The Effect of Meeting Rates on Matching Outcomes*

Sonia Jaffe† Simon Weber‡

April 30, 2017

Abstract

We extend the classic matching model of Choo and Siow (2006) to allow for the possibility that rate at which potential partners meet affects their probability of matching. We investigate the implications on estimated match surplus and supermodularity.

1 Introduction

Many matching markets – marriage markets, student-to-school matching, etc – have been analyzed using models in the style of Choo and Siow (2006). These models allow for some unobserved heterogeneity and account for how changes in the number of agents of each type affect matching patterns. However, all agents on the other side of the market are assumed to be potential match partners – agents do not *meet* (and therefore cannot meet at different rates).

We extend a standard matching framework to incorporate the idea of agents meeting; this changes the relationship between match surplus and matching patterns. The standard Choo and Siow (2006) framework over estimates the systematic match surplus for groups with many individuals because it does not account for the fact that a given individual will meet more people from a more populous group. Also, if individuals are more likely to meet others of their own type, the matching surplus is less supermodular than otherwise estimated.

Our model is most closely related to Dupuy and Galichon (2014); in their continuous version of the Choo and Siow setting, men only have access to a set of acquaintances, which is a random subset of the whole population of women. An extension in Menzel (2015), which considers a non-transferable utility matching model also allows for the idea of meeting a subset of agents. However, neither of these papers allow the probability of meeting to vary with type or consider how changes in meeting rates affect match patterns. Our paper is also related to the diverse literature on the assortativeness of marriage,[1] particularly more recent

*The authors are grateful for the insightful comments of Pierre-Andre Chiappori, Alfred Galichon, Scott Kominers, and workshop participants at the University of Chicago.

†University of Chicago spj@uchicago.edu.

‡University of Chicago and Sciences Po simonweber@uchicago.edu.

[1] Papers in sociology and demography (e.g., Schwartz and Mare, 2005; Kalmijn, 1991) have used statistical methods such as log-linear models to provide measures of assortativeness and analyze their trends. Some recent work in economics has focused on the impact on between-household inequalities (Eika et al., 2017; Greenwood et al., 2014).

work that uses structural matching models to estimate assortativeness (e.g., Chiappori et al., 2016). Many papers in this literature focus on education (Chiappori et al., 2016; Eika et al., 2017), a trait for which availability effects could be potentially important. (It is reasonable to think that many people meet potential partners in school or at work where people of their education level are disproportionately represented.) Therefore, it is difficult to know whether measures of assortativeness purefly reflect preferences or also include effects of availability.[2]

While most of this literature have shown an increase in the assortativeness of marriage over the past few decades in the US,[3] changes in 'meeting' frequencies because of population shifts could be one hypothesis to explain the trend. However, we show that a simple effect where individuals are x times as likely to meet an individual of their own type *does not* generate changes in assortativeness over time, so it cannot explain the observed changes in marriage patterns. We consider an alternative model where individuals of a given type (e.g. education level) meet some people (e.g. at work) only from their own group and meet others (e.g. at bars) in proportion to their numbers in the population; importantly, the more potential partners they meet at work, the fewer they seek to meet in the general population. In this model, changes in relative populations can generate changes in the assortativeness of the match.

2 Model

We present two ways for thinking about how differences in meeting rates may affect matching patterns.

Sub-types

The mechanics of Choo and Siow (2006) (CS) require a mass of agents of each potential partner types, so instead of thinking of agents meeting individuals, we model agents as meeting partners of unobservably different subtypes. In our marriage market, a man k and a woman ℓ are categorized by an observable type $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ respectively. Additionally, each type is divided into equal-sized unobservable sub-types $i_k \in \mathcal{I}$ and $j_\ell \in \mathcal{J}$. Types may vary by size, we assume that the mass of each subtypes is fixed to 1, so differences in type populations is reflected in differences in the number of sub-types; there are n_x subtypes of men within type x (i.e. $\{i|x_i = x\}$) and m_y subtypes of women within y, where n_x and m_y are finite for all x, y.

Each subtype of men only meet women from some subtypes and a man cannot match with a woman he does not meet.[4] The set of subtypes of women that men in subtype i meet

[2] There is, however, evidence that individuals have at least some preference for partners with similar education levels, e.g. in Bruze (2011).

[3] An exception is Gihleb and Lang (2016).

[4] Equivalently, assume that the idiosyncratic part of the match is infinitely negative for women of subtypes

make up their choice set $C^i = \cup_{y \in \mathcal{Y}} c_y^i$, where c_y^i is a random subset of $j \in y$. We assume that meeting is reciprocal so that $i \in C^j = \cup_{x \in \mathcal{X}} c_x^j$ if and only if $j \in C^i$. The number of meetings varies only by type so $|c_y^i| = a_{x_i y}$. The number of subtypes of x that a woman of type y meets, b_{xy}, is constrained by an adding up constraint: $m_y b_{xy} = a_{xy} n_x$.

Each man of subtype i must choose whether to remain single or match with a woman of one of the $a_{x_i} = \sum_{y \in \mathcal{Y}} a_{x_i y}$ subtypes that he meets. The non-random component of preferences depends only on type so all sub-types within a type are ex-ante identical. They vary only in which other subtypes they meet and their random preference draws for those subtypes. When man k of subtype i matches with a woman of subtype j, his utility is

$$\alpha_{x_i y_j} - \tau_{x_i y_j} + \epsilon_{kj}$$

where $\alpha_{x_i y_j}$ is a systematic component of i's utility, $\tau_{x_i y_j}$ is an equilibrium transfer paid and ϵ_{kj} is a random component to utility that depends on the subtype of the partner. Similarly, if woman ℓ of subtype j chooses a man from subtype i from her $b_{y_j} = \sum_{x \in \mathcal{X}} b_{xy}$ options, her utility has a type-based systematic component, a transfer, and random component:

$$\gamma_{x_i y_j} + \tau_{x_i y_j} + \eta_{i\ell}.$$

Without loss of generality, we normalize the systematic component of being single to zero so if a man or a woman chooses to remain single, they receive utilities ϵ_{k0} and $\eta_{0\ell}$, respectively.If the random components of utility are independently drawn from an Extreme Value Type I distribution, the mass of matches between men i and women j if they meet is

$$\mu_{ij} = \frac{\exp(\alpha_{x_i y_j} - \tau_{x_i y_j})}{1 + \sum_{g \in C^i} \exp(\alpha_{x_i y_g} - \tau_{ig})}, \qquad \mu_{ij} = \frac{\exp(\gamma_{x_i y_j} + \tau_{x_i y_j})}{1 + \sum_{k \in C^j} \exp(\gamma_{x_i y_k} + \tau_{ik})}.$$

The mass of single men of subtype i and single women of subtype j are

$$\mu_{i0} = \frac{1}{1 + \sum_{j \in C^i} \exp(U_{x_i y_j} - \tau_{x_i y_j})}, \qquad \mu_{0j} = \frac{1}{1 + \sum_{i \in C^j} \exp(V_{x_i y_j} + \tau_{x_i y_j})}.$$

In Choo and Siow terms, these are supply and demand equations. Therefore, at equilibrium these give

$$\mu_{ij}^2 = \exp(\alpha_{x_i y_j} + \gamma_{x_i y_j}) \mu_{i0} \mu_{0j}$$

Since subtypes are ex-ante identical, we expect them to have the same probability of being single. Therefore, the mass of single men of type x is $\mu_{x0} = \mu_{i0} n_x$ for all $i : x_i = x$ (and

not in that group.

similarly for women). The mass of matches between a man i women of type y is

$$\mu_{iy} = \sum_{j \in c_y^i} \mu_{ij} = a_{x_i y} \left(\exp(\alpha_{x_i y_j} + \gamma_{x_i y_j}) \frac{\mu_{0y}}{m_y} \frac{\mu_{x0}}{n_x} \right)^{\frac{1}{2}}.$$

The expected number of matches between men of type x and woman of type y is

$$\begin{aligned} \mu_{xy} = \sum_{i | x_i = x} \mu_{iy} &= n_x a_{x_i y} \left(\exp(\alpha_{x_i y_j} + \gamma_{x_i y_j}) \frac{\mu_{0y}}{m_y} \frac{\mu_{x0}}{n_x} \right)^{\frac{1}{2}} \\ &= \left(a_{xy} b_{xy} \exp(\alpha_{x_i y_j} + \gamma_{x_i y_j}) \mu_{0y} \mu_{x0} \right)^{\frac{1}{2}}, \end{aligned} \tag{1}$$

where the second line follows from the adding up constraint $n_y b_{xy} = a_{xy} m_x$. Equation (1) generalizes the aggregate matching function from Choo and Siow (2006) to take into account meeting rates.

Familiarity

Instead of thinking of meeting as restricting the choice set, we can think of it as increasing an individual's preference for a certain type of person. This is similar in spirit to Mourifie and Siow (2015) who model peer effects by having the utility of a man of type x married with a woman of type y depends positively on the number of matches between those types in the population. In our model, we have instead that a man's utility of matching depends on the number of women of that type that he meets. The utility of man k of matching with a woman of type y when he meets $a_{x_k y}$ women of that type is

$$\alpha_{x_k y} - \tau_{x_k y} + \ln(a_{x_k y}) + \epsilon_{ky}.$$

for woman ℓ, the utility of matching with a man of type x is

$$\gamma_{x y_\ell} + \tau_{x y_\ell} + \ln(b_{x y_\ell}) + \eta_{x\ell}.$$

These give match probabilities

$$\mu_{xy} = \left(a_{xy} b_{xy} \exp(\alpha_{x_i y_j} + \gamma_{x_i y_j}) \mu_{0y} \mu_{x0} \right)^{\frac{1}{2}} \tag{1'}$$

which are the same as in the model based on sub-types.

Equivalence

Since the remainder of the analysis relies only on the equation for μ_{xy} found in Equations (1) and (1'), the two frameworks for conceptualizing the model have all of the same implications. For the sake of simplicity, we mostly stick to the language of the familiarity model for the rest of the paper.

2.1 Surplus

The total systematic surplus when a man of type x and a woman of type y match is

$$\phi_{xy} \equiv \alpha_{xy} + \gamma_{xy} = \log\left(\frac{\mu_{xy}^2}{\mu_{x0}\mu_{0y}}\right) - \log(a_{xy}b_{xy}).$$

The CS model will over estimate the surplus between groups that are very likely to meet each other because it attributes the high match rate to high surplus instead of to a higher rate of meeting. The CS estimate of surplus from the matching frequencies is

$$\phi_{xy}^{CS} = \log\left(\frac{\mu_{xy}^2}{\mu_{x0}\mu_{0y}}\right) = \phi_{xy} + \log(a_{xy}b_{xy}).$$

If we are thinking of a_{xy} and b_{xy} as the number of subtypes a person meets within a type, then correlation of random errors within a type decreases the usefullness and the effect of meeting additional subtypes of a given type. If ρ is the correlation ($E[\epsilon_{ij}\epsilon_{ik}] = \rho$ if $y_j = y_k$),[5] then

$$\phi_{xy}^{CS} = \log\left(\frac{\mu_{xy}^2}{\mu_{x0}\mu_{0y}}\right) = \phi_{xy} + (1-\rho)\log(a_{xy}b_{xy}).$$

Note that perfect correlation ($\rho = 1$) corresponds to the CS case of only getting one random draw per type. See Appendix A.

2.2 Assortativeness

The assortativeness of the match surplus is a measure of how much more surplus is generated when two agents of type A match and two agents of type B match compared to having two matches where one agent is A and one is type B. Letting $x = y = t$ and $x' = y' = t'$, with $t \neq t'$, the supermodularity across these two types is

$$SM_{t,t'} \equiv \phi_{t,t} + \phi_{t',t'} - \phi_{t,t'} - \phi_{t',t}.$$

The different meeting rates make the supermodularity corresponding to a given matching pattern different from the standard CS model. The CS measurement of supermodularity is

$$SM_{t,t'}^{CS} = \log\left(\frac{\mu_{tt}\mu_{t't'}}{\mu_{t't}\mu_{tt'}}\right) = SM_{t,t'} - \log\left(\frac{a_{t't'}b_{t't'}}{a_{t't}b_{t't}}\frac{a_{tt}b_{tt}}{a_{tt'}b_{tt'}}\right). \quad (2)$$

If people are more likely to meet others of their own type, then CS over-estimates the supermodularity of the matching surplus.

[5]Since ϵ is mean zero and variance 1, $\rho = \frac{E[\epsilon_{ij}-\mu_{\epsilon_{ij}}][\epsilon_{jk}-\mu_{\epsilon_{jk}}]}{\sigma_{\epsilon_{ik}}\sigma_{\epsilon_{jk}}} = E[\epsilon_{ij}\epsilon_{ik}]$.

2.3 Measurement

The meeting frequencies cannot be non-parametrically identified separately from the match surplus using only data on matching patterns. However, with logical parameterizations of the frequencies, one can look at how meeting patterns affect the matching and the estimated surplus.

Note that since we have normalized the standard deviation of the random errors to one, the surplus is measured in standard deviation units. If the surplus of matching with a college educated partner is .5 higher than with a high school educated partner, that means it is equivalent to moving up .5 standard deviations in the within-type distribution of match utilities. This means that, if errors are uncorrelated, increasing the number of meetings between two types of agents by 1% while decreasing their match surplus by .01 standard deviation would have no effect.

3 Paramaterizing the meeting rates

We consider three models for how agents meet potential partners and discuss their implications for matching patterns and surplus measurement.

3.1 Baseline: Random meeting

If meetings were completely random, an individual would meet people of each type in proportion to their numbers in the population.

$$a_{xy} = a_x \frac{m_y}{M} \qquad\qquad b_{xy} = b_y \frac{n_x}{N}.$$

The adding up constraint ($a_{xy} \cdot n_x = b_{xy} \cdot m_y$) for each pair of types requires that $a_x \cdot N = b_y \cdot M$ for every x and y, so the number of potential partners an agent meets cannot vary by type, $a_x = a \; \forall x$ and $b_y = b \; \forall y$ with $\frac{a}{b} = \frac{M}{N}$.

In this case, a model that did not account for the number of meetings between types would over estimate the surplus of matching with someone from a large group because it would not take into account that individuals were meeting more people from that group. (If an individual's preference does not vary across members of a group, it matters less how many of them she meets.) The meeting rate a and the average surplus of matching are not separately identified in a single market; however, looking at changes over time, one can adjust for changes in the populations m_y and n_x and see how that affects the estimated change in surplus for those groups.

3.2 Model 1: Increased probability of meeting own type

Since we are particularly interested in assortative matching, we want to allow for the possibility that individuals are more like to meet potential partners of their own type (e.g. education level). Let γ be the additional likelihood of meeting someone of the same type. If we use

$$\gamma_{xy} = \begin{cases} \gamma & x = y, \\ 0 & x \neq y, \end{cases}$$

then we have meeting frequencies

$$a_{xy} = a\frac{m_y}{M}(1+\gamma_{xy}) \cdot \theta \qquad\qquad b_{xy} = b\frac{n_x}{N}(1+\gamma_{xy}) \cdot \theta,$$

where $\theta = \left(1 + \gamma \sum_x \frac{m_x n_x}{NM}\right)^{-1}$ is a multiplier to keep the average number of women a man meets equal to a.[6]

These match frequencies give

$$\phi^{CS}_{xy} - \phi_{xy} = \log\left(\theta^2 ab\right) + \log\left(\frac{m_y}{M}\right) + \log\left(\frac{n_x}{N}\right) + 2\log(1+\gamma) \cdot 1\{x = y\}. \qquad (3)$$

The first three terms come from the random matching. These terms may cause variation in ϕ^{CS}_{xy} over time even if the fundamental surplus is unchanged, because the population counts m_x and n_y can vary over time (and θ depends on these counts). However, since they do not effect the surplus specific to a given pair, they net out when estimating the supermodularity.

The last term in Equation (3) only applies for matches between the same type. It therefore affects the estimated supermodularity of the match surplus. Using Equation (2), if $x = y = t$ and $x' = y' = t'$, the observed measure supermodularity is

$$SM^{CS}_{t',t} = SM_{t',t} + 4\log(1+\gamma). \qquad (4)$$

If γ is positive and the econometrician assumes it is zero, that will lead to an overestimate of the supermodularity of the matching surplus. However, if γ has not changed overtime, it cannot explain the changes overtime in the measured supermodularity because it does not interact with the population counts that vary over time.

An informal explanation of the increasing educational assortativeness in marriage matching is that since more women are going to college, college educated men are more likely to meet college educated women. However unless there has also been an *increase in the relative probability* of meeting a given college educated woman verse a given woman without a college education, the model shows this simple explanation is insufficient.

[6]The fact that a and b do not vary by type is again a result of the adding up constraint.

3.3 Model 2: Two ways of meeting

An extension of the idea that more college educated men are meeting more college educated women in college is that they then have less reason to meet women other places. Say that a man i meets γ people from a population that are all his type – e.g. at school or work – and $\frac{m_{x_i}}{m_{x_i}+n_{x_i}}$ of them are female. He then meets the rest of the a total women in proportion to their presence in the general population – e.g. at a bar, at the DMV. The fraction he meets from the general population is

$$q_x = 1 - \frac{\gamma}{a}\frac{m_x}{n_x + m_x}.$$

Similarly, women meet a total of b men; $\gamma\frac{m_{x_i}}{m_{x_i}+n_{x_i}}$ men from only their own type and fraction

$$r_y = 1 - \frac{\gamma}{b}\frac{n_y}{n_y + m_y}$$

from the general population.

Those from the general population are met in proportion to their presence in the population weighted by the number of people they are looking to meet. This gives meeting frequencies

$$a_{xy} = aq_x\frac{m_y r_y}{\sum_{y'} m_{y'} r_{y'}} \qquad\qquad b_{xy} = br_y\frac{n_x q_x}{\sum_{x'} n_{x'} q_{x'}}$$

when $x \neq y$. The first term aq_x is the remaining number of women a man of type x wants to meet in the general population. The fraction is the number of men that women of type y are looking to meet, $bm_y r_y$, divided by the number of men that all women are looking to meet, $b\sum_{y'} m_{y'} r_{y'}$.

When $x = y$,

$$\begin{aligned} a_{xx} &= a\left(q_x\frac{m_x r_x}{\sum_{y'} m_{y'} r_{y'}} + \frac{\gamma}{a}\frac{m_x}{n_x + m_x}\right) & b_{xx} &= b\left(r_x\frac{n_x q_x}{\sum_{x'} n_{x'} q_{x'}} + \frac{\gamma}{b}\frac{m_y}{n_y + m_y}\right) \\ &= aq_x\frac{m_x r_x}{\sum_{y'} m_{y'} r_{y'}}\left(1 + \frac{\gamma}{a}\frac{\sum_{y'} m_{y'} r_{y'}}{(n_x + m_x) q_x r_x}\right) & &= br_x\frac{n_x q_x}{\sum_{x'} n_{x'} q_{x'}}\left(1 + \frac{\gamma}{b}\frac{\sum_{x'} n_{x'} q_{x'}}{(n_x + m_x) q_x r_x}\right). \end{aligned}$$

A man of type x meets $\gamma\frac{m_x}{n_x+m_x}$ women of his own type at work, but also meets $aq_x\frac{m_x r_x}{\sum_{y'} m_{y'} r_{y'}}$ of them in the general population.[7]

[7] The meeting frequencies satisfy the adding up constraint $a_{xy} n_x = b_{xy} m_y$ because

$$\frac{1}{a}\sum_y m_y r_y = \frac{N}{M}\frac{1}{b}\left(M - \frac{\gamma}{a}\sum_x \frac{m_x n_x}{m_x + n_x}\right) = \frac{1}{b}\left(N - \frac{\gamma}{a\frac{M}{N}}\sum_x \frac{m_x n_x}{m_x + n_x}\right) = \frac{1}{b}\sum_x n_x q_x.$$

When calculating supermodularity, the terms of the form $aq_x \frac{m_y r_y}{\sum_{y'} m_{y'} r_{y'}}$ will cancel, but the ones in parentheses will not; therefore, there is an interaction between the meeting rate $\frac{\gamma}{b}$ and population counts. Let $\theta = \sum_x n_x q_x = b/a \sum_y m_y r_y$, which captures the overall amount of meeting in the general population. If $x = y = t$ and $x' = y' = t'$, supermodularity is

$$SM^{CS}_{t't',tt} - SM_{t't',tt} = 2 \log \left(\left(1 + \frac{\gamma}{b} \frac{\theta}{(n_t + m_t) q_t r_t} \right) \left(1 + \frac{\gamma}{b} \frac{\theta}{(n_{t'} + m_{t'}) q_{t'} r_{t'}} \right) \right). \quad (5)$$

Unlike in the previous models, changing populations of types can affect the Choo and Siow measure of supermodularity. Population counts have both a direct effect on the measured supermodularity and an indirect effect via θ. The direct effect of moving a woman to a group with fewer women is to decrease the effect of meeting rates on the assortativeness of the match, (thereby decreasing the measured supermodularity). The intuition is that when there are few women of a given type, most of them go to partners of the same type; as the number of women of that type grows they still disproportionately meet with their own type, but the marginal share is smaller than the average share. The decrease in the extent to which men disproportionately meet women of their own type causes a decrease in the extent to which they disproportionately match to women of their own type – a decrease in the assortativeness of the match. Formally, if there are the same number of men of types t' and t and more women of type t than of type t', moving a woman from t to t' will decrease $SM^{CS}_{t',t}$. (See Appendix B.)

The indirect effect of population counts on the supermodularity is via θ. The effect of shifting a woman from t to t' on θ is

$$d\theta = -\frac{\gamma}{a} \left(\left(\frac{n_{t'}}{n_{t'} + m_{t'}} \right)^2 dm_{t'} + \left(\frac{n_t}{n_t + m_t} \right)^2 dm_t \right),$$

so moving a woman between t and t' decreases θ whenever

$$-\left(\left(\frac{n_t}{n_t + m_t} \right)^2 (-1) + \left(\frac{n_{t'}}{n_{t'} + m_{t'}} \right)^2 (1) \right) < 0$$

$$\left(\frac{n_t}{n_t + m_t} \right) < \left(\frac{n_{t'}}{n_{t'} + m_{t'}} \right) \Rightarrow \left(\frac{m_t}{n_t + m_t} \right) > \left(\frac{m_{t'}}{n_{t'} + m_{t'}} \right)$$

The indirect effect of shifting a woman from t to t' depends on the relative gender shares of the two types. Making the gender shares within types more balanced indirectly decreases the effect of meeting rates on the match probabilities. If there are the same number of men of both types,[8] then the direct effect and the indirect effect via the amount of searching

[8] This is a sufficient, but not necessary condition.

for partners in the general population (θ) work in the same direction: both indicate that shifting a woman from a type that has more women to a type that has fewer will decrease the assortativeness of the match.

4 Conclusion

We adapt the Chow and Siow matching model to allow for potential partners to meet each other at different rates. We show that both random meeting and meeting where one is more likely to meet one's own type change the relationship between the match surplus function and the equilibrium matching. Having a higher probability of meeting one's own type increases the assortativeness of the observed match, but not in a way that interacts with population counts; therefore it cannot explain changes in assortativeness over time. We develop an alternative model where meeting additional potential partners of one's own type makes a person seek fewer potential partners of other types. In this model the meeting rates interact with population counts and could lead to changes in match assortativenss over time. It predicts that assortative mating will generally decrease when the ratios within genders become more equal (holding match surplus fixed).

References

Bruze, G. (2011). Marriage choices of movie stars: Does spouse's education matter? *Journal of Human Capital*, 5(1):1–28.

Chiappori, P.-A., Salanié, B., and Weiss, Y. (2016). Partner choice and the marital college premium: Analyzing marital patterns over several decades. *Working Paper*.

Choo, E. and Siow, A. (2006). Who marries whom and why. *Journal of Political Economy*, 114(1):175–201.

Dupuy, A. and Galichon, A. (2014). Personality traits and the marriage market. *Journal of Political Economy*, 122(6):1271–1319.

Eika, L., Mogstad, M., and Zafar, B. (2017). Educational assortative mating and household income inequality. *Working Paper*.

Gihleb, R. and Lang, K. (2016). Educational homogamy and assortative mating have not increased. *Working Paper*.

Greenwood, J., Guner, N., Kocharkov, G., and Santos, C. (2014). Marry your like: Assortative mating and income inequality. *American Economic Review*, 104(5):348–353.

Kalmijn, M. (1991). Status homogamy in the United States. *American Journal of Sociology*, 97(2):496–523.

Menzel, K. (2015). Large matching markets as two-sided demand systems. *Econometrica*, 83(3):897–941.

Mourifie, I. and Siow, A. (2015). Cohabitation versus marriage: Marriage matching with peer effects. *Working Paper*.

Schwartz, C. R. and Mare, R. D. (2005). Trends in educational assortative marriage from 1940 to 2003. *Demography*, 42(4):621–646.

A Nested Logit

Let the correlation of the random utility shocks for the subtypes within each type be ρ; if $\rho = 1$ then all the sub-types within a type are equivalent and if $\rho = 0$ then the sub-types within a type are as different from each other as they are from sub-types of a different type. In this case the match probabilities give

$$\begin{aligned}\frac{\mu_{ij}}{\mu_{i0}} &= \exp\left(\frac{\alpha_{x_i y_j} - \tau_{x_i y_j}}{1-\rho}\right)\left(\sum_{j' \in c^i_{y_j}} \exp\left(\frac{\alpha_{x_i y_j} - \tau_{x_i y'_j}}{1-\rho}\right)\right)^{-\rho} \\ &= \exp\left(\frac{\alpha_{ij} - \tau_{x_i y_j}}{1-\rho}\right)^{1-\rho} a_{x_i y_j}^{-\rho}.\end{aligned}$$

Combining with the equivalent formula for women, we get

$$\mu_{ij}^2 = \exp(\alpha_{x_i y_j} + \gamma_{x_i y_j})\mu_{i0}\mu_{0j} a_{x_i y_j}^{-\rho} b_{x_i y_j}^{-\rho}.$$

Aggregating, as done in the text, gives

$$\mu_{xy} = \left(a_{xy}^{1-\rho} b_{xy}^{1-\rho} \exp(\alpha_{xy} + \gamma_{xy})\mu_{0y}\mu_{x0}\right)^{\frac{1}{2}}.$$

This gives the surplus formula

$$\phi_{xy} = \log\left(\frac{\mu_{xy}^2}{\mu_{0y}\mu_{x0}}\right) - (1-\rho)\log(a_{xy}b_{xy}) = \phi_{xy}^{CS} - (1-\rho)\log(a_{xy}b_{xy}).$$

B Effect of population counts on θ

The increased assortativeness due to meeting (from Equation (5)) is

$$2\log\left(1 + \frac{\gamma}{b}\frac{\theta}{f(n_t, m_t)}\right) + 2\log\left(1 + \frac{\gamma}{b}\frac{\theta}{f(n_{t'}, m_{t'})}\right)$$

where $f(n_t, m_t) = (n_t + m_t)q_t r_t > 0$.

Consider moving a woman from t to t'. This will decrease the assortativeness whenever

$$\frac{-f_2(n_t, m_t)}{f^2(n_t, m_t) + \frac{\theta\gamma}{b} f(n_t, m_t)}(-1) + \frac{-f_2(n_{t'}, m_{t'})}{f^2(n_{t'}, m_{t'}) + \frac{\theta\gamma}{b} f(n_{t'}, m_{t'})}(1) < 0,$$

that is when

$$\frac{f_2(n_{t'}, m_{t'})}{f^2(n_{t'}, m_{t'}) + \frac{\theta\gamma}{b} f(n_{t'}, m_{t'})} > \frac{f_2(n_t, m_t)}{f^2(n_t, m_t) + \frac{\theta\gamma}{b} f(n_t, m_t)}. \tag{6}$$

Expanding $f(\cdot,\cdot)$ (and recalling that $aN = bM$), we have

$$f = (n_t + m_t) - \frac{\gamma}{b} n_t - \frac{\gamma}{b}\frac{N}{M} m_t + \frac{N}{M}\left(\frac{\gamma}{b}\right)^2 \frac{n_t m_t}{n_t + m_t},$$

so

$$f_2 = 1 - \frac{\gamma}{b}\frac{N}{M} + \frac{N}{M}\left(\frac{\gamma}{b}\frac{n_t}{n_t + m_t}\right)^2 > 0,$$

$$f_{22} = -2\frac{N}{M}\left(\frac{\gamma}{b}\right)^2 \frac{n_t^2}{(m_t + n_t)^3} < 0.$$

The function f is increasing in the number of women m, and the derivative of f with respect to m is decreasing in m. Therefore, if $n_t = n_{t'}$, moving a woman from t to t' decreases assortativeness (Equation (6) holds) whenever $m_t > m_{t'}$. That is, assortativeness will decrease as we split women more evenly across the two groups.