A Theory Paper on Rank List Length

By Tyler Hoppenfeld

Your abstract here, please.

I. Introduction

Centralized matching markets themselves encompass only a small fraction of the labor market, however the study of centralized matching markets plausibly sheds light on the broader matching of employees to jobs. Furthermore, there are a variety of markets, including nearly every academic job market and most entry-level job markets for new college graduates, where there is no obvious reason why a centralized once-a-year clearing house could not work, and yet do not have this structure. A better understanding of the strengths and limits of centralized markets could help us understand what inefficiencies exist in these markets, whether more efficient solutions are possible, and possibly allow for the design of more efficient markets.

Theory tells us that there is no matching mechanism that reliably produces a stable outcome wherein both sides have the incentive to report their preferences accurately (Roth 1985), however in practice many matching markets persist with both sides (apparently) reporting their preferences honestly (Roth and Peranson 1999). This result can be explained by the fact that in many of these markets the set of stable matchings is actually quite small. For example, in the National Resident Matching Program (NRMP) matching of recently graduated physicians to resident positions, all but approximately 10 of 20,000 participants have a unique partner in the set of all stable matchings (Roth and Peranson 1999), and so in all but a very small number of cases any misrepresentation of preferences would lead to a less-prefered outcome. Roth and Peranson posit that the small set of stable matchings (SSM) of these markets is likely driven by the correlated nature of preferences, and by the fact that for highly correlated preferences, the SSM is very small. In a parallel argument, they note that finite preference lists, as is the case with the NRMP, lead to a very small SSM as well.

The concept of stable matchings is intuitively appealing in the same way the concept of the core is, but it also has empirical support as an equilibrium. In a review of systems that assign trainee physicians to hospitals in Great Brittan, the markets that produced stable matchings persisted with broad participation, while the markets that did not produce stable matchings failed, with large numbers of applicants and hospitals preempting the market and contracting before the market (Roth 2003).

Roth and Peranson (1999) use simulation to demonstrate these results, and (Immorlica and Mahdian (2005) and then Kojima and Pathak (2009) build on this with the theoretical result that as the number of market participants grows, the SSM becomes arbitrarily small as a fraction of market size. This theoretical work takes short preference lists as a feature of preferences, and does not admit a preference structure where a positive measure of the potential matches can all be highly desirable to a positive measure of the market participants.

This paper extends these theoretical results to examine short rank lists in a setting that allows for more realistic preference formation and short rank lists. In this setting I show that short rank lists are not sufficient to ensure a small SSM. Further, I characterize a partition over the universe of agents, and explore how that partition can be used to understand the size of the SSM and the magnitude of the difference in utility between the best and worst stable matchings an agent can receive.

II. Model

A. Preliminary Definitions

Let there be M men and W women, who together form the universe of participants $U = W \cup M$. Each man has a strict preference relation over the set of women he could be matched with and being unmatched, which we denote as being matched with himself. Therefore we say that he has preferences $\succ_m$ over the set $W \cup \{m\}$. Women have likewise have preferences $\succ_w$ over $M \cup \{w\}$. A market is tuple Γ $(M, W, \succ_M, \succ_W)$ where $\succ_M = (\succ_m)_{m \in M}$ and $\succ_W = (\succ_w)_{w \in W}$.

A matching ϕ is a mapping from U onto itself such that (i) for every m, $|\phi(m)| = 1$, $\phi(m) \in W \cup \{m\}$ and for every w, $|\phi(w)| = 1$, $\phi(w) \in M \cup \{w\}$, (ii) $\phi(m) = w$ iff $\phi(w) = m$. That is, a matching is simply a one-to-one assignment of men to women. We also denote the set of all people matched in ϕ as $\{\phi\}$.

We say a matching ϕ is blocked by a pair $\{m, w\}$ if $w \succ_m \phi(m)$ and $m \succ_w \phi(w)$. It is individually rational if $\forall a \in U, \phi(a) \succeq a$. A matching is stable if it is individually rational and not blocked.

B. Deferred Acceptance

The deferred acceptance mechanism is a mainstay of this paper. In the man-proposing version of this mechanism, the algorithm iteritavely selects a man m who proposes to his most preferred woman who has not yet rejected him, unless he has been rejected by all women who he prefers to being unmatched. If she prefers him to her current tenative assignment (and prefers him to being unmatched), she holds his proposal and rejects her current assignment (if one exists). Otherwise, she rejects m. This process repeats until all men are either tentatively matched or have run out of proposals to make, and arrives at a stable matching that is weakly preferred by all men to any other stable matching. We call this the man-optimal stable match, because it is weakly preferred by all men to any other stable match. A parallel result exists for the woman-proposing algorithm.

C. Market Participants

The market consists of n men and n women, indexed $\{m_1, m_2, ..., m_n\}$ and women indexed $\{w_1, w_2, ..., w_n\}$

Each man's preference list by first ranking all women according to their index number, so w_1 is his favorite, w_n his least favorite, and so on. With probability ϵ man m_i transposes the rankings of w_{i-1} and w_i, with probability ϵ he transposes the rankings of w_{i-1} and w_i, and with probability $1-2\epsilon$ his rank list is unchanged. The women form their preferences in the same way.

In this way, we arrive at a preference structure where agents all broadly agree on which matches are desirable, but have idiosyncratic preference variation.

III. Analysis

To analyze this market we first need two separate mathematical results, one to partition the market into smaller segments that we can analyze separately, and one to help us identify cases with several stable matches.

A. An Algorithm to Partition the Market

To begin, call $U_0 = M \cup W$, define $p_1(m)$ is man m's first choice among U_0, and likewise for women.

Next choose m, the man in U_0 with the lowest index number, and select $w = p_1(m)$, $m' = p_1(w)$, and so on, until a cycle is found. Call the members of that cycle C^*. The algorithm proceeds as follows:

1) define the correspondence R from U_0 to $\mathcal{P}(U_0)$ such that $\forall a \in U_0, R(a) = \emptyset$
2) for each $a \in C^*$, alter R by replacing $R(a)$ with $R(a) \cup \{p_1(m)\}$
3) for each $a \in U_0$, and $b \in U_0$, if $b \in R(a)$ replace $R(b) = \{a\} \cup R(b)$
4) for each $a \in U_0$ such that $R(a) \neq \emptyset$, call c their least preferred member of $R(a)$. Now, for each $d \in U_0 | d \succ_a c$, replace $R(a) = R(a) \cup d$
5) if R has been altered in step 3 or 4, return to step 3, otherwise proceed
6) call C_0 the set of all agents for whom $R(a) \neq \emptyset$
7) Perform man-proposing deferred acceptance on C_0, and call the result ϕ_{m0}

To find the next partition, call $U_1 = U_0 \setminus \{\phi_{m0}\}$, and repeat the entire process to find C_1 and ϕ_{m1}

LEMMA 1: $\{\phi_{mi}\} = \{\phi_{wi}\}$

PROOF:
This is an application of the rural hospitals theorem, which says that the set of unmatched agents is the same in all stable matchings (Roth 1986).

PROPOSITION 1: $\phi_m = \{\phi_{m0}, \phi_{m1}, ...\}$ *is identical to the man proposing deferred acceptance matching* Φ_m

PROOF:

First we prove this for ϕ_{m0}, proceeding by contradiction. Call $w = \phi_{m0}(m) \neq \Phi_m(m)$

Call $m' = \Phi_m(w)$

Either $m' \succ_w m$ or $m \succ_w m'$. We will handle each case separately.

First suppose $m' \succ_w m$

- Since w is always matched to her favorite man of all who ever propose to her, m' must never have proposed to w. Since men propose in descending order of prefernce until all acceptable options are exhausted, m' must be matched to $w' = \phi_{m0}(m')$ where $w' \succ_{m'} w$
- call $m'' = \Phi_m(w')$
- as the pair $\{m', w'\}$ are not a blocking pair for Φ_m, $m'' \succ_{w'} m'$
- we can continue as above to form the infinate sequence $\{m, w, m', w', m'', ...\}$ contained within the finite set C_0. Since this is an infinite sequence on a finite set, we have a cycle such that each man prefers the woman he is matched with in ϕ_{m0}, and each women prefers her match in Φ_m
- this cycle represents a pareto improvement over Φ_m for men, and thus it cannot be a stable matching on U. So there must be a man in $U \setminus C_0$ such that he forms a blocking pair with one of the women in this cycle.
- the man should have been added to C_0 at step 4, thus a contradiction

Now suppose $m \succ_w m'$

- call $w^a = \Phi_m(m)$
- as $\{m, w\}$ do not block ϕ_{m0}, $w^a \succ_m w$
- call $m^a = \phi_{m0}(w^a)$
- now as $\{m, w^a\}$ do not block Φ_m, $m^a \succ_{w^a} m$
- now call $w^b = \Phi_m(m^a)$ and continue as above to form an infinite sequence on the finite set C_0
- we now have a cycle where Φ_m is a pareto improvement over ϕ_{m0} for men
- Since this is a stable match in U but not in C_0, there must be a man in $C_0 \setminus U$ such that he forms a blocking pair with one of the women in this cycle.
- $C_0 \subseteq U$, so $C_0 \setminus U = \emptyset$, and so we have a contradiction

Thus ϕ_{m0} is a subset of Φ_m

We now note that after removing ϕ_{m0} from U, the same logic holds, and the proof continues by induction.

PROPOSITION 2: *If there is a man* $m \in \{C_i \setminus \phi_i\}$ *then there is no woman* $w \in \{C_i \setminus \phi_i\}$

PROOF:

As all participants rank all other participants, if there were an unmatched man and woman, they would wish to match with each other.

PROPOSITION 3: ϕ_{mi} *depends only on the preferences held by the agents in* $\{C_0 \cup C_1 \cup ... \cup C_i\}$

PROOF:

Given proposition 1, it is sufficient to show that there is no set of preferences that could be held by $U \setminus \{C_0 \cup C_1 \cup ... \cup C_i\}$ that would change the makeup of C_i.

Since no member of C^* holds any member of $U \setminus \{C_0 \cup C_1 \cup ... \cup C_i\}$ as their first choice, no other preference profile held by members of $U \setminus \{C_0 \cup C_1 \cup ... \cup C_i\}$ would alter the composition of C^*.

Since steps 3 and 4 of the algorithm in subsection III.A only depend on the preferences of individuals for whom $R(a) \neq \emptyset$, no preferences held by any member of the set $U \setminus \{C_0 \cup C_1 \cup ... \cup C_i\}$ influences the workings of the algorithm.

Thus, the preferences of $U \setminus \{C_0 \cup C_1 \cup ... \cup C_i\}$ are irrelevant to ϕ_{mi}

B. Analyzing the model

We begin by partitioning the market as in subsection III.A .

There are 16 possible preference relation sets for the participants $\{m_1, m_2, w_1, w_2\}$. I have enumerated them in the appendix. For the simplest case, (Type A) in which no preference transpositions have occurred, $C_0 = \{m_1, w_1\}, \{\phi_0\} = \{m_1, w_1\}$, and the preference transpositions for $\{m_2, m_3, w_2, w_3\}$ are drawn from the same distribution as $\{m_1, m_2, w_1, w_2\}$. For eleven cases (Type B) $\{\phi_0\} \cup \{\phi_1\} = \{m_1, m_2, w_1, w_2\}$, and $\{m_3, m_4, w_3, w_4\}$ are drawn from the same distribution as $\{m_1, m_2, w_1, w_2\}$. For two cases (Type C) $\{\phi_1\} = \{m_1, m_2, w_1, w_2\}$, $\{m_3, m_4, w_3, w_4\}$ are drawn from the same distribution as $\{m_1, m_2, w_1, w_2\}$, and notably in these two cases multiple stable matchings are possible.

Finally, there are two cases (Type D) where $\{\phi_0\} = \{m_1, w_1\}$, but where the preference transpositions for $\{m_2, m_3, w_2, w_3\}$ are drawn from a restricted

As the partition algorithm progresses through the rank lists it transitions through these cases as a Markov chain with the transition matrix:

$$\Pi = \begin{bmatrix} (1-\epsilon)^4 & \epsilon^4 + 4\epsilon^2(1-\epsilon)^2 + 6\epsilon(1-\epsilon)^3 & 2\epsilon^2(1-\epsilon)^2 & 2\epsilon(1-\epsilon)^3 \\ (1-\epsilon)^4 & \epsilon^4 + 4\epsilon^2(1-\epsilon)^2 + 6\epsilon(1-\epsilon)^3 & 2\epsilon^2(1-\epsilon)^2 & 2\epsilon(1-\epsilon)^3 \\ (1-\epsilon)^4 & \epsilon^4 + 4\epsilon^2(1-\epsilon)^2 + 6\epsilon(1-\epsilon)^3 & 2\epsilon^2(1-\epsilon)^2 & 2\epsilon(1-\epsilon)^3 \\ \frac{(1-\epsilon)^4}{3\epsilon^2(1-\epsilon)^2+4\epsilon(1-\epsilon)^3+(1-\epsilon)^4} & \frac{2\epsilon^2(1-\epsilon)^2+2\epsilon(1-\epsilon)^3}{3\epsilon^2(1-\epsilon)^2+4\epsilon(1-\epsilon)^3+(1-\epsilon)^4} & \frac{\epsilon^2(1-\epsilon)^2}{3\epsilon^2(1-\epsilon)^2+4\epsilon(1-\epsilon)^3+(1-\epsilon)^4} & \frac{2\epsilon(1-\epsilon)^3}{3\epsilon^2(1-\epsilon)^2+4\epsilon(1-\epsilon)^3+(1-\epsilon)^4} \end{bmatrix}$$

and the stationary distribution:

$$\left[\frac{14\epsilon^5-17\epsilon^4+8\epsilon^3-10\epsilon^2+6\epsilon-1}{4\epsilon^5-10\epsilon^4+6\epsilon^3-2\epsilon-1} \quad \frac{2\epsilon^6-16\epsilon^5+17\epsilon^4-10\epsilon^3+10\epsilon^2-6\epsilon}{4\epsilon^5-10\epsilon^4+6\epsilon^3-2\epsilon-1} \quad \frac{-2\epsilon^6+2\epsilon^5+2\epsilon^3-2\epsilon^2}{4\epsilon^5-10\epsilon^4+6\epsilon^3-2\epsilon-1} \quad \frac{4\epsilon^5-10\epsilon^4+6\epsilon^3+2\epsilon^2-2\epsilon}{4\epsilon^5-10\epsilon^4+6\epsilon^3-2\epsilon-1}\right]$$

Now I turn my attention to type C. Consider the sub-case in which $w_2 \succ_{m_1} w_1$ and $m_2 \succ_{w_2} m_2$. With probability $p = \epsilon\ w_3 \succ_{m_2} w_2$, in which case $C_0 = \{m_1, w_1, m_2, w_2, w_3\}$, and $\phi_0 = \{(m_1 : w_2), (m_2 : w_1)\}$ is the only stable matching. With probability $p = 1-\epsilon\ w_2 \succ_{m_2} w_3$, in which case $C_0 = \{m_1, w_1, m_2, w_2\}$, and $\phi_{m0} = \{(m_1 : w_2), (m_2 : w_1)\}$ and $\phi_{w0} = \{(m_1 : w_1), (m_2 : w_2)\}$ are both stable matchings. Likewise, in the subcase where $w_2 \succ_{m_2} w_1$ and $m_2 \succ_{w_1} m_2$, $m_3 \succ_{w_2} m_2$ with probability $p = 1-\epsilon$, in which case $C_0 = \{m_1, w_1, m_2, w_2\}$, and $\phi_{m0} = \{(m_1 : w_2), (m_2 : w_1)\}$ and $\phi_{w0} = \{(m_1 : w_1), (m_2 : w_2)\}$ are both stable matchings.

Now we calculate the fraction of agents with more than one stable matching.

First, note that, on average, each Markov chain step advances through the preference lists by the distance:

$$\frac{2\left(-2\epsilon^6+2\epsilon^5+2\epsilon^3-2\epsilon^2\right)}{4\epsilon^5-10\epsilon^4+6\epsilon^3-2\epsilon-1} + \frac{4\epsilon^5-10\epsilon^4+6\epsilon^3+2\epsilon^2-2\epsilon}{4\epsilon^5-10\epsilon^4+6\epsilon^3-2\epsilon-1} + \frac{14\epsilon^5-17\epsilon^4+8\epsilon^3-10\epsilon^2+6\epsilon-1}{4\epsilon^5-10\epsilon^4+6\epsilon^3-2\epsilon-1} + \frac{2\left(2\epsilon^6-16\epsilon^5+17\epsilon^4-10\epsilon^3+10\epsilon^2-6\epsilon\right)}{4\epsilon^5-10\epsilon^4+6\epsilon^3-2\epsilon-1}$$

and so the fraction of agents who have multiple stable equilibria:

$$\frac{2(1-\epsilon)\left(-2\epsilon^6+2\epsilon^5+2\epsilon^3-2\epsilon^2\right)}{\left(\frac{2\left(-2\epsilon^6+2\epsilon^5+2\epsilon^3-2\epsilon^2\right)}{4\epsilon^5-10\epsilon^4+6\epsilon^3-2\epsilon-1} + \frac{4\epsilon^5-10\epsilon^4+6\epsilon^3+2\epsilon^2-2\epsilon}{4\epsilon^5-10\epsilon^4+6\epsilon^3-2\epsilon-1} + \frac{14\epsilon^5-17\epsilon^4+8\epsilon^3-10\epsilon^2+6\epsilon-1}{4\epsilon^5-10\epsilon^4+6\epsilon^3-2\epsilon-1} + \frac{2\left(2\epsilon^6-16\epsilon^5+17\epsilon^4-10\epsilon^3+10\epsilon^2-6\epsilon\right)}{4\epsilon^5-10\epsilon^4+6\epsilon^3-2\epsilon-1}\right)(4\epsilon^5-10\epsilon^4+6\epsilon^3-2\epsilon-1)}$$

As a sanity check, I plot the relationship between the fraction of agents with multiple stable matchings, see figure III.B

IV. Next steps

It seems that after consolidating the progress here I should move on in two fronts: relaxing the constraints on preference formation and analyzing the strategic behavior of market participants.

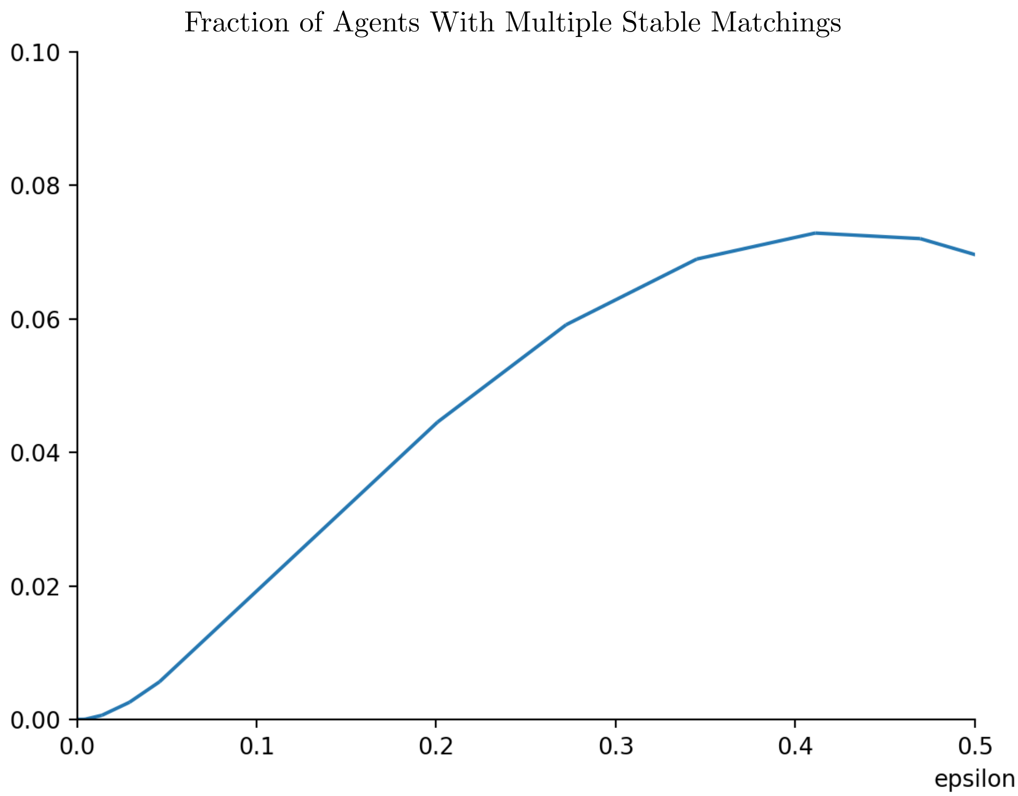
REFERENCES

Roth, Alvin E. 1986. "On the allocation of residents to rural hospitals: a general property of two-sided matching markets." *Econometrica: Journal of the Econometric Society*, 425—-427.

Mathematical Appendix

A1. Permutations of preferences for $\{m_1, m_2, w_1, w_2\}$

I denote preferences with a four-bit binary number, in which a zero in the first position indicates that $w_1 \succ_{m_1} w_2$ while a one in the first position indicates that


Fraction of Agents With Multiple Stable Matchings
0.10
0.08
0.06
0.04
0.02
0.00
0.0
0.1
0.2
0.3
0.4
0.5
epsilon

$w_1 \prec_{m_1} w_2$. Likewise the second through fourth positions indicate the preferences of m_2, , w_1,and w_2.

Type A

$(0,0,0,0)$: $C_0 = \{m_1, w_1\}, \{\phi_0\} = \{m_1, w_1\}$. Preferences for $\{m_2, m_3, w_2, w_3\}$ are unrestricted.

Type B

$(0,0,1,0)$: $C_0 = \{m_2, w_1\}, \{\phi_0\} = \{m_2, w_1\}, C_1 = \{m_1, w_2\}, \{\phi_1\} = \{m_1, w_2\}$. Preferences for $\{m_3, m_4, w_3, w_4\}$ are unrestricted.

$(0,0,1,1)$: $C_0 = \{m_2, w_1\}, \{\phi_0\} = \{m_2, w_1\}, C_1 = \{m_1, w_2\}, \{\phi_1\} = \{m_1, w_2\}$. Preferences for $\{m_3, m_4, w_3, w_4\}$ are unrestricted.

$(0,1,0,1)$: $C_0 = \{m_1, w_1\}, \{\phi_0\} = \{m_1, w_1\}, C_1 = \{m_2, w_2\}, \{\phi_1\} = \{m_2, w_2\}$. Preferences for $\{m_3, m_4, w_3, w_4\}$ are unrestricted.

$(0,1,1,1)$: $C_0 = \{m_2, w_2\}, \{\phi_0\} = \{m_2, w_2\}, C_1 = \{m_1, w_1\}, \{\phi_1\} = \{m_1, w_1\}$. Preferences for $\{m_3, m_4, w_3, w_4\}$ are unrestricted.

$(1,0,0,0)$: $C_0 = \{m_1, w_2\}, \{\phi_0\} = \{m_1, w_2\}, C_1 = \{m_2, w_1\}, \{\phi_1\} = \{m_2, w_1\}$. Preferences for $\{m_3, m_4, w_3, w_4\}$ are unrestricted.

$(1,0,1,0)$: $C_0 = \{m_1, w_2\}, \{\phi_0\} = \{m_1, w_2\}, C_1 = \{m_2, w_1\}, \{\phi_1\} = \{m_2, w_1\}$. Preferences for $\{m_3, m_4, w_3, w_4\}$ are unrestricted.

$(1,0,1,1)$: $C_0 = \{m_2, w_1\}, \{\phi_0\} = \{m_2, w_1\}, C_1 = \{m_1, w_2\}, \{\phi_1\} = \{m_1, w_2\}$. Preferences for $\{m_3, m_4, w_3, w_4\}$ are unrestricted.

$(1,1,0,0)$: $C_0 = \{m_1, w_2\}, \{\phi_0\} = \{m_1, w_2\}, C_1 = \{m_2, w_1\}, \{\phi_1\} = \{m_2, w_1\}$. Preferences for $\{m_3, m_4, w_3, w_4\}$ are unrestricted.

$(1,1,0,1)$: $C_0 = \{m_2, w_2\}, \{\phi_0\} = \{m_2, w_2\}, C_1 = \{m_1, w_1\}, \{\phi_1\} = \{m_1, w_1\}$. Preferences for $\{m_3, m_4, w_3, w_4\}$ are unrestricted.

$(1,1,1,0)$: $C_0 = \{m_1, w_2\}, \{\phi_0\} = \{m_1, w_2\}, C_1 = \{m_2, w_1\}, \{\phi_1\} = \{m_2, w_1\}$. Preferences for $\{m_3, m_4, w_3, w_4\}$ are unrestricted.

$(1,1,1,1)$: $C_0 = \{m_2, w_2\}, \{\phi_0\} = \{m_2, w_2\}, C_1 = \{m_1, w_1\}, \{\phi_1\} = \{m_1, w_1\}$. Preferences for $\{m_3, m_4, w_3, w_4\}$ are unrestricted.

Type C

$(1,0,0,1)$: $C_0 = \{m_1, w_1, m_2, w_2\}$ or $C_0 = \{m_1, w_1, m_2, w_2, w_3\}$ if $w_3 \succ_{m_2} w_2$, $\{\phi_0\} = \{m_1, w_1, m_2, w_2\}$. Preferences for $\{m_3, m_4, w_3, w_4\}$ are unrestricted.

$(0,1,1,0)$: $C_0 = \{m_1, w_1, m_2, w_2\}$ or $C_0 = \{m_1, w_1, m_2, w_2, m_3\}$ if $m_3 \succ_{w_2} m_2$, $\{\phi_0\} = \{m_1, w_1, m_2, w_2\}$. Preferences for $\{m_3, m_4, w_3, w_4\}$ are unrestricted.

Type D

$(0,0,0,1)$: $C_0 = \{m_1, w_1\}, \{\phi_0\} = \{m_1, w_1\}$. Preferences for $\{m_2, m_3, w_2, w_3\}$ fit the pattern $(\cdot, \cdot, 0, \cdot)$.

$(0,1,0,0)$: $C_0 = \{m_1, w_1\}, \{\phi_0\} = \{m_1, w_1\}$. Preferences for $\{m_2, m_3, w_2, w_3\}$ fit the pattern $(0, \cdot, \cdot, \cdot)$.