Matching Function Equilibria with Partial Assignment: Existence, Uniqueness and Estimation

Liang Chen, Eugene Choo, Alfred Galichon, Simon Weber

Matching Function Equilibria with Partial Assignment: Existence, Uniqueness and Estimation

Liang Chen* Eugene Choo† Alfred Galichon‡ Simon Weber§

February 4, 2021

Abstract

In this paper, we argue that models coming from a variety of fields share a common structure that we call matching function equilibria with partial assignment. This structure revolves around an aggregate matching function and a system of nonlinear equations. This encompasses search and matching models, matching models with transferable, non-transferable and imperfectly transferable utility, and matching with peer effects. We provide a proof of existence and uniqueness of an equilibrium as well as an efficient algorithm to compute it. We show how to estimate parametric versions of these models by maximum likelihood. We also propose an approach to construct counterfactuals without estimating the matching functions for a subclass of models. We illustrate our estimation approach by analyzing the impact of the elimination of the Social Security Student Benefit Program in 1982 on the marriage market in the United States.

Keywords: matching function, maximum likelihood estimation, counterfactuals, matching, transferable utility, imperfectly transferable utility

JEL codes: C1, C78, I2, J12.

*Economics and Management School, Wuhan University. Email: liang.chen@whu.edu.cn
†Division of Social Science, Yale-NUS Colege. Email: eugene.choo@yale-nus.edu.sg
‡New York University and Courant Institute. Email: ag133@nyu.edu or galichon@cims.nyu.edu
§University of York. Email: simon.weber@york.ac.uk

1 Introduction

Social scientists and demographers have long been fascinated with the workings of the marriage market. In particular, how does the distribution of marriages between heterogenous agents get affected by demographic changes in the number of available men and women, or some exogenous changes that affects the desirability of certain matches relative to remaining single. One of the key challenge confronting researchers is to understand the role of preferences and how that differs from the pressures of relative scarcity of heterogenous agents in the marriage market. Most of the empirical settings involve aggregate data of final matches of who marries whom.[1] Confronted with aggregate data, demographers and economists have heavily relied on *aggregate matching functions*. Such functions allow the researcher to relate the number of matches between two partners of given characteristics (such as age or education) to their respective supply of single individuals in the population.[2] One drawback of this earlier approach is the absence of behavioral foundation for these matching functions. Consequently, it is difficult to interpret the model parameters or think about estimation.

Choo and Siow (2006b) proposed an approach to estimate the aggregate matching surplus using an equilibrium transferable utility model of marriage where agents have unobserved and heterogenous taste for partners of known types. Using only aggregate data, their approach rationalizes a matching function where preference parameters are primitives of the behavioral model. Unlike some matching markets where preference data may be directly collected, it is not immediately clear how data on final matches could be used to identify preferences. The approach borrows many ideas from the structural IO literature where estimating model primitives from equilibrium models using limited aggregate data is a norm. These identified primitives are important inputs when considering counterfactual experiments such as how the distribution of marriage responds to exogenous demographic shocks or government intervention. Chiappori, Salanié, and Weiss (2017) and Galichon and Salanié (2020), among others build on this and relax some of the assumptions in the Choo and Siow (2006b) model.

This paper unifies the themes in these papers and makes four contributions.

[1]While there are some examples where preference ordering are directly solicited from marriage market participants, these cases are rare. See Hitsch, Hortaçsu, and Ariely (2010) for example.

[2]A standard matching function is the harmonic marriage matching function used by demographers (see e.g. Qian and Preston (1993) and Schoen (1981)).

Our first contribution is to define a class of models, that we call *matching function equilibria*, that share a common structure with, in its center, a behaviorally coherent aggregate matching function and a system of nonlinear equations that balances available individuals for each type. We argue that our approach is not anecdotal. It is striking to see that many settings in the matching literature can be characterized by an aggregate matching function and a system of nonlinear equations.[3] In this paper, we focus on models with partial assignment, i.e. in which unassigned agents are allowed (models with full assignment are discussed in the companion paper Chen et al. (2020)). We recall the proof of existence and uniqueness of a solution of Galichon, Kominers, and Weber (2019), without claiming novelty here. These existence and uniqueness results are important for policy analysis as they ensure a unique counterfactual equilibrium when considering the effects of policy changes (see our third and fourth contributions below).

Our second contribution is to show how to estimate the matching functions parameters by maximum likelihood using two computational approaches: (i) the nested approach; and (ii) the MPEC (mathematical program with equilibrium constraints) approach. We provide analytical expressions of the gradient of the log-likelihood for both approaches for efficient estimation. Furthermore, we show that for models featuring a constant return to scale matching function, one can obtain close-form formulas for the confidence intervals of estimates.

Our third contribution is to show that we can conduct counterfactual experiments without estimating the model parameters. We call this the Parameter-Free approach. All the papers in the literature have taken the standard approach where we first estimate the structural parameters followed by computation of the counterfactual equilibrium using the estimated parameters. We show that for a subclass of matching function equilibrium models whose matching functions satisfy two properties, (i) homogeneity and (ii) multiplicative separability in parameters, we can compute the counterfactual equilibrium without knowledge of the estimated parameters. These two properties allow us to reformulate these equilibrium

[3]This includes the search and matching model of Shimer and Smith (2000), Non Transferable Utility models of Dagsvik (2000) and Menzel (2015), Transferable Utility models of Choo and Siow (2006b) and Siow (2008), Imperfectly Transferable Utility models of Galichon, Kominers, and Weber (2019), demand models of Berry, Levinsohn, and Pakes (1995), bilateral trade models of Head and Mayer (2014), non additive random utility models, and models of the labor market with full assignment, among others.

matching models into a system of nonlinear equations, with the ratio of counterfactual to observed equilibria as unknowns. The new system of non-linear equations is free from structural parameters. We show that this new system of equations has a unique solution in terms of the ratio of the counterfactual to observed equilibria which allow us to compute the counterfactual equilibrium accordingly. While this Parameter-Free approach has limitations in terms of scope of models it applies to, it has the computational advantage of not requiring the estimation of the structural parameters in the matching functions.

Our final contribution is to apply our proposed estimation approaches to analyze the impact of the elimination of the Social Security Student Benefit Program in 1982 on the marriage market using the Transferable Utility model in Choo and Siow (2006b) (CS hereafter). In the United States, the Social Security Student Benefit Program established in 1965 provided financial aid for children of deceased, disabled or retired workers to attend college.[4] The elimination of this benefit program in 1982 is one of the largest policy changes in college students' financial aid in the United States. The number of college beneficiaries is estimated to have dropped from about 600,000 in 1981 to 66,000 in 1986. The average total monthly payment fell substantially from about \$196 million in 1981 to \$26 million in 1986. Dynarski (2003) show that this policy change had a significant causal effect on reducing students' college attendance and completion among eligible students.

Our goal in this empirical application is to estimate the impact of the policy change on the marriage distribution by computing the counterfactual marriage distribution in 1987/88 had the Social Security Student Benefit Program not been eliminated by using our parametric and the Parameter-Free approaches. Using causal effect estimates from the education literature, and our estimates of aid eligibility computed from the 1986 Current Population Survey, we construct a counterfactual distribution of available single men and women by age and education. We show that in this counterfactual scenario, there will be around 17,000 (3%) fewer marriages among high school graduates and 10,000 (3%) more marriages among college graduates in 1987/88. Interestingly, marriages between college educated men and high school educated women would also increase by around 4000 (2.7%) while the change for

[4]Recipients need to be not married and enrolled in college fulltime. The financial aid is provided up until the semester the recipient turns 22 years of age.

marriages between college educated women and high school educated men is negligible. This is probably due to the social norms of men preferring spouses who are not more educated than themselves, which is embedded in the preference parameters.

Our paper relates to the literature on matching functions used by demographers, e.g. Qian and Preston (1993) and Schoen (1981). However, as we argue above, these matching functions lack behavioral micro-foundation. We look for such foundations in the two-sided matching literature, starting from the canonical models of matching with transferable utility, e.g. Koopmans and Beckmann (1957), Shapley and Shubik (1971), and Becker (1973), and non transferable utility, e.g. Gale and Shapley (1962). In particular, we rely on recent advances on the structural estimation of these models, which exploits heterogeneity in tastes for identification purposes, e.g Dagsvik (2000), Menzel (2015), and Choo and Siow (2006b). We argue that these models share a common structure, that revolves around an aggregate matching function and a system of nonlinear equations. These models, and their variants, provides us with a behaviorally coherent matching function, as pointed out in Choo and Siow (2006b), Choo and Siow (2006a), Chiappori, Salanié, and Weiss (2017) and Mourifié (2018).

However, we also go beyond the TU and NTU cases. We argue in Section 2.2 that this structure can be found in a variety of models. Therefore, our paper also relates to the literature on matching with imperfectly transferable utility, e.g. Galichon, Kominers, and Weber (2019), and search and matching model, e.g. Shimer and Smith (2000) and Goussé, Jacquemet, and Robin (2017). To the best of our knowledge, ours is the first paper to point out that the matching function structure is found in a surprisingly large number of models.

We provide the complete econometric toolbox to take these models to the data. In particular, we show how to estimate parametric versions of these models using a nested or MPEC approach. Thus our paper relates to the MPEC literature, e.g. Dubé, Fox, and Su (2012), Su and Judd (2012) and Pang, Su, and Lee (2015). We show how conduct counterfactual experiments by using both the parametric and Parameter-Free approaches, and provide a number of computational techniques to ensure tractability.

Organization of the paper. Section 2 introduces and characterizes matching function equilibrium models with partial assignment. Section 3 outlines the estimation by maximum likelihood and provides an analytic expression for the gradient of the log-likelihood. We also provide the formulas for computing the confidence intervals of the parameter estimates in models with homogeneous matching functions. Section 4 outlines the Parameter-Free approach for estimating the counterfactual equilibrium for a subclass of matching function equilibrium models. Section 5 illustrates our approaches by investigating the impact of the elimination of the Social Security Student Benefit Program in 1982 on the marriage market in the United States. Section 6 concludes. All proofs of our main results can be found in the Appendix. Additional results, numerical experiments, details about the data construction and additional figures are available in the Online Appendix.

2 Matching Function Equilibrium Models with partial assignment

In this section, we show that many behavioural models in the matching literatures can be characterized by an aggregate matching function and a system of nonlinear equations. We focus on situations in which we allow for unassigned agents, and provide a formal definition for this class of matching models, that we label *matching function equilibrium* models with partial assignment. The full assignment case (in which unassigned agents are not allowed) is dealt with in Chen et al. (2020). We then recall the general proof of existence and uniqueness of the equilibrium in these models.

Consider a general marriage market where two populations of men (indexed by $i \in I$) and women (indexed by $j \in J$) meet and may form heterosexual pairs.[5] We assume that men (resp. women) can be gathered in groups of similar characteristics, or types, $x \in \mathcal{X}$ (resp. $y \in \mathcal{Y}$), with $|\mathcal{X}|$ (resp. $|\mathcal{Y}|$) denoting the number of types for men (resp. women). The total mass of men of type x (resp. women of type y) is denoted n_x (resp. m_y). We define the sets, $\mathcal{X}_0 = \mathcal{X} \cup \{0\}$, and $\mathcal{Y}_0 = \mathcal{Y} \cup \{0\}$, which include singlehood as an option. The set of married pairs is given by $\mathcal{XY} = \mathcal{X} \times \mathcal{Y}$, and the set of household types, $\mathcal{XY}_0 = \mathcal{X} \times \mathcal{Y} \cup \mathcal{X} \times \{0\} \cup \{0\} \times \mathcal{Y}$. The mass of marriages between men of type x and women of type

[5] The setup and notations can be adapted to any two-sided one-to-one matching markets, e.g., replacing men by workers and women by firms.

y is denoted by μ_{xy}, while the mass of single men of type x and the mass of single women of type y are denoted by μ_{x0} and μ_{0y}, respectively. We can now define matching function equilibrium models with partial assignment below.

2.1 Definition

In the partial assignment case, one shall assume that μ_{xy} is a deterministic function of the masses of unassigned type x and y agents, denoted μ_{x0} and μ_{0y}, respectively.

Definition 1. Matching function equilibrium with partial assignment
A matching function equilibrium model with partial assignment determines the mass of xy pairs, μ_{xy}, and the masses of unassigned agents, μ_{x0} and μ_{0y}, by an aggregate matching function (or generalized gravity equation) which relates the former to the latter by

$$\mu_{xy} = M_{xy}(\mu_{x0}, \mu_{0y}) \tag{1}$$

where μ_{x0} and μ_{0y} are determined by a system of nonlinear accounting equations,

$$\begin{cases} n_x = \mu_{x0} + \sum_{y \in \mathcal{Y}} M_{xy}(\mu_{x0}, \mu_{0y}), \ \forall x \in \mathcal{X} \\ m_y = \mu_{0y} + \sum_{x \in \mathcal{X}} M_{xy}(\mu_{x0}, \mu_{0y}), \ \forall y \in \mathcal{Y} \end{cases}. \tag{2}$$

Note that the matching equilibrium in these models is fully characterized by the system of nonlinear equations (2) with unknowns (μ_{x0}, μ_{0y}).

We now provide some examples to show that the defined matching function equilibrium model encompasses many behavioural models in the matching and related literatures.

2.2 Examples

a) A model of search and matching.

Consider the search and matching model of Shimer and Smith (2000). Adapting notation and setting to allow for discrete types, Equation (1) (from page 347) of the paper reads $\sum_{y \in \mathcal{Y}} \mu_{xy} = \rho/\delta \sum_{y \in \mathcal{Y}} \mu_{x0}\mu_{0y}A(x, y)$, where A is an indicator function which equals one if man x who meets woman y agree to match, and ρ and δ denote the meeting and separation rates, respectively. Thus, the aggregate matching function is

$$M(\mu_{x0}, \mu_{0y}) = \frac{\rho}{\delta}\mu_{x0}\mu_{0y}A(x, y) \tag{3}$$

which must satisfy the accounting equations in (2). This approach has been successfully applied to the marriage market by Goussé, Jacquemet, and Robin (2017), with a similar aggregate matching function.

b) Models of matching with Non Transferable Utility (NTU).

Many models of two-sided matching describe markets in which agents have heterogenous preferences over their potential partners and face a choice problem with random utility shocks. In a Non-Transferable Utility (NTU) framework, it is assumed that when a man of type x and woman of type y decide to match, they receive payoffs α_{xy} and γ_{xy}, respectively, as well as an idiosyncratic component which is often assumed to enter additively. In this NTU setting, Dagsvik (2000) and Menzel (2015) provide a tractable expression for the aggregate matching function when the number of market participants grows large. Equation (4.3) (on page 925) of Menzel (2015) reads,

$$M(\mu_{x0}, \mu_{0y}) = \mu_{x0}\mu_{0y}\exp(\alpha_{xy} + \gamma_{xy}). \tag{4}$$

Galichon and Hsieh (2017) study a similar NTU model but introduce the novel solution concept of equilibrium under rationing-by-waiting and study its properties.

c) Models with Transferable Utility (TU).

The previous expression is remarkably similar to the well-known aggregate matching function obtained in the seminal contribution of Choo and Siow (2006b) in the Transferable Utility setting. In this case, α_{xy} and γ_{xy} represent pre-transfer utilities, where the equilibrium payoff of man i of type x, married with a woman of type y, is $u_i = \alpha_{xy} - \tau_{xy} + \epsilon_{iy}$. The equilibrium payoff is the sum of a systematic component ($U_{xy} \equiv \alpha_{xy} - \tau_{xy}$, where τ_{xy} is the transfer from the type x man to the type y woman) and an idiosyncratic part (ϵ). Similarly, the payoff of woman j of type y, married to a man of type x, is $v_j = \gamma_{xy} + \tau_{xy} + \eta_{iy}$ where the systematic part is denoted V_{xy}. Choo and Siow (2006b) adopts an extreme value random utility model of McFadden (1974), and show that the systematic components of equilibrium payoffs can be recovered from the observed matching probabilities, that is $U_{xy} = \log \frac{\mu_{xy}}{\mu_{x0}}$ and $V_{xy} = \log \frac{\mu_{xy}}{\mu_{0y}}$. Let $\Phi_{xy} = \alpha_{xy} + \gamma_{xy}$ denote the total surplus from a (x, y) match. It follows from Equation (11) (on page 181) of that paper,

$$M(\mu_{x0}, \mu_{0y}) = \sqrt{\mu_{x0}\mu_{0y}} \exp(\frac{\Phi_{xy}}{2}) \tag{5}$$

This matching function has several interesting properties: it is homogenenous of degree 1, satisfying constant returns to scale, and the effect of μ_{x0} and μ_{0y} is symmetric.

d) Models with Imperfectly Transferable Utility (ITU).

Building on this literature, Galichon, Kominers, and Weber (2019) (GKW hereafter) provide a general framework for matching with Imperfectly Transferable Utility, that encompass both the classical fully- and non-transferable utility settings[6]. In particular, they show that in equilibrium, the systematic utilities, U_{xy} and V_{xy}, satisfy a key feasibility condition, $D_{xy}(U_{xy}, V_{xy}) = 0$. Among the properties of the distance function D, it is worth noting that $D(u + a, v + a) = a + D(u, v)$. Hence, when taste shocks are assumed to be logit (as in Choo and Siow (2006b)), the feasibility condition rewrites as, $D_{xy}(\log \mu_{xy} - \log \mu_{x0}, \log \mu_{xy} - \log \mu_{0y}) = 0$, which combined with the property on D, gives

$$M(\mu_{x0}, \mu_{0y}) = \exp\left(-D_{xy}(-\log \mu_{x0}, -\log \mu_{0y})\right). \tag{6}$$

This matching function also satisfies homogeneity of degree 1 but not necessarily symmetry in μ_{x0} and μ_{0y}.

e) Harmonic mean matching function.

Schoen (1981) analyzes the marriage market using a matching function which is based on the harmonic mean. Interestingly, the ITU-logit framework introduced above allows us

[6]For a study of ITU settings using revealed preferences techniques instead of matching functions, see Cherchye et al. (2017).

to recover a micro-founded version of that matching function. Assume that whenever a man of type x is matched with a woman of type y, they bargain to split their income (normalized to 2) into private consumption for the man and the woman denoted by (c^a_{xy}) and (c^b_{xy}), respectively. Assume that the utility payoffs received by the man and the woman are $u = \alpha_{xy} + \tau_{xy} \log c^a_{xy}$ and $v = \gamma_{xy} + \tau_{xy} \log c^b_{xy}$, respectively. Assuming the budget constraint is $c^a_{xy} + c^b_{xy} \leq 2$, one can verify that the distance function is given by $D_{xy}(u, v) = \tau_{xy} \log \left((\exp\left((u - \alpha_{xy})/\tau_{xy}\right) + \exp\left((v - \gamma_{xy})/\tau_{xy}\right))/2 \right)$, so by equation (6) gives us the matching function

$$M(\mu_{x0}, \mu_{0y}) = \left[\frac{\exp\left(\frac{-\alpha_{xy}}{\tau_{xy}}\right)}{2} \times \mu_{x0}^{-1/\tau_{xy}} + \frac{\exp\left(\frac{-\alpha_{xy}}{\tau_{xy}}\right)}{2} \times \mu_{0y}^{-1/\tau_{xy}} \right]^{-\tau_{xy}} \tag{7}$$

Whenever $\tau_{xy} = 1$, we recover the harmonic mean matching function, as in Equation 1 (on page 281) in Qian (1998) (up to some multiplicative constants).

f) Models with peer effects.

Mourifié and Siow (2017) assumes that the utility of man i of type x, matched with a woman of type y, is given by $\alpha_{xy} - \tau_{xy} + \psi_x \ln \mu_{xy} + \epsilon_{iy}$, where α_{xy} is the pre-transfer utility from the match, τ_{xy} is the transfer made by the man to its partner, and ϵ_{iy} an idiosyncratic component drawn from an Extreme Value Type-I distribution. Similarly, the payoff for woman j of type y, matched with a man of type x is, $\gamma_{xy} + \tau_{xy} + \Psi_y \ln \mu_{xy} + \eta_{xj}$. This is exactly the model of Choo and Siow (2006b), except for the introduction of the peer effects terms, $\psi_x \ln \mu_{xy}$ and $\Psi_y \ln \mu_{xy}$. Further computation yields

$$M(\mu_{x0}, \mu_{0y}) = \mu_{x0}^{a_{xy}} \mu_{0y}^{b_{xy}} \exp\left(\frac{\Phi_{xy}}{2 - \psi_x - \Psi_y} \right), \tag{8}$$

the Cobb Douglas aggregate matching function. If we impose the restrictions, $a_{xy} = a$, $b_{xy} = b$, and $a + b = 1$, we recover the matching function of Chiappori, Salanié, and Weiss (2017). Note that in its general formulation, this matching function does not satisfy homogeneity of degree 1 nor symmetry.

2.3 Existence and Uniqueness

Under mild conditions on the matching function, there exists a unique equilibrium matching with partial assignment. These conditions are as follows:

Assumption 1. *The aggregate matching function,* $M_{xy}\left(\mu_{x0},\mu_{0y}\right)$, *satisfies the following three conditions*

(i) $M_{xy}:\left(\mu_{x0},\mu_{0y}\right)\mapsto M_{xy}\left(\mu_{x0},\mu_{0y}\right)$ *is* continuous.

(ii) $M_{xy}:\left(\mu_{x0},\mu_{0y}\right)\mapsto M_{xy}\left(\mu_{x0},\mu_{0y}\right)$ *is* weakly isotone.

(iii) For each $\mu_{x0}>0$, $\lim_{\mu_{0y}\to 0^+} M_{xy}\left(\mu_{x0},\mu_{0y}\right)=0$. *Similarly, for each* $\mu_{0y}>0$, $\lim_{\mu_{x0}\to 0^+} M_{xy}\left(\mu_{x0},\mu_{0y}\right)=0$.

In section 2.2, we provided several examples of aggregate matching functions borrowed from the matching literature. It is easy to show that conditions (i)-(iii) are met in all examples. It turns out that these very mild conditions on M_{xy} are sufficient to prove the existence and uniqueness result in Theorem 1 below. This extends a result in GKW beyond the case when the matching function is homogenous of degree 1. We complement this result with an algorithm which provides an efficient way of solving these equations.

Theorem 1 (Galichon, Kominers, and Weber (2019))**.** *Under Definition 1 and Assumption 1, there exists a unique equilibrium matching with partial assignment, given by* $\mu^*_{xy}=M_{xy}(\mu^*_{x0},\mu^*_{0y})$, *for any* $x\in\mathcal{X}$ *and* $y\in\mathcal{Y}$, *where the pair of vectors* $(\mu^*_{x0})_{x\in\mathcal{X}}$ *and* $(\mu^*_{0y})_{y\in\mathcal{Y}}$ *is the unique solution to the system*

$$\begin{cases} n_x=\mu_{x0}+\sum_{y\in\mathcal{Y}} M_{xy}(\mu_{x0},\mu_{0y}) \\ m_y=\mu_{0y}+\sum_{x\in\mathcal{X}} M_{xy}(\mu_{x0},\mu_{0y}) \end{cases}.$$

Proof. See Appendix A. ∎

The proof of existence is constructive. It relies on an iterative procedure that we call *Iterative Projective Fitting procedure* (IPFP), which has been used in various fields under different names (see Galichon, Kominers, and Weber (2015))[7]. GKW introduce this algorithm in their ITU-logit setting. The proof of uniqueness relies on Berry, Gandhi, and Haile

[7] For instance it is also known as the RAS algorithm (Kruithof (1937); see also Idel (2016)).

(2013). Because this algorithm turns out to be very useful in practice, we provide a formal description of it below.

Algorithm 1. *The Generalized Iterative Projection Fitting Procedure (Galichon, Kominers, and Weber (2015)) works as follow*

Step 0	*Fix the initial value of* μ_{0y}*, at* $\mu_{0y}^0 = m_y$*.*
Step $2t+1$	*Keep the values* μ_{0y}^{2t} *fixed. For each* $x \in \mathcal{X}$*, solve for the value,* μ_{x0}^{2t+1} *of* μ_{x0}*, such that the equality,* $\sum_{y \in \mathcal{Y}} M_{xy}(\mu_{x0}, \mu_{0y}^{2t}) + \mu_{x0} = n_x$ *holds.*
Step $2t+2$	*Keep the values* μ_{x0}^{2t+1} *fixed. For each* $y \in \mathcal{Y}$*, solve for the value,* μ_{0y}^{2t+2} *of* μ_{0y}*, such that equality,* $\sum_{x \in \mathcal{X}} M_{xy}(\mu_{x0}^{2t+1}, \mu_{0y}) + \mu_{0y} = m_y$ *holds.*

The algorithm terminates when, $\sup_y |\mu_{0y}^{2t+2} - \mu_{0y}^{2t}| < \epsilon$*, where* ϵ *is a sufficiently small positive value.*

3 Maximum Likelihood Estimation

In this section, we show how to conduct parametric inference on the matching function equilibrium models with partial assignment, and how the structural parameters can be estimated using maximum likelihood. To reduce computational time, we provide an analytic expression for the gradient of the likelihood, as well as formulas to compute confidence intervals when the matching function is homogeneous of degree 1.

Let us assume that M_{xy} belongs to a parametric family M_{xy}^θ, where there is a unique value, θ_0 which rationalizes the observed data. Our goal is to estimate the parameters θ_0 in M_{xy}^θ. Note that for given masses n_x and m_y, denoted by (n, m), and a given parameter θ, we can obtain the (unique) equilibrium masses of singles (μ_{x0}, μ_{0y}) (using Algorithm 1). We can then compute the predicted mass of marriages between type x men, and type y women, μ_{xy}^θ, using our aggregate matching function M_{xy}^θ. These quantities are all we need to construct a likelihood.

3.1 The Likelihood

We first define the probability of forming a match pair $(x, y) \in \mathcal{XY}_0$-type of household. Using the quantities (μ_{x0}, μ_{0y}), we define the vector of model predicted matching frequencies, $\Pi_{xy}(\theta, n, m)$ as

$$\Pi_{xy}(\theta, n, m) = \frac{M^{\theta}_{xy}(\mu_{x0}, \mu_{0y})}{1' M^{\theta}_{xy}(\mu_{x0}, \mu_{0y})} \tag{9}$$

for all $xy \in \mathcal{XY}_0$, where $1' M^{\theta}_{xy}(\mu_{x0}, \mu_{0y})$ is the predicted total number of households. If we observe the matching, $\hat{\mu} = (\hat{\mu}_{xy}, \hat{\mu}_{x0}, \hat{\mu}_{0y})$ from the data, then the log-likelihood is given by,

$$l\left(\hat{\mu}|\theta, \mu_{x0}, \mu_{0y}, n, m\right) = \sum_{xy \in \mathcal{XY}_0} \hat{\mu}_{xy} \log \Pi_{xy}\left(\theta, n, m\right). \tag{10}$$

In practice, n and m are replaced by their efficient estimator $\hat{n}$, and $\hat{m}$ that can be computed from the observed matching, $(\hat{\mu}_{xy}, \hat{\mu}_{x0}, \hat{\mu}_{0y})$. Hence, the maximum likelihood estimator solves the following problem

$$\max_{\theta, \mu_{x0}, \mu_{0y}} l\left(\hat{\mu}, \hat{n}, \hat{m}|\theta, \mu_{x0}, \mu_{0y}\right), \tag{11}$$

$$\text{subject to} \qquad \begin{cases} \hat{n}_x = \mu_{x0} + \sum_{y \in \mathcal{Y}} M^{\theta}_{xy}(\mu_{x0}, \mu_{0y}) \\ \hat{m}_y = \mu_{0y} + \sum_{x \in \mathcal{X}} M^{\theta}_{xy}(\mu_{x0}, \mu_{0y}) \end{cases}.$$

We rewrite the constraints as $G(\theta, \mu_{x0}, \mu_{0y}) = 0$. We propose two computational approaches to solve this estimation problem.

3.2 Computation

3.2.1 The nested approach

The first approach is to get rid of the constraints, $G(\theta, \mu_{x0}, \mu_{0y}) = 0$, and maximize over μ_{x0} and μ_{0y} by solving for the equilibrium $(\mu^{\theta}_{x0}, \mu^{\theta}_{0y})$, for any value of θ. From Theorem 1 above, we know that such an equilibrium always exists and is unique. From these unique values of μ^{θ}_{x0} and μ^{θ}_{0y}, μ^{θ}_{xy} is deduced from M^{θ}_{xy} and the log-likelihood can be computed.

Estimation proceeds as follow: (i) fix a value of θ ; (ii) solve the system of equations (2) and obtain the unique μ^{θ}_{x0} and μ^{θ}_{0y} ; (iii) deduce μ^{θ}_{xy} from $M^{\theta}_{xy}(\mu^{\theta}_{x0}, \mu^{\theta}_{0y})$ and compute

$\Pi_{xy}(\theta, n, m)$ according to (9) ; (iv) compute the log-likelihood in Equation (10).

This approach has the advantage that by construction, μ^{θ}_{x0} and μ^{θ}_{0y} solves $G(\theta, \mu_{x0}, \mu_{0y}) = 0$ for any value of θ. Hence, we can apply the Implicit Function Theorem to compute the gradient of the unconstrained likelihood, $l\left(\hat{\mu}, \hat{n}, \hat{m}|\theta\right)$. Most of the current available methods update the parameters θ at each iteration using the gradient of the objective function. Numerical approximation of the gradient may be very time consuming since evaluating the log-likelihood requires that we solve the system of equations (2). In Theorem 2 below, we provide an analytic expression of the gradient, which is particularly useful for applied work.

Theorem 2 (Gradient of the log-likelihood)**.** *Let N^h denote the predicted number of households. Then, the derivative of the predicted frequency of a match pair $(x, y) \in \mathcal{X}\mathcal{Y}_0$-type household with respect to θ^k is given by,*

$$\partial_{\theta^k}\Pi_{xy} = \frac{\partial_{\theta^k}\mu_{xy}}{N^h} + \frac{\mu_{xy}}{N^h \times N^h}\sum_{xy\in\mathcal{X}\mathcal{Y}_0}\partial_{\theta^k}\mu_{xy}$$

where $\partial_{\theta^k}\mu_{xy} = \partial_{\mu_{x0}}M^{\theta}_{xy}\left(\mu^{\theta}_{x0}, \mu^{\theta}_{0y}\right)\partial_{\theta^k}\mu_{x0} + \partial_{\mu_{0y}}M^{\theta}_{xy}\left(\mu^{\theta}_{x0}, \mu^{\theta}_{0y}\right)\partial_{\theta^k}\mu_{0y} + \partial_{\theta^k}M_{xy}\left(\mu^{\theta}_{x0}, \mu^{\theta}_{0y}\right)$, *whenever* $xy \in \mathcal{X}\mathcal{Y}$, *and*

$$\begin{pmatrix}\partial_{\theta^k}\mu_{x0}\\ \partial_{\theta^k}\mu_{0y}\end{pmatrix} = \Delta^{-1}\begin{pmatrix}c^k\\ d^k\end{pmatrix}$$

otherwise, where $c^k_x = -\sum_{y\in\mathcal{Y}}\partial_{\theta^k}M_{xy}\left(\mu^{\theta}_{x0}, \mu^{\theta}_{0y}\right)$, *and* $d^k_y = -\sum_{x\in\mathcal{X}}\partial_{\theta^k}M_{xy}\left(\mu^{\theta}_{x0}, \mu^{\theta}_{0y}\right)$, *and* Δ *is expressed blockwise by*

$$\Delta = \begin{pmatrix} diag(1+\sum^{\theta}_{y\in\mathcal{Y}}\partial_{\mu_{x0}}M^{\theta}_{xy}) & \left(\partial_{\mu_{0y}}M^{\theta}_{xy}\right)_{xy} \\ \left(\partial_{\mu_{x0}}M^{\theta}_{xy}\right)_{yx} & diag(1+\sum_{x\in\mathcal{X}}\partial_{\mu_{0y}}M^{\theta}_{xy}) \end{pmatrix}. \tag{12}$$

Proof. See Appendix A. ∎

3.2.2 The MPEC approach

In our second approach, we rewrite Problem (11) as the Lagrangian

$$\min_{\lambda\in R^k}\max_{u\in R^k,\theta\in R^d} l\left(\theta, u, v\right) + \lambda G\left(\theta, \mu_{x0}, \mu_{0y}\right)$$

where λ is the Lagrange multiplier associated with the constraint $G(\theta, u, v) = 0$, and $u = (-\log(\mu_{x0}))_{x\in\mathcal{X}}$, and $v = (-\log(\mu_{0y}))_{y\in\mathcal{Y}}$. The first order conditions are therefore

$$\begin{aligned} Z_1(\theta, u, v, \lambda) &= 0 = \partial_\theta l(\theta, u, v) + \lambda \partial_\theta G(\theta, u, v), \\ Z_2(\theta, u, v, \lambda) &= 0 = \partial_{u,v} l(\theta, u, v) + \lambda \partial_{u,v} G(\theta, u, v), \text{ and} \\ Z_3(\theta, u, v, \lambda) &= 0 = G(\theta, u, v), \end{aligned}$$

which defines a map Z,

$$\begin{array}{ccc} \mathbb{R}^d \times \mathbb{R}^{|\mathcal{X}|} \times \mathbb{R}^{|\mathcal{Y}|} \times \mathbb{R}^{|\mathcal{X}|+|\mathcal{Y}|} & \to & \mathbb{R}^d \times \mathbb{R}^{|\mathcal{X}|} \times \mathbb{R}^{|\mathcal{Y}|} \times \mathbb{R}^{|\mathcal{X}|+|\mathcal{Y}|} \\ (\theta, u, v, \lambda) & \to & Z = (Z_1(\theta, u, v, \lambda), Z_2(\theta, u, v, \lambda), Z_3(\theta, u, v, \lambda)). \end{array}$$

Maximizing the likelihood is equivalent to finding the root of Z. In general, numerical methods will require the knowledge of the Jacobian of Z, which is given by:

$$JZ = \begin{pmatrix} \partial_\theta^2 l(\theta, u, v) + \lambda \partial_\theta^2 G(\theta, u, v) & \partial_\theta \partial_{u,v} l(\theta, u, v) + \lambda \partial_\theta \partial_{u,v} G(\theta, u, v) & \partial_\theta G(\theta, u, v) \\ \partial_\theta \partial_{u,v} l(\theta, u, v) + \lambda \partial_\theta \partial_{uv} G(\theta, u, v) & \partial_{u,v}^2 l(\theta, u, v) + \lambda \partial_{u,v}^2 G(\theta, u, v) & \partial_{u,v} G(\theta, u, v) \\ \partial_\theta G(\theta, u, v) & \partial_{u,v} G(\theta, u, v) & 0 \end{pmatrix} \tag{13}$$

We discuss the advantages and disadvantages of each approach in Online Appendix B.1, and provide numerical experiments as well.

3.3 Estimation in large markets

In this section, we shall discuss the difference between homogeneous and non-homogeneous models in terms of estimation, and provide formulas to compute the confidence intervals for the homogenous case.

3.3.1 Homogeneous and non-homogeneous matching functions

Recall that in equilibrium, the scarcity constraints from equation (2) are satisfied, that is $n_x = \mu_{x0}^* + \sum_{y\in\mathcal{Y}} M_{xy}\left(\mu_{x0}^*, \mu_{0y}^*\right)$ and $m_y = \mu_{0y}^* + \sum_{x\in\mathcal{X}} M_{xy}\left(\mu_{x0}^*, \mu_{0y}^*\right)$ for some matching function M_{xy}. Consider the equivalent system of equations, with rescaled quantities, i.e

$$\frac{n_x}{K} = \frac{\mu^*_{x0}}{K} + \sum_{y \in \mathcal{Y}} \widetilde{M}_{xy}\left(\frac{\mu^*_{x0}}{K}, \frac{\mu^*_{0y}}{K}\right) \text{ and } \frac{m_y}{K} = \frac{\mu^*_{0y}}{K} + \sum_{x \in \mathcal{X}} \widetilde{M}_{xy}\left(\frac{\mu^*_{x0}}{K}, \frac{\mu^*_{0y}}{K}\right)$$

where we introduce the new matching function $\widetilde{M}_{xy}(a,b) \equiv \frac{1}{K} M_{xy}(Ka, Kb)$, for any given $K > 0$. When the model is homogeneous, $\widetilde{M} = M$ for any $K > 0$. However, in the non-homogeneous case, there is no guarantee that even when K grows large, the matching function $\widetilde{M}$ will converge to a non trivial stable matching function $\bar{M}$. In addition, homogeneous models allow us to work with quantities that can be interpreted as frequencies instead of masses. Indeed, take $K = N^\star$, the total number of household in equilibrium. Hence

$$\frac{n_x}{N^\star} = \frac{\mu^*_{x0}}{N^\star} + \sum_{y \in \mathcal{Y}} M_{xy}\left(\frac{\mu^*_{x0}}{N^\star}, \frac{\mu^*_{0y}}{N^\star}\right) \text{ and } \frac{m_y}{N^\star} = \frac{\mu^*_{0y}}{N^\star} + \sum_{x \in \mathcal{X}} M_{xy}\left(\frac{\mu^*_{x0}}{N^\star}, \frac{\mu^*_{0y}}{N^\star}\right). \quad (14)$$

Let $\zeta_x = \frac{n_x}{N^*}$ and $\zeta_y = \frac{m_y}{N^*}$. Then $(\pi^*_{x0}, \pi^*_{0y}) = (\frac{\mu^*_{x0}}{N^*}, \frac{\mu^*_{0y}}{N^*})$ is the unique solution to the system of equation $AM(\pi_{x0}, \pi_{0y}) = \zeta$.[8] In this case, the key novelty is that the inputs of the aggregate matching function are interpreted as the frequencies of single men of type x and single women of type y in the population of households instead of masses, while the output is interpreted as the frequency of married couples of type xy in the population of households instead of the mass. This is no longer the case with non-homogeneous models, as there is no guarantee that the output of the aggregate matching function can be interpreted as matching frequencies when the inputs are also frequencies. Intuitively, this is also the reason why we can only compute confidence intervals for homogeneous models, which we provide in the next section.

In practice, homogeneous models imply that we observe the matching frequencies $\hat{\pi}$ from the data and estimate θ by maximum likelihood where the likelihood is given by,

$$l\left(\hat{\pi}|\theta, \zeta\right) = \sum_{xy \in \mathcal{X}\mathcal{Y}_0} \hat{\pi}_{xy} \log \Pi_{xy}\left(\theta, \zeta\right), \quad (15)$$

and ζ will be replaced by its estimator $A\hat{\pi}$. This does not change the estimation, as $\Pi_{xy}(\theta, \zeta)$

[8]For convenience, we rewrite the system of equations in (2) as $AM(\mu_{x0}, \mu_{0y}) = \begin{pmatrix} n_x \\ m_y \end{pmatrix}$, where A is an $(|\mathcal{X}| + |\mathcal{Y}|) \times (|\mathcal{X}||\mathcal{Y}| + |\mathcal{X}| + |\mathcal{Y}|)$ matrix.

is homogeneous of degree 0 in ζ by the homogeneity property. It means that in the homogenous case, working with the observed $\hat{\pi}$ is sufficient to carry on estimation, which is not the case in the non-homogeneous setting, where we must work with $\hat{\mu}$. Asymptotically, the noise added by $\hat{\pi}$ is normally distributed around π, so that $N^{1/2}(\hat{\pi} - \pi) \sim \mathcal{N}(0, V_\pi)$, where $V_\pi = diag(\pi) - \pi\pi'$. However, in the non-homogeneous case, we have instead, $N^{1/2}(\hat{\mu} - \mu) \sim \mathcal{N}(0, V_\mu)$, where $V_\mu = diag(N\mu) - \mu\mu'$.

3.3.2 Confidence intervals for homogeneous models

We now provide formulas to compute confidence intervals estimates when the matching function is homogeneous of degree 1. Note that these formulas are provided in Galichon, Kominers, and Weber (2019) for imperfectly transferable utility matching models with logit heterogeneity in taste. In that paper, it is shown that such models lead to a matching function that is homogeneous of degree 1, as recalled in Section 2.2(d). Therefore, if we can show that any matching function that is homogeneous of degree 1 is in fact equivalent to a matching model with ITU and logit unobserved heterogeneity as introduced in GKW, then we can make use of their results. This is proven in the following theorem.

Theorem 3. *(i) Consider an ITU-logit model as introduced in GKW. Then, the associated aggregate matching function is homogenous of degree 1. (ii) Consider a matching function equilibrium model and assume that its aggregate matching function satisfies Assumption 1 and homogeneity of degree 1. Then, this is equivalent to a matching model with ITU and logit unobserved heterogeneity as introduced in GKW.*

Proof. See Appendix A. ∎

The equivalence between the matching function equilibrium models with homogeneity of degree 1 and ITU-logit models in GKW suggests that we can simply make use of GKW's results for computing confidence intervals. Note that in this case, ζ is estimated by $A\hat{\pi}$ and thus, as noted earlier, doing so introduces additional noise in the estimates of θ, so that the computation of the variance-covariance matrix of θ cannot rely on the standard formulas. For the sake of clarity, we reproduce the results from GKW to compute the variance-covariance matrix V_θ in close form, and give a detailed proof of the result.

Theorem 4 (Confidence Intervals)**.** *When θ is estimated by maximum likelihood as described in Section 3 and homogeneity is satisfied, then $V_\theta = (\mathcal{I}_{11})^{-1} + \mathcal{I}_{11}^{-1}\mathcal{I}_{12}AV_\pi A'\mathcal{I}_{12}'\mathcal{I}_{11}^{-1}$ where we denote $\mathcal{I}_{11} = -(D_\theta \log \Pi)' \, diag(\pi)(D_\theta \log \Pi)$, and $\mathcal{I}_{12} = (D_\theta \log \Pi)' \, diag(\pi)(D_\zeta \log \Pi)$ and $V_\pi = diag(\pi) - \pi\pi'$.*

Proof. See appendix A. ∎

Note that $D_\theta \log \Pi$ and $D_\zeta \log \Pi$ have analytic expressions, which we give in the proof.

4 Estimation of Counterfactuals

One key advantage of estimating behavioural structural model is that it allows researchers to conduct counterfactual experiments for policy analysis. This is particularly relevant in the context of the marriage market, as researchers are often interested in how the marriage distribution changes in response to a change in the number of available individuals due to policy changes, such as changes to birth control policies, tax policies, divorce laws, financial aid program, and so on.

In this section, we introduce two methods to conduct counterfactual experiments: (i) the parametric approach, which relies on previously estimated structural parameters; (ii) the *parameter-free approach*, which works whenever the matching functions are homogeneous and multiplicatively separable in parameters.

With some abuse of notation, let the marriage distribution and available individuals at the observed equilibrium be denoted $\mu^* = (\mu^*_{xy}, \mu^*_{x0}, \mu^*_{0y})$ and (n_x, m_y) for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. Consider a counterfactual policy change that shifts the number of available individuals from n_x and m_y to n'_x and m'_y for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, while leaving the parameter θ unchanged. One main goal in the counterfactual analysis is to estimate the new equilibrium marriage distribution, denoted by $\mu^{*\prime} = (\mu^{*\prime}_{xy}, \mu^{*\prime}_{x0}, \mu^{*\prime}_{0y})$, under the counterfactual available individuals, (n'_x, m'_y).

4.1 The Parametric Approach

The parametric approach for the counterfactual analysis in the structural econometrics literature follows a two-step procedure: (i) estimate the parameters, $\hat{\theta}$, from the observed matches μ^*, by using the proposed nested or MPEC approaches in Section 3 ; (ii) compute the counterfactual equilibrium unmarrieds $(\mu_{x0}^{*\prime}, \mu_{0y}^{*\prime})$ under (n_x', m_y') by solving the system (16) below and obtain the counterfactual equilibrium matches $\mu_{xy}^{*\prime}$ using $\mu_{xy}^{*\prime} = M_{xy}^{\hat{\theta}}(\mu_{x0}^{*\prime}, \mu_{0y}^{*\prime})$.

$$\begin{cases} n_x' = \mu_{x0} + \sum_{y \in \mathcal{Y}} M_{xy}^{\hat{\theta}}(\mu_{x0}, \mu_{0y}) \\ m_y' = \mu_{0y} + \sum_{x \in \mathcal{X}} M_{xy}^{\hat{\theta}}(\mu_{x0}, \mu_{0y}) \end{cases} \tag{16}$$

Note that the existence and uniqueness results in Section 2 ensure that there is a unique solution to the system (16), and therefore, we have a unique prediction for the marriage distribution in the counterfactual case. The main advantage of the parametric approach is its generality. It applies to all matching function equilibrium models that satisfy the conditions given in Definition 1. However, as we see the simulation results in Online Appendix B.1, estimating the parameters $\hat{\theta}$ using the nested and MPEC approaches involves solving the nonlinear system (16) at each iteration. Moreover, the time for solving the nonlinear system (16) and the number of iterations required increase with the market size. Therefore, the estimation procedure can be computationally intensive and thus very time consuming, in particular when the market size is getting large.

4.2 The Parameter-Free Approach

In this section, we propose an alternative new approach, what we call the *Parameter-Free approach*, to compute counterfactuals in the context of matching function equilibrium models defined by Definition 1. We show that although the parameter-free approach applies only to a subset of models in which the associated matching functions take the multiplicative homogeneous form, it has a computational advantage as it only requires solving the system (2) once. We begin by making two more restrictions on the matching functions:

Assumption 2. *The matching function is multiplicatively separable in parameters with the form* $M_{xy}^{\theta}(\mu_{x0}, \mu_{0y}) = f(\theta) g(\mu_{x0}, \mu_{0y})$.

Assumption 3. *The matching function $M^\theta_{xy}(\mu_{x0}, \mu_{0y})$ is a homogeneous function in (μ_{x0}, μ_{0y}).*

Assumption 2 restricts the matching function to be multiplicatively separable in parameters, while Assumption 3 further restricts the matching function to be a homogeneous function. The two assumptions seems to be restrictive, but it is easy to check that most examples in Section 2.2 satisfy these assumptions.

Our parameter-free approach uses the ratio of the marriage distribution at the counterfactual equilibrium relative to that at the observed equilibrium. In particular, it consists of two steps. In the first step, by Assumptions 2 and 3, we show that the ratio of the matching function $M^\theta_{xy}(\mu_{x0}, \mu_{0y})$ evaluated at the counterfactual equilibrium relative to the observed equilibrium is free of model parameters. In the second step, we substitute this ratio of matches into the system (2) to generate a system in terms of ratios of unmarrieds at the counterfactual equilibrium relative to the observed equilibrium. This new system of equations in terms of ratios of unmarrieds at the two equilibria is also free of model parameters. We show that this new system has a unique solution in terms of the ratio of unmarrieds. We subsequently use this equilibrium ratio of unmarrieds to calculate the changes in the matching distributions between the observed and counterfactual equilibria. We now formally present these two steps.

Let us introduce the notation $\tilde{z} = z'/z$, which for any variable z denotes the ratio of the counterfactual equilibrium quantities to the observed quantities. Consider taking the ratio of matching function, $\mu_{xy} = M^\theta_{xy}(\mu_{x0}, \mu_{0y})$, evaluated at $(\mu^{*\prime}_{x0}, \mu^{*\prime}_{0y})$ under the counterfactual equilibrium and (μ^*_{x0}, μ^*_{0y}) under the observed equilibrium, which yields

$$\tilde{\mu}_{xy} = \frac{\mu^{*\prime}_{xy}}{\mu^*_{xy}} = \frac{M^\theta_{xy}(\mu^{*\prime}_{x0}, \mu^{*\prime}_{0y})}{M^\theta_{xy}(\mu^*_{x0}, \mu^*_{0y})} = \frac{g(\mu^{*\prime}_{x0}, \mu^{*\prime}_{0y})}{g(\mu^*_{x0}, \mu^*_{0y})} = g(\tilde{\mu}_{x0}, \tilde{\mu}_{0y}), \tag{17}$$

where the third equality is derived from Assumption 2 and the fourth equality is from the homogenous Assumption 3.

We next introduce $\tilde{\mu}_{xy}$ into the system (2) by expressing them in terms of ratios of counterfactual quantities relative to observed quantities, $(\tilde{\mu}_{xy}, \tilde{\mu}_{x0}, \tilde{\mu}_{0y}, \tilde{n}_x, \tilde{m}_y)$. Recall the system (2), i.e $n_x = \mu_{x0} + \sum_{y\in\mathcal{Y}} M^\theta_{xy}(\mu_{x0}, \mu_{0y})$ and $m_y = \mu_{0y} + \sum_{x\in\mathcal{X}} M^\theta_{xy}(\mu_{x0}, \mu_{0y})$. Evaluating the system at the quantities under both counterfactual and observed equilibria, dividing them, and substituting (17) into it, we obtain

$$\begin{cases} \tilde{n}_x = p_{x0}\tilde{\mu}_{x0} + \sum_{y\in\mathcal{Y}} p_{xy} \cdot g(\tilde{\mu}_{x0}, \tilde{\mu}_{0y}) \\ \tilde{m}_y = q_{0y}\tilde{\mu}_{0y} + \sum_{x\in\mathcal{X}} q_{xy} \cdot g(\tilde{\mu}_{x0}, \tilde{\mu}_{0y}) \end{cases}, \tag{18}$$

where the parameters $p_{x0} = \mu^*_{x0}/n_x$ and $p_{xy} = \mu^*_{xy}/n_x$ are the observed equilibrium probabilities that men of type x remain single and marry with type a woman of type y, respectively. Similarly, $q_{0y} = \mu^*_{0y}/m_y$ and $q_{xy} = \mu^*_{xy}/m_y$ are the probabilities that type y women remain single and marry type x men, respectively. This new system of equations is free of the parameters θ. This system has $|\mathcal{X}| + |\mathcal{Y}|$ number of nonlinear equations with $|\mathcal{X}| + |\mathcal{Y}|$ unknowns $(\tilde{\mu}_{x0}, \tilde{\mu}_{0y})$. If there exists a unique solution for $(\tilde{\mu}_{x0}, \tilde{\mu}_{0y})$ in the new system (18), we can then obtain $(\tilde{\mu}_{x0}, \tilde{\mu}_{0y})$ by solving the system only once. Using the obtained $(\tilde{\mu}_{x0}, \tilde{\mu}_{0y})$, the new counterfactual equilibrium quantities, $(\mu^{*\prime}_{x0}, \mu^{*\prime}_{0y}, \mu^{*\prime}_{xy})$ can be computed accordingly from the definition, $\tilde{z} \equiv z'/z$, and equation (17).

The final issue left for our approach to work is to show that the system (18) has a unique solution in $(\tilde{\mu}_{x0}, \tilde{\mu}_{0y})$. Indeed, we show that under Assumptions 1 , 2, and 3, there exist a unique solution in $(\tilde{\mu}_{x0}, \tilde{\mu}_{0y})$ for the system (2). This result is formally stated in Theorem 5.

Theorem 5. *Under Assumptions 1, 2, and 3, there exists a unique solution of the system (18) with unknowns $(\tilde{\mu}_{x0}, \tilde{\mu}_{0y})$ for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$.*

Proof. See Appendix A. ∎

The proof of existence relies on a revised procedure based on Algorithm 1, while the proof of uniqueness similarly relies on Berry, Gandhi, and Haile (2013).

The estimation procedure of our parameter-free approach works as follows: (i) compute the changes $(\tilde{n}_x, \tilde{m}_y)$, and the probabilities (p_{x0}, p_{xy}) and (q_{x0}, q_{xy}), and then obtain $\tilde{\mu}_{xy}$ using (17) ; (ii) compute the changes in unmarrieds $(\tilde{\mu}_{x0}, \tilde{\mu}_{0y})$ by solving (18), and then compute $(\mu^{*\prime}_{x0}, \mu^{*\prime}_{0y}, \mu^{*\prime}_{xy})$ accordingly.

Our approach highlights a previously undocumented new property in the matching literature, that ratios in the equilibrium numbers of unmarried men and women can be inferred from the observed matching equilibrium by using the system (18), which themselves are free of the model parameters for some class of matching models. The first paper that proposed this 'hat' approach is Eaton, Dekle, and Kortum (2007). While this approach is quite commonly used in the trade literature, it does not typically generate a transformed model that is free of structural parameters. Unlike our case here, parameters such as the elasticity of substitution typically remain in the 'hat' model and these parameters need to be calibrated or estimated.

Note that comparing to the parametric approach, the parameter-free approach has an advantage in terms of computational time as it does not require the estimation of model parameters. However, it is also worth noting that the parameter-free approach only applies to models whose associated matching functions are homogeneous and multiplicatively separable in parameters. It seems that the two assumptions are restrictive, but most matching in examples of Section 2.2 satisfy these two assumptions and our parameter-free approach applies to the models. For example, it is trivial to show that matching functions in the TU models of CS, NTU models of Menzel, and BLP models are homogeneous and multiplicatively separable in parameters.

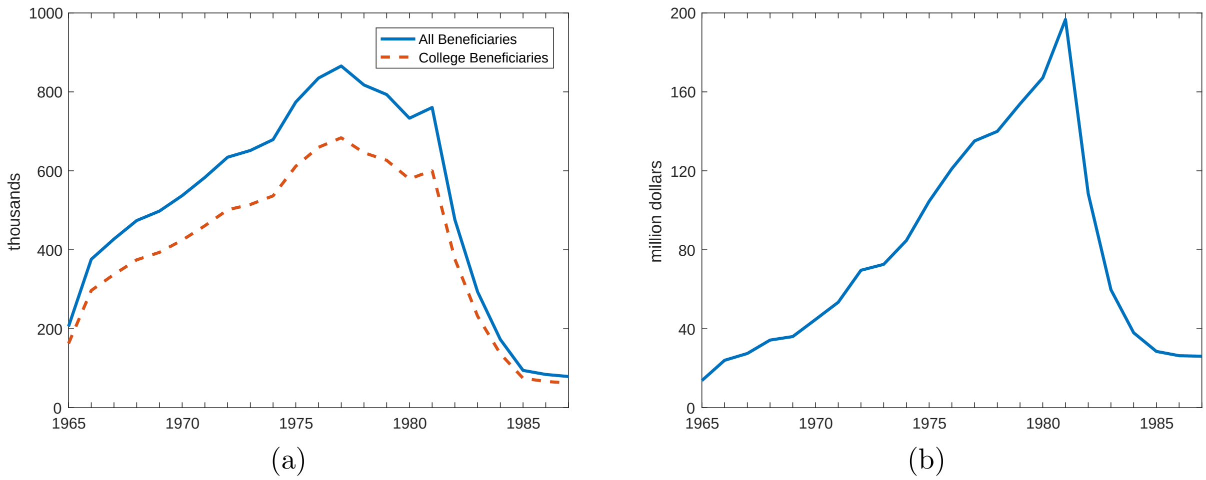
5 Empirical Application - The 1982 Elimination of the Social Security Student Benefit Program

The goal in this section is to investigate how the elimination of the Social Security Student Benefit Program in 1982 has affected the 1987/88 age-education marriage distributions in the United States. We do so by employing the CS matching model and the proposed parametric and parameter-free approaches to counterfactual analysis.

5.1 The background

Under the 1939 Amendments to the Social Security Act, the children of deceased, disabled, and retired Social Security beneficiaries could receive Social Security payments until they reach the age of eighteen. In 1965, these payments were extended to persons up to twenty two years of age still enrolled as full-time college students. The Social Security student benefits were paid to eligible college students as monthly lump sums. The benefits were extremely

Figure 1: Number of beneficiaries and average monthly benefits payment



Panel (a): Number of social security beneficiaries and college student beneficiaries; Panel (b): Average monthly benefit payment. Source: Social Security Administration (SSA) Research Note #11.

generous, especially considering the cost of public four-year colleges and universities at that time.[9] In the peak year of 1977, there were about $900,000$ benefit recipients. In the peak pay-out year of 1981, about $2.4 billion were paid as student benefits.[10]

In 1981, Congress voted to eliminate the Social Security Student Benefits Program from 1982 onwards.[11] Since then, the number of student benefit recipients and the program spending dropped dramatically. As shown in Figure 1a, the number of student beneficiaries dropped from around $760,000$ in 1981 to $84,000$ recipients in 1986. The number of college student beneficiaries is estimated to have dropped from about $600,000$ to $66,000$ recipients in 1986.[12] The amount paid to eligible students was reduced immediately after the elimination of the program in 1981 (see Figure 1b). The average monthly payment fell substantially from about $196 million in 1981 to $26 million in 1986.

Except for the introduction of the Pell Grant program in the early 1970's and the various

[9]The average annual benefit in 1980 paid to a child of deceased parent was $\$6,700$ while the average tuition and fees for public four-year colleges and universities was $\$1,900$.

[10]These statistics are obtained from "Research Note #11: The History of Social Security Student Benefits published by the Historian's Office".

[11]According to the 'Social Security Administration Research Note #11', "Benefits paid to post-secondary students ages 18-21 are to be phased-out; The phase-out is to be completed by April 1985; Benefits to elementary and/or secondary school students older than 18, are to end in August 1982."

[12]We do not have access to annual share of college student beneficiaries. According to a 1977 Social Security Administration (SSA) survey, about 79% student beneficiaries are in post-secondary institutions.

G.I. Bills, the elimination of the Social Security Student Benefit program is the largest policy change in financial aid for college students. The impacts of various financial aid programs on students' college attendance and completion has been well studied in the literature.[13] For the Social Security Student Benefit program, Dynarski (2003) found that the elimination of the program has had a large significant causal effect on students' college attendance and completion. It is well known that education is a primary attribute in the marriage market. The marriage matching distribution has important implications on fertility and population growth, labor-force participation of women, income inequality, etc (see e.g., Becker (1991)). Our goal is to understand how the 1982 elimination of the Social Security Student Benefit program affects the marriage matching distribution through its impact on students' college attendance and completion.

5.2 The data

We attempt to answer this question by estimating the counterfactual marriage distribution in 1987/88 had the Social Security Student Benefit program not been eliminated and compare it with the observed 1987/88 marriage matching distribution using the CS model. Since the CS model is static, we need to choose a specific year for our analysis. The year 1987/88 is sufficiently far along after the policy change to allow those most affected by the policy to reach a marriageable age. The elimination of aid in 1982 would most likely affect the then and soon-to-be high school seniors. It will also affect those individuals attending college and those who were considering going to college in the near future.[14]

For this exercise, we require three data inputs:

i) the observed number of available single men and women by age and education as a result of the elimination of the financial aid program, (n_x, m_y), for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$,

[13] See e.g., Manski and Wise (1983) and Kane (1994) on Pell grant introduced in 1973, Reyes (1997) on the Middle Income Student Assistance Act, which eliminated the income cutoff for the Guaranteed Student Loan, and Angrist (1993) on War II G.I. Bills, which provided a generous monthly stipend to veterans in college.

[14] The year 1988 also happens to be the last year for which we have access to educational attainment of newly weds. We also wanted to minimize the effect of other educational policy that came into effect towards the end of the 1980s that would have confounded our results, such as 'The Emergency Immigration Education Act' of 1984.

ii) the counterfactual number of available single men and women by age and education had the financial aid program not been eliminated, (n'_x, m'_y), for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$,

iii) the observed flow of new marriages by age and education as a result of the elimination of the financial aid program, μ.

Using these inputs, our empirical framework allows us to construct the counterfactual marriage distribution, μ', i.e. the marriage distribution by age and education had the financial aid program not been eliminated in 1981.

Data on the flow of new marriages as a result of the financial aid program elimination (that is, item iii)) is constructed using the 1987/88 Vital Statistics marriage records obtained from the National Bureau of Economic Research data website.[15] Before 1989, the Vital Statistics tracks new marriages by educational attainment for 22 reporting states.[16] However, from 1989 onwards, information on educational attainment of newly weds are no longer recorded. Our analysis will focus on these 22 reporting states. We construct item i), the observed number of available single men and women as a result of the financial aid program elimination using the 1986 U.S. Current Population Survey (CPS) for these 22 reporting states. More details about the data construction can be found in Online Appendix B.2.

Table 1: Numbers of available single male and female in millions from the 1986 CPS for the 22 reporting states

	Male	Female
High school or less (HS)	8.79	10.41
	(63.3%)	(65.3%)
College (Col)	4.24	4.72
	(30.5%)	(29.6%)
Graduate school (GS)	0.86	0.80
	(6.2%)	(5.1%)
Total	13.88	15.94

Percentage of total in parenthesis

Individuals are differentiated by their age and educational attainment. We divide edu-

[15]Like in CS, the flow of new marriages in constructed by taking a two year average of the new marriages in 1987 and 1988. This helps reduce the number of marriage pairs with zero new marriages.

[16]These 22 states include California, Connecticut, Hawaii, Illinois, Kansas, Kentucky, Louisiana, Maine, Mississippi, Missouri, Montana, Nebraska, New Hampshire, New York, North Carolina, Rhode Island, Tennessee, Utah, Vermont, Virginia, Wisconsin and Wyoming.

cational attainment into three levels - high school diploma or less, some years of college or college degree, and graduate school. Where convenient, we will refer to these three groups with the abbreviation HS, Col and GS respectively. Table 1 provides count statistics for our sample of single individuals by education in 1986. These are constructed by taking the average of unmarried individuals from the twelve CPS monthly surveys in 1986. There are around 16 million single women and around 14 million single men between the ages of 16 and 75 in our sample from the 22 reporting states. There are dramatically fewer single adults with Col and GS educational attainment compared to those with HS education. Around 63% of single men and 65% of single women has qualification up to a high school diploma. Only around 30% of single men and women have some years in a college or a college degree and only 6% of single men and 5% of single women have post-college qualification.

Table 2: Number of marriages by education in thousands from the 1987/88 Vital Statistics for the 22 reporting states

		Female		
		High School	College	Graduate School
Male	High School	573.96	167.71	11.35
	College	153.47	303.81	34.10
	Graduate School	14.40	53.21	40.39

Table 2 tabulates the 1987/88 marriage distribution by education groups for the 22 reporting states. There is an average of 1.36 million marriages over this two year period. As evident from the table, there is strong assortative matching by education groups. With the exception of men with graduate school qualification, each of the remaining 5 groups of single men and women are most likely to marry a spouse with the same educational attainment. Men with graduate school qualification are more likely to marry college educated than graduate school educated women. This pattern does not hold for women with graduate school qualification.[17]

Our methodology also requires data on (n'_x, m'_y), the supply of single men and women had the student benefit program not been eliminated (that is, item ii) above). Since this

[17] With age ranging from 16 to 75 years and the 3 education groups, we have 180 types (or age-education combination) of both men and women. Since individuals who are younger than 23 years of age rarely completed graduate school education, we exclude individuals younger than 23 with graduate school qualification. This reduces the number of types from 180 to 173 for both men and women.

counterfactual is unobserved to the econometrician, we use the estimates from Dynarski (2003) to construct the counterfactual changes in available single men and women. Using data from the CPS, Dynarski (2003) proxies the benefits eligibility by the death of a parent during the individual's childhood. The author employs the difference-in-differences framework to analyze the impact of the elimination of the Social Security Student Benefit program on the probability of college attendance and completion for students who were eligible to the program. Let the causal effect on college attendance and completion be denoted by γ and δ, respectively. Dynarski (2003) finds that eliminating the financial aid program on average reduces the probability of attending college by about $\hat{\gamma} = 24.3\%$ and the probability of completing any year of college by $\hat{\delta} = 16.1\%$ for the eligible students.

Table 3: Effect on single pool had the student benefit program not been eliminated.

Age in 82	Age in 86	Aid Effects
14	18	$\rho_{18} \cdot \gamma$
15	19	$\rho_{18} \cdot \gamma + \rho_{19} \cdot \delta$
16	20	$\rho_{18} \cdot \gamma + \rho_{19} \cdot \delta + \rho_{20} \cdot \delta$
17	21	$\rho_{18} \cdot \gamma + \rho_{19} \cdot \delta + \rho_{20} \cdot \delta + \rho_{21} \cdot \delta$
18	22	$\rho_{18} \cdot \gamma + \rho_{19} \cdot \delta + \rho_{20} \cdot \delta + \rho_{21} \cdot \delta$
19	23	$\rho_{19} \cdot \delta + \rho_{20} \cdot \delta + \rho_{21} \cdot \delta$
20	24	$\rho_{20} \cdot \delta + \rho_{21} \cdot \delta$
21	25	$\rho_{21} \cdot \delta$

Since these estimates are for eligible students and we do not observe actual benefit recipients in our data, we need to also compute the fraction of the population who are eligible for these benefits. The program did not differentiate between male or female recipients. Hence, we construct the counterfactual numbers of available singles without distinguishing their gender. Using the similar approach as in Dynarski (2003), we proxy the proportion of age i benefit eligible individuals (males or females), ρ_i, by the fraction of age i individuals whose father are deceased, retired or disable in their cohort from the 1980 U.S. census. We assume that individuals (males or females) are in high school till they are 18 and that a college degree takes 4 years.

Consider now a counterfactual setting where the financial aid program had not been eliminated. Eighteen years old high school seniors in 1986 would have been fourteen years old high schoolers in 1982 when the program was eliminated. Our estimates suggest that an additional $\rho_{18}\gamma$ proportion of high school graduates would have attended college had the program not been eliminated. As for nineteen year olds in 1986, a $\rho_{18}\gamma$ fraction of these individuals would have attended college when they were eighteen years old and $\rho_{19}\delta$ of them would not have dropped out of college that year. Hence, an additional $(\rho_{18}\gamma + \rho_{19}\delta)$ of the nineteen years old high school graduates would have attended college in 1986. We repeat this calculation for individuals aged between eighteen and twenty five years old in our 1986 supply of single men and women. Table 3 tabulates the calculations of changes to the population of single high school graduates by age in the counterfactual setting.

Table 4: Available number of single men and women between ages 18 and 22 (millions)

		Male	
	CPS in 1986	Counterfactual Policy	% Change
High school or less	2.52	2.36	-6.12%
College	1.23	1.39	12.50%
		Female	
	CPS in 1986	Counterfactual Policy	% Change
High school or less	2.05	1.93	-5.90%
College	1.28	1.40	9.47%

Table 4 compares the observed number of single men and women between the ages of 18 and 22 with the counterfactuals computed using the procedure just outlined. While the estimated causal effect in Dynarski (2003) was statistically significant on those eligible for the benefits, the overall effect of the program elimination on the number of single men and women remains modest due to the small fraction of eligible individuals in the population. Our calculation suggests that the number of college graduated men and women between the ages of 18 and 22 would have increased by approximately 136,000 and 121,000 respectively. This represents an increase of around 12.5% and 9.47% more college educated men and women aged between the ages of 18 and 22, respectively. Figure 4 in Online Appendix B.3 shows the observed and counterfactual available single men and women by age between 16 and 22 in 1986.

5.3 Implementations and specifications

Using the above data, we now compare the following three approaches to counterfactual analysis.

1) The first approach is the parametric approach by using a nonparametric specification of the CS model. Following the nonparametric approach proposed in Choo and Siow (2006b), we first nonparametrically estimate the joint surplus Φ_{xy} in the matching function from the CS model. We then solve for the counterfactual unmarrieds (μ'_{x0}, μ'_{0y}) by substituting the estimated joint surpluses into the system (16), and calculate matches μ'_{xy} from the matching function accordingly.

2) The second approach is the parametric approach by using a parametric specification of the CS model. We first use the nested maximum likelihood approach proposed in Section 3 to estimate a parameterized version of the joint surpluses Φ_{xy} in the CS model.[18] Once the joint surpluses has been estimated, we can solve for the counterfactual unmarrieds and matches similarly as in the first approach.

   In this appoach, we parametrize Φ_{xy} as follows. Let the type of a man be defined as $x = (x_a, x_e)$, where $x_a \in \{16, \cdots, 75\}$ (we relabel it as $x_a \in \{1, \cdots, 60\}$) denotes the man's age and $x_e \in \{hs, col, grad\}$ (we relabel it as $x_e \in \{1, 2, 3\}$) denotes the man's education level. Similarly, we define the type of a woman as $y = (y_a, y_e)$ with y_a the age of the woman relabelled as $y_a \in \{1, \cdots, 60\}$ and y_e her education relabelled as $y_e \in \{1, 2, 3\}$. Our parametric specification for Φ_{xy} is given by,

$$\Phi^{\theta}_{xy} = \theta_0 + \sum_{i=2}^{60} \theta^{ma}_i \mathbb{1}\{x_a = i\} + \sum_{j=2}^{3} \theta^{me}_j \mathbb{1}\{x_e = j\} + \sum_{i=2}^{60} \theta^{wa}_i \mathbb{1}\{y_a = i\} + \sum_{j=2}^{3} \theta^{we}_j \mathbb{1}\{y_e = j\}$$
$$+ \sum_{i=1}^{59} \theta^{mwa}_i \mathbb{1}\{|x_a - y_a| = i\} + \sum_{j=1}^{2} \theta^{mwe}_j \mathbb{1}\{|x_e - y_e| = j\}, \tag{19}$$

   where the second and third terms capture the surplus from men's traits, the fourth and fifth terms capture the surplus from women's traits, and the sixth and seventh terms

[18] We could also estimate the model using moment matching techniques. For a discussion on the use of maximum likelihood and moment matching techniques in matching models, see Galichon and Salanié (2020).

capture the surplus from the interactions between the traits of men and women.[19] This specification requires us to estimate 184 parameters in $\boldsymbol{\theta}$.

3) The third approach is the parameter-free approach presented in Section 4.2. The implementation follows the steps in Section 4.2 closely. We first derive the changes of matches in terms of changes in unmarrieds using the matching function in CS, $\tilde{\mu}_{xy} = \sqrt{\tilde{\mu}_{x0}\tilde{\mu}_{0y}}$. We then solve for $(\tilde{\mu}_{x0}, \tilde{\mu}_{0y})$ by substituting $\tilde{\mu}_{xy}$ into the system (18), and then calculate $\tilde{\mu}_{xy}$ and the counterfactual matches, $\mu'_{x0}, \mu'_{x0}, \mu'_{xy}$, accordingly.

Among the three approaches, the two parametric approaches (1) and (2) require the estimation of the joint surpluses of each match type while the parameter-free approach does not. Moreover, while the steps of the parametric approach (1) and the parameter-free approach (3) differ, it is worth highlighting that the counterfactual distributions estimated from both procedures are identical. The transformation of the matching equilibrium model in the parameter-free approach allows us to by-pass the estimation of the non-parametric joint marital surplus. Hence, we will focus our attention on the joint matching surpluses, and counterfactual marriage distributions estimated using the nonparametric (approach (1)) and parametric (approach (2)) marriage surplus specifications.

5.4 The estimated joint surpluses

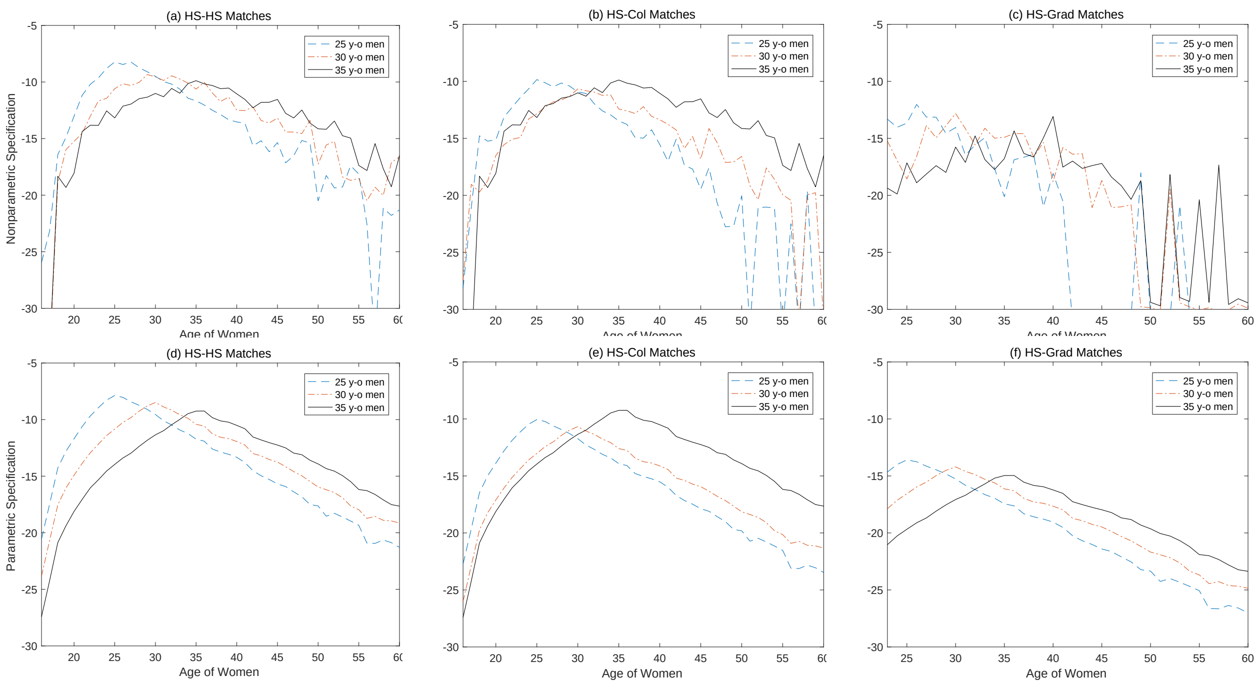
We report and compare the joint surpluses estimated from both approaches (1) and (2). In Figure 2, we focus on the joint surplus estimates for high school men matched with women with different education levels[20]. The subfigures in the first row display the estimated joint surpluses for high school men at three different ages using the nonparametric specification, while the subfigures in the second row graph the corresponding estimates coming from the parametric specification. These figures show two main features of the estimated joint surpluses of high school men: (1) the joint surplus estimates exhibit strong assortativeness in education and age under both the nonparametric and parametric specifications. The couples

[19]For instance, the joint surplus for a match with man's traits $x = (x_a = 6, x_e = 2)$ and woman's trait $y = (y_a = 8, y_e = 3)$ is given by $\Phi_{xy}(\boldsymbol{\theta}) = \theta_0 + \theta_6^{ma} + \theta_2^{me} + \theta_8^{wa} + \theta_3^{we} + \theta_2^{mwa} + \theta_1^{mwe}$.

[20]The features of the joint surplus estimates for college and graduate men are similar, and can be found in Figures 5 and 6 in Online Appendix B.3.

with similar education and age generally obtain the highest surplus. (2) the joint surpluses estimated from the parametric specification are smoother than those estimated from non-parametric specification, especially for marriages involving old GS educated women. This is due to the fact that we observe very thin cells for the marriages between high school men with old or GS educated women,[21] which gives us imprecise nonparametric estimates for the joint surpluses.

Figure 2: Estimated Φ_{xy} for high school men



5.5 The changes in marriage distributions due to the policy change

We now estimate the counterfactual marriage distributions by using the two parametric approaches (i.e. the nonparametric and parametric marriage surplus specifications). We then compare the estimated counterfactual marriage distributions with the observed distribution for the 22 reporting states, to obtain the changes in marriage distributions due to the Social Security Benefit Program in 1982. While the elimination of the aid program would have affected the marriage distribution of the whole United States, our analysis is unfortunately

[21]There are about 63% zero cells for the number of marriages between high school men at 25,30, or 35 years old and women older than 55 years old, and for the marriages between high school men of the three ages with graduate school women, there are about 49.7% zero cells in the observed data.

confined to the 22 reporting states for which we have data on the flow of new marriages by age and education attainment.

Figure 3 displays the numbers of single individuals by age and education in both the observed and counterfactual marriage distributions. It shows that the numbers of single individuals in the two counterfactual marriage distributions estimated from the two specifications are similar for the HS and Col degree individuals, while the numbers for single GS educated individuals differ in the two counterfactual distributions. This difference in the latter is likely due to the imprecise estimates of the joint surplus involving GS educated individuals. When comparing the numbers of singles in the counterfactual marriage distributions with those in the observed marriage distribution, we find that eliminating the financial aid program would lead to less HS and more Col educated singles in equilibrium. This is consistent with the impacts of the financial aid program on the available numbers of individuals, shown in Figure 4 in Online Appendix B.3.

Figure 3: Counterfactual and observed numbers of single men and women

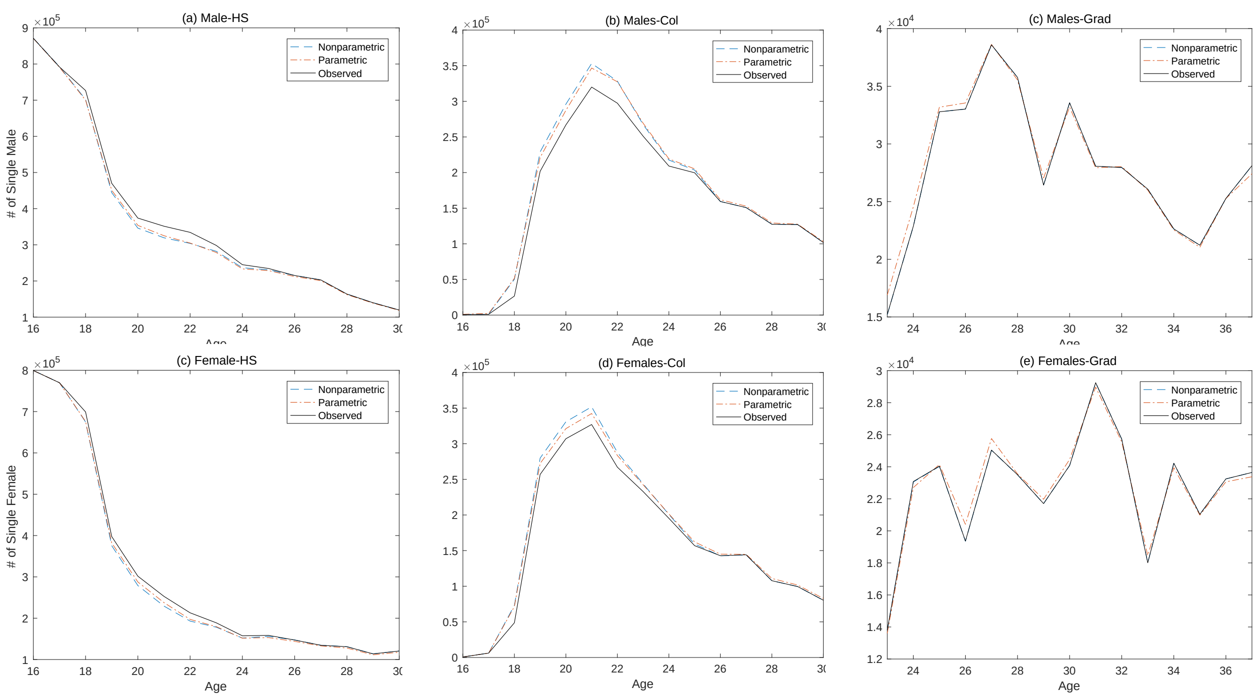


Table 5 tabulates the changes to the marriage distribution by education in the counterfactual marriage distributions estimated using both the nonparametric and parametric specifications.[22] We will use the convention HS-Col to refer to marriages between high school educated men and college educated women. Focusing on the nonparametric specification in the top panel of Table 5, we can see that the high degree of positive assortative matching implies that the biggest changes happen on HS-HS and Col-Col matches: the number of HS-HS marriages would fall by around 16798 (2.93%) matches while Col-Col marriages increase by around 8909 (2.93%) . A large change is seen among Col-HS marriages, which would increase by around 3985 (2.6%) compared to the modest decrease of 634 (0.38%) marriages among HS-Col marriages. This is probably a reflection of social norms of men preferring spouses who are not more educated than themselves, which is embedded in the preference parameters.

In the bottom panel of Table 5, we perform the same computations as above, but using the parametric marital surplus specification. The qualitative patterns remain similar for the changes to the marriages between individuals with HS or/and Col education. However, the changes in the number of the matches involving GS individuals differ significantly across the two different specification. This is a reflection of large differences in estimates across the two specifications in the preference parameters for GS educated individuals.

Table 5: Changes in Number of New Marriages by Education

Nonparametric Specification				
		Female		
		High School	College	Grad. School
Male	High School	-16798.1 (-2.93%)	-634.3 (-0.38%)	-31.4 (-0.28%)
	College	3985.0 (2.60%)	8909.7 (2.93%)	189.5 (0.06%)
	Grad. School	4.5 (0.03%)	86.8 (0.16%)	7.0 (0.02%)
Parametric Specification				
		Female		
		High School	College	Grad. School
Male	High School	-16982.6 (-2.96%)	3185.7 (1.90%)	-1267.2 (-11.16%)
	College	3432.2 (2.24%)	11749.5 (3.87%)	-706.6 (-0.23%)
	Grad. School	1078.3 (7.49%)	-3301.8 (-6.21%)	2370.1 (5.87%)

In Online Appendix B.3, we also provide graphs that display the changes in the marriage

[22] Recall that the nonparametric specification and parameter-free approach provide identical estimates of the counterfactual marriage distribution. We thus omit the results from the parameter-free approach in Table 5.

distribution for the four education pairs by age estimated from the nonparametric specification.[23] The graphs are consistent with the numbers reported in Table 5. In addition, they also show some patterns in the changes of numbers between pairs with different ages.

6 Conclusion

In the context of matching models, researchers are often interested in how the equilibrium matching distribution would change in response to a change in the structure of the market (e.g. for marriage markets, changes in the number of available men or women). We study this question in matching function equilibrium matching models, a class of matching models that's characterized by a matching function and a system of demographic constraints. We point out that a surprisingly large number of models in the matching literature belong to this class. In this paper, we focus on the partial assignment case. We show how one can parametrically estimate the matching functions of these models by maximum likelihood; we provide efficient computing techniques, an analytic expression for the gradient of the log-likelihood, and formulas to compute confidence intervals.

We study counterfactuals from policy changes that change the number of available men and women on the market but is assumed to leave the matching surplus parameters unchanged. We show how to compute the counterfactual equilibrium matching distribution when the structural parameters of the matching function have been previously estimated. In addition, we show that for a certain subclass of matching function equilibrium models, the counterfactual distributions are identified without estimating the structural parameters. We illustrate our framework by analyzing the impact of the elimination of the Social Security Student Benefit Program in 1982 on college attendance and the marriage market. We show that, had the policy not been abandoned, there would have been around 17,000 (3%) fewer marriages among high school graduates and 10,000 (3%) more marriages among college graduates in 1987/88, in the 22 reporting states for which we have data.

[23]The results for the changes estimated from parametric specification are similar, which is omitted.

References

[1] Joshua D Angrist. “The effect of veterans benefits on education and earnings”. In: *Industrial and Labor Relations Review* 46.4 (1993), pp. 637–652.

[2] Gary S Becker. “A theory of marriage: Part I”. In: *Journal of Political Economy* 81.4 (1973), pp. 813–846. ISSN: 0226740854.

[3] Gary S. Becker. *A Treatise On The Family*. Harvard University Press, 1991, pp. 304–304.

[4] Steven Berry, Amit Gandhi, and Philip Haile. “Connected substitutes and invertibility of demand”. In: *Econometrica* 81.5 (2013), pp. 2087–2111.

[5] Steven Berry, James Levinsohn, and Ariel Pakes. “Automobile prices in market equilibrium”. In: *Econometrica* (1995), pp. 841–890.

[6] Liang Chen et al. “Existence and Uniqueness in Matching FunctionEquilibria with Full Assignment”. In: *working paper* (2020).

[7] Laurens Cherchye et al. “Household consumption when the marriage is stable”. In: *American Economic Review* 107.6 (June 2017), pp. 1507–34.

[8] Pierre-André Chiappori, Bernard Salanié, and Yoram Weiss. “Partner choice, investment in children, and the marital college premium”. In: *American Economic Review* 107.8 (Aug. 2017), pp. 2109–67.

[9] Eugene Choo and Aloysius Siow. “Estimating a marriage matching model with spillover effects”. In: *Demography* 43.3 (2006), pp. 463–490.

[10] Eugene Choo and Aloysius Siow. “Who marries whom and why”. In: *Journal of Political Economy* 114.1 (2006), pp. 175–201.

[11] John K Dagsvik. “Aggregation in matching markets”. In: *International Economic Review* 41.1 (2000), pp. 27–58.

[12] Jean-Pierre Dubé, Jeremy T. Fox, and Che-Lin Su. “Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation”. In: *Econometrica* 80.5 (2012), pp. 2231–2267.

[13] Susan M. Dynarski. "Does aid matter? Measuring the effect of student aid on college attendance and completion". In: *American Economic Review* 93.1 (Mar. 2003), pp. 279–288.

[14] Jonathan Eaton, Robert Dekle, and Samuel Kortum. "Unbalanced trade". In: *American Economic Review* 97.2 (2007), pp. 351–355.

[15] D Gale and L S Shapley. "College admissions and the stability of marriage". In: *The American Mathematical Monthly* 69.1 (Jan. 1962), pp. 9–15.

[16] Alfred Galichon and Yu-Wei Hsieh. "A theory of decentralized matching markets without transfers". In: *Working Paper* (2017).

[17] Alfred Galichon, Scott Duke Kominers, and Simon Weber. "Costly concessions: An empirical framework for matching with imperfectly transferable utility". In: *Journal of Political Economy* 127.6 (2019), pp. 2875–2925.

[18] Alfred Galichon, Scott Duke Kominers, and Simon Weber. "The Nonlinear Bernstein-Schrödinger Equation in Economics". In: *International Conference on Networked Geometric Science of Information*. Springer. 2015, pp. 51–59.

[19] Alfred Galichon and Bernard Salanié. "Cupid's invisible hand: Social surplus and identification in matching models". In: *Available at SSRN 1804623* (2020).

[20] Marion Goussé, Nicolas Jacquemet, and Jean-Marc Robin. "Marriage, labor supply, and home production". In: *Econometrica* 85.6 (2017), pp. 1873–1919. ISSN: 1468-0262.

[21] Keith Head and Thierry Mayer. "Gravity equations: Workhorse,toolkit, and cookbook". In: *Handbook of International Economics*. Vol. 4. Elsevier, 2014. Chap. Chapter 3, pp. 131–195.

[22] Gunter J. Hitsch, Ali Hortaçsu, and Dan Ariely. "Matching and sorting in online dating". In: *American Economic Review* 100.1 (2010), pp. 130–63.

[23] Martin Idel. "A review of matrix scaling and Sinkhorn's normal form for matrices and positive maps". In: *arXiv preprint arXiv:1609.06349* (2016).

[24] Thomas J Kane. "College entry by blacks since 1970: The role of college costs, family background, and the returns to education". In: *Journal of Political Economy* 102.5 (1994), pp. 878–911.

[25] T.C. Koopmans and M.J. Beckmann. "Assignment problems and the location of economic activities". In: *Econometrica* 25.1 (1957), pp. 53–76.

[26] J Kruithof. "Telefoonverkeersrekening". In: *De Ingenieur* 52 (1937), E15–E25.

[27] Charles F Manski and David A Wise. *College choice in America*. Harvard University Press, 1983.

[28] D. McFadden. "Conditional Logit Analysis of Qualitative Choice Behavior." In: *Frontiers in Econometrics*. Ed. by Paul Zarembka. New York: Academic Press, 1974.

[29] Konrad Menzel. "Large matching markets as two-Sided demand systems". In: *Econometrica* 83.3 (2015), pp. 897–941.

[30] Ismael Mourifié. "A marriage matching function with flexible spillover and substitution patterns". In: *Economic Theory* (Sept. 2018). ISSN: 1432-0479.

[31] Ismael Mourifié and Aloysius Siow. "The cobb douglas marriage matching function: Marriage matching with peer and scale effects". In: *Working Paper* (2017).

[32] Jong-Shi Pang, Che-Lin Su, and Yu-Ching Lee. "A constructive approach to estimating pure characteristics demand models with pricing". In: *Operations Research* 63.3 (2015), pp. 639–659.

[33] Zhenchao Qian. "Changes in assortative mating: The impact of age and education, 1970-1990". In: *Demography* 35.3 (1998), pp. 279–292.

[34] Zhenchao Qian and Samuel H Preston. "Changes in american marriage, 1972 to 1987: Availability and forces of attraction by age and education". In: *American Sociological Review* (1993), pp. 482–495.

[35] Suzanne Louise Reyes. "Educational opportunities and outcomes: The role of the guaranteed student loan". In: *Unpublished manuscript* (1997).

[36] R Schoen. "The harmonic mean as the basis of a realistic two-sex marriage model." In: *Demography* 18.2 (1981), pp. 201–216.

[37] Lloyd S Shapley and Martin Shubik. "The Assignment game I: The core". In: *International Journal of Game Theory* 1.1 (1971), pp. 111–130.

[38] Robert Shimer and Lones Smith. "Assortative matching and search". In: *Econometrica* 68.2 (2000), pp. 343–369.

[39] Aloysius Siow. "How does the marriage market clear? An empirical framework". In: *Canadian Journal of Economics* 41.4 (2008), pp. 1121–1155.

[40] Che-Lin Su and Kenneth L. Judd. "Constrained optimization approaches to estimation of structural models". In: *Econometrica* 80.5 (2012), pp. 2213–2230.

A Proofs

A.1 Proof of Theorem 1

Proof. Part (i). Proof of Existence. The proof of existence relies on Algorithm 1. See Galichon, Kominers, and Weber (2019) for a full proof.

Part (ii). Proof of Uniqueness. The proof of uniqueness relies on Berry, Gandhi, and Haile (2013). First, introduce the quantities $u_x = \mu_{x0}$ and $u_y = -\mu_{0y}$, and construct the full vector $u = (\{u_x\}_{x \in \mathcal{X}}, \{u_y\}_{y \in \mathcal{Y}})$. Solving the system of equations (2) is equivalent to finding the root of the following system

$$\begin{aligned} \sigma_x(u) &= u_x + \sum_y M_{xy}(u_x, -u_y) - n_x \\ \sigma_y(u) &= u_y - \sum_x M_{xy}(u_x, -u_y) + m_y \end{aligned} \tag{20}$$

Assumption 1 in Berry, Gandhi, and Haile (2013) is satisfied as σ is defined over the Cartesian product of intervals $\Pi_{x \in \mathcal{X}}[0, n_x]\Pi_{y \in \mathcal{Y}}[-m_y, 0]$. Finally, introduce $\sigma_0(u) = 1 - \sum_{x \in \mathcal{X}} \sigma_x(u) - \sum_{y \in \mathcal{Y}} \sigma_y(u)$.

Note that $\forall z' \neq z$, $\sigma_z(u)$ is weakly decreasing in $u_{z'}$ from the weakly isotony of M, and that $\sigma_0(u)$ is strictly decreasing in u_z $\forall z \in \mathcal{X} \cup \mathcal{Y}$. Indeed,

$$\sigma_0(u) = 1 + \sum_x n_x - \sum_y m_y - \sum_x u_x - \sum_y u_y$$

which is strictly decreasing in any entry of the vector u. We can conclude that Assumption 2 from Berry, Gandhi, and Haile (2013) is satisfied, as well as Assumption 3 by application of Lemma 1. Hence, σ is inverse isotone, which provides uniqueness. Indeed, assume that $\sigma(u) = \sigma(u')$, so that $\sigma(u) \leq \sigma(u')$ and $\sigma(u) \geq \sigma(u')$ which implies by inverse isotony that $u \leq u'$ and $u \geq u'$, hence $u = u'$. Therefore, there is a unique root u^* to the system of equation (20). Hence, there is a unique solution to system (2), with $\mu^*_{x0} = u^*_x$ and $\mu^*_{0y} = -u^*_y$. QED. ∎

A.2 Proof of Theorem 2

Proof. The expression for $\partial_{\theta^k}\Pi_{xy}$ follows immediately from the fact that $\Pi_{xy} = \frac{\mu_{xy}}{1'\mu_{xy}}$ and that $1'\mu_{xy} = \sum_{x\in\mathcal{X}} n_x + \sum_{y\in\mathcal{Y}} m_y - \sum_{xy\in\mathcal{XY}} \mu_{xy}$.

By the Implicit Function Theorem in (2), one has

$$\begin{aligned}
\partial_{\theta^k}\mu_{x0} + \sum_{y\in\mathcal{Y}} (\partial_{\theta^k}\mu_{x0}\partial_{\mu_{x0}}M_{xy} + \partial_{\theta^k}\mu_{0y}\partial_{\mu_{0y}}M_{xy}) &= -\sum_y \partial_{\theta^k}M_{xy} \\
\partial_{\theta^k}\mu_{0y} + \sum_{x\in\mathcal{X}} (\partial_{\theta^k}\mu_{x0}\partial_{\mu_{x0}}M_{xy} + \partial_{\theta^k}\mu_{0y}\partial_{\mu_{0y}}M_{xy}) &= -\sum_x \partial_{\theta^k}M_{xy}
\end{aligned}$$

which can be written using the expression of Δ (12) as

$$\Delta \begin{pmatrix} \partial_{\theta^k}\mu_{x0} \\ \partial_{\theta^k}\mu_{0y} \end{pmatrix} = \begin{pmatrix} c^k \\ d^k \end{pmatrix}$$

and Δ being a strictly diagonally dominant matrix, is invertible, QED. ∎

A.3 Proof of Theorem 3

Proof. (i) First, recall that in the ITU-logit model, the aggregate matching function is given by $M_{xy}(\mu_{x0}, \mu_{0y}) = \exp\left(-D_{xy}(-\log(\mu_{x0}), -\log(\mu_{0y})\right)$ where the distance function D_{xy} is defined by $D_{xy}(u_x, v_y) = \min\{z \in \mathbb{R} : (u_x - z, v_y - z) \in \mathcal{F}_{xy}\}$ and where $\mathcal{F}_{xy}$ is the bargaining set for the xy pair. By definition, $D_{xy}(u_x + a, v_y + a) = D_{xy}(u_x, v_y) + a$, which implies $M_{xy}(\lambda\mu_{x0}, \lambda\mu_{0y}) = \lambda M_{xy}(\mu_{x0}, \mu_{0y})$.

(ii) Second, let us introduce the mapping $D_{xy}(u,v) = -\log M_{xy}\left(e^{-u}, e^{-v}\right)$ (in the following, we will drop the indices for convenience). We will show that D is the distance function associated with some proper bargaining set (see GKW for a definition) and that M is the associated aggregate matching function when the idiosyncratic component of individual payoffs is logit.

Step 1. We begin by constructing the bargaining set $\mathcal{F}$ as follows

$$\mathcal{F} = (u,v) \in \mathbb{R}^2 : D(u,v) \leq 0$$

Step 2. Let us show that the set $\mathcal{F}$ is a proper bargaining set. First, note that assumption 1 does not ensure that $\mathcal{F}$ is non-empty. However, this is not much of a concern in our setting: there will simply be no match between the two corresponding individuals in equilibrium (only a mild additional assumption is required to obtain non-emptiness: $M_{xy}(\mu_{x0}, \mu_{0y})$ is bounded below by 1 as μ_{x0} (μ_{0y}) approaches infinity while μ_{0y} (μ_{x0}) is bounded below by 0 ; as a matter of fact, it is satisfied on all of our introductory examples). Closedness follows from the continuity of M by assumption 1 -(i). From assumption 1 -(ii), we can deduce that $\mathcal{F}$ is lower comprehensive. Indeed, assume that $(u,v) \in \mathcal{F}$. By construction, $D(u,v) \leq 0$. Take (u', v') with $u' \leq u$ and $v' \leq v$. By weak isotonicity of M, we have $D(u', v') \leq D(u,v) \leq 0$, hence $(u', v') \in \mathcal{F}$. Finally, we can show that $\mathcal{F}$ is bounded above. Indeed, assume $u_n \to +\infty$ and v_n bounded below, then for n large enough, $M(u_n, v_n) < 1$ by assumption 1-(iii), so that $D(u_n, v_n) > 0$, that is $(u_n, v_n) \notin \mathcal{F}$ (the same reasoning applied with $v_n \to +\infty$ and u_n bounded below).

Step 3. Let us now show that D is the distance function associated with the bargaining set $\mathcal{F}$. The distance from the point (u,v) to the frontier is the value z such that $(u-z, v-z)$ belongs to the frontier of the bargaining set. By construction, $D(u-z, v-z) = 0$ but homogeneity of degree one implies that $D(u+a, v+a) = D(u,v) + a$, therefore $z = D(u,v)$.

Step 4. GKW showed that in equilibrium, individuals receive a payoff that is the sum of two components: a systematic component that depends on the observable characteristics of the partners, denoted respectively for men and women, U_{xy} and V_{xy} ; and an idiosyncratic component ϵ_{iy} and η_{xj}. Assume that when a man of type x meet with a woman of type y, they decide upon a utility wedge w and receive $U_{xy} = -D_{xy}(0, -w)$ and $V_{xy} = -D_{xy}(w, 0)$. Note that the functions U and V, as defined here, explicitly represent the bargaining frontier. Assuming logit heterogeneities, the systematic component of utility can be recovered from the marriage patterns by the usual formulas $U_{xy} = \log \frac{\mu_{xy}}{\mu_{x0}}$ and $V_{xy} = \log \frac{\mu_{xy}}{\mu_{0y}}$. The equilibrium condition in GKW is $D_{xy}(U_{xy}, V_{xy}) = 0$, that is $D_{xy}(\log \mu_{xy} - \log \mu_{x0}, \log \mu_{xy} - \log \mu_{0y}) = 0$ which yields to the aggregate matching function

$$M_{xy}(\mu_{x0}, \mu_{0y}) = \exp\left(-D_{xy}\left(-\log \mu_{x0}, -\log \mu_{0y}\right)\right) = \mu_{xy}$$

This concludes the proof. ∎

A.4 Proof of Theorem 4

In the following, $\Pi\left(\theta, \zeta\right)$ denotes the predicted frequencies given θ and the frequencies ζ. We also introduce the rescaling operator $\Pi^{\theta}(\pi_0) = \frac{M^{\theta}(\pi_0)}{1' M^{\theta}(\pi_0)}$, and note that$\Pi\left(\theta, \zeta\right) = \Pi^{\theta} \circ \left(AM^{\theta}\right)^{-1}(\zeta)$.

Part (i). A first expression. In the maximization of the log-likelihood, the first order conditions with respect to θ are $\hat{\pi}' \partial_{\theta} \ln \Pi\left(\theta, A\hat{\pi}\right) = 0$ that we denote $F\left(\hat{\theta}, \hat{\pi}\right) = 0$. From a serie of Taylor expansions around the true value of θ and π, we can then deduce that $\left(\hat{\theta} - \theta\right) = -\left(D_{\theta}F\right)^{-1}\left(D_{\pi}F\right)\left(\hat{\pi} - \pi\right)$ where we use the notation D for the Jacobian matrix. Note that asymptotically, $N^{1/2}\left(\hat{\pi} - \pi\right) \sim \mathcal{N}\left(0, V_{\pi}\right)$ where $V_{\pi} = diag\left(\pi\right) - \pi\pi'$. Hence, it follows that

$$N^{1/2}\left(\hat{\theta} - \theta\right) \Rightarrow \mathcal{N}\left(0, V_{\theta}\right)$$

where $V_{\theta} = \left(D_{\theta}F\right)^{-1}\left(D_{\pi}F\right)V_{\pi}\left(D_{\pi}F\right)'\left(\left(D_{\theta}F\right)'\right)^{-1}$.

Part (ii). Analytic expressions for each component. Let us begin with $D_{\theta}F$. Note that by definition, we have $\Pi' 1 = 1$. Hence, $\Pi' \partial_{\theta^i} \log \Pi = 0$ and $\Pi' \partial^2_{\theta^i \theta^j} \log \Pi + \partial_{\theta^j} \Pi' \partial_{\theta^i} \log \Pi = 0$.

Finally, we get$\pi'\partial^2_{\theta^i\theta^j}\log\Pi = -\pi'\partial_{\theta^i}\log\Pi'\partial_{\theta^j}\log\Pi$ so

$$D_\theta F = -D_\theta \log\Pi' diag(\pi) D_\theta \log\Pi$$

and $D_\theta \log\Pi$ can be obtained from our results on the gradient of the log-likelihood.

We may now turn to $D_\pi F$. We have

$$D_\pi F := (D_\theta \log\Pi)' - (D_\theta \log\Pi)' (D_\zeta \Pi) A$$

where, as before, $D_\theta \log\Pi$ appears in Theorem 2. We obtain $D_\zeta\Pi$ as

$$D_\zeta \Pi = \left(D_{\pi_0}\Pi^\theta\right)\left(\left(D_{\pi_0} A M^\theta\right)^{-1}\right)$$

The expressions above allow us to prove the announced formula for V_θ. We have

$$\begin{aligned}
V_\theta &= (D_\theta F)^{-1}(D_\pi F) V_\pi (D_\pi F)' \left((D_\theta F)'\right)^{-1} \\
&= T^{-1}(D_\pi F) V_\pi (D_\pi F)' (T')^{-1} \\
&= T^{-1}\left((D_\theta \log\Pi)' - (D_\theta \log\Pi)'(D_\zeta\Pi) A\right) V_\pi \left((D_\theta \log\Pi)' - (D_\theta \log\Pi)'(D_\zeta\Pi) A\right)' (T')^{-1} \\
&= T^{-1}\left(-(D_\theta \log\Pi)' + (D_\theta \log\Pi)'(D_\zeta\Pi) A\right) V_\pi \left(A'(D_\zeta\Pi)'(D_\theta \log\Pi) - (D_\theta \log\Pi)\right) (T')^{-1} \\
&= T^{-1}(D_\theta \log\Pi)'(D_\zeta\Pi) A V_\pi A' (D_\zeta\Pi)'(D_\theta \log\Pi)(T')^{-1} + H
\end{aligned}$$

where $H = G + G' + T^{-1}(D_\theta \log\Pi)' V_\pi (D_\theta \log\Pi)(T')^{-1}$ and is composed of three terms, and two of them are symmetric. Let us start with these symmetric terms:

$$G = T^{-1}(D_\theta \log\Pi)' V_\pi A' (D_\zeta\Pi)'(D_\theta \log\Pi)(T')^{-1}$$

this is

$$T^{-1}(D_\theta \log\Pi)' A' diag(\pi)(D_\zeta\Pi)'(D_\theta \log\Pi)(T')^{-1} - T^{-1}(D_\theta \log\Pi)' \pi\pi' A' (D_\zeta\Pi)'(D_\theta \log\Pi)(T')^{-1}.$$

The first term in this sum is 0 because $(D_\theta \log \Pi)' A' = 0$. The second term is

$$\begin{aligned} & T^{-1}(D_\theta \Pi)' diag(\mu)^{-1} \mu \zeta' (D_\zeta \Pi)' (D_\theta \log \Pi)(T')^{-1} \\ = \; & T^{-1}(D_\theta \Pi)' 1 \zeta' (D_\zeta \Pi)' (D_\theta \log \Pi)(T')^{-1} \end{aligned}$$

where $(D_\theta \Pi)' 1 = 0$. Hence,

$$\begin{aligned} H &= T^{-1}(D_\theta \log \Pi)' V_\pi (D_\theta \log \Pi)(T')^{-1} \\ &= (T')^{-1} \end{aligned}$$

We obtain

$$V_\theta = (\mathcal{I}_{11})^{-1} + \mathcal{I}_{11}^{-1}\mathcal{I}_{12} A V_\pi A' \mathcal{I}_{12}' \mathcal{I}_{11}^{-1}$$

where $\mathcal{I}_{11} = -(D_\theta \log \Pi)' diag(\pi)(D_\theta \log \Pi)$ and $\mathcal{I}_{12} = (D_\theta \log \Pi)' diag(\pi)(D_\zeta \log \Pi)$. This concludes the proof.

A.5 Proof of Theorem 5

Proof. Part (i). Proof of Existence. The proof of existence relies on a revised procedure based on Algorithm 1, which is stated below. See Galichon, Kominers, and Weber (2019) for a full proof.

Algorithm 2. *The revised Algorithm works as follow*

Step 0	*Fix the initial value of* $\hat{\mu}_{0j}$ *at* $\hat{\mu}_{0j}^0 = \hat{f}_j/q_{0j}$ *for all* $j \in \mathcal{J}$ *and* $\hat{\mu}_{i0}$ *at* $\hat{\mu}_{i0}^0 = \hat{m}_i/p_{i0}$ *for all* $i \in \mathcal{I}$.
Step $2t+1$	*Keep* $\hat{\mu}_{0j}^{2t}$ *for all* $j \in \mathcal{J}$ *fixed. Solve* $\hat{\mu}_{i0}^{2t+1}$ *of* $\hat{\mu}_{i0}$ *for all* $i \in I$ *sequentially such that the equality* $\hat{m}_y = q_{0y}\hat{\mu}_{0y} + \sum_{x\in\mathcal{X}} q_{xy} \cdot m(\hat{\mu}_{x0}, \hat{\mu}_{0y})$*, holds.*
Step $2t+2$	*Keep* $\hat{\mu}_{i0}^{2t+1}$ *for all* $i \in \mathcal{I}$ *fixed. Solve* $\hat{\mu}_{0j}^{2t+2}$ *of* $\hat{\mu}_{0j}$ *for all* $j \in J$ *sequentially such that the equality* $\hat{n}_x = p_{x0}\hat{\mu}_{x0} + \sum_{y\in\mathcal{Y}} p_{xy} \cdot m(\hat{\mu}_{x0}, \hat{\mu}_{0y})$*, holds.*

The algorithm terminates when $\sup_y |\mu_{0y}^{2t+2} - \mu_{0y}^{2t}| < \epsilon$.

Part (ii). Proof of Uniqueness. The proof of uniqueness relies on Berry, Gandhi, and

Haile (2013) similarly as it in the proof of Theorem 1. ∎

B Online Appendix

B.1 Simulations

In this section, we conduct simulations to investigate the numerical performance of the nested and MPEC approaches for maximum likelihood estimation. In the nested approach, it is crucial (i) to be able to solve system (2) and (ii) to do so in an efficient way, for the sake of minimizing computation time. Therefore, we first investigate the performance of the IPFP algorithm and Newton Descent method for solving the system (2) before turning our attention to the nested and MPEC approaches.

B.1.1 Solving system (2) for $(\mu^{\theta}_{x0}, \mu^{\theta}_{0y})$

Theorem 1 and Algorithm 1 address both capability and efficiency concerns of solving system (2). However in practice, Algorithm 1 is not necessarily the most efficient way to solve for $(\mu^{\theta}_{x0}, \mu^{\theta}_{0y})$. When the Jacobian of system (2) is known, it can be solved very efficiently using Newton descent methods, which we recall below.

Algorithm 3. *Rewrite the system of nonlinear equations (2) as* $\sigma(\mu_0) = 0$*, where* $\mu_0 = (\mu_{x0}, \mu_{0y})$*. The Newton's Descent method works as follows*

Step 0	*Fix the initial value of* μ_0 *at* $\mu_0^0 = (n', m')'$.
Step t	*Given* μ_0^{t-1}*, solve* $J\sigma(\mu_0^{t-1})\delta = -\sigma(\mu_0^{t-1})$*, where* $J\sigma(\mu_0^{t-1})$ *is the Jacobian matrix at* μ_0^{t-1}*. Update* $\mu_0^t = \mu_0^{t-1} + \delta$

The algorithm terminates when $\sup_y |\mu_0^{t+1} - \mu_0^t| < \epsilon$.

To benchmark these different methods, we consider the Exponentially Transferable Utility model in GKW. The aggregate matching function is given by

$$M_{xy}(\mu_{x0}, \mu_{0y}) = \exp(-D_{xy}(-\log \mu_{x0}, -\log \mu_{0y}))$$

where $D_{xy}(u, v) = \tau_{xy} \log\left((\exp((u - \alpha_{xy})/\tau_{xy}) + \exp((v - \alpha_{xy})/\tau_{xy}))/2\right)$. We draw the types x and y from two uniform distributions, assume that $\alpha_{xy} = xy$ and $\gamma_{xy} = xy$, and fix $\tau_{xy} = \tau = 1, \forall xy \in \mathcal{XY}$. In the experiment, we vary $|\mathcal{X}|$, the number of types on the men side of the market, and fix $|\mathcal{Y}| = 1.5|\mathcal{X}|$. Finally, we assume that $n_x = m_y = 1$, $\forall x \in \mathcal{X}$, $\forall y \in \mathcal{Y}$.

Table 6 below summarizes the number of iteration and computation time averaged over 50 replications of the experiment for three numerical methods: the IPFP algorithm described in Algorithm 1, its parallelized version, and the Newton's method described in Algorithm 3.

Table 6: IPFP and Newton method

	IPFP		par. IPFP		Newton	
Mkt. Size	Iter.	Time	Iter.	Time	Iter.	Time
10	14.64	0.36	14.64	1.45	26.42	0.01
50	9.2	1.17	9.2	1.12	27.08	0.09
100	8	2.16	8	1.33	23.22	0.29
200	7	4.35	7	2.43	23.62	1.15
300	7	7.21	7	3.99	25.9	3.08
500	6	13.54	6	7.4	31.38	12.12
1000	6	40.92	6	18.32	37.56	81.53
2000	5	106.93	5	47.77	24.02	680.16
5000	5	497.02	5	199.14	26	7934.81

This table raises three comments. First, to improve the computational efficiency of the Newton's Descent method, we provide the analytic expression of the Jacobian matrix of system (2). Such analytic expression is not always available, in which case the Jacobian must be approximated numerically, which will greatly increase computation time for this method (at least for large markets). Second, there is no guarantee of convergence when using the Newton's Descent Algorithm. We notice no such issues when performing this simple experiment, but nonconvergence may well be an issue with more complex models. For these two reasons, and since Newton's method performs only better for smaller market, this method is not our preferred algorithm. Furthermore, it should be added that the IPFP Algorithm is very suitable for parallel computing. The gains are negative for small markets, but as market size grows, we manage to reduce computation time by a factor of two. The parallel IPFP runs on four processors (which is what is currently available on most high-end personal computers). This suggests that performance could be further improved when running on computing clusters.

B.1.2 Estimation

We test the numerical performance of our maximum likelihood estimator, using the nested and MPEC approaches. The setup of the experiments remain the same as before, but we assume $\alpha_{xy} = \alpha \times x \times y$ and $\gamma_{xy} = \gamma \times x \times y$, where α and γ are arbitrarily chosen. Given $\theta_0 = (\alpha, \gamma)$, we compute the equilibrium matching μ^{θ_0} using the IPFP algorithm and set $\hat{\mu} = \mu^{\theta_0}$. Then, we test if we are able to recover θ_0 from the observed $\hat{\mu}$ using our maximum likelihood estimator in this correctly specified case. The results are reported in Table 7 below.

Table 7: Estimation

	Nested Approach			MPEC		
Mkt. Size	Iter.	Time	% Failure	Iter.	Time	% Failure
10	18.64	5.4	6	15.91	5.1	8
50	21.6	21.12	0	16.2	8.01	20
100	23.34	42.05	0	18.05	17.58	22
200	26.98	101.67	0	36.1	102.56	16
300	26.8	179.25	0	20.64	121.44	16
500	25.67	343.16	4	27.27	549.76	26

First, a word of caution, the nested approach we implement here relies on a simple version of the IPFP algorithm, so its performance can be further improved using the parallel version. It is difficult to interpret the results in Table 7. The MPEC algorithm seems to perform better for small market sizes as it converges faster to the correct value of θ. For larger markets, however, the IPFP approach does better in some cases, for example with 100 or 500 men. Note that the number of iterations is relatively similar across methods, but performing one iteration can be computationally burdensome in the MPEC case. Indeed, the nested approach only requires solving for the equilibrium matching using the IPFP algorithm and computing the gradient as in Theorem 2. The MPEC approach, on the other hand, requires the computation of the Jacobian matrix in equation (13), which involves second order and cross partial derivatives. Although we do have analytic expressions for these components, it can still be cumbersome to compute due to the size of these objects. For example, in the case with 500 men, 750 women and two parameters, the Jacobian matrix is a 2502×2502 matrix. Finally, Table 7 illustrates a common issue with Newton-like methods, that is, non-convergence.

B.2 Data Construction

As discussed in the main text, we require three data inputs to implement our approach:

i) the number of available single men and women by age and education had the financial aid program not been eliminated, (n_x, m_y), for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$,

ii) the number of available single men and women by age and education as a result of the elimination of the financial aid program, (n_x, m_y), for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, and

iii) the flow of new marriages by age and education as a result of the elimination of the financial aid program, μ'.

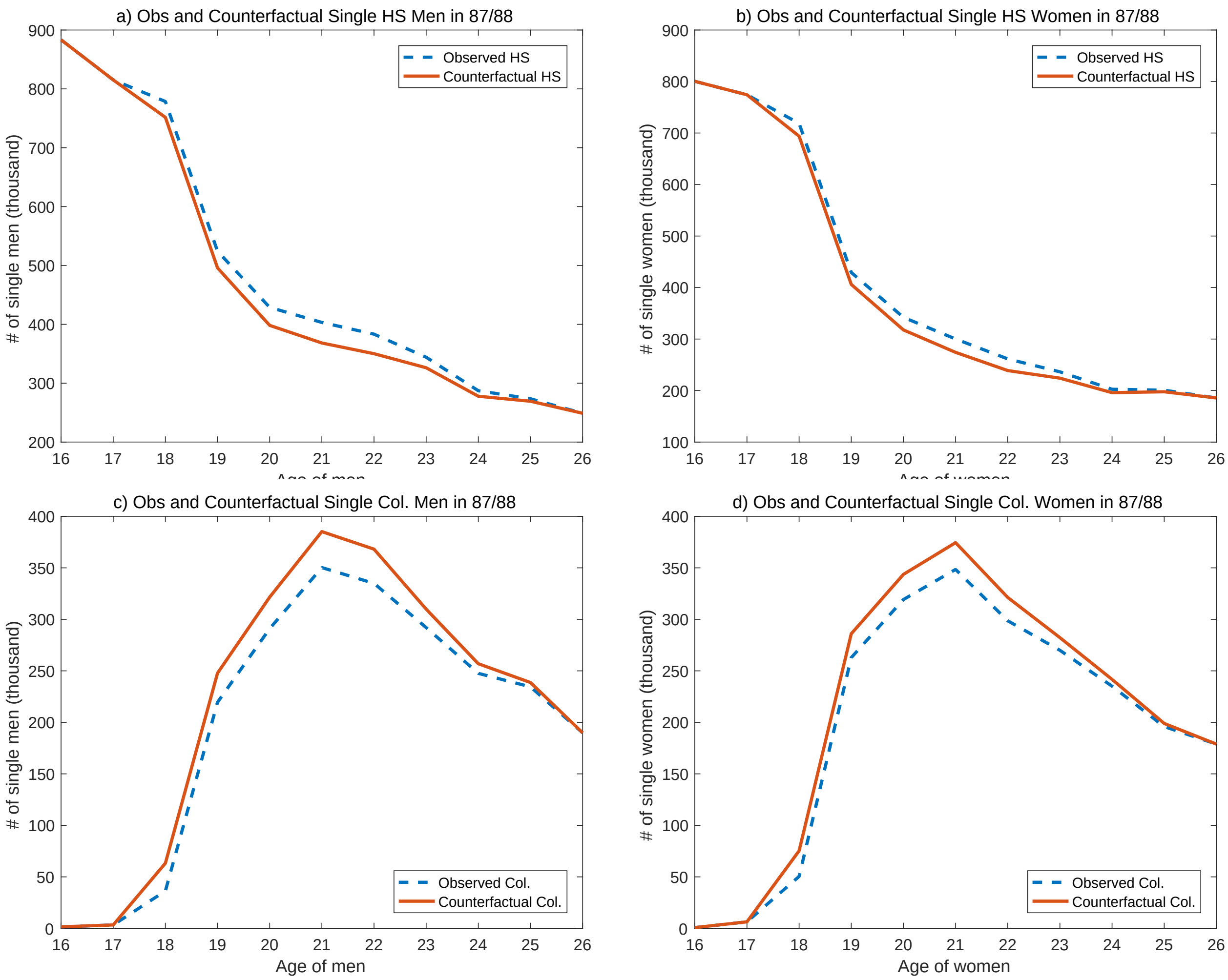
The flow of new marriages as a result of the policy change μ' (item iii) above) is collected from the Vital Statistics in 1987/88 obtained from the National Bureau of Economic Research data website. Vital Statistics recorded the education and age of married couples until 1988 for 22 reporting states. The 22 reporting states are California, Connecticut, Hawaii, Illinois, Kansas, Kentucky, Louisiana, Maine, Mississippi, Missouri, Montana, Nebraska, New Hampshire, New York, North Carolina, Rhode Island, Tennessee, Utah, Vermont, Virginia, Wisconsin, Wyoming.

The number of available single men and women by age and education after the policy change, (n_x, m_y) (item ii) above are constructed from the Integrated Public-Use Microdata (IPUMS hereafter) files of the U.S. CPS data. The sample used in this study is the monthly data in 1986. In order for the CPS data to match the marriage data from the Vital Statistics, our sample comprises only of individuals from the 22 states reporting states. We take an average of 12 monthly CPS surveys for the 22 matching states in 1986 to construct the yearly available population vectors.

The age range studied is between 16 and 75 years old. The education information is obtained from the variable "`EDUC`" in the US CPS data. The education attainment is divided into three groups: high school graduate or less, some years of college or college graduate, more than college. There are 180 possible age-education combinations from the 60 age groups and 3 education levels. We exclude 5 groups - these are individuals who are less than 23 years of age with more than college education. This leaves us with 173 types of men and women. The variable "`marst`" in IPUMS CPS data provides us with marital status information. It distinguishes an individual either married, separated, divorced, widowed or never married/single. We consider separated, divorced, widowed, never married/single individuals as available single individuals in the marriage market and calculated the number of available men and women for each type by adding the weight from each sample in the dataset.

B.3 Additional figures

Figure 4: Changes in available HS and Col single men and women



These figures graph the observed and counterfactual available single men and women by ages between 16 and 22 in 1986. In the counterfactual scenario where the aid program was not eliminated, we expect there to be more available single college graduates and fewer available single high school graduates. As expected, our calculations create a wedge between the observed and counterfactual number of available single individuals between the ages of eighteen and twenty five.

Figure 5: Estimated Φ_{xy} of college men

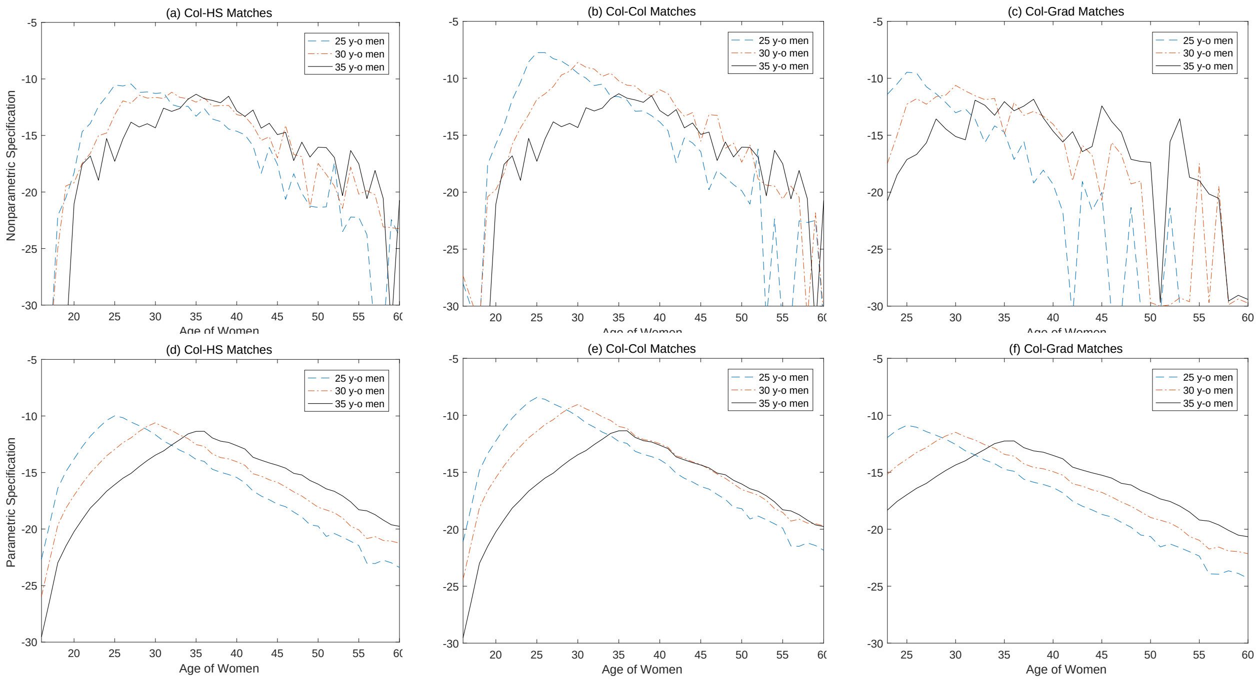


Figure 6: Estimated Φ_{xy} of graduate men

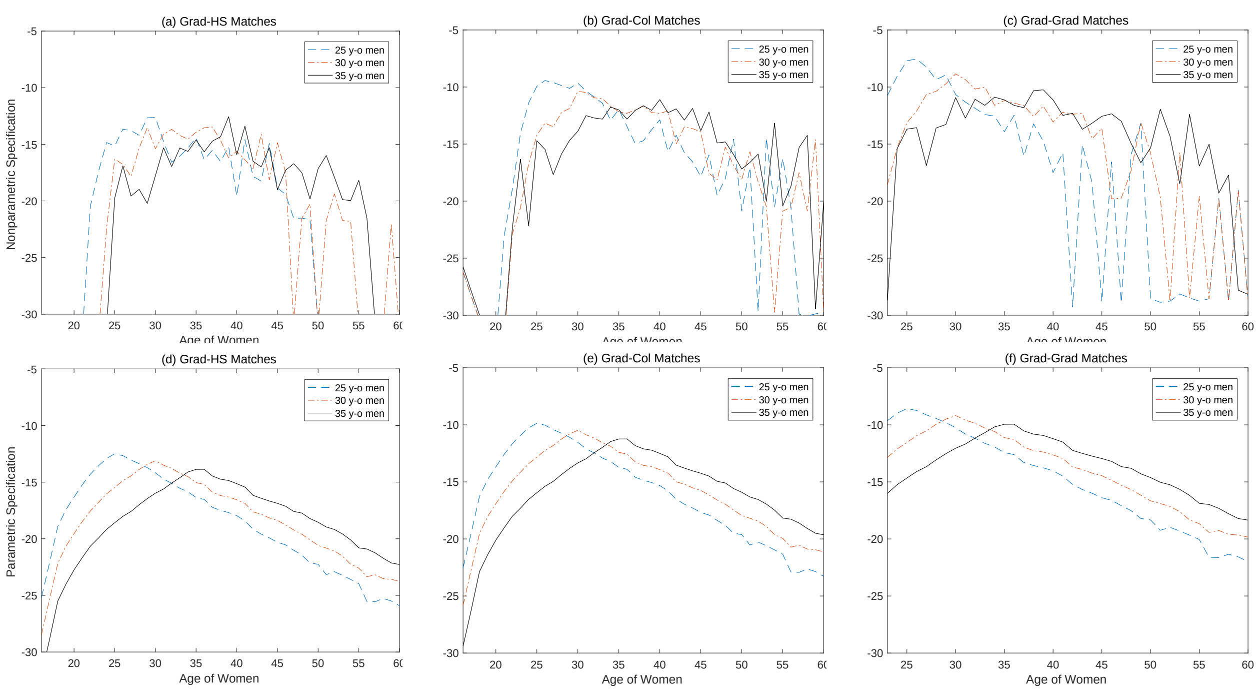
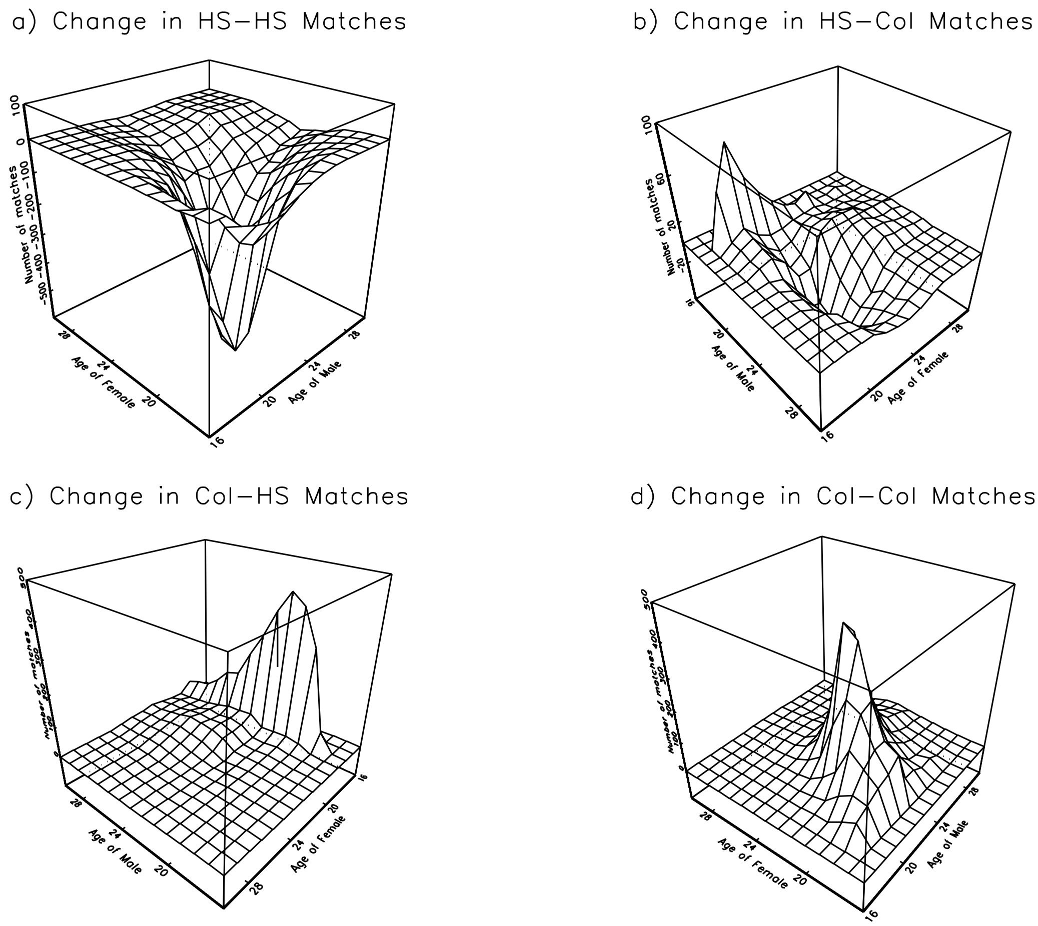


Figure 7: Changes in to marriage distribution by education and age



These figures graph the changes in the marriage distribution for the four education pairs by age estimated from the nonparametric specification. Figure 7a and 7d show the changes for HS-HS and Col-Col matches, respectively. Consistent with the numbers reported in Table 5, Figure 7a shows that continuing the Social Security Benefit Program would have decreased the number of HS-HS marriages for all age pairs. Strong positive assortative matching by age also means that the biggest decrease is experienced by similarly aged young couples. The decrease becomes smaller as the age gap between husbands and wives increases. We see a similar but opposite effect amongst Col-Col marriages. The continuation of the financial aid program would have increased the number of new Col-Col marriages in 1987/88. The biggest increase would occur among similarly aged young couples. Figure 7b graphs the changes for new HS-Col marriages. Unlike the changes in HS-HS and Col-Col marriages, the CS model predicts that the increase in the number of single Col individuals and the decrease in the number of HS individuals would benefit young men and disadvantage older men with HS qualification. More specifically, marriages for young men between the ages of 18 to 21 with HS qualification would increase. However the model also predicts that marriages for men with HS qualification, older than 21 years old would fall. As for changes in new Col-HS marriages, Figure 7c suggests an overall increase for all age pairs. The biggest increase is experienced by young single women with HS qualifications.