Labeling Dependence in Separable Matching Markets

Alfred Galichon[1] **Bernard Salanié**[2]

December 9, 2019

[1]Economics Department, FAS and Mathematics Department, Courant Institute, New York University; e-mail: ag133@nyu.edu. Galichon's research has received funding from NSF DMS-1716489.
[2]Department of Economics, Columbia University; e-mail: bsalanie@columbia.edu.

Abstract

Independence of irrelevant alternatives (IIA) has been much studied in single-agent decision problems. We explore its extension to models of two-sided choice and perfectly transferable utility. We start with models with a separable logit structure, à la Choo and Siow (2006). We first show that this model satisfies a weak version of IIA. On the other hand, we conjecture that no separable model satisfies a stronger version of IIA. We then exhibit a two-sided version of the "blue bus/red bus" paradox, which shows that the separable logit model is not robust to irrelevant relabeling.

Keywords: matching, IIA.

JEL codes: C78, D61, C13.

Introduction

Independence of irrelevant alternatives (IIA) has figured prominently in at least three different contexts in economics. Its first appearance dates back to Nash's paper on the bargaining problem (Nash (1950)). When discussing the representation of choice by a utility function, Nash's property 7 required that an optimal choice cannot become non-optimal when the choice set is restricted to a subset that contains it. In formal terms, let $C(S)$ denote the set of optimal choices set from a set of alternatives. Nash required that if $C(S) \subset T \subset S$, then $C(T) = C(S)$. While this condition is necessary and sufficient for C to be represented by a binary relation, it may not be acyclic[1]. In the monograph that gave birth to social choice theory, Arrow (1951) defined independence of irrelevant alternatives as imposing that social preferences between a pair of alternatives only depend on the collection of individual preferences over that pair. The next step was taken by Luce and Raiffa (1957) and Luce (1959) with decision-making under risk. Luce called independence from irrelevant alternatives his Axiom 1 (p. 6), which requires that probabilistic choice satisfy:

$$\text{if } R \subset S \subset T, \quad \text{then } P_T(R) = P_S(R)P_T(S)$$

where $P_T(S)$ is the probability that choice from T belongs to S. Luce described it as "one reasonable possibility" to extend to probabilistic choice the notion that adding new alternatives should not change preferences between existing alternatives. As is now well-known, Luce's IIA has stark consequences for stochastic choice: it implies the existence of a positive function v such that for any finite set S and any alternative $x \in S$,

$$P_S(x) = \frac{v(x)}{\sum_{y \in S} v(y)}.$$

If we define $u(x) = \ln v(x)$, then choice probabilities take the familiar "multinomial logit" form

$$P_S(x) = \frac{\exp(u(x))}{\sum_{y \in S} \exp(u(y))}.$$

[1]Arrow (1959) showed that if the domain of the choice correspondence C contains all finite subsets, then IIA is necessary and sufficient for the existence of a transitive complete binary relation that represents C.

The aim of this note is to explore the extension of IIA to a class of models of two-sided choice, i.e. matching models. Indeed, matching models with a random utility structure have recently gained popularity (in the wake of Dagsvik (2000) and Choo and Siow (2006)) and it is important to understand the implications carried by the random utility structure. The literature has already pointed out that the logit specification implies implausible substitution patterns, just as it does in one-sided choice models (Graham, 2013; Choo, 2015). Extensions have been proposed to remedy this—see Mourifié and Siow (2017), Mourifié (2019) and Galichon and Salanié (2019). Our focus here is different: we will show that in addition to constraining cross-elasticities, these models are prone to well-known paradoxes tied up to the logit structure.

We will focus on one-to-one bipartite matching models with perfectly transferable utility. In addition, we will restrict our analysis to markets in which the joint surplus is "separable" in the sense of Chiappori, Salanié, and Weiss (2017) and Galichon and Salanié (2019). Separability is a concept developed with empiricists in mind, following the pioneering work of Choo and Siow (2006). It assumes that conditional on observed "types", the characteristics of the agents that are unobserved by the analyst do not interact in the production of joint surplus. We consider this as a first step towards a more general inquiry into IIA and related properties in matching models.

Section 1 defines our setup and notation. We present two definitions of IIA for matching models in section 2. Section 3 shows that weak IIA holds in the Choo and Siow (2006) model as well as in a specific, nested extension. Section 4 is directly motivated by the well-known blue bus/red bus example in the tradition of Debreu (1960). It shows that irrelevant relabeling of types by the analyst may lead to incorrect predictions in the logit model.

1 Separable Matching with Transferable Utility

In all of the following, we consider frictionless bipartite matching with perfectly transferable utility (TU). Each match is formed of two partners drawn from separate populations. For simplicity, we call one the "husband" and one the "wife". Husbands are drawn from a population of men indexed by $i \in \mathcal{I}$, and wives from a population of women indexed by $j \in \mathcal{J}$. We call i the *identity* of a man, as opposed to his *type* $x \in \mathcal{X}$. Identities are observed by all participants in the market. On the other hand, the econometrician only observes *types*, which partition the set of identities. We will write $i \in x$ or $x_i = x$ to denote that the man of identity i has type x. The corresponding notation for women will be $j \in y$ or $y_j = y$, where $y \in \mathcal{Y}$. We do not restrict the sets $\mathcal{I}, \mathcal{J}, \mathcal{X}$ and $\mathcal{Y}$ at this stage. We will denote $F_{\mathcal{I}}$ (resp. $F_{\mathcal{J}}$) the cumulative distribution function of i (resp. j.)

Note here that some of the type information available to the econometrician may well be payoff-irrelevant. This will matter in section 4.

A match between a man i and a woman j generates a *joint surplus* $\tilde{\Phi}_{ij}$. This is shared between the two partners so that they achieve individual match surpluses $\tilde{U}_{ij}, \tilde{V}_{ij} = \tilde{\Phi}_{ij} - \tilde{U}_{ij}$, over and above the utilities they get by remaining single. We denote singlehood as "partner 0", and we will use the notation $\mathcal{X}_0 = \mathcal{X} \cup \{0\}$ and $\mathcal{Y}_0 = \mathcal{Y} \cup \{0\}$.

A *matching* is a collection of numbers $0 \leq \mu_{ij} \leq 1$ that represent the probability that man i and woman j are matched. It must be feasible:

$$\forall i, \quad \int \mu_{ij} dF_{\mathcal{J}}(j) \leq 1 \quad \text{and} \quad \forall j, \quad \int \mu_{ij} dF_{\mathcal{I}}(i) \leq 1.$$

Our equilibrium concept is stable matching. A stable matching is a feasible matching that maximizes total joint surplus

$$\int \int \mu_{ij} dF_{\mathcal{I}}(i) dF_{\mathcal{J}}(j).$$

Associated to the feasibility constraints are multipliers $\tilde{u}_i$ and $\tilde{v}_j$; the corresponding first-order conditions are

$$\tilde{\Phi}_{ij} \leq \tilde{u}_i + \tilde{v}_j \quad \text{for all } i, j,$$

with equality for any match that has non-zero probability in equilibrium ($\mu_{ij} > 0$). Such an equilibrium match must split the surplus in such a way that

$$\tilde{\Phi}_{ij} = \tilde{u}_i + \tilde{v}_j$$

for some admissible values of the multipliers.

Since econometricians only observe types, their data can only consist of

- the numbers n_x (resp. m_y) of available men (resp. women) for every type $x \in \mathcal{X}$ (resp. $y \in \mathcal{Y}$)
- for each pair (x, y) the numbers μ_{xy} of matches between men of type x and women of type y.

They could also observe some statistic on the joint surplus of (x, y) matches. This could be the distribution of the joint surplus of all observed (x, y); or, less ambitiously, its average value, or some statistic like the number of divorces that brings some information on outcomes. We shall assume away such information in this paper. We will denote μ_{x0} the number of single men of type x, which is obtained by subtracting the total number of their matches from n_x.

Moreover, we will assume that *conditional on observed types*, interactions between the unobserved characteristics of the partners do not create (or destroy) joint surplus. More precisely, we state:

Assumption 1 (Separability)**.** *Equivalently:*

(i) the joint surplus from a match between man $i \in x$ and woman $j \in y$ can be written as

$$\tilde{\Phi}_{ij} = \Phi_{xy} + \varepsilon_y^i + \eta_x^j.$$

(ii) if men i and i' both have type x and women j and j' both have type y, then

$$\tilde{\Phi}_{ij} + \tilde{\Phi}_{i'j'} = \tilde{\Phi}_{ij'} + \tilde{\Phi}_{i'j}.$$

Separability was defined in Chiappori, Salanié, and Weiss (2017) and underlies the general analysis of TU models of Galichon and Salanié (2019). We refer the reader to these papers for a discussion. What matters most here is that separability allows us to decompose the two-sided equilibrum into a family of one-sided discrete choice problems with endogenous, type-dependent prices. More formally, there exists a decomposition $\Phi_{xy} = U_{xy} + V_{xy}$ of the types-driven part of joint surplus such that in equilibrium[2],

- man i is matched with a partner of the type y that maximizes $(U_{x_i t} + \varepsilon^i_t)$ over t
- woman j is matched with a partner of the type x that maximizes $(V_{zy_j} + \eta^j_z)$ over z.

The analogy with the one-sided problem is clear; but the crucial difference is that the type-dependent parameters U_{xy} and V_{xy} are endogenous, unlike the "mean utilities" of the random utility models. This will require adapting the definition of IIA.

2 IIA in Matching Models

In matching with transferable utilities, as in any equilibrium model, prices and allocations reflect value and scarcity. Intuition suggests that if for instance a man i belongs to a type x that is highly valued, in the sense that the values of Φ_{xy} are high for all potential types of partners y, then this man will do relatively well on the marriage market. Since the market-clearing prices are reflected in the decomposition $\boldsymbol{\Phi} = \boldsymbol{U} + \boldsymbol{V}$, this is simply saying that such men tend to marry women who are also highly valued, and to appropriate a large share of the joint surplus in their marriage. The same conclusion obtains if type x is relatively rare, that is if n_x is relatively small.

Now consider the ratio μ_{xy}/μ_{xt} for two types of women $y \neq t$. Dividing through by n_x, this can also be written as the ratio of the probabilities that any given man of type x will marry a woman of type y or t. We will denote $\mu^{\mathcal{X}}_{y|x}$ the probability that a man of type x marries a woman of type y (and $\mu^{\mathcal{Y}}_{x|y}$ the probability that a woman of type y marries a man of type x). Then μ_{xy}/μ_{xt} is the odds ratio for men of type x,

$$R^{\mathcal{X}}(y, t; x) = \frac{\mu^{\mathcal{X}}_{y|x}}{\mu^{\mathcal{X}}_{t|x}}.$$

[2]With obvious adaptations to account for unmatched partners.

In a one-sided model, IIA requires that this ratio be independent of the set of women in the marriage market, as long as this set includes some women of type y and some women of type t. The previous paragraph suggests that this cannot hold in a matching market: at a minimum, the odds ratio must depend on the numbers of women of types y and t.

Remember that the primitives of a separable matching model are the numbers of men and women of each type (n_x) and (m_y); the values of the type-dependent joint surplus (Φ_{xy}); and the distributions of the unobserved terms (ε^i_y) and (η^j_x). We propose a weaker definition of IIA:

Property 1 (Weak IIA for separable matching with transferable utilities)**.** *Fix the parameters* (Φ_{xy}) *and the distributions of the unobserved terms* ε^i_y *and* η^j_x.

The model satisfies weak IIA *if and only if for all types of men* x *and* z *in* $\mathcal{X}$ *and all types of women* y *and* t *in* $\mathcal{Y}$*, the double odds ratio*

$$\frac{R^{\mathcal{X}}(y,t;x)}{R^{\mathcal{X}}(y,t;z)}$$

is independent of all subpopulation sizes (n_x) *and* (m_y).

Note that this ratio is simply

$$\frac{\mu_{xy}\mu_{zt}}{\mu_{xt}\mu_{zy}}.$$

Let us define the odds ratio $R^{\mathcal{Y}}(x,z;y)$ for women of type y as $\mu^{\mathcal{Y}}_{x|y}/\mu^{\mathcal{Y}}_{z|y}$. Then the double odds ratio in Property 1 can also be written as

$$\frac{R^{\mathcal{Y}}(x,z;y)}{R^{\mathcal{Y}}(x,z;t)}$$

so that the definition applies to both sides of the market.

It is not obvious a priori that there exist separable matching models in which Property 1 holds; but as we will see in section 3, the Choo and Siow (2006) model satisfies it, and so does one specific class of nested logit models.

Weak IIA implies more specific restrictions; for instance, the elasticity of $\Pr_S(x)$ with respect to the mean utility $u(y)$ of an element y of S is the same for all $y \in S - \{x\}$. This is a very unappealing restriction to impose on a demand system. Partly as a consequence, the empirical literature moved away from the multinomial logit model; it adopted variants in which IIA does not hold, such as mixed multinomial logit. We will focus in section 4 on a more fundamental paradox, inspired by the famous blue bus-red bus example.

A valid criticism of the notion of IIA we have adopted in property 1 is that since it does not include the option of remaining unmatched, it leaves out one of the marital options. To remedy this, we introduce a strong IIA property as follows:

Property 2 (Strong IIA for separable matching with transferable utilities)**.** *Choose the parameters* (Φ_{xy}) *and the distributions of the unobserved terms* ε_y^i *and* η_x^j*.*

The resulting model satisfies strong IIA *if and only if the following two sets of conditions are met:*

(i) for all types of men x *and* z *in* $\mathcal{X}$ *and all men's marital options* y *and* t *in* $\mathcal{Y}_0$*, the double odds ratio*

$$\frac{R^{\mathcal{X}}(y,t;x)}{R^{\mathcal{X}}(y,t;z)} = \frac{\mu_{xy}\mu_{zt}}{\mu_{xt}\mu_{zy}}$$

is independent of all subpopulation sizes (n_x) *and* (m_y)*; and*

(ii) for all types of women y *and* t *in* $\mathcal{Y}$ *and all women's marital options* x *and* z *in* $\mathcal{X}_0$*, the double odds ratio*

$$\frac{R^{\mathcal{Y}}(x,z;y)}{R^{\mathcal{Y}}(x,z;t)} = \frac{\mu_{xy}\mu_{zt}}{\mu_{xt}\mu_{zy}}$$

is independent of all subpopulation sizes (n_x) *and* (m_y)*.*

This is a strong concept; so strong that even the baseline model of Choo and Siow does not satisfy it as we shall next see.

3 IIA in Separable Models

Choo and Siow (2006) added two assumptions to the separable model: that markets are "large" and that the unobserved terms of Assumption 1 are iid draws from a standard type-I extreme value distribution.

They showed that it generates a very convenient multinomial logit form for the separable model. More precisely, the Choo and Siow model has

$$\mu_{y|x} = \frac{\exp(U_{xy})}{\sum_t \exp(U_{xt})}$$
$$\mu_{x|y} = \frac{\exp(V_{xy})}{\sum_t \exp(V_{ty})}$$
$$U_{xy} + V_{xy} = \Phi_{xy}$$

where the (U_{xy}) and the (V_{xy}) are the equilibrium quantities defined in section 1. In addition, in all models of this class we have

$$\frac{\mu_{xy}\mu_{zt}}{\mu_{xt}\mu_{zy}} = \exp\left(\left(\Phi_{xy} + \Phi_{zt} - \Phi_{xt} - \Phi_{zy}\right)/2\right),$$

which is independent of the numbers of men and women. As a consequence, Property 1 holds in all of them.

Proposition 1. *The Choo and Siow model satisfies weak IIA as defined in Property 1.*

Note that the Choo and Siow model is not the only separable model that satisfies weak IIA. Consider a separable model where the heterogeneity is a nested logit on both sides of the market, with only two nests: one for singlehood, another one for all other marital options. The coefficient of the nests is λ on the side of men, γ on the side of women; the Choo-Siow model obtains for $\lambda = \gamma = 1$. It is not hard to see that

$$\mu_{xy} = \frac{\mu_{x0}^{1/(\lambda+\gamma)}}{\left(m_y - \mu_{0y}\right)^{\frac{1-\lambda}{\lambda+\gamma}}} \frac{\mu_{0y}^{1/(\lambda+\gamma)}}{\left(m_y - \mu_{0y}\right)^{\frac{1-\gamma}{\lambda+\gamma}}} \exp\left(\frac{\Phi_{xy}}{\lambda+\gamma}\right).$$

This implies that $\log \mu_{xy} - \frac{\Phi_{xy}}{\lambda+\gamma}$ is additively separable between x and y;

$$\frac{\mu_{xy}\mu_{zt}}{\mu_{xt}\mu_{zy}} = \exp\left(\left(\Phi_{xy} + \Phi_{zt} - \Phi_{xt} - \Phi_{zy}\right)/(\lambda+\gamma)\right),$$

and this model too satisfies weak IIA. We have not found any other separable model for which weak IIA holds. Conversely, the dynamic extension of the Choo and Siow (2006) model by Choo (2015) shares its logit structure but it does not exhibit weak IIA, as shown by inspecting his formula (3.1). This is due to the combination of time discounting and a finite horizon.

It is straightforward to see that Choo and Siow's model does *not* satisfy strong IIA. Indeed, using the definition with $y = 0$ in Choo and Siow's model gives

$$\frac{R^{\mathcal{X}}(0,t;x)}{R^{\mathcal{X}}(0,t;z)} = \frac{\mu_{x0}}{\mu_{z0}}\frac{\mu_{zt}}{\mu_{xt}} = \sqrt{\frac{\mu_{x0}}{\mu_{z0}}}\exp\left((\Phi_{zt} - \Phi_{xt})/2\right)$$

which clearly depends on the subpopulation sizes. A similar calculation shows that the nested logit model described above does not satisfy IIA for any value of λ and γ. In fact, we conjecture that no separable model can satisfy strong IIA. The intuition is simple. Adding one man to the population of type x must reduce the average equilibrium utility of this type of men. In separable models, this must reduce all $\mu^X_{t|z}$ for all $z, t > 0$ and increase the probabilities of singlehood for all men[3]. Therefore all ratios μ_{zt}/μ_{z0} must decrease; but they should decrease most for men of type x.

4 A modified blue-bus/red-bus example

In his review of Luce (1959), Debreu (1960) showed how IIA leads to counterintuitive predictions. His example used classical music recordings; we will give it in the form popularized by McFadden (1974) (p. 113) as the "blue bus/red bus example". In this story, commuters initially can only go to work with their car or with a blue bus; and a third of them choose to take the bus. Suppose that the bus company adds red buses to its fleet, and the population has no color preferences; then one would not expect the proportion of bus trips to change. But according to IIA, car trips should still be twice more frequent than trips with blue buses, and also than trips with red buses. This is only possible if the proportion of car trips becomes one half. To put it differently, IIA suggests that 25% of car commuters should start taking the bus simply because of a color change that (we assumed) means nothing to them.

[3]This follows from the formulæ in section IV.A of Galichon and Salanié (2017).

What is really at stake in this story is that simply *splitting* options should not change market shares. What matters to commuters is the essence of a bus, not its payoff-irrelevant attributes.

Matching markets are two-sided by their very nature; this is reflected in the endogeneous nature of the decomposition $\boldsymbol{\Phi} \equiv \boldsymbol{U}+\boldsymbol{V}$. Still, it is not hard to construct illustrations similar to Debreu's example. Because "mean utilities" are endogenous, notation and characterization take more work; but the intuition is similar. The following example shows that the multinomial logit model of Choo and Siow (2006) indeed generates similar choice paradoxes.

Let x and y consist of education, with two levels C (college) and N (no college), and suppose that the matrix $\boldsymbol{\Phi}$ of Assumption 1 has

$$\exp(\Phi_{NN}/2) = a \; ; \; \exp(\Phi_{NC}/2) = \exp(\Phi_{CN}/2) = b \; ; \; \exp(\Phi_{CC}/2) = c,$$

where a, b, c are arbitrary positive numbers.

Call this the *original model.* Now let us distinguish two types of college graduates: those (C_e) whose Commencement fell on an even-numbered day and those (C_o) for whom it was on an odd-numbered day. We will assume that this difference is payoff-irrelevant, so that

$$\exp(\Phi_{NC_i}/2) = \exp(\Phi_{C_i n}/2) = a \;\text{ for } i = e, o$$

and $\exp(\Phi_{C_i C_j}) = c$ for $i, j = e, o$. We will also assume that the population of college graduates is split evenly across Commencement days: $n_{C_e} = n_{C_o}$ and $m_{C_e} = m_{C_o}$.

We now show that adding the irrelevant Commencement distinction in this *revised model* is "equivalent" to changing the joint surplus Φ of the original model. More precisely:

Proposition 2. *In the revised model, define* $\mu_{CC} = \sum_{i,j=e,o} \mu_{C_i C_j}$ *the total number of matches between college-educated partners;* $\mu_{CN} = \mu_{C_e,N} + \mu_{C_o,N}$ *the total number of matches between college-educated men and non-college women (and symmetrically* μ_{NC}*); and* $\mu_{C0} = \mu_{C_e,0} + \mu_{C_o,0}$ *the total number of college-educated men who remain single (and symmetrically* μ_{0C}*).*

In equilibrium, these numbers are identical to the equilibrium matching patterns of the original model after substituting $\mathbf{\Phi}'$ *to* $\mathbf{\Phi}$*, where*

$$\exp(\Phi'_{NN}/2) = a \; ; \; \exp(\Phi'_{NC}/2) = \exp(\Phi'_{CC}/2) = b\sqrt{2} \; ; \; \exp(\Phi'_{CN}/2) = 2c.$$

Proof: see the Appendix.

It is easy to check that substituting $\mathbf{\Phi}'$ to $\mathbf{\Phi}$ does not affect $\mathcal{C} = ac/b^2$. On the other hand, it does change equilibrium marriage patterns. Suppose for instance that $n_C = m_C$ and $n_N = m_N$: there are as many men as women in either education group. Then all equations are symmetric in gender, and we must have $\mu_{C0} = \mu_{0C}$ and $\mu_{N0} = \mu_{0N}$, both in the original and in the revised model. The equations for men in the original model simplify to:

$$n_C = \mu_{C0}(1+c) + b\sqrt{\mu_{N0}\mu_{C0}}$$
$$n_N = \mu_{N0}(1+a) + b\sqrt{\mu_{N0}\mu_{C0}}.$$

Suppose moreover that the college-educated are half of the population in each gender: $n_N = n_C \equiv n$ and $m_N = m_C \equiv m$, so that by subtraction $\mu_{C0}(1+c) = \mu_{N0}(1+a)$. Then we obtain in the original model

$$\mu_{C0} = \mu_{0C} = \sqrt{\frac{1+a}{1+c}}\frac{n}{b+\sqrt{(1+a)(1+c)}}$$
$$\mu_{N0} = \mu_{0N} = \sqrt{\frac{1+c}{1+a}}\frac{n}{b+\sqrt{(1+a)(1+c)}}$$
$$\mu_{NN} = a\mu_{N0}$$
$$\mu_{NC} = \mu_{CN} = \frac{bn}{b+\sqrt{(1+a)(1+c)}}$$
$$\mu_{CC} = c\mu_{C0}.$$

In the revised model, $b \to b\sqrt{2}$ and $c \to 2c$. Remember that a, b and c are exponentials and therefore positive. Since μ_{C0} is a decreasing function of both b and c, it must be lower in the revised model. μ_{NC} is an increasing function of $b/\sqrt{1+c}$; but

$$b\sqrt{2}/\sqrt{1+2c} > b/\sqrt{1+c}$$

and μ_{NC} must be higher. Since $\mu_{N0} + \mu_{NN} + \mu_{NC}$ is fixed at m and $\mu_{NN}/\mu_{N0} = a$ is unchanged, it follows that both μ_{NN} and μ_{N0} must be lower. Finally,

$$\mu_{CC} = n - \mu_{C0} - \mu_{CN} = \frac{n}{b + \sqrt{(1+a)(1+c)}} \left(\sqrt{1+c} - \frac{1}{\sqrt{1+c}} \right) = \frac{n}{b + \sqrt{(1+a)(1+c)}} \frac{c}{\sqrt{1+c}}.$$

We have just seen that $\frac{c}{\sqrt{1+c}}$ increases by a factor of more than $\sqrt{2}$. The denominator $b + \sqrt{(1+a)(1+c)}$ increases by less, since $\sqrt{1+c}$ increases by a factor smaller than $\sqrt{2}$ and nb by just $\sqrt{2}$. Therefore μ_{CC} must increase.

To recapitulate:

- There are fewer college graduate singles. This is the equivalent of more people taking the bus in Debreu (1960): mere payoff-irrelevant splits increase probabilities of choice.
- More surprisingly, there are also fewer non-college singles; but the fall in singles is smaller than for college graduates.
- There are more matches between N and C, fewer between N and N, and more between C and C.
- Since the expected utility is simply minus the logarithm of the probability of singlehood in the CS model, expected utilities increase at each level of education.

These are clearly unappealing properties: since the Commencement date is irrelevant to all market participants, a “more reasonable” model would imply none of these changes. This example shows that the Choo and Siow model is not robust to irrelevant labeling:

splitting college graduation according to the payoff-irrelevant parity of the Commencement day changes matching patterns μ and expected utilities for college graduates. In econometric terms, this results in a misspecified model and inconsistent estimates.

This begs a question: are there any separable models where taking onboard irrelevant type labels does not cause misspecification? Define *Generalized Random Coefficients models* (GRC) to be those separable models for which there exist real-valued random variables ε_i and η_j and functions ζ_{xy} and ξ_{xy} with

$$\forall x \in \mathcal{X}, \forall y \in \mathcal{Y}_0,\ \varepsilon_y^i = \zeta_{xy}(\varepsilon_i) \text{ and } \forall y \in \mathcal{Y}, \forall x \in \mathcal{X}_0, \eta_x^j = \xi_{xy}(\eta_j),$$

In these models, the random utilities associated to any pair of alternatives are perfectly dependent. GRC models can be viewed as an extension of the random coefficient models popular in empirical IO, with the important restriction that separability rules out an "idiosyncratic" term of the form ν_{ij}. The GRC models extend the Random Scalar Coefficients models of Galichon and Salanié (2019), in which the functions ζ_{xy} and ξ_{xy} are linear and the random variables ε_i and η_j are scalar.

It is easily seen that Generalized Random Coefficients models do *not* satisfy IIA. On the other hand, they are robust to the "excessive type-splitting" in our Commencement day example, as long as the econometrician allows for a flexible specification of the Φ, ξ and ζ components that makes it possible for the "commencement day" type to have a zero coefficient.

Concluding Remarks

We have conjectured that no separable model can satisfy strong IIA. In contrast, Dagsvik (2000) proposed a model which is *not* separable. Its "marriage matching function" is

$$\mu_{xy} = \mu_{x0}\mu_{0y} \exp(\Phi_{xy}),$$

where μ_{x0} and μ_{0y} are adjusted by the marginal constraints

$$
\begin{aligned}
n_x &= \mu_{x0} + \sum_{y \in \mathcal{Y}} \mu_{x0}\mu_{0y} \exp(\Phi_{xy}) \\
m_y &= \mu_{0y} + \sum_{x \in \mathcal{X}} \mu_{x0}\mu_{0y} \exp(\Phi_{xy}).
\end{aligned}
$$

This model has very different properties from separable models. To cite just one major difference: as all subpopulations scale up homothetically, marriage rates increase. Thus while all separable models have constant returns to scale, Dagsvik's exhibits increasing returns to scale.

It is easy to verify that strong IIA holds in Dagsvik's model. We suspect that it may be the only tractable empirical matching model that satisfies strong IIA.

References

Arrow, K. (1959): "Rational Choice Functions and Orderings," *Economica*, 26, 121–127.

Chiappori, P.-A., B. Salanié, and Y. Weiss (2017): "Partner Choice, Investment in Children, and the Marital College Premium," *American Economic Review*, 107, 2109–67.

Choo, E. (2015): "Dynamic Marriage Matching: An Empirical Framework," *Econometrica*, 83, 1373–1423.

Choo, E., and A. Siow (2006): "Who Marries Whom and Why," *Journal of Political Economy*, 114, 175–201.

Dagsvik, J. (2000): "Aggregation in Matching Markets," *International Economic Review*, 41, 27–58.

Debreu, G. (1960): "Review of R. D. Luce, *Individual choice behavior: A theoretical analysis*," *American Economic Review*, 50, 186–188.

Decker, C., E. Lieb, R. McCann, and B. Stephens (2012): "Unique Equilibria and

Substitution Effects in a Stochastic Model of the Marriage Market," *Journal of Economic Theory*, 148, 778–792.

Galichon, A., and B. Salanié (2017): "The Econometrics and Some Properties of Separable Matching Models," *American Economic Review Papers and Proceedings*, 107, 251–255.

Galichon, A., and B. Salanié (2019): "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models," Columbia University mimeo.

Graham, B. (2013): "Uniqueness, Comparative Static, And Computational Methods for an Empirical One-to-one Transferable Utility Matching Model," *Structural Econometric Models*, 31, 153–181.

Luce, R. D. (1959): *Games and Decisions*. New York: Wiley.

Luce, R. D., and H. Raiffa (1957): *Individual Choice Behavior*. New York: Wiley.

McFadden, D. (1974): "Conditional Logit Analysis of Qualitative Choice Behaviour," in *Frontiers of Econometrics*, ed. by P. Zarembka. Academic Press: New York.

Mourifié, I. (2019): "A Marriage Matching Function with Flexible Spillover and Substitution Patterns," *Economic Theory*, 67, 421–461.

Mourifié, I., and A. Siow (2017): "The Cobb Douglas Marriage Matching function: Marriage Matching with Peer and Scale Effects," University of Toronto mimeo.

Nash, J. (1950): "The Bargaining Problem," *Econometrica*, 18, 155–162.

Appendix

A Proof of Proposition 2

The results in Choo and Siow show that given numbers (n_C, n_N, m_C, m_N) of men and women, the equilibrium in the original model has

$$n_C = \mu_{C0} + \mu_{CN} + \mu_{CC} \tag{A.1}$$
$$n_N = \mu_{N0} + \mu_{NN} + \mu_{NC}$$
$$m_C = \mu_{0C} + \mu_{NC} + \mu_{CC}$$
$$m_N = \mu_{0N} + \mu_{NN} + \mu_{CN}$$
$$\mu_{CN} = b\sqrt{\mu_{C0}\mu_{0N}} \tag{A.2}$$
$$\mu_{NC} = b\sqrt{\mu_{N0}\mu_{0C}}$$
$$\mu_{CC} = c\sqrt{\mu_{C0}\mu_{0C}} \tag{A.3}$$
$$\mu_{NN} = a\sqrt{\mu_{N0}\mu_{0N}}.$$

Take college-educated men for instance. Equation (A.1) requires that the number of college-educated men who remain single and who marry any type of woman must add up to the number of college-educated men available. Substituting equations (A.2) and (A.3) gives

$$\mu_{C0} + \left(b\sqrt{\mu_{0N}} + c\sqrt{\mu_{0C}}\right)\sqrt{\mu_{C0}} = n_C.$$

As noted by Graham (2013) and Decker, Lieb, McCann, and Stephens (2012), this is a quadratic equation in $\sqrt{\mu_{C0}}$, given the numbers of single women μ_{0N} and μ_{0C}. They showed that the whole system can be rewritten as the four coupled quadratic equations in the square roots of the numbers of singles:

$$\begin{aligned} n_C &= \mu_{C0} + \left(b\sqrt{\mu_{0N}} + c\sqrt{\mu_{0C}}\right)\sqrt{\mu_{C0}} \\ n_N &= \mu_{N0} + \left(a\sqrt{\mu_{0N}} + b\sqrt{\mu_{0C}}\right)\sqrt{\mu_{N0}} \\ m_C &= \mu_{0C} + \left(b\sqrt{\mu_{N0}} + c\sqrt{\mu_{C0}}\right)\sqrt{\mu_{0C}} \\ m_N &= \mu_{0N} + \left(a\sqrt{\mu_{N0}} + b\sqrt{\mu_{C0}}\right)\sqrt{\mu_{0N}}. \end{aligned} \tag{A.4}$$

Now let us introduce the payoff-irrelevant Commencement distinction. Under our assumptions, the quadratic equation that defines the equilibrium for college-educated men whose Commencement was on an even day take the form

$$\mu_{C_e0} + \sqrt{\mu_{C_e0}}\left(c\left(\sqrt{\mu_{0C_e}} + \sqrt{\mu_{0C_o}}\right) + b\sqrt{\mu_{0N}}\right) = n_{C_e} = n_C/2.$$

and that for college-educated men whose Commencement was on an odd day is

$$\mu_{C_o0} + \sqrt{\mu_{C_o0}}\left(c\left(\sqrt{\mu_{0C_e}} + \sqrt{\mu_{0C_o}}\right) + b\sqrt{\mu_{0N}}\right) = n_{C_o} = n_C/2.$$

These two equations are identical; and since they have only one feasible root, they imply $\mu_{C_e0} = \mu_{C_o0}$. Similarly, $\mu_{0C_e} = \mu_{0C_o}$. For men as for women, there are just as many single college graduates in both Commencement groups.

This in turn implies that $\mu_{C_iC_j} = c\sqrt{\mu_{C_i0}\mu_{0C_j}}$ cannot depend on $i, j = e, o$; and that $\mu_{C_iN} = b\sqrt{\mu_{C_i0}\mu_{0N}}$ and μ_{NC_i} cannot depend on $i = e, o$.

Putting things together, and using the notation defined in the statement of the Proposition, we obtain

$$\begin{aligned} \mu_{CC} &= 4c\sqrt{\frac{\mu_{C0}}{2}\frac{\mu_{0C}}{2}} = 2c\sqrt{\mu_{C0}\mu_{0C}} \\ \mu_{CN} &= 2b\sqrt{\frac{\mu_{C0}}{2}\mu_{0N}} = b\sqrt{2}\sqrt{\mu_{C0}\mu_{0N}} \\ \mu_{NC} &= 2b\sqrt{\mu_{N0}\frac{\mu_{0C}}{2}} = b\sqrt{2}\sqrt{\mu_{N0}\mu_{0C}} \\ \mu_{NN} &= a\sqrt{\mu_{N0}\mu_{0N}}. \end{aligned}$$

This gives the equilibrium equations

$$\begin{aligned} n_C &= \mu_{C0} + b\sqrt{2}\sqrt{\mu_{C0}\mu_{0N}} + 2c\sqrt{\mu_{C0}\mu_{0C}} \\ n_N &= \mu_{N0} + a\sqrt{\mu_{N0}\mu_{0N}} + b\sqrt{2}\sqrt{\mu_{N0}\mu_{0C}} \end{aligned} \tag{A.5}$$

and two symmetric equations for women. But the system (A.4) had

$$\begin{aligned} n_C &= \mu_{C0} + b\sqrt{\mu_{C0}\mu_{0N}} + c\sqrt{\mu_{C0}\mu_{0C}} \\ n_N &= \mu_{N0} + a\sqrt{\mu_{N0}\mu_{0N}} + b\sqrt{\mu_{N0}\mu_{0C}}. \end{aligned} \tag{A.6}$$

Comparing (A.5) and (A.6) shows that adding the irrelevant Commencement distinction changes the equilibrium matching patterns as if we had changed the joint surplus $\mathbf{\Phi}$ to $\mathbf{\Phi}'$.