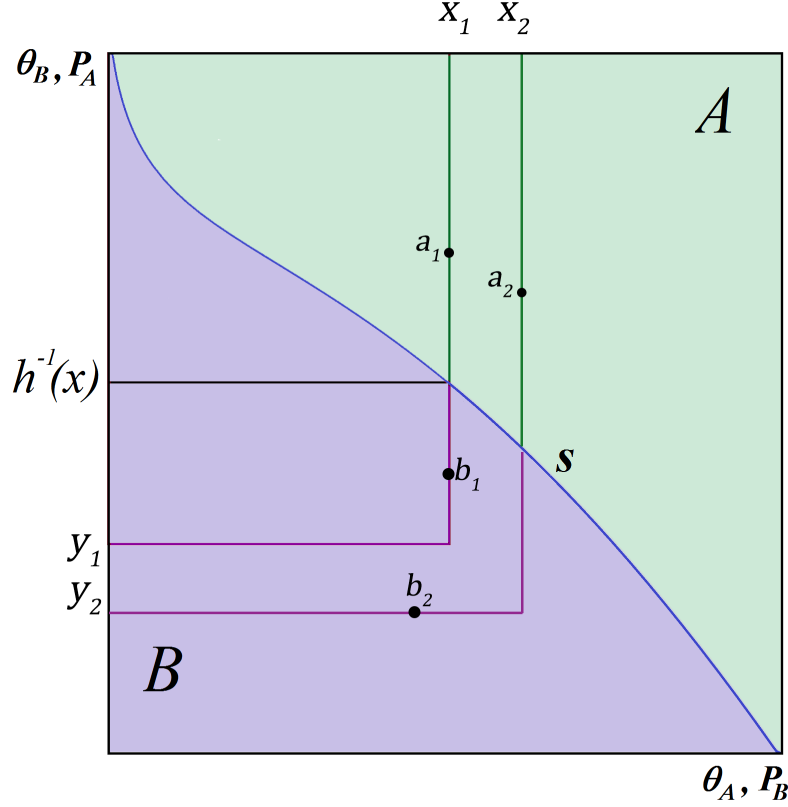


FIGURE 3.6. Proposition 8, case 1 example: $1 - F_{B^Y}(y) \leq F_{A^X}(x)$ for both (x_1, y_1) and (x_2, y_2) .

the single peaked algorithm reduces to the horizontal preference algorithm. In fact this is the case. In the horizontal preference case, the distributions have support only on the diagonal, where preference equals trait.

From here, we remove the overlap and can now easily draw a monotonically decreasing curve $s = (x, h(x))$ that separates the two sides, and by letting $h(x) = F_B^{-1}(1 - F_A(x))$, we have the appropriate matches and an equal mass of matched out agents at every step in the process, as desired.

Similarly, Becker's NTU model with vertical preferences is also a special case of single-peaked preferences, namely, when everyone's preference is for higher traits. If there is a maximal trait $\bar{\theta}$, we can simply set preferences to $\bar{\theta}$. This affords a simple graphical representation, with the distributions varying along trait with support only at preference $\bar{\theta}$. Choosing $s = (x, h(x))$ and letting $h(x) = F_B^{-1}(F_A(x))$, we have the appropriate matches and an equal mass of matched out agents at every step in the process and the two distributions are separated by s , and we have the familiar positive assortative matching for vertical preferences with nontransferable utility.

5. INTERPRETATION AND EMPIRICAL IMPLICATIONS

5.1. Interpretation. These models are amenable to some interpretation. While in the vertical case we have positively assortative matching (PAM), and in the horizontal case we have PAM in the overlap and negatively assortative matching (NAM) in the remainders, in the equal-marginals case of single peaked preferences, we have two modes of matching that encompass these previous cases. First, we have a converse form of PAM over trait and preference in the overlap region, where we have positive assortment in A trait and B preference and in B trait and A preference—that is, increase in one parameter corresponds to increase in the other parameter in one's match. Note that this is more than just PAM—the matches have exactly reversed

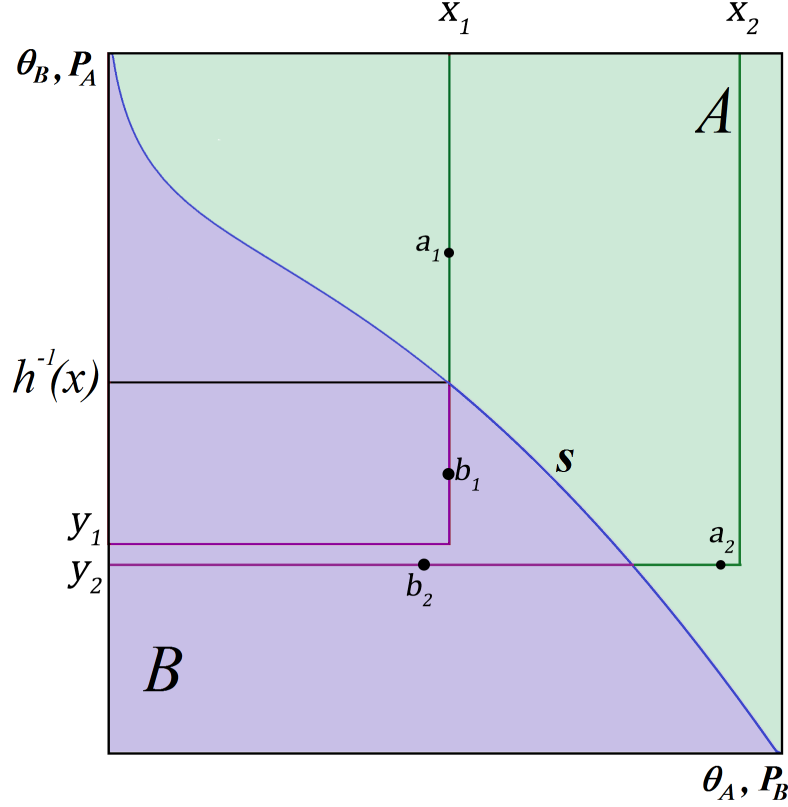


FIGURE 3.7. Proposition 8, case 2 example: $1 - F_{B|Y}(y) \leq F_{A|X}(x)$ for (x_1, y_1) only.

trait and preference. We can also see that the standard PAM of the horizontal model is actually a special case of this converse PAM, where only the fact that trait and preference are equal ensures that A trait equals B trait. For the remainders, we have either PAM in trait only (not in preference) if the separating curve s has a positive slope, and NAM in trait only if s has a negative slope. This again corresponds to the horizontal and vertical cases, with the line in the vertical case having a positive slope, while the horizontal case has a negative slope.

The more general model of proposition 8 is a bit more complicated, but also yields an intuitive interpretation. Without loss of generality, assume a strictly decreasing h and A at the top right, with higher mass on side A as in figure 3.5. In this situation, B agents can be thought of as being in shortage at $(x, h^{-1}(x))$ in figure 3.5, as there aren't enough of them to match to A agents as in the equal marginals outcome. As such, and given that matching utilities for each agent in a match need not correlate in any way, we'd expect that B agents will often be able to leverage this scarcity to get a matching outcome more favorable to their side, and this is in fact the case. Notice that the entire region β_1 in figure 3.5 gets perfect matches (from their perspective), while β_2 agents have a similar matching outcome to B agents in the equal marginals case. A agents, on the other hand, match to B agents whose trait is further from their preferences than the agents with trait $h^{-1}(x)$ that they'd match to in the equal marginals case, so they are worse off. In terms of assortment, we see that, for the overlap region the result is the same as the equal-marginals case, while in the remainders we have some of the agents exhibiting PAM or NAM in trait as before (e.g. agents matching from α to β_2), while the surplus agents of α match to β_1 agents with converse PAM in surplus trait and shortage preference only (there is no PAM in shortage trait and surplus preference, as in the perfect matching stage).

FIGURE 4.1. Horizontal preferences in the single peaked preference framework.

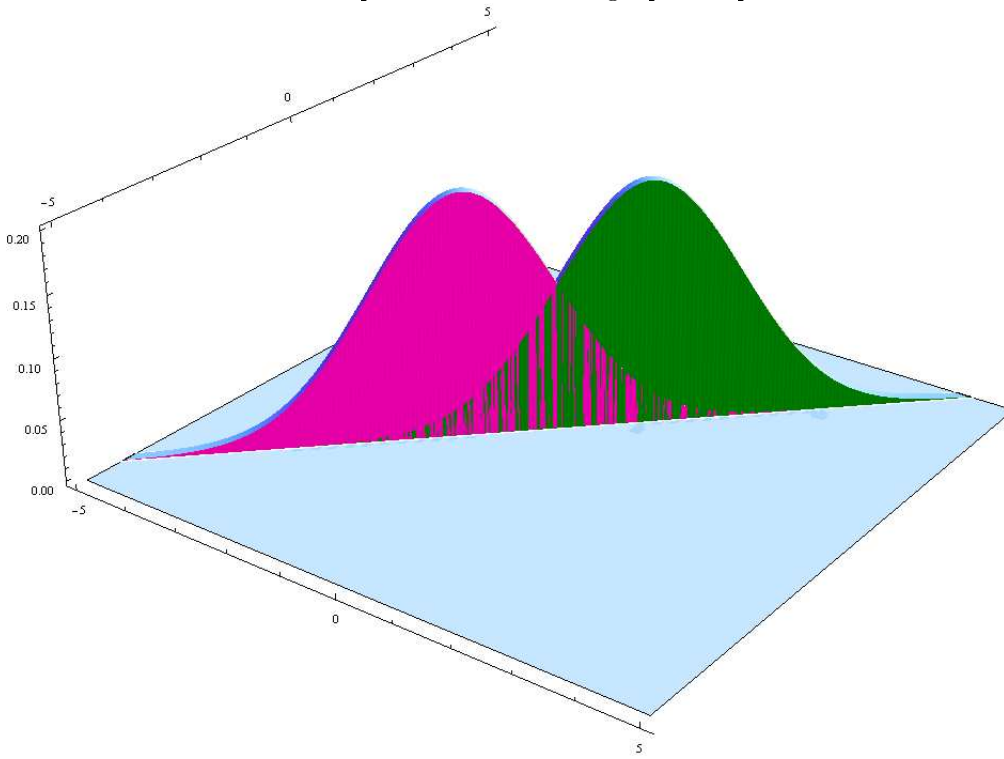


FIGURE 4.2. Horizontal matching in the single peaked preference framework

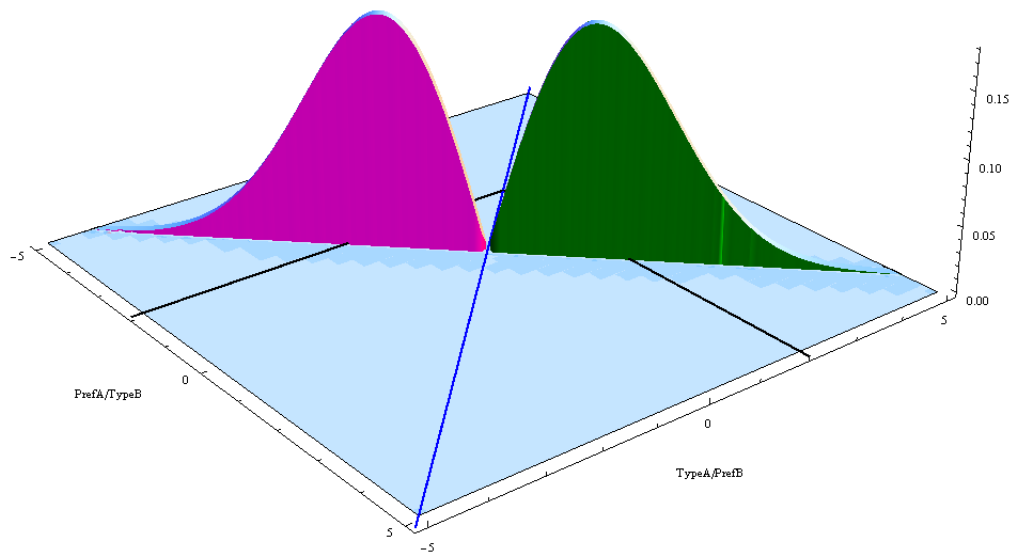
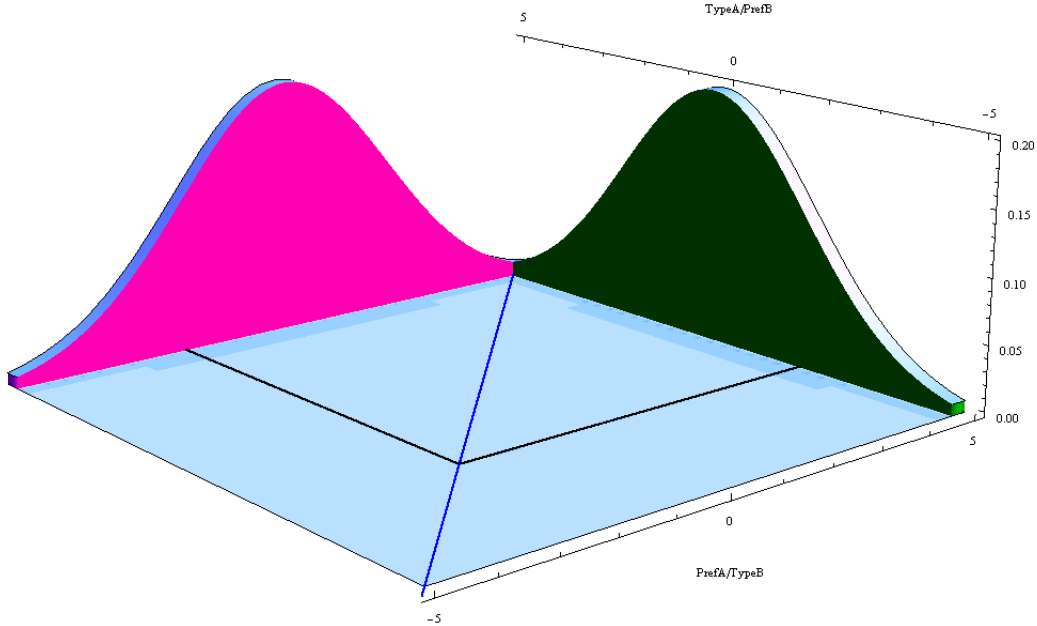


FIGURE 4.3. Vertical preferences in the single peaked preference framework.



5.2. Empirical Implications. The empirical implications of this model are also fairly straightforward. Again, without loss of generality assume a strictly decreasing h and A at the top right, with higher mass on side A as in figure 3.5. As a 's trait increases, the matching function that gives the distribution of possible matches exhibits distributions where the maximum preference of matches is equal to a 's trait and thus increasing (β_2 in figure 3.5 is the set of possible matches with preferences less than a 's trait). Note that all potential matches for an A agent with trait x must have the same trait, y , unless their preference is equal to a 's trait, and that this y is decreasing in x . The distribution of possible matches also includes B agents whose trait is greater than y but less than the minimum preference of a agents with trait x in A' , β_1 in figure 3.5. This region's upper and lower bounds in B agent trait decrease in x , while the B agent's preference is of course increasing in x as it is equal to x . Finally, we can expect a mass point of perfect matches where preference and trait are reversed from the first stage of the matching process. If we only observe trait, we would expect to see a mass point at the minimum trait a matches to, and as a 's trait increases, we would expect that mass point to move downward.

6. EXAMPLE APPLICATION

6.1. Model Preliminaries. So far we have developed an algorithm for finding stable matchings, but haven't used it in any economic application. In this section, we will work through a simple model with a price setting firm and compare the results when we assume vertical preferences and when we allow a broader class of single peaked preferences. Specifically, we'll consider a two sided market model with a monopolist firm running a matching platform and charging a fixed entry fee, and a two sided matching market where agents can either pay and use the platform to search for matches, getting their perfect information match, or stay off the platform and get a fixed expected payoff, normalized to zero, by matching without the information of the platform. The idea here is to have a stylized model of a firm that allows agents to choose from a larger set of matches than can be obtained off the platform. Another more palatable assumption would be to have off platform agents match randomly to other off platform agents, but this will make solving the model much more difficult, so for this example we'll proceed with the zero utility outside option.

To make things concrete, suppose agent trait is BMI, and agents uniformly prefer lower BMI's, so that we can shoehorn this problem into a vertical framework. For analytical ease, assume BMI ranges from 0 to 1, and utility of matching is given by $U_a(a, b) = B - \theta_b$, where a and b are trait and preference vectors and θ_b is the match's trait and B is the benefit of a perfect match. Similarly, we have $U_b(a, b) = B - \theta_a$. Suppose

there are uniform distributions of agents of measure 1 over $[0,1]$ on each side. Then agents will match to their own trait on the platform and, for a price p , the utility of joining the platform $U_a(a) = B - \theta_a - p$.

6.2. Derivation. The structure of the utility function makes it clear that there will be a cutoff C below which agents join the platform and above which they refuse, as on-platform utility is decreasing in own trait. Specifically, we have equality between joining and not when

$$0 = B - C - p$$

$$\implies p = B - C$$

Assume $B < 2$ so the firm's strategy isn't a corner solution. Then the firm wants to maximize

$$\Pi(C) = 2(B - C)C$$

so the first order condition is

$$\frac{\partial \Pi}{\partial C} = 2B - 4C = 0$$

$$\implies C = B/2$$

And we have

$$p = B/2$$

$$\Pi = B^2/2$$

This is clearly an equilibrium of the game, as no consumer would like to deviate from or to the platform, and the firm is maximizing profit. To remove unwanted equilibria where, e.g., no one joins the platform, consider the following equilibrium refinement: we consider only equilibria that can result from all agents joining the platform, the firm then setting the price, then agents leaving the platform if their match is worse than the outside option and forevermore assuming each agent's current strategy will be repeating in future periods. In this scenario, all agents join, the firm sets the price derived above, then everyone above the cutoffs leaves. After this step, no one has an incentive to change their strategy.

Now suppose that agents are bivariate uniformly distributed over BMI's from 0 to 1 and over peak preferences from 0 to L . Suppose also that $2L \leq B \leq 2 - 2L$. Now we have a nontrivial single peaked problem. Utility for an A agent is now $U_a(a, b) = B - |\theta_b - P_a|$. Then agents who enter the platform with $\theta < L$ match to their ideal as $f_A(\theta, P) = f_B(P, \theta)$ for $(P, \theta) < (L, L)$. For the remainder, agents match to their own trait if they join the platform and get utility $U_a(a) = B - \theta_a + P_a - p$. We will also find a cutoff here, but one that varies in both trait and peak preference:

$$0 = B - C + P - p$$

$$\implies C = B - p + P$$

so the cutoff is a linear, decreasing function of own preference. Then, for a given price p , the mass of agents who join the platform on a given side is given by

$$\begin{aligned} \frac{1}{L} \int_0^L B - p + P \, dP \\ = (B - p) + L/2 \end{aligned}$$

so profit is given by

$$\Pi(p) = 2((B - p) + L/2)p$$

and the first order condition for price is

$$\frac{\partial \Pi}{\partial p} = 2B + L - 4p = 0$$

$$\implies B/2 + L/4 = p$$

We can now compute the following:

$$C(P) = B - B/2 - L/4 + P$$

$$C(P) = B/2 - L/4 + P$$

$$\Pi = 2((B - B/2 - L/4) + L/2)(B/2 + L/4)$$

$$\Pi = 2(B/2 + L/4)^2$$

6.3. Analysis. We can now compare this simple model with and without nontrivial single peaked preferences. It should be noted that comparing profit, cutoffs, and prices is fairly questionable here. We can see that price and profit increase as the range of peak preferences increase, while the cutoffs are also higher on average. However, this is largely an artifact of the outside option having constant utility for all agents. We would expect (and solving the same model with an outside option of a random match from the off platform agents on the other side confirms) that agents with peak preferences for higher trait (less demanded) matches get higher expected utility from their random match from the set of off platform agents, so the higher profits due to higher match utility for higher peak preference agents is attenuated by the better outside option for those same agents.

What we can compare is the general structure of the outcomes. In the vertical case we have the standard assortative-in-trait outcome for agents on the platform, and a cutoff below which the more desirable agents are willing to pay to ensure they get an equally desirable match, while undesirable agents have less incentive to join as they wouldn't receive a good match under perfect information. In the nontrivial single-peaked case, the agents below (L,L) get a perfect match, with converse assortment in all parameters. We also have a region below a cutoff where agents match assortatively in trait only. However, we see that the cutoff depends on both trait and preference, so there is no unique cutoff in trait. Agents with a higher peak preference are willing to accept a higher trait match than their lower peak peers, so we have an intervals over trait where everyone joins, then fewer and fewer agents of that trait join, and final where no one joins. While finding closed forms for the cutoffs in formulations of this problem with random off platform matches is difficult or impossible, they generally exhibit similar qualitative features, with cutoffs that vary in peak preference and regions with converse assortment. It should be noted that the way the cutoffs vary in peak preference don't necessarily conform to the cutoff increasing in peak preference result here, and generally depend on the functional form of the disutility for more distant matches.

7. CONCLUSION

This paper derived an algorithm for finding matching outcomes in a generalization of some environments that have previously been explored in the matching literature, names Becker's vertical model and Clark's horizontal model. By allowing for a large range of single peaked distributions, this algorithm can be used to explore the behavior of matching markets and profit seeking matching platforms under more relaxed assumptions on preferences, as seen in section 5. An obvious extension of this thread of research is to explore more theoretical models like the one it section 5. This paper analyzes an intentionally simple model, with "just so" distributions that make the math quite simple. It remains to be seen, however, what behavior would be found in less trivial cases.

There are several other plausible extensions for this model which will not be explored fully in this paper but may be worth further consideration. For example, other than the restriction placed on preferences, a major shortcoming of the original homophilic model is that it assumes that agents are able to match perfectly. This seems quite unrealistic, and prevents analysis of many interesting aspects of matching markets. To address this, we can retain the homophilic model's structure and add noise to the type signal perceived by the opposing agents. Aside from dropping the somewhat unpalatable assumption that agents can perfectly choose from an infinite menu of options, this provides a basis for the exploration of issues involving matching accuracy. For example, what price can a dating website charge given the quality of its matching service? What is the welfare loss when matching is not exact? It may also be possible to use this framework to model adverse selection in dating markets, where instead of a noisy signal individuals observe a reported signal.

Another useful extension would be to add vertical differentiation. All of these possibilities involve significant complications, but could be very useful.

Aside from extensions to the baseline model, further study could include determining how these models relate to analogous finite matching models or models with search. For example, in a framework where agents have horizontal preferences and have one potential match drawn from the distribution each period and can accept or reject and wait, and where there is no replacement—that is, as agents match out, the distribution loses their mass—as the discount factor goes to 1 the model should asymptotically approach the horizontal matching outcome described in section 2. The argument, loosely, is that as time discounting goes to zero, the intervals over which agents are willing to match and their potential match is willing to match to them asymptotically approach the degenerate interval of their free-choice match. First, the overlap will match out, and when the the overlap mass approaches zero surplus overlap agents the ϵ amount of time discounting will become significant as the wait time to a match approaches infinity, and surplus agents will finally accept their bad matches. Agents in areas where only one distribution has support will match to agents around their free-choice match from the beginning. Of course, this argument must be formalized. In the case where there is replacement in the distribution, such that the distribution is constant across time, all agents in the intersection of the supports will match to an interval asymptotically approaching their own type since there will always be a chance of getting a near perfect match bounded above zero, and agents outside the support intersection will match outward, with the negative assortment seen before for the remainder distributions. This too requires formalization.

REFERENCES

- [1] Becker, Gary S, 1973. "A Theory of Marriage: Part I," *Journal of Political Economy*, University of Chicago Press, vol. 81(4), pages 813-46, July-Aug.
- [2] Clark, S. (2003), "Matching and Sorting with Horizontal Heterogeneity", ESE Discussion Paper 98,
- [3] Clark, S. (2006). The uniqueness of stable matchings. *Contributions in Theoretical Economics*, 6(1), 1-28.
- [4] Clark, S. (2007). Matching and sorting when like attracts like.
- [5] College Admissions and the Stability of Marriage D. Gale and L. S. Shapley *The American Mathematical Monthly* , Vol. 69, No. 1 (Jan., 1962), pp. 9-15 Published by: Mathematical Association of America
- [6] Beauty Is a Beast, Frog Is a Prince: Assortative Matching with Nontransferabilities Author(s): Patrick Legros and Andrew F. Newman Reviewed work(s): Source: *Econometrica*, Vol. 75, No. 4 (Jul., 2007), pp. 1073-1102
- [7] "Two-Sided Matching with Spatially Differentiated Agents," Klumpp, Tilman, *Journal of Mathematical Economics* 45, 376-390, 2009.