Partial identification in matching models for the marriage market*

Cristina Gualdani[†] Shruti Sinha[‡]

December 2020

Abstract

We study partial identification of the preference parameters in models of one-to-one matching with perfectly transferable utilities, without imposing parametric distributional restrictions on the unobserved heterogeneity and with data on one large market. We provide a tractable characterisation of the identified set, under various classes of nonparametric distributional assumptions on the unobserved heterogeneity. Using our methodology, we re-examine some of the relevant questions in the empirical literature on the marriage market which have been previously studied under the Multinomial Logit assumption.

Keywords: One-to-One Matching, Marriage Market, Transfers, Stability, Partial Identification, Nonparametric Identification, Linear Programming, Econometrics.

*We would like to thank Pierre-André Chiappori and Thierry Magnac for their useful comments. We would also like to thank the seminar and conference participants at Toulouse School of Economics, the University of Warwick, IAAE 2019, EEA-ESEM 2019, the Bristol-TSE Econometrics Workshop, the 2019 Cowles Foundation conference on Matching: Optimal Transport and Beyond, and the 2019 Network Econometrics Juniors' Conference at Northwestern University. Rossi Abi-Rafeh provided excellent research assistance. We acknowledge funding from the French National Research Agency (ANR) under the Investments for the Future (Investissements d'Avenir) program, grant ANR-17-EURE-0010. An earlier version of the paper was previously circulated under the title "Partial Identification and Inference in One-to-One Matching Models with Transfers".

†Email: cristina.gualdani@tse-fr.eu, Toulouse School of Economics, University of Toulouse Capitole, Toulouse, France.

‡Email: shruti.sinha@tse-fr.eu, Toulouse School of Economics, University of Toulouse Capitole, Toulouse, France.

1 Introduction

Matching markets are two-sided markets, where agents on each side of the market have preferences over matching with agents on the other side. For example, social interactions lead individuals to find marital partners, production tasks are assigned to workers, and auctions sort buyers with sellers. While the economic theory of matching models has been around for more than five decades, the literature on empirical matching models is relatively recent (see Chiappori and Salanié, 2016, for a review).

An important strand of this literature focuses on the one-to-one matching model, in which every agent forms at most one match. Each possible match generates a surplus (hereafter, match surplus). In the framework where utilities are perfectly transferable, agents can share the match surplus with their partners without frictions. Since Becker (1973), the one-to-one matching model with perfectly transferable utilities (hereafter, 1to1TU) has been extensively used in household economics to represent the marriage market (see Chiappori, 2017, for a review).[1] In particular, researchers have exploited the 1to1TU model to estimate the systematic part of the match surplus. Recovering the systematic match surplus is useful in itself, for example, to investigate sorting patterns and how they change over time, to learn about the complementarities and substitutabilities of partner characteristics, to assess the efficiency and welfare implications of the status-quo assignment, and to measure the impact of pre-marital decisions on the match surplus sharing rule.

In Section 2, we report a survey of the most relevant empirical papers that use the 1to1TU model to represent the marriage market. All such papers proceed under strong parametric distributional restrictions on the unobserved heterogeneity. These restrictions amount to imposing i.i.d. standard Extreme Value Type I taste shocks, independently distributed from covariates. These restrictions make the 1to1TU model *just* identified with data on *one large* market and allow one to *point* identify the systematic match surplus via standard Multinomial Logit formulas (Choo and Siow, 2006).

The motivation for using the Multinomial Logit 1to1TU model is computational simplicity. However, this framework may lead to paradoxes contrasting with economic sense. For example, it is well known that the one-sided Multinomial Logit model is inherently linked to the IIA axiom and severely restricts cross-elasticities. The same holds in two-sided markets and causes undesirable comparative static predictions, as explained in Graham (2013a) and Galichon and Salanié (2019). Further, independence of taste shocks from covariates can induce biases if there is underlying endogeneity or conditional heteroskedasticity.

The fact that widespread empirical practices rest on the Multinomial Logit 1to1TU model raises a number of questions. Does the 1to1TU model retain any identifying power on the sys-

[1] The 1to1TU model has also been used to study matching of CEOs to firms (Chen, 2017), matching of academics to offices (Baccara, İmrohoroğlu, Wilson, and Yariv, 2012), merging of banks (Akkus, Cookson, and Hortaçsu, 2016), formation of research alliances (Mindruda, Moeen, and Agarwal, 2016), and collaboration between academics and firms (Mindruda, 2013; Banal-Estañol, Macho-Stadler, and Pérez-Castrillo, 2018).

tematic match surplus without restrictions on the taste shock distribution, when we have data on one large market? If not, is it still possible to recover some information on the systematic match surplus under nonparametric assumptions on the taste shock distribution? How are the answers to relevant policy questions driven by the parametric assumptions typically imposed on the taste shock distribution? The contribution of our paper is to address these issues and, by doing so, offer methodological guidance for researchers who wish to consider more robust alternatives to (or, do sensitivity checks of) the Multinomial Logit 1to1TU model.

In Section 4, we start by observing that, if the taste shock distribution is not assumed to be *fully* known by the researcher, then the 1to1TU model is *under*-identified with data on one large market (Galichon and Salanie, 2020). We show that, in the absence of any restriction on the taste shock distribution, the under-identification issue is severe, as the 1to1TU model is completely uninformative about the systematic match surplus. Formally, this means that, for every possible value of the systematic match surplus, there exists a taste shock distribution which when combined with that value of the systematic match surplus, rationalises the data.

We proceed by investigating whether the 1to1TU model retains any information on the systematic match surplus under various classes of nonparametric assumptions on the taste shock distribution (for instance, independence of taste shocks from covariates, quantile restrictions, symmetry restrictions, and identically distributed marginals). Answering this question poses the challenge of tractably characterising the identified set of the systematic match surplus. We do that by extending the linear programming computational approach of Torgovitsky (2019) (also known as PIES[2]) to our framework. For a given value of the systematic match surplus, this method changes the search over the space of infinite-dimensional cumulative distribution functions into a search over a space of cumulative distribution functions evaluated at a finite number of points. The latter search can be written as a simple linear programming problem. Further, note that the analyst would need to solve the linear programming problem for every admissible value of the systematic match surplus. Usually, in the partial identification literature, this is carried out by generating a grid of points to approximate the parameter space and then repeating the exercise of interest for each grid point. The difficulty of doing so increases with the size of the grid which, in the 1to1TU model, increases exponentially with the cardinality of the covariates' support, hence leading quickly to a computational bottleneck. We solve this issue by showing that the parameter space can be ex-ante partitioned into a finite number of subsets such that, for each subset, every value belonging to that subset gives rise to the same linear programming problem. Therefore, the analyst has to solve the linear programming problem only once for each subset.

In Section 6, we use our methodology to re-examine some of the relevant questions in the empirical literature on the marriage market that have been previously answered by relying on the Multinomial Logit 1to1TU model. A key question that stands out in this literature is whether educational sorting (i.e., the tendency of agents to marry someone with similar or, alternatively, very different education levels) has changed over time. Answering this question is important because

[2]That is, Partial Identification by Extending Subdistributions.

educational sorting may have a crucial impact on inequality by determining family formation and intergenerational transmission of human capital (Kremer, 1997; Fernández and Rogerson, 2001; Fernández, Guner, and Knowles, 2005; Heckman and Mosso, 2014; Dupuy and Weber, 2019; Eika, Mogstad, and Zafar, 2019; Chiappori, Costa-Dias, Crossman, and Meghir, 2020; Ciscato and Weber, 2020).

Assessing changes in educational sorting is a difficult task because it requires us to disentangle the effect of changes in the marginal probability distribution of education categories from potential structural changes in the match surplus. In fact, men and especially women have become more educated over time. Hence, people with similar education levels are mechanically more likely to marry. We thus want to detect changes in educational sorting that occur in excess of changes that would naturally occur because of such distributional variations in education.

The literature proposes two approaches to measure changes in educational sorting. The first amounts to using indices of sorting which are based on comparing the empirical match probabilities to a counterfactual world where matching happens randomly (Fernández and Rogerson, 2001; Greenwood, Guner, and Kocharkov, 2003; Liu and Lu, 2006; Greenwood, et al., 2014; Abbott, et al., 2019; Eika, Mogstad, and Zafar, 2019; Shen, 2019). The second amounts to using a structural model of the marriage market, such as the 1to1TU model, to estimate individual preferences and analyse how they change over time. The second approach has been implemented by Siow (2015), Chiappori, Salanié, and Weiss (2017), and Chiappori, Costa-Dias, and Meghir (2020), based on the Multinomial Logit 1to1TU model.

Both approaches suggest that, on average, positive educational sorting has increased in the U.S. in the past decades. However, there is some debate around this trend when we look closer at each education category. For instance, Eika, Mogstad, and Zafar (2019) find that positive educational sorting has declined among the highly educated and increased among the less educated. Instead, Chiappori, Salanié, and Weiss (2017) find that positive educational sorting has increased in each education category, particularly among college graduates. Similar results are obtained by Chiappori, Costa-Dias, and Meghir (2020). Note that while the conclusions achieved via the structural approach have the advantages of explicitly linking the observed marital patterns with the economic determinants of sorting, such findings may be driven by the Multinomial Logit structure. We thus use our methodology to assess whether those conclusions are robust and obtainable also under nonparametric assumptions on the taste shock distribution.

Using data from the American Community Survey between years 1940 and 1966, we implement our methodology under three alternative and increasingly restrictive classes of nonparametric assumptions on the taste shock distribution. We find that, up to year 1950, the 1to1TU model is completely uninformative about the direction of educational sorting. The 1to1TU model starts to unambiguously reveal the presence of positive educational sorting in year 1955 within some education categories. In year 1966, the presence of positive educational sorting is confirmed within every education category and seems especially evident within the most (least) educated fraction of the population. Therefore, the 1to1TU model is progressively more informative over time about the direction of educational sorting. However, we find no evidence that positive educational sorting has increased (or decreased) over time within each education category, under any of the three classes of distributional assumptions considered. This is because the estimated identified intervals for the relevant structural parameters remain always unbounded on the right-hand-side. This differs from the empirical findings of Chiappori, Salanié, and Weiss (2017), which are based on the same data. Our results suggest that those findings may be driven by the Multinomial Logit structure.

In Section 6, we consider another relevant question in the empirical literature on the marriage market and obtain conclusions similar to the above. In particular, as discussed by Chiappori, Iyigun, and Weiss (2009) and Chiappori, Salanié, and Weiss (2017), the increase in educational sorting makes a higher stock of human capital more valuable on the marriage market. As a consequence, they predict an increase in the expected maximum payoff an agent can receive in the marriage market due to achieving a college degree ("marital college premium"), especially among women. Their empirical findings corroborate such a prediction in the U.S., based on the Multinomial Logit 1to1TU model. By contrast, the estimated identified sets do not show any evidence that the marital college premium has increased over time, under any of the three classes of nonparametric distributional assumption considered. Further, the 1to1TU model seems to be particularly uninformative about the women's side.

In what follows, Section 2 reviews the literature, Section 3 introduces the model, Section 4 discusses identification, Section 5 presents simulations, Section 6 describes the empirical applications, and Section 7 concludes.

2 Literature review

This section reviews some related literature in addition to the references of Section 1.

Applications of the Multinomial Logit 1to1TU model to the marriage market The Multinomial Logit 1to1TU model has been introduced by Choo and Siow (2006) and since then has become popular in the empirical research on the marriage market. Several papers use it to learn whether matching preferences are positive assortative by age, education, geographical location, etc. (Choo and Siow, 2006; Botticini and Siow, 2011; Bruze, Svarer, and Weiss, 2015; Choo, 2015;[3] Siow, 2015; Galichon, Kominers, and Weber, 2019[4]). Other papers use it to assess which of the partner characteristics are complements/substitutes in the production of the systematic match surplus and their relative strengths (Dupuy and Galichon, 2014;[5] Ciscato, Galichon, Goussé, 2020).

[3] Bruze, Svarer, and Weiss (2015) and Choo (2015) incorporate dynamic aspects into the framework of Choo and Siow (2006).

[4] Galichon, Kominers, and Weber (2019) extend the framework of Choo and Siow (2006) to imperfectly transferable utilities.

[5] Dupuy and Galichon (2014) extend the framework of Choo and Siow (2006) to continuous covariates.

The Multinomial Logit 1to1TU model has been frequently used to investigate the evolution over time of the link between education levels and marriage market outcomes. In particular, the existing literature has studied questions like how educational sorting in the marriage market has changed over time (Siow, 2015; Chiappori, Salanié, and Weiss, 2017; Chiappori, Costa-Dias, and Meghir, 2020; Chiappori, Costa-Dias, Crossman, and Meghir, 2020) and how the marital college premium has changed over time (Chiappori, Iyigun, and Weiss, 2009; Chiappori, Salanié, and Weiss, 2017).

Other papers adopt the Multinomial Logit 1to1TU model to measure the effect on marital choices of exogenous events, changing the distribution of individual characteristics on each side of the market, such as the famine in China between 1958 and 1961 (Brandt, Siow, and Carl, 2016).

The Multinomial Logit 1to1TU model is often incorporated into bigger structural models. Examples of these include collective household models with marriage and labor supply (Choo and Seitz, 2013); life cycle models of education, marriage, labor supply, and consumption (Chiappori, Costa-Dias, and Meghir, 2018); collective household models with marriage, labor supply, home production choices, and joint taxation (Gayle and Shephard, 2019[6]); and collective household models with marriage, fertility decisions, and child socialisation choices (Bisin and Tura, 2020). Mourifié and Siow (2020) extend the Multinomial Logit 1to1TU model to allow for peer effects and cohabitation.

Note that, in all the papers cited above, the essence of the identification arguments pertaining to marital choices exploits data on one large market.

Econometrics of the 1to1TU model Galichon and Salanié (2020) investigate identification in the 1to1TU model when one dispenses with the Multinomial Logit structure of Choo and Siow (2006). They show that the 1to1TU model is *just* identified (and, thus, the systematic match surplus is *point* identified) with data on *one large* market under the assumption that the taste shock distribution is *fully* known by the analyst. They also provide closed form expressions for the systematic match surplus in the case of some parametric distributional families. Their findings can allow one to introduce correlation among the taste shocks and conditional heteroskedasticity. However, fully knowing the taste shock distribution requires either such distribution to be parameter-free, or the analyst to fix the value of each parameter entering it, which can be ad hoc and restrictive.

The existing literature offers three ways to introduce unknown parameters in the taste shock distribution, so as to circumvent the exact identification result of Galichon and Salanié (2020), while maintaining point identification. The first amounts to *further parameterising* the systematic match surplus in order to free some degrees of freedom.[7] The second amounts to exploiting variations of matching patterns across a *few large cohorts* which feature different distributions of

[6]Gayle and Shephard (2019) allow for imperfectly transferable utilities.

[7]Note that, in the 1to1TU model of Choo and Siow (2006) and Galichon and Salanié (2020), the systematic match surplus is already parameterised because the observed characteristics of agents are discrete.

covariates, independent matching processes, identical systematic match surplus up to some drifts or linear/quadratic trends, and identical taste shock distribution (Chiappori, Salanié and Weiss, 2017). The third amounts to exploiting variations of matching patterns across *many i.i.d. markets* (Fox, 2010; Fox, 2018; Fox, Yang, and Hsu, 2018; Sinha, 2018), as in the empirical IO tradition.[8] All three approaches are appealing because, by preserving point identification, they can offer precise answers to relevant policy questions. However, it is hard to gauge the robustness merits of the first two approaches because they introduce unknown parameters in the taste shock distribution at the cost of further restricting other parts of the model. In particular, parameterising the systematic match surplus (for instance, via a linear index) may not be desirable because it does not necessarily translate into an analogous parameterisation of the equilibrium systematic payoffs gained by each individual which, as we will see, are key objects for the identification analysis. Also, restricting the evolution of the systematic match surplus across cohorts may be problematic as the systematic match surplus evolves according to complicated underlying dynamics. Lastly, the third approach can be difficult to implement practically because in most datasets it is unclear as to what truly defines i.i.d. markets. For instance, the majority of the empirical applications of the third approach assumes that consecutive years represent i.i.d. markets, which can be often hard to justify (Baccara, İmrohoroğlu, Wilson, and Yariv, 2012; Mindruda, 2013; Akkus, Cookson, and Hortaçsu, 2016; Mindruda, Moeen, and Agarwal, 2016; Chen, 2017; Banal-Estañol, Macho-Stadler, and Pérez-Castrillo, 2018).

Recent advances in the partial identification literature have pointed out an alternative route to avoid parametric assumptions on the taste shock distribution, without adding any additional restriction on the systematic match surplus, and while remaining within a *one large* market framework. In particular, Graham (2011; 2013b) shows that if the taste shocks are i.i.d., then the signs of some complementarities between the spouses' observed characteristics are identified. Fox (2018) bounds the systematic match surplus under the assumption that the taste shocks are exchangeable across the observed characteristics of the potential partners. Our paper contributes to this strand of the literature by constructing the identified set of the systematic match surplus without requiring the taste shocks to be i.i.d. or exchangeable, which can both be strong assumptions.

Econometrics of partially identified models More generally, this paper is related to the literature on partial identification in applied research (Ho and Rosen, 2017; Molinari, 2020). Further, our approach can also be used to do a sensitivity analysis of the empirical findings obtained under the Multinomial Logit 1to1TU model. In fact, our methodology allows one to construct the identified set of the systematic match surplus under increasingly restrictive nonparametric assumptions on the taste shock distribution. Note that here we consider discrete relaxations of the Multinomial Logit 1to1TU model. Continuous relaxations of parametric distributional assumptions on the unobserved heterogeneity are studied by Christensen and Connault (2019).

[8] With data on many i.i.d. markets, one can fully dispense with the parametric restrictions on the taste shock distribution.

3 The model

This section describes the 1to1TU model that has been previously studied in Choo and Siow (2006) and Galichon and Salanié (2020).

We refer to agents on one side of the market as men and to agents on the other side of the market as women but the analysis is not restricted to the marriage market. The 1to1TU model relies on four main assumptions which we outline in what follows.

Assumption 1. (*Large market*) There is a two-sided market. One side of the market is populated by an uncountably infinite set of men, $\mathcal{I}$, with measure $d\tilde{\mu}_{\mathcal{I}}$. The other side is populated by an uncountably infinite set of women, $\mathcal{J}$, with measure $d\tilde{\mu}_{\mathcal{J}}$. The two sides of the market are stochastically independent. $\diamond$

Assumption 2. (*Finite number of observed types*) Each man $i \in \mathcal{I}$ is characterised by a type, X_i, with finite support, $\mathcal{X}$. The mass of men of type $x \in \mathcal{X}$ is denoted by m_x. Each woman $j \in \mathcal{J}$ is characterised by a type, Y_j, with finite support, $\mathcal{Y}$. The mass of women of type $y \in \mathcal{Y}$ is denoted by w_y. Without loss of generality, we normalise the total mass of agents to 1, i.e., $\sum_{x\in\mathcal{X}} m_x + \sum_{y\in\mathcal{Y}} w_y = 1$. The realisations of X_i and Y_j are observed by the researcher and all agents. Lastly, we define the sets of partner types that are available to men and women by $\mathcal{Y}_0 \equiv \mathcal{Y} \cup \{0\}$ and $\mathcal{X}_0 \equiv \mathcal{X} \cup \{0\}$, respectively, where "0" represents the option to not match. $\diamond$

Assumption 3. (*Taste shocks*) Each man $i \in \mathcal{I}$ is endowed with a $|\mathcal{Y}_0| \times 1$ vector of taste shocks, $\epsilon_i \equiv (\epsilon_{iy}\ \forall y \in \mathcal{Y}_0)$, where ϵ_{iy} denotes the idiosyncratic preference of man i for marrying a woman of type $y \in \mathcal{Y}_0$. Conditional on $X_i = x$ and for each $x \in \mathcal{X}$, ϵ_i has cumulative distribution function (hereafter, CDF) F_x. F_x is absolutely continuous with respect to the Lebesgue measure and has support in $\mathbb{R}^{|\mathcal{Y}_0|}$, where $|\mathcal{Y}_0|$ denotes the cardinality of $\mathcal{Y}_0$. Each woman $j \in \mathcal{J}$ is endowed with a $|\mathcal{X}_0| \times 1$ vector of taste shocks, $\eta_j \equiv (\eta_{xj}\ \forall x \in \mathcal{X}_0)$, where η_{xj} denotes the idiosyncratic preference of woman j for marrying a man of type $x \in \mathcal{X}_0$. Conditional on $Y_j = y$ and for each $y \in \mathcal{Y}$, η_j has CDF G_y. G_y is absolutely continuous with respect to the Lebesgue measure and has support in $\mathbb{R}^{|\mathcal{X}_0|}$. The realisations of ϵ_i and η_j are not observed by the researcher. $\diamond$

Assumption 4. (*Separability*) A match between man $i \in \mathcal{I}$ of type $x \in \mathcal{X}$ and woman $j \in \mathcal{J}$ of type $y \in \mathcal{Y}$ generates a match surplus defined as

$$\tilde{\Phi}_{ij} \equiv \Phi_{xy} + \epsilon_{iy} + \eta_{xj},$$

where $\Phi \equiv (\Phi_{xy}\ \forall (x, y) \in \mathcal{X} \times \mathcal{Y})$ is the systematic match surplus. The payoff of man $i \in \mathcal{I}$ from remaining unmatched is

$$\tilde{\Phi}_{i0} \equiv \epsilon_{i0}.$$

The payoff of woman $j \in \mathcal{J}$ from remaining unmatched is

$$\tilde{\Phi}_{0j} \equiv \eta_{0j}.$$

◇

Assumption 1 outlines the *one large market* framework, as discussed in Section 2. The restriction on the stochastic independence of the two sides of the market is not crucial for our discussion and can be relaxed.

Assumption 2 requires each agent to belong to one type. There are finite number of types, which are defined by the Cartesian product of all the individual characteristics observed by the researcher. Note that Assumption 2 does not rule out types being defined by multi-dimensional covariates. However, we emphasise that most of the empirical applications on the marriage market cited in Section 2 define types using only a one-dimensional covariate, such as education or age. This is because, even in the Multinomial Logit 1to1TU model, having types determined by multi-dimensional covariates often leads to empirical match probabilities close to zero which can cause identification issues. For the same reason, even in the Multinomial Logit 1to1TU model, it is desirable to consider $\mathcal{X}$ and $\mathcal{Y}$ with relatively small cardinalities.

Assumption 3 requires each agent to have idiosyncratic marital preferences over the types of the potential partners and not over their identities. It implies that women (men) of the same type are perfect substitutes for a man (woman).

Assumption 4 imposes that the match surplus is the sum of two components. One is the systematic match surplus, that is determined by the types of the potential partners. The other is the sum of the taste shocks of the potential partners. In particular, the latent heterogeneity entering the match surplus equation does not consist of an ij-indexed term. Instead, it is modelled through the sum of two terms, $\epsilon_{iy} + \eta_{xj}$, each of which only depends on the type of the potential partner. Assumption 4 is typically referred to as "separability". For a thorough discussion of Assumption 4, see Chiappori (2017).

A matching consists of

(i) A measure $d\tilde{\mu}$ on the set $\mathcal{I} \times \mathcal{J}$, such that the marginal of $d\tilde{\mu}$ over $\mathcal{I}$ ($\mathcal{J}$) is $d\tilde{\mu}_{\mathcal{I}}$ ($d\tilde{\mu}_{\mathcal{J}}$).

(ii) A set of payoffs, $\{\tilde{U}_i\}_{i\in\mathcal{I}}$ and $\{\tilde{V}_j\}_{j\in\mathcal{J}}$, such that

$$\tilde{U}_i + \tilde{V}_j = \tilde{\Phi}_{ij} \quad \forall (i,j) \in \text{supp}(d\tilde{\mu}).$$

(Chiappori, McCann, and Nesheim, 2010; Chiappori, McCann, and Pass, 2020). That is, a matching consists of a match assignment and a match surplus sharing rule. A match assignment is a description of who is matched with whom. A match surplus sharing rule tells us how the match surplus is divided between spouses. Such division of the match surplus relies on endogenously determined transfers, ensuring that every agent maximises her utility and the market clears.

A matching, $d\tilde{\mu}, \{\tilde{U}_i\}_{i\in\mathcal{I}}, \{\tilde{V}_j\}_{j\in\mathcal{J}}$, is stable when no agent has an incentive to change her partner, i.e.

$$\tilde{U}_i \geq \tilde{\Phi}_{i0} \; \forall i \in \mathcal{I},$$
$$\tilde{V}_j \geq \tilde{\Phi}_{0j} \; \forall j \in \mathcal{J},$$
$$\tilde{U}_i + \tilde{V}_j \geq \tilde{\Phi}_{ij} \; \forall (i,j) \in \mathcal{I} \times \mathcal{J}.$$

The first two sets of inequalities imply that married agents would not prefer being single. The last set of inequalities states that no man and woman can get a strictly higher match surplus by matching together than what they get under $d\tilde{\mu}$. Lastly, it can be shown that a stable matching exists under mild continuity assumptions (Villani, 2009). Moreover, in the limit of continuous and atomless populations, the stable matching is generically unique (Gretsky, Ostroy and Zame, 1992). Importantly for the identification analysis, the resulting equilibrium mass of couples where the man is of type x and the woman is of type y is unique for every (x, y). In what follows, we denote this equilibrium mass of couples by

$$\mu_{xy} \; \forall (x,y) \in \mathcal{Z} \equiv \mathcal{X}_0 \times \mathcal{Y}_0 \setminus \{0,0\}.$$

4 Identification

Our discussion of identification is divided into five sections. In Section 4.1, we describe which data are considered available to the researcher. In Section 4.2, we present the policy-relevant parameters. In Sections 4.3, 4.4, and 4.5, we illustrate our identification arguments.

4.1 Data

We assume that the market of interest has already reached the stable matching. In other words, as the analyst collects more data, the asymptotic fiction is that the researcher learns more about the already established stable matching, without altering it. For identification, we assume that the analyst knows $\{\mu_{xy}\}_{(x,y)\in\mathcal{Z}}$, $\{m_x\}_{x\in\mathcal{X}}$, and $\{w_y\}_{y\in\mathcal{Y}}$. For estimation and inference, we will replace $\{\mu_{xy}\}_{(x,y)\in\mathcal{Z}}$, $\{m_x\}_{x\in\mathcal{X}}$, and $\{w_y\}_{y\in\mathcal{Y}}$ with consistent sample analogues, resulting from sampling at random from the market at the individual level or at the household level.

In what follows,

$$p_{xy} \equiv \frac{\mu_{xy}}{\sum_{(x,y)\in\mathcal{Z}} \mu_{xy}},$$

denotes the equilibrium proportion of couples where the man is of type x and the woman is of type y, for every $(x, y) \in \mathcal{Z}$. We also define

$$p_{y|x} \equiv \frac{\mu_{xy}}{m_x} \quad \text{and} \quad p_{x|y} \equiv \frac{\mu_{xy}}{w_y},$$

as the equilibrium probability of marrying a woman of type $y \in \mathcal{Y}_0$ conditional on being a man of type $x \in \mathcal{X}$, and the equilibrium probability of marrying a man of type $x \in \mathcal{X}_0$ conditional on being a woman of type $y \in \mathcal{Y}$, respectively. Lastly, we define

$$p_x \equiv \frac{m_x}{\sum_{x \in \mathcal{X}} m_x} \quad \text{and} \quad p_y \equiv \frac{w_y}{\sum_{y \in \mathcal{Y}} w_y},$$

as the proportion of men of type $x \in \mathcal{X}$ and the proportion of women of type $y \in \mathcal{Y}$, respectively.

4.2 Structural parameters of interest

From an economic perspective, our main interest lies in recovering the systematic match surplus, Φ. In fact, (partially) identifying Φ allows one to answer three important questions considered in the marriage market literature.

First, Φ can be used to learn which of the spouses' observed characteristics are complements/substitutes in the production of the systematic match surplus. For instance, suppose that $\mathcal{X}$ collects r education levels and $\mathcal{Y}$ collects l income categories. Without loss of generality, order the r education levels from lowest to highest and label them $\mathcal{X} \equiv \{1, ..., r\}$. Similarly, order the l income categories from lowest to highest and label them $\mathcal{Y} \equiv \{1, ..., l\}$. For any $(x, y), (\tilde{x}, \tilde{y}) \in \mathcal{X} \times \mathcal{Y}$ with $x > \tilde{x}$ and $y > \tilde{y}$, consider the cross-difference operator,

$$D_{xy,\tilde{x}\tilde{y}}(\Phi) \equiv \Phi_{xy} + \Phi_{\tilde{x}\tilde{y}} - \Phi_{x\tilde{y}} - \Phi_{\tilde{x}y}. \tag{1}$$

$D_{xy,\tilde{x}\tilde{y}}(\Phi)$ measures how the incremental (dis)value of marrying a more-educated man evolves as the income of the woman increases. Hence, the vector $(D_{xy,\tilde{x}\tilde{y}}(\Phi)\ \forall (x, y), (\tilde{x}, \tilde{y}) \in \mathcal{X} \times \mathcal{Y}, x > \tilde{x}, y > \tilde{y})$ captures the intensity of the complementarity/substitutability between a man's education and a woman's income in the production of the systematic match surplus. Similar ideas can be extended to the case where types are defined by the Cartesian product of multi-dimensional covariates. Assessing complementarities/substitutabilities and their relative strengths is important to discover the key drivers of the gains to matching. For a review of papers using the Multinomial Logit 1to1TU model to study complementarities/substitutabilities in marital decisions, see the first paragraph of Section 2. Complementarities/substitutabilities are also central in the analysis of Fox (2010) and Graham (2011; 2013b).

Second, complementarities are inherently related to the concept of positive assortativeness, i.e., the tendency of agents to match with similar people. Investigating positive assortativeness and its evolution over time has been a focus of empirical research since Becker (1973) because it can be crucial to understand the sources of inequality in intergenerational outcomes. For instance, if the education level of parents affects their children's school attainment and marriage is positively assortative by education, then inequality in the next generation may be higher. In the 1to1TU model one can detect positive assortativeness within a given market and evaluate changes in positive assortativeness across markets (for instance, across cohorts) by estimating the supermodular core of Φ,

$$D(\Phi) \equiv (D_{xx,\tilde{x}\tilde{x}}(\Phi)\ \forall x, \tilde{x} \in \mathcal{X}, x > \tilde{x}), \tag{2}$$

within each market. If each component of the vector $D(\Phi)$ is positive, then there is positive assortative matching. Further, if the components of the vector $D(\Phi)$ increase across markets, then one can conclude that positive assortative matching increases across markets as well. For a review of papers using the Multinomial Logit 1to1TU model to study how positive assortativeness in the marriage market has evolved over time, see the first paragraph of Section 2. Further details and references are in Sections 1 and 6.

Third, Φ can be used to answer some normative welfare questions. For instance, as discussed by Graham (2011; 2013b), Φ allows the analyst to characterise how alternative assignments (or policies which cause re-assignments) alter the average systematic match surplus, while leaving the availability of resources (i.e., $\{p_x\}_{x\in\mathcal{X}}$ and $\{p_y\}_{y\in\mathcal{Y}}$) constant. In particular, the average reallocation effect (ARE) of Graham, Imbens, and Ridder (2014) measures the impact of implementing an alternative assignment, $\{p^a_{xy}\}_{(x,y)\in\mathcal{Z}}$, on the average systematic match surplus relative to the status-quo assignment, $\{p_{xy}\}_{(x,y)\in\mathcal{Z}}$. The ARE is defined as

$$ARE \equiv \sum_{(x,y)\in\mathcal{Z}} \Phi_{xy}(p^a_{xy} - p_{xy}).$$

Alternative assignments of interest can be positive/negative assortative assignments and random assignments. Computing the ARE for such alternative assignments can help to better evaluate the consequences on welfare and inequality of the status-quo assignment. The efficiency of the status-quo assignment can also be assessed by computing the maximum reallocation effect (MRE), analysed by Bhattacharya (2009), which measures the maximum gain in the average systematic match surplus available through reallocations. The MRE is defined as

$$MRE \equiv \max_{\{p^a_{xy}\}_{(x,y)\in\mathcal{Z}}} \sum_{(x,y)\in\mathcal{Z}} \Phi_{xy}(p^a_{xy} - p_{xy}).$$

For an empirical application of these ideas to the relationship between parental schooling and children's educational attainment, see Graham (2013c).

In addition to recovering Φ, our methodology also permits the analyst to (partially) identify how Φ is shared between spouses. In particular, let U_{xy} be the part of Φ_{xy} that is gained by a man of type x when matching with a woman of type y. Let V_{xy} be the part of Φ_{xy} that is gained by a woman of type y when matching with a man of type x.[9] Knowing $U \equiv (U_{xy}\ \forall(x,y)\in\mathcal{X}\times\mathcal{Y}_0)$ and $V \equiv (V_{xy}\ \forall(x,y)\in\mathcal{X}_0\times\mathcal{Y})$ is useful to measure the impact of pre-marital decisions on marriage market outcomes. For instance, consider the men's side and suppose that $\mathcal{X}$ collects r education levels, which are ordered from lowest to highest and labelled $\mathcal{X}\equiv\{1,...,r\}$ with r indicating the college degree. Let $\bar{U}_x$ be the expected payoff that man i of type $x\in\mathcal{X}$ gets when marrying, i.e.,

$$\bar{U}_x \equiv E_{F_x}(\max_{y\in\mathcal{Y}_0} U_{xy} + \epsilon_{iy}|X_i = x).$$

[9] U_{xy} and V_{xy} reappear below in Proposition 1.

For any $\tilde{x} \in \mathcal{X}$ with $\tilde{x} < x$, the difference $\bar{U}_x - \bar{U}_{\tilde{x}}$ denotes the gain in expected utility obtained by reaching education level x instead of $\tilde{x}$. Therefore, it represents the marital *education* premium. When $x = r$, such a difference is called the marital *college* premium. These quantities have received particular interest because they capture the value of human capital on the marriage market (Chiappori, Iyigun, and Weiss, 2009; Chiappori, Salanié, and Weiss, 2017). As shown in Galichon and Salanié (2020), the marital education premium is equal to

$$\bar{U}_x - \bar{U}_{\tilde{x}} = \sum_{y \in \mathcal{Y}_0} p_{y|x} U_{xy} - \sum_{y \in \mathcal{Y}_0} p_{y|\tilde{x}} U_{\tilde{x}y} + E_{F_x}(\epsilon_{iy_i^*} | X_i = x) - E_{F_{\tilde{x}}}(\epsilon_{i\tilde{y}_i^*} | X_i = \tilde{x}), \tag{3}$$

where $y_i^* \in \mathcal{Y}_0$ is the optimal choice of man i of type x and $\tilde{y}_i^* \in \mathcal{Y}_0$ is the optimal choice of man i of type $\tilde{x}$. Note from (3) that computing the marital education premium requires the specification of $\{F_x\}_{x \in \mathcal{X}}$. Thus, our methodology is insufficient to compute the marital education premium. Nevertheless, our methodology allows the analyst to evaluate quantities that contribute to the marital education premium. In particular, we will provide bounds for the difference

$$C_{x\tilde{x}}(U) \equiv \sum_{y \in \mathcal{Y}_0} p_{y|x} U_{xy} - \sum_{y \in \mathcal{Y}_0} p_{y|\tilde{x}} U_{\tilde{x}y}, \tag{4}$$

which represents the change in the average systematic payoffs due to reaching education level x instead of $\tilde{x}$. In turn, such bounds will help us make certain conclusions on $\bar{U}_x - \bar{U}_{\tilde{x}}$. Further details are in our empirical application in Section 6. In what follows,

$$C(U) \equiv (C_{x\tilde{x}}(U) \ \forall x, \tilde{x} \in \mathcal{X}, x > \tilde{x}) \quad \text{and} \quad C(V) \equiv (C_{y\tilde{y}}(V) \ \forall y, \tilde{y} \in \mathcal{Y}, y > \tilde{y}). \tag{5}$$

4.3 Two multinomial choice models

Galichon and Salanié (2020) provide a key result for our identification analysis which relies on the separability restriction stated in Assumption 4.[10]

Proposition 1. (*Galichon and Salanié, 2020*) Given the primitives Φ, $\{p_x\}_{x \in \mathcal{X}}$, $\{p_y\}_{y \in \mathcal{Y}}$, $\{F_x\}_{x \in \mathcal{X}}$, $\{G_y\}_{y \in \mathcal{Y}}$ generating the stable matching $d\tilde{\mu}$, $\{\tilde{U}_i\}_{i \in \mathcal{I}}$, $\{\tilde{V}_j\}_{j \in \mathcal{J}}$, there exist vectors

$$U \equiv (U_{xy} \ \forall (x,y) \in \mathcal{X} \times \mathcal{Y}_0) \in \mathbb{R}^{|\mathcal{X} \times \mathcal{Y}_0|} \text{ and } V \equiv (V_{xy} \ \forall (x,y) \in \mathcal{X}_0 \times \mathcal{Y}) \in \mathbb{R}^{|\mathcal{X}_0 \times \mathcal{Y}|},$$

such that

$$\tilde{U}_i = \max_{y \in \mathcal{Y}_0} (U_{xy} + \epsilon_{iy}) \qquad \forall i \in \mathcal{I} \text{ of type } x \in \mathcal{X}, \forall x \in \mathcal{X}, \tag{6}$$

$$\tilde{V}_j = \max_{x \in \mathcal{X}_0} (V_{xy} + \eta_{xj}) \qquad \forall j \in \mathcal{J} \text{ of type } y \in \mathcal{Y}, \forall y \in \mathcal{Y}, \tag{7}$$

$$U_{xy} + V_{xy} = \Phi_{xy},\ U_{x0} = 0,\ V_{0y} = 0 \qquad \forall (x,y) \in \mathcal{X} \times \mathcal{Y}. \tag{8}$$

[10]This result also appears in previous working paper versions of Galichon and Salanié (2020), in Chiappori, Salanié, Tillman, and Weiss (2008), and in Chiappori, Salanié, and Weiss (2017).

◇

Proposition 1 allows us to rewrite the framework of Section 3 as two separate one-sided multinomial choice models linked by market-clearing transfers that are implicitly embedded into the vectors U and V. Such an alternative representation of the problem is useful as it immediately suggests a way to investigate the identification of Φ. That is, the researcher can study separate identification of U and V from (6) and (7) using various restrictions on the unobserved heterogeneity, and then obtain identification results for Φ through (8).

Further, Galichon and Salanié (2020) prove that if $\{F_x\}_{x\in\mathcal{X}}$ and $\{G_y\}_{y\in\mathcal{Y}}$ are fully known by the analyst, then the 1to1TU model is just identified and, thus, Φ is point identified. Note that fully knowing $\{F_x\}_{x\in\mathcal{X}}$ and $\{G_y\}_{y\in\mathcal{Y}}$ requires either such conditional CDFs to be parameter-free, or the analyst to fix the value of each parameter entering them. In particular, a widespread practice in the empirical literature amounts to assuming that the taste shocks are i.i.d. standard Extreme Value Type I , independently distributed from types, so that Φ can be recovered via standard Multinomial Logit arguments applied to each side of the market (Choo and Siow, 2006). As discussed in Sections 1 and 2, imposing a Multinomial Logit structure or, more generally, completely specifying $\{F_x\}_{x\in\mathcal{X}}$ and $\{G_y\}_{y\in\mathcal{Y}}$ can be ad hoc and restrictive. Our objective is examining the consequences on the identification of Φ when parametric assumptions on the unobserved heterogeneity are relaxed.

4.4 The extent of under-identification

As shown by Galichon and Salanié (2020), if $\{F_x\}_{x\in\mathcal{X}}$ and $\{G_y\}_{y\in\mathcal{Y}}$ are not assumed to be fully known by the analyst, then the 1to1TU model is under-identified. As a first step, this section investigates the extent of under-identification by answering the following question: does the 1to1TU model, with data on one large market, retain some identifying power on Φ without imposing any restriction on $\{F_x\}_{x\in\mathcal{X}}$ and $\{G_y\}_{y\in\mathcal{Y}}$? Proposition 2 below claims that, in the absence of any restriction on $\{F_x\}_{x\in\mathcal{X}}$ and $\{G_y\}_{y\in\mathcal{Y}}$, the 1to1TU model is completely uninformative about Φ. That is, for every possible value of Φ, there exist some $\{F_x\}_{x\in\mathcal{X}}$ and $\{G_y\}_{y\in\mathcal{Y}}$ which when combined with that value of Φ, rationalise the empirical conditional choice probabilities, $\{p_{y|x}\}_{(x,y)\in\mathcal{X}\times\mathcal{Y}_0}$ and $\{p_{x|y}\}_{(x,y)\in\mathcal{X}_0\times\mathcal{Y}}$.

Before presenting Proposition 2, we introduce some useful notation. Let $\mathcal{U}$, $\mathcal{V}$, and Θ denote the parameter spaces of U, V, and Φ, respectively, i.e.,

$$\begin{aligned}
&\mathcal{U} \equiv \{(U_{xy}\ \forall(x,y)\in\mathcal{X}\times\mathcal{Y}_0)\in\mathbb{R}^{|\mathcal{X}\times\mathcal{Y}_0|} : U_{x0}=0\ \forall x\in\mathcal{X}\},\\
&\mathcal{V} \equiv \{(V_{xy}\ \forall(x,y)\in\mathcal{X}_0\times\mathcal{Y})\in\mathbb{R}^{|\mathcal{X}_0\times\mathcal{Y}|} : V_{0y}=0\ \forall y\in\mathcal{Y}\},\\
&\Theta \equiv \mathbb{R}^{|\mathcal{X}\times\mathcal{Y}|}.
\end{aligned}$$

Further, let $\mathcal{F}$ and $\mathcal{G}$ be the function spaces of all admissible $\{F_x\}_{x\in\mathcal{X}}$ and $\{G_y\}_{y\in\mathcal{Y}}$, respectively.[11]

[11] Note that one element of $\mathcal{F}$ is a family of $|\mathcal{X}|$ conditional CDFs, $\{F_x\}_{x\in\mathcal{X}}$. Similarly, one element of $\mathcal{G}$ is a

Lastly, for any $y \in \mathcal{Y}_0$, $U \in \mathcal{U}$, and $\{F_x\}_{x\in\mathcal{X}} \in \mathcal{F}$, let $\kappa(U, F_x, y)$ be the predicted probability of marrying a woman of type $y \in \mathcal{Y}_0$ conditional on being a man of type $x \in \mathcal{X}$, i.e.,

$$\kappa(U, F_x, y) \equiv \lambda_{F_x}(\{(e_y\ \forall y \in \mathcal{Y}_0) \in \mathbb{R}^{|\mathcal{Y}_0|} : U_{xy} + e_y \geq U_{x\tilde{y}} + e_{\tilde{y}}\ \forall \tilde{y} \in \mathcal{Y}_0 \setminus \{y\}\}), \tag{9}$$

where λ_{F_x} is the measure associated with F_x. Similarly, for any $x \in \mathcal{X}_0$, $V \in \mathcal{V}$, and $\{G_y\}_{y\in\mathcal{Y}} \in \mathcal{G}$, let $\kappa(V, G_y, x)$ be the predicted probability of marrying a man of type $x \in \mathcal{X}_0$ conditional on being a woman of type $y \in \mathcal{Y}$, i.e.,

$$\kappa(V, G_y, x) \equiv \lambda_{G_y}(\{(e_x\ \forall x \in \mathcal{X}_0) \in \mathbb{R}^{|\mathcal{X}_0|} : V_{xy} + e_x \geq V_{\tilde{x}y} + e_{\tilde{x}}\ \forall \tilde{x} \in \mathcal{X}_0 \setminus \{x\}\}), \tag{10}$$

where λ_{G_y} is the measure associated with G_y.

Proposition 2. (*Under-identification of* Φ) Given $\Phi \in \Theta$, consider any $(U, V) \in \mathcal{U} \times \mathcal{V}$ such that

$$U_{xy} + V_{xy} = \Phi_{xy} \qquad \forall (x, y) \in \mathcal{X} \times \mathcal{Y}. \tag{11}$$

Then, there exist $\{F_x\}_{x\in\mathcal{X}} \in \mathcal{F}$ and $\{G_y\}_{y\in\mathcal{Y}} \in \mathcal{G}$ such that

$$p_{y|x} = \kappa(U, F_x, y) \qquad \forall (x, y) \in \mathcal{X} \times \mathcal{Y}_0, \tag{12}$$

$$p_{x|y} = \kappa(V, G_y, x) \qquad \forall (x, y) \in \mathcal{X}_0 \times \mathcal{Y}. \tag{13}$$

$\diamond$

Proposition 2 makes it clear that one needs to impose some distributional restrictions on the unobserved heterogeneity for the 1to1TU model to be potentially informative about Φ.[12,13]

4.5 Adding nonparametric assumptions on unobserved heterogeneity

In this section, we ask ourselves whether the 1to1TU model retains any identifying power on Φ under various classes on nonparametric distributional assumptions on the unobserved heterogeneity, so as to still be able to address relevant policy matters while maintaining a certain degree of robustness. To answer this question, we adopt a computational approach. In particular, we start from observing that if $\{F_x\}_{x\in\mathcal{X}}$ and $\{G_y\}_{y\in\mathcal{Y}}$ are not assumed to be fully known by the analyst, then the 1to1TU model is under-identified and, in turn, Φ is partially identified. Hence, we provide a methodology to practically construct the identified set of Φ under various classes of nonparametric distributional assumptions on the unobserved heterogeneity.

family of $|\mathcal{Y}|$ conditional CDFs, $\{G_y\}_{y\in\mathcal{Y}}$.

[12]Note that the point identification result in Galichon and Salanié (2020) establishes the converse of Proposition 2. That is, given $\Phi \in \Theta$, $\{F_x\}_{x\in\mathcal{X}} \in \mathcal{F}$, and $\{G_y\}_{y\in\mathcal{Y}} \in \mathcal{G}$, there exists $(U, V) \in \mathcal{U} \times \mathcal{V}$ such that (11)-(13) are satisfied.

[13]Note that Proposition 2 still holds even if we impose scale normalisations on U and V, as discussed in the proof in Appendix A.1. Note also that appropriate location normalisations have been already imposed by Proposition 1, i.e., $U_{x0} = V_{0y}$ for every $(x, y) \in \mathcal{X} \times \mathcal{Y}$.

Intuitively, the identified set of Φ is the set of values of Φ such that there exists U, V, $\{F_x\}_{x\in\mathcal{X}}$, and $\{G_y\}_{y\in\mathcal{Y}}$ that satisfy (11)-(13). We denote it by Θ^*. Note that, by Proposition 1, we can construct Θ^* by separately focusing on each side of the market. In fact, first, we can construct the identified set of U (hereafter, denoted by $\mathcal{U}^*$), i.e., the set of values of U such that there exists $\{F_x\}_{x\in\mathcal{X}}$ that satisfies (12). Then, we can construct the identified set of V (hereafter, denoted by $\mathcal{V}^*$), i.e., the set of values of V such that there exists $\{G_y\}_{y\in\mathcal{Y}}$ that satisfies (13). Finally, we can obtain Θ^* from (11). In what follows, we explain how to construct $\mathcal{U}^*$. Similar arguments have to be implemented for constructing $\mathcal{V}^*$.

Recall that in multinomial choice models what matter are differences in utilities. Therefore, as a preliminary step, we rewrite the identification problem using the taste shock differences. Without loss of generality, we label the women's types as $\mathcal{Y} \equiv \{1, ..., r\}$. Let

$$\Delta\epsilon_i \equiv (\epsilon_{i1} - \epsilon_{i0}, ..., \epsilon_{ir} - \epsilon_{i0}, \epsilon_{i1} - \epsilon_{i2}, ..., \epsilon_{i1} - \epsilon_{ir}, \epsilon_{i2} - \epsilon_{i3}, ..., \epsilon_{i2} - \epsilon_{ir}, ..., \epsilon_{ir-1} - \epsilon_{ir}), \tag{14}$$

be the vector of differences between every pair of taste shocks of man $i \in \mathcal{I}$, with length $d \equiv \binom{r+1}{2}$.

We highlight three important facts. First, each $\{F_x\}_{x\in\mathcal{X}} \in \mathcal{F}$ determines a corresponding family of d-dimensional conditional CDFs of $\Delta\epsilon_i$, which we denote by $\{\Delta F_x\}_{x\in\mathcal{X}} \in \Delta\mathcal{F}$. Second, the first r components of $\Delta\epsilon_i$ can be arbitrary, while the remaining $(d-r)$ components are linear combination of the first r components. Hence, for every $x \in \mathcal{X}$, ΔF_x has support contained in the region

$$\begin{aligned}\mathcal{B} \equiv \{(b_1, b_2, ..., b_d) \in \mathbb{R}^d :\ & b_{r+1} = b_1 - b_2, b_{r+2} = b_1 - b_3, ..., b_{2r-1} = b_1 - b_r,\\ & b_{2r} = b_2 - b_3, ..., b_{3r-3} = b_2 - b_r, ...,\\ & b_d = b_{r-1} - b_r\}.\end{aligned}$$

Third, for each $(x, y) \in \mathcal{X} \times \mathcal{Y}_0$, the function $\kappa(U, F_x, y)$ defined in (9) can be equivalently expressed, with slight abuse of notation, by using ΔF_x in place of F_x. In light of these three facts and for a given $U \in \mathcal{U}$, verifying whether there exists $\{F_x\}_{x\in\mathcal{X}} \in \mathcal{F}$ solving (12) is equivalent to verifying whether

$$\exists\ \{\Delta F_x\}_{x\in\mathcal{X}} \in \Delta\mathcal{F} \text{ s.t. } p_{y|x} = \kappa(U, \Delta F_x, y)\ \forall (x, y) \in \mathcal{X} \times \mathcal{Y}_0 \text{ and } \lambda_{\Delta F_x}(\mathcal{B}) = 1\ \forall x \in \mathcal{X},$$

where $\lambda_{\Delta F_x}$ is the measure associated with ΔF_x.

Moreover, one may want to impose various nonparametric restrictions on $\{\Delta F_x\}_{x\in\mathcal{X}}$ in order to obtain informative bounds on U (and, ultimately, Φ), as discussed above. We denote by $\Delta\mathcal{F}^\dagger \subset \Delta\mathcal{F}$ the restricted collection of families of conditional CDFs. We describe later which classes of nonparametric restrictions on $\{\Delta F_x\}_{x\in\mathcal{X}}$ are considered.

Thus, to sum up, our objective is to construct the identified set of U, defined as

$$\mathcal{U}^* \equiv \{U \in \mathcal{U} : \exists \ \{\Delta F_x\}_{x\in\mathcal{X}} \in \Delta\mathcal{F}^\dagger \text{ s.t. } p_{y|x} = \kappa(U, \Delta F_x, y) \ \forall (x, y) \in \mathcal{X} \times \mathcal{Y}_0 \text{ and } \lambda_{\Delta F_x}(\mathcal{B}) = 1 \ \forall x \in \mathcal{X}\}.$$

We organise the discussion in three steps discussed in Sections 4.5.1, 4.5.2, and 4.5.3, respectively. For any given $U \in \mathcal{U}$, in Section 4.5.1 we explain how to verify whether there exists $\{\Delta F_x\}_{x\in\mathcal{X}} \in \Delta\mathcal{F}^\dagger$ such that $p_{y|x} = \kappa(U, \Delta F_x, y)$ for every $(x, y) \in \mathcal{X} \times \mathcal{Y}_0$. In Section 4.5.2, we provide a result which reduces the computational burden of repeating the first step for every $U \in \mathcal{U}$. In Section 4.5.3, we explain how to impose the condition $\lambda_{\Delta F_x}(\mathcal{B}) = 1$ for every $x \in \mathcal{X}$, which we call the degeneracy condition.

Lastly, we introduce some useful notation adopted in the forthcoming arguments. $\bar{\mathbb{R}}$ denotes the extended real line. 0_d is the $d \times 1$ vector of zeros. $\Delta\epsilon_{i,l}$ is the l-th component of $\Delta\epsilon_i$ and $\Delta F_{x,l}$ is the l-th marginal of ΔF_x, where $l \in \{1, ..., d\}$. $\Delta\epsilon_i^y$ is the $(r-1) \times 1$ subvector of $\Delta\epsilon_i$ collecting the taste shock differences that are relevant when choosing $y \in \mathcal{Y}_0$. For instance, consider $r = 2$ (and, hence, $d = 3$). When choosing 0, man i evaluates $\epsilon_{i1} - \epsilon_{i0}$ and $\epsilon_{i2} - \epsilon_{i0}$. Thus, given the definition of $\Delta\epsilon_i$ in (14), $\Delta\epsilon_i^0 \equiv (\epsilon_{i1} - \epsilon_{i0}, \epsilon_{i2} - \epsilon_{i0})$. Similarly, $\Delta\epsilon_i^1 \equiv (\epsilon_{i1} - \epsilon_{i0}, \epsilon_{i1} - \epsilon_{i2})$ and $\Delta\epsilon_i^2 \equiv (\epsilon_{i2} - \epsilon_{i0}, \epsilon_{i1} - \epsilon_{i2})$. We denote by ΔF_x^y the distribution of $\Delta\epsilon_i^y$ conditional on $X_i = x$, which can be obtained by appropriately marginalising ΔF_x.

4.5.1 First step

As part of the construction of $\mathcal{U}^*$, the analyst has to find whether, for any given $U \in \mathcal{U}$,

$$\exists \ \{\Delta F_x\}_{x\in\mathcal{X}} \in \Delta\mathcal{F}^\dagger \text{ s.t. } p_{y|x} = \kappa(U, \Delta F_x, y) \ \forall (x, y) \in \mathcal{X} \times \mathcal{Y}_0. \tag{15}$$

Without parametric restrictions on the agents' unobserved heterogeneity, (15) is an infinite-dimensional existence problem. We use Theorem 1 in Torgovitsky (2019) to transform (15) into a linear programming problem. Essentially, the theorem states that, under various classes of nonparametric restrictions on $\{\Delta F_x\}_{x\in\mathcal{X}}$, verifying (15) is equivalent to finding whether a linear system has at least one solution. In what follows, we provide an intuition of the theorem and the list of nonparametric restrictions on $\{\Delta F_x\}_{x\in\mathcal{X}}$ that are considered in our simulations and empirical application. We give a more formal statement of the theorem in Appendix B.

For simplicity, let $r = 2$ (hence, $d = 3$) and $\Delta\mathcal{F}^\dagger = \Delta\mathcal{F}$, so that (15) can be more explicitly rewritten as[14]

$$\exists \ \{\Delta F_x\}_{x\in\mathcal{X}} \in \Delta\mathcal{F} \text{ s.t.}$$
$$p_{1|x} = 1 + \Delta F_x(-U_{x1}, \infty, U_{x2} - U_{x1}) - \Delta F_x(\infty, \infty, U_{x2} - U_{x1}) - \Delta F_x(-U_{x1}, \infty, \infty) \ \forall x \in \mathcal{X},$$
$$p_{2|x} = \Delta F_x(\infty, \infty, U_{x2} - U_{x1}) - \Delta F_x(\infty, -U_{x2}, U_{x2} - U_{x1}) \ \forall x \in \mathcal{X},$$
$$p_{0|x} = \Delta F_x(-U_{x1}, -U_{x2}, \infty) \ \forall x \in \mathcal{X}. \tag{16}$$

Note that, for every $x \in \mathcal{X}$, (16) depends on the values of ΔF_x at a *finite* number of 3-tuples. We collect such 3-tuples in the following three sets:

$$\mathcal{A}_{x,1,U} \equiv \{-U_{x1}, \infty, -\infty\},$$
$$\mathcal{A}_{x,2,U} \equiv \{-U_{x2}, \infty, -\infty\},$$
$$\mathcal{A}_{x,3,U} \equiv \{U_{x2} - U_{x1}, \infty, -\infty\},$$

where $\mathcal{A}_{x,1,U}$ collects the elements at which ΔF_x is evaluated along the first dimension, $\mathcal{A}_{x,2,U}$ collects the elements at which ΔF_x is evaluated along the second dimension, and $\mathcal{A}_{x,3,U}$ collects the elements at which ΔF_x is evaluated along the third dimension. We add $-\infty$ to each set because the value of one-dimensional CDFs at $-\infty$ is known and equal to 0 by definition. Lastly, we define $\mathcal{A}_{x,U} \equiv \mathcal{A}_{x,1,U} \times \mathcal{A}_{x,2,U} \times \mathcal{A}_{x,3,U}$. Thus, (16) can be equivalently written as

$$\forall x \in \mathcal{X}, \ \exists \ \Delta\bar{F}_x^U : \mathcal{A}_{x,U} \to \mathbb{R} \text{ s.t.}$$
$$p_{1|x} = 1 + \Delta\bar{F}_x^U(-U_{x1}, \infty, U_{x2} - U_{x1}) - \Delta\bar{F}_x^U(\infty, \infty, U_{x2} - U_{x1}) - \Delta\bar{F}_x^U(-U_{x1}, \infty, \infty), \tag{17}$$
$$p_{2|x} = \Delta\bar{F}_x^U(\infty, \infty, U_{x2} - U_{x1}) - \Delta\bar{F}_x^U(\infty, -U_{x2}, U_{x2} - U_{x1}), \tag{18}$$
$$p_{0|x} = \Delta\bar{F}_x^U(-U_{x1}, -U_{x2}, \infty), \tag{19}$$
$$\text{and } \{\Delta\bar{F}_x^U\}_{x\in\mathcal{X}} \text{ can be “extended” to a proper family of conditional CDFs in } \Delta\mathcal{F}\,. \tag{20}$$

(17)-(20) reveal that verifying (16) is equivalent to first finding whether a system of 3 linear equations, (17)-(19), has a solution with respect to the finite-domain function $\Delta\bar{F}_x^U$, and second ensuring that such a solution can be extended to a proper conditional CDF as required by (20), for every $x \in \mathcal{X}$.

Using fundamental results in copula theory, in particular Sklar's Theorem, it can be shown that verifying whether $\Delta\bar{F}_x^U$ can be extended to a proper conditional CDF amounts to checking

[14] See Appendix C for details.

if it satisfies the following linear system:

$$\Delta\bar{F}_x^U(-\infty,t,q)=0 \qquad \forall(t,q)\in\mathcal{A}_{x,2,U}\times\mathcal{A}_{x,3,U}, \tag{21}$$
$$\Delta\bar{F}_x^U(t,-\infty,q)=0 \qquad \forall(t,q)\in\mathcal{A}_{x,1,U}\times\mathcal{A}_{x,3,U}, \tag{22}$$
$$\Delta\bar{F}_x^U(t,q,-\infty)=0 \qquad \forall(t,q)\in\mathcal{A}_{x,1,U}\times\mathcal{A}_{x,2,U}, \tag{23}$$
$$\Delta\bar{F}_x^U(\infty,\infty,\infty)=1, \tag{24}$$
$$0\leq\Delta\bar{F}_x^U(t,q,r)\leq 1 \qquad \forall(t,q,r)\in\mathcal{A}_{x,U}, \tag{25}$$
$$\begin{aligned}&[-\Delta\bar{F}_x^U(t,q,r)+\Delta\bar{F}_x^U(t',q,r)+\Delta\bar{F}_x^U(t,q',r)-\Delta\bar{F}_x^U(t',q',r)+\\ &\Delta\bar{F}_x^U(t,q,r')-\Delta\bar{F}_x^U(t',q,r')-\Delta\bar{F}_x^U(t,q',r')+\Delta\bar{F}_x^U(t',q',r')]\geq 0\end{aligned} \qquad \begin{aligned}&\forall(t,q,r),(t',q',r')\in\mathcal{A}_{x,U}\\ &\text{s.t. } (t,q,r)\leq(t',q',r').\end{aligned} \tag{26}$$

Specifically, bearing in mind the properties defining CDFs, (21)-(24) ensure that $\Delta\bar{F}_x^U$ is equal to 0 when at least one of its arguments is $-\infty$ and equal to 1 when all its arguments are ∞. (25) guarantees that the range of $\Delta\bar{F}_x^U$ is a subset of $[0,1]$. (26) requires $\Delta\bar{F}_x^U$ to be a 3-increasing function, i.e., for each pair of 3-tuples in $\mathcal{A}_{x,U}$ which are comparable, (t,q,r) and (t',q',r'), the volume of the 3-dimensional box with vertices from $\{t,t'\}\times\{q,q'\}\times\{r,r'\}$ is positive. Further, if $\Delta\bar{F}_x^U$ satisfies (26), then it is called a 3-dimensional subdistribution. By merging (17)-(19) with (21)-(26), a simple linear programming problem is obtained in which we must verify feasibility with respect to $\{\Delta\bar{F}_x^U\}_{x\in\mathcal{X}}$.

The above methodology remains valid under various classes of nonparametric restrictions on $\{\Delta F_x\}_{x\in\mathcal{X}}$, which can be simply imposed on $\{\Delta\bar{F}_x^U\}_{x\in\mathcal{X}}$ as linear constraints. In particular, in our simulations and empirical application, we will consider the following restrictions (not necessarily all maintained simultaneously):

Assumption 5. (*Nonparametric assumptions on* $\{\Delta F_x\}_{x\in\mathcal{X}}$)

5.1. $\Delta\epsilon_i$ is independent of X_i, i.e.,

$$\Delta F_x(a)=\Delta F_{\tilde{x}}(a)\ \forall a\in\bar{\mathbb{R}}^d,\ \forall x,\tilde{x}\in\mathcal{X}.$$

5.2. Conditional on X_i and for each $l\in\{1,...,d\}$, $\Delta\epsilon_{i,l}$ has a distribution symmetric at 0, i.e.,

$$\Delta F_{x,l}(a)+\Delta F_{x,l}(-a)=1\ \forall a\in\bar{\mathbb{R}},\ \forall x\in\mathcal{X}.$$

5.3. Conditional on X_i, $\{\Delta\epsilon_{i,l}\}_{l\in\{1,\ldots,d\}}$ are identically distributed, i.e.,

$$\Delta F_{x,l}(a)=\Delta F_{x,\tilde{l}}(a)\ \forall a\in\bar{\mathbb{R}},\ \forall l,\tilde{l}\in\{1,...,d\},\ \forall x\in\mathcal{X}.$$

5.4. Conditional on X_i, $\{\Delta\epsilon_i^y\}_{y\in\mathcal{Y}_0}$ are identically distributed, i.e.,

$$\Delta F_x^y(a)=\Delta F_x^{\tilde{y}}(a)\ \forall a\in\bar{\mathbb{R}}^{r-1},\ \forall y,\tilde{y}\in\mathcal{Y}_0,\ \forall x\in\mathcal{X}.$$

5.5. Conditional on X_i, $\Delta\epsilon_i$ has a distribution with median zero, i.e.,

$$\Delta F_x(0_d) = \frac{1}{2} \; \forall x \in \mathcal{X}.$$

$\diamond$

We refer the reader to Assumption A in Torgovitsky (2019) (which we also report in Appendix B) for other nonparametric distributional assumptions on the taste shock differences that can be accommodated. Finally, we remark that the above methodology does not allow to impose nonparametric assumptions directly on the conditional CDFs of the original taste shocks, i.e., on $\{F_x\}_{x\in\mathcal{X}}$. Given that in multinomial choice models what matters are differences in utilities (and not absolute levels), this is the price we pay for having an approach that flexibly works under several classes of nonparametric distributional restrictions.

4.5.2 Second step

As part of the construction of $\mathcal{U}^\star$, the analyst has to verify the feasibility with respect to $\{\Delta\bar{F}^U_x\}_{x\in\mathcal{X}}$ of the linear programming problem described in Section 4.5.1 for every $U \in \mathcal{U}$. Typically, this is done in the partial identification literature by constructing a grid of points to approximate $\mathcal{U}$, and then repeating the exercise of interest for each grid point. The difficulty of implementing such an approach increases with the size of the grid, which in turn, increases exponentially with r, hence quickly leading to a computational bottleneck. In what follows, we give a characterisation of $\mathcal{U}$ so that the issue of verifying feasibility for every $U \in \mathcal{U}$ is reduced to verifying feasibility for a handful of $U \in \mathcal{U}$. We first provide an intuition of the result and then a more formal statement.

For simplicity, let $r = 2$ and $\Delta\mathcal{F}^\dagger = \Delta\mathcal{F}$, as in Section 4.5.1. The only piece of the linear programming problem of Section 4.5.1 that might induce different feasibility sets for different values of U is (26), which requires the function $\Delta\bar{F}^U_x$ to be 3-increasing. This is because different values of U might generate different pairs of comparable 3-tuples in $\mathcal{A}_{x,U}$ and, in turn, 3-dimensional boxes with vertices located at different positions (in other words, "differently ordered") in $\mathbb{R}^d$. Such observation immediately suggests that if two values of U generate pairs of 3-tuples in $\mathcal{A}_{x,U}$ that are "similarly ordered" in $\mathbb{R}^d$, then they should lead to the same constraints on $\Delta\bar{F}^U_x$ of the type (26). In turn, if this holds $\forall x \in \mathcal{X}$, then the linear programming problems for such two values of U should have the same feasibility set.

We now explain what is meant by "similarly ordered". Recall the definition of sets $\mathcal{A}_{x,1,U}$, $\mathcal{A}_{x,2,U}$, $\mathcal{A}_{x,3,U}$, and $\mathcal{A}_{x,U}$ in Section 4.5.1. For every $l \in \{1, 2, 3\}$, list the 3 elements of $\mathcal{A}_{x,l,U}$ in a *column vector*, $\alpha_{x,l,U}$, so that

$$\alpha_{x,1,U} \equiv (-U_{x1}, \infty, -\infty)^\top, \alpha_{x,2,U} \equiv (-U_{x2}, \infty, -\infty)^\top, \alpha_{x,3,U} \equiv (U_{x2} - U_{x1}, \infty, -\infty)^\top.$$

Similarly, list the twenty-seven 3-tuples of set $\mathcal{A}_{x,U}$ in a 27×3 matrix, $\alpha_{x,U}$, which we report in Appendix C. Lastly, list the 27 elements of the image set of $\Delta\bar{F}^U_x$ in a column vector, $f_{x,U}$.

Consider any $U, \tilde{U}$ in $\mathcal{U}$. Note that, for every $l \in \{1,2,3\}$, both $\alpha_{x,l,U}$ and $\alpha_{x,l,\tilde{U}}$ have the smallest element listed in the third row, the biggest element listed in the second row, and the intermediate element listed in the first row, for every $x \in \mathcal{X}$. This implies that the rows of $\alpha_{x,U}$ that are comparable in the sense of (26) have the same row-indices as the comparable rows of $\alpha_{x,\tilde{U}}$, for every $x \in \mathcal{X}$. Therefore, the rows of $f_{x,U}$ which are restricted by (26) have the same row indices as the restricted rows of $f_{x,\tilde{U}}$, for every $x \in \mathcal{X}$. As a result, the feasible set of the linear programming for U is equal to that for $\tilde{U}$.

The above idea can be generalised as follows. For each $l \in \{1, ..., d\}$, let k denote the cardinality of $\mathcal{A}_{x,l,U}$. Let Π_1 be the set of all possible permutations without repetition of $\{1, ..., k\}$ and let $\Pi_2 \equiv \{<, =\}^{k-1}$. Lastly, let

$$\pi : \bar{\mathbb{R}}^k \to \Pi_1 \times \Pi_2,$$

where

$$\pi(\alpha) \equiv (\pi_1(\alpha), \pi_2(\alpha)), \ \forall \alpha \in \bar{\mathbb{R}}^k.$$

In particular, $\pi_1(\alpha)$ sorts the k elements of α from smallest to largest and reports their positions in the original vector. $\pi_2(\alpha)$ reports the relational operators, $<$ or $=$, among the sorted elements of α. When α contains multiple elements with the same value, then we can adopt any convention on which element to sort first. We call $\pi(\alpha)$ the "π-ordering" of α. For instance, suppose $\alpha = (100, 99, \infty)$. Then, $\pi(\alpha) = \{(2,1,3), (<,<)\}$. Suppose $\alpha = (5, 5, -\infty)$. Then, $\pi(\alpha) = \{(3,2,1), (<,=)\}$. By following arguments similar to those outlined in the previous paragraph, it can be shown that if $U, \tilde{U}$ in $\mathcal{U}$ have the same π-ordering for every $l \in \{1, ..., d\}$ and $x \in \mathcal{X}$, then the feasible set of the linear programming for U is equal to the feasible set of the linear programming for $\tilde{U}$. Proposition 3 summarises such a finding.

Proposition 3. (*Simplifying grid search over* $\mathcal{U}$) Let $\Delta\mathcal{F}^\dagger$ satisfy any of the restrictions listed in Assumption 5. Take any $U, \tilde{U}$ in $\mathcal{U}$ such that $\pi(\alpha_{x,l,U}) = \pi(\alpha_{x,l,\tilde{U}})$ for every $l \in \{1, ..., d\}$ and $x \in \mathcal{X}$. Then,

$$\exists \ \{\Delta F_x\}_{x \in \mathcal{X}} \in \Delta\mathcal{F}^\dagger \text{ s.t. } p_{y|x} = \kappa(U, \Delta F_x, y) \ \forall (x,y) \in \mathcal{X} \times \mathcal{Y}_0,$$

if and only if

$$\exists \ \{\Delta F_x\}_{x \in \mathcal{X}} \in \Delta\mathcal{F}^\dagger \text{ s.t. } p_{y|x} = \kappa(\tilde{U}, \Delta F_x, y) \ \forall (x,y) \in \mathcal{X} \times \mathcal{Y}_0.$$

$\diamond$

Remark A.1 in Appendix A explains how Proposition 3 is implemented in practice.

4.5.3 Third step

Recall from Section 4.5 that the analyst has to ensure that $\{\Delta F_x\}_{x\in\mathcal{X}} \in \Delta\mathcal{F}^\dagger$, that solves $p_{y|x} = \kappa(U, \Delta F_x, y)$ for every $(x, y) \in \mathcal{X} \times \mathcal{Y}_0$, has support contained in the region $\mathcal{B}$, i.e.,

$$\lambda_{\Delta F_x}(\mathcal{B}) = 1 \; \forall x \in \mathcal{X}. \tag{27}$$

We start with providing an equivalent characterisation of (27) in terms of zero measure conditions on boxes in $\mathbb{Q}^d$, where $\mathbb{Q} \subset \mathbb{R}$ is the set of rational numbers. The underlying intuition is that if $\lambda_{\Delta F_x}(\mathcal{B}) = 1$, then any d-dimensional box in $\mathbb{R}^d$ not intersecting the region $\mathcal{B}$ has probability measure zero. Conversely, it is sufficient to impose such a zero probability measure condition for all the d-dimensional boxes in $\mathbb{Q}^d$ not intersecting the region $\mathcal{B}$ to satisfy $\lambda_{\Delta F_x}(\mathcal{B}) = 1$.

Proposition 4. (*Degeneracy condition*) For any $(\hat{b}, \tilde{b}) \in \mathbb{R}^2$, consider the d-dimensional boxes in $\mathbb{R}^d$

$$B_{t,p,q}(\hat{b}, \tilde{b}) \equiv \{(z_1, ..., z_d) \in \mathbb{R}^d\colon \; z_t > \hat{b} + \tilde{b}, \;\; z_p \leq \hat{b}, \;\; z_q \leq \tilde{b}\},$$

and

$$Q_{t,p,q}(\hat{b}, \tilde{b}) \equiv \{(z_1, ..., z_d) \in \mathbb{R}^d\colon \; z_t \leq \hat{b} + \tilde{b}, \;\; z_p > \hat{b}, \;\; z_q > \tilde{b}\},$$

$\forall t \in \{1, ..., r-1\}$ and $\forall (p, q) \in \{(t+1, r+1), (t+2, r+2), ..., (r, d)\}$. For each $\{\Delta F_x\}_{x\in\mathcal{X}} \in \Delta\mathcal{F}^\dagger$ and $x \in \mathcal{X}$,

$$\begin{aligned} &\lambda_{\Delta F_x}(B_{t,p,q}(\hat{b}, \tilde{b})) = \lambda_{\Delta F_x}(Q_{t,p,q}(\hat{b}, \tilde{b})) = 0, \\ &\qquad \forall t \in \{1, ..., r-1\}, \; \forall (p, q) \in \{(t+1, r+1), (t+2, r+2), ..., (r, d)\}, \forall (b, \tilde{b}) \in \mathbb{Q}^2, \end{aligned} \tag{28}$$

if and only if $\lambda_{\Delta F_x}(\mathcal{B}) = 1$. $\diamond$

We add two key remarks. First, (28) can be more explicitly rewritten using ΔF_x. For instance, if $r = 2$, then (28) is equivalent to

$$\begin{aligned} &\Delta F_x(\infty, \hat{b}, \tilde{b}) - \Delta F_x(\hat{b} + \tilde{b}, \hat{b}, \tilde{b}) = 0, \\ &\Delta F_x(\hat{b} + \tilde{b}, \infty, \infty) - \Delta F_x(\hat{b} + \tilde{b}, \hat{b}, \infty) - \Delta F_x(\hat{b} + \tilde{b}, \infty, \tilde{b}) + \Delta F_x(\hat{b} + \tilde{b}, \hat{b}, \tilde{b}) = 0, \end{aligned} \tag{29}$$

where both equations are linear in ΔF_x.[15] Second, given our first remark, (28) constitutes a countably *infinite* number of equations that are linear in ΔF_x. As such, we cannot incorporate all of them in the linear programming problem of Section 4.5.1. To operationalise (28), we suggest to consider only a *finite* number of 2-tuples, $(b, \tilde{b})$, from $\mathbb{Q}^2$. In practice, this means that the obtained identified set is an outer set of the sharp one. It can be made arbitrarily close to being sharp by considering an arbitrarily large number of 2-tuples, depending on the available computational resources. In our simulations and empirical application, we have drawn such 2-tuples at random. We have also constructed the identified set while considering various numbers of 2-tuples and

[15] See Appendix C for details.

obtained negligible differences among the resulting regions.

In Appendix D, we provide an example of a linear programming problem which lists all the conditions in Section 4.5.1 and Section 4.5.3.

5 Simulations

In this section, we implement the methodology described in Section 4.5 using simulated data. In order to ensure that the volume of our identified set is not improperly inflated relative to the point identified case, we impose some scale normalisations.[16] In particular, when Assumption 5.1 (independence between $\Delta\epsilon_i$ and X_i) is not imposed, we divide each of $\{U_{xy}\}_{y\in\mathcal{Y}}$ by $|U_{x1}|$ for every $x \in \mathcal{X}$ and each of $\{V_{xy}\}_{x\in\mathcal{X}}$ by $|V_{1y}|$ for every $y \in \mathcal{X}$. Hence, the scale normalisations are

$$U_{x1} \in \{-1, 1\} \; \forall x \in \mathcal{X} \quad \text{and} \quad V_{1y} \in \{-1, 1\} \; \forall y \in \mathcal{Y}. \tag{30}$$

Instead, when Assumption 5.1 is imposed, we divide each element of U by $|U_{11}|$ and each element of V by $|V_{11}|$. Hence, the scale normalisations are

$$U_{11} \in \{-1, 1\} \quad \text{and} \quad V_{11} \in \{-1, 1\}. \tag{31}$$

We refer the reader to Appendix E.1 for a thorough discussion of (30) and (31).

As a first exercise, we investigate the identifying power of the 1to1TU model when we consider various combinations of the distributional restrictions listed in Assumption 5. In particular, we fix $\mathcal{X} = \mathcal{Y} \equiv \{1, ..., r\}$ with $r = 2$, $\Phi^{\text{true}}_{11} = \Phi^{\text{true}}_{22} = 4$, and $\Phi^{\text{true}}_{12} = \Phi^{\text{true}}_{21} = 1$, where the superscript "true" highlights the true parameter values. We simulate data under three DGPs, which differ in $\{p_x\}_{x\in\mathcal{X}}$, $\{p_y\}_{y\in\mathcal{Y}}$, $\{F_x\}_{x\in\mathcal{X}}$, and $\{G_y\}_{y\in\mathcal{Y}}$:

(DGP1) $\{\epsilon_{iy}\}_{y\in\mathcal{Y}_0}$ are i.i.d., where ϵ_{iy} is distributed independently from X_i, as standard Extreme Value Type I . The proportion of men of type 1 is 1/2. Analogous assumptions are imposed on the women's side.

(DGP2) ϵ_i is distributed independently of X_i as a normal mixture, with 2 equally weighted components. The mean and variance-covariance matrix of each mixture component are reported in Appendix E.2. Analogous assumptions are imposed on $\{G_y\}_{y\in\mathcal{Y}}$. The proportion of men of type 1 is 1/6. The proportion of women of type 1 is 1/5.

(DGP3) ϵ_i is distributed as a normal mixture, with 2 equally weighted components, whose variance-covariance matrix varies across $x \in \mathcal{X}$. The mean and variance-covariance ma-

[16] Note that appropriate location normalisations have been already imposed by Proposition 1, i.e., $U_{x0} = V_{0y} = 0$ for every $(x, y) \in \mathcal{X} \times \mathcal{Y}$.

trix of each mixture component are reported in Appendix E.2. The proportion of men of type 1 is 1/2. Analogous assumptions are imposed on the women's side.

We emphasise that, in each of the three DGPs above, if $\{F_x\}_{x\in\mathcal{X}}$ and $\{G_y\}_{y\in\mathcal{Y}}$ are assumed to be fully known by the analyst, then U, V, and Φ are point identified, as shown by Galichon and Salanié (2020). We consider the seven specifications of distributional assumptions summarised in Table 1. In Table 2, we report the true values and the identified sets of U, V, and Φ under

Assumptions	[1]	[2]	[3]	[4]	[5]	[6]	[7]
5.1	✓	✓					
5.2		✓	✓	✓		✓	
5.3		✓		✓		✓	
5.4					✓	✓	✓
5.5							✓

Table 1: Assumptions on the unobserved heterogeneity maintained in the different specifications.

the three DGPs above and for each specification of Table 1.[17] We also report the true values and the identified sets of $D_{22,11}(\Phi)$, $C_{21}(U)$, and $C_{21}(V)$, defined in Section 4.2. Moreover, we report the values of U, V, Φ, $D_{22,11}(\Phi)$, $C_{21}(U)$, and $C_{21}(V)$ that are recovered by using the Multinomial Logit formulas of Choo and Siow (2006) (hereafter, CS values). We distinguish between the case when Assumption 5.1 is imposed ("w/ 5.1") and the case when Assumption 5.1 is not imposed ("w/o 5.1") because, as highlighted by (30) and (31), these two cases entail different scale normalisations.[18] Further, note that the true values and the CS values of U, V, Φ, $D_{22,11}(\Phi)$, $C_{21}(U)$, and $C_{21}(V)$ that are reported in Table 2 are also normalised according to (30) and (31), in order to ensure that they belong to our identified sets. For instance, the normalised true value of Φ_{22} under Assumption 5.1 equals $\frac{U_{22}^{\text{true}}}{|U_{11}^{\text{true}}|}+\frac{V_{22}^{\text{true}}}{|V_{11}^{\text{true}}|}$ and, hence, may be different from the non-normalised true value, $\Phi_{22}^{\text{true}}=4$. Lastly, note that in DGP3 we do not consider specifications [1] and [2] of Table 1. This is because DGP3 features conditional heteroskedasticity and, hence, [1] and [2] would be misspecified.

We highlight a few facts. First, in each of the three DGPs considered, specifications [5], [6], and [7] deliver the tightest bounds for U, V, Φ, $C_{21}(U)$, and $C_{21}(V)$. Moreover, specifications [5], [6], and [7] generate numerically identical bounds. Also specifications [3] and [4] generate numerically identical bounds.

Second, in none of the cases considered, the identified set of $D_{22,11}(\Phi)$ is bounded on both sides. This is because, there is always at least one component of Φ whose identified set is unbounded on at least one side. In particular, under DGP1 and DGP3, the upper bound of $D_{22,11}(\Phi)$ is always ∞ because the upper bound of U_{22} or V_{22} is always ∞.

Third, in DGP1 and DGP3 the sign of $D_{22,11}(\Phi)$ is recovered unambiguously most of the times. Instead, in DGP2, the sign of $D_{22,11}(\Phi)$ is never identified. As discussed in Section 4.2, detecting

[17] To compute U^{true} and V^{true} under DGP2 and DGP3, we use the formula provided by Proposition 2 in Galichon and Salanie (2020). More details are in Appendix E.3.

[18] Consequently, the corresponding identified sets are not necessarily nested.

the sign of $D_{22,11}(\Phi)$ is important in itself because it reveals the direction of assortativeness. Graham (2011; 2013b) shows that if the taste shocks are i.i.d., then the sign of $D_{22,11}(\Phi)$ is identified. Our simulations highlight that i.i.d.-ness is not a necessary condition.

Fourth, in none of the cases considered, the identified sets of $C_{21}(U)$ and $C_{21}(V)$ are bounded on both sides. In particular, under DGP1 and DGP3, the upper bound of $C_{21}(U)$ and $C_{21}(V)$ is always ∞ because the upper bound of U_{22} and V_{22} is always ∞. If types represent education levels, such a result reveals that the identified set of the marital education premium will also be unbounded on at least one side (see Equation (3)).

Fifth, note that in DGP1 and DGP3 the signs of $C_{21}(U)$ and $C_{21}(V)$ are recovered unambiguously under specifications [5], [6], and [7]. Instead, in DGP2, the signs of $C_{21}(U)$ and $C_{21}(V)$ are never identified.

Further, DGP1 and DGP3 produce identified sets of U, V, Φ, and $D_{22,11}(\Phi)$ that are identical. This is because DGP1 and DGP3 generate very similar empirical match probabilities.

Lastly, in DGP2 and DGP3 the assumption that the taste shocks are i.i.d. standard Extreme Value Type I is misspecified. This implies that the CS values are different, sometimes quite significantly, from the true values of U, V, Φ, $D_{22,11}(\Phi)$, $C_{21}(U)$, and $C_{21}(V)$. Nevertheless, the CS values always belong to our identified sets. This is by construction, because the collections of conditional CDFs contemplated by the specifications in Table 1 include the Extreme Value family.

As a second exercise, we investigate how the identifying power of the 1to1TU model varies as the number of types increases. In particular, we simulate data under three DGPs, featuring $r = 3$, $r = 4$, and $r = 5$ for both sides of the market, respectively. In each DGP, $\{\epsilon_{iy}\}_{y \in \mathcal{Y}_0}$ are i.i.d., where ϵ_{iy} is distributed independently from X_i, as standard Extreme Value Type I . Types are represented in equal proportions. $\Phi_{xx} = 4$ and $\Phi_{x\tilde{x}} = 1$ for every $x, \tilde{x} \in \mathcal{X}$. In Tables 3 and 4, we report the normalised true values and the identified sets of U, V, Φ, $D(\Phi)$, $C(U)$, and $C_{21}(V)$, under such three DGPs and for specifications [5] and [6] of Table 1.

Overall, the findings of Table 2 are confirmed. We highlight a few additional facts. First, differently from Table 2, here specification [6] leads to bounds tighter than those produced by specification [5] in some cases. For instance, see the identified set of U_{13} when $r = 3$.

Second, in none of the cases considered, the identified sets of $D(\Phi)$, $C(U)$, and $C(V)$ are bounded on both sides. Most of the times, we recover unambiguously the signs of the $D(\Phi)$'s components under specification [6]. We identify the signs of the $C(U)$ and $C(V)$'s components less often. In fact, in many cases (in all cases for $r = 5$), the identified sets of $C(U)$ and $C(V)$ are unbounded on both sides.

	Specifications from Table 1	U_{11}	U_{12}	U_{21}	U_{22}	V_{11}	V_{12}	V_{21}	V_{22}	Φ_{11}	Φ_{12}	Φ_{21}	Φ_{22}	$D_{22,11}(\Phi)$	$C_{21}(U)$	$C_{21}(V)$
w/ 5.1 DGP1	True & CS	1	0.3062	0.1761	0.8674	1	0.2226	0.2764	0.9595	2	0.5289	0.4525	1.8269	2.8456	-0.1399	-0.0427
	[1]	$\{-1,1\}$	$(-\infty,\infty)$	$(-\infty,0.9]$	$(-\infty,\infty)$	$\{-1,1\}$	$(-\infty,0.9]$	$(-\infty,\infty)$	$(-\infty,\infty)$	$\{-2,2\}$	$(-\infty,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$
	[2]	1	$(-\infty,0.9]$	$(-\infty,0.9]$	$[0.1,\infty)$	1	$(-\infty,0.9]$	$(-\infty,0.9]$	$[0.1,\infty)$	2	$(-\infty,1.8]$	$(-\infty,1.8]$	$[0.2,\infty)$	$[0.4,\infty)$	$[-4.1048,\infty)$	$[-4.0832,\infty)$
w/o 5.1 DGP1	True & CS	1	0.3062	1	4.9262	1	1	0.2764	4.3103	2	1.3062	1.2764	9.2365	8.6539	2.9329	2.5393
	[3]	1	$(-\infty,0.8]$	$\{-1,1\}$	$[0.2,\infty)$	1	$\{-1,1\}$	$(-\infty,0.8]$	$[0.2,\infty)$	2	$(-\infty,1.8]$	$(-\infty,1.8]$	$[0.4,\infty)$	$[0.8,\infty)$	$[-0.8987,\infty)$	$[-0.8872,\infty)$
	[4]	1	$(-\infty,0.8]$	$\{-1,1\}$	$[0.2,\infty)$	1	$\{-1,1\}$	$(-\infty,0.8]$	$[0.2,\infty)$	2	$(-\infty,1.8]$	$(-\infty,1.8]$	$[0.4,\infty)$	$[0.8,\infty)$	$[-0.8987,\infty)$	$[-0.8872,\infty)$
	[5]	1	$[-0.8,0.8]$	1	$[1.2,\infty)$	1	1	$[-0.8,0.8]$	$[1.2,\infty)$	2	$(0.2,1.8]$	$(0.2,1.8]$	$[2.4,\infty)$	$[0.8,\infty)$	$[0.1530,\infty)$	$[0.1724,\infty)$
	[6]	1	$[-0.8,0.8]$	1	$[1.2,\infty)$	1	1	$[-0.8,0.8]$	$[1.2,\infty)$	2	$(0.2,1.8]$	$(0.2,1.8]$	$[2.4,\infty)$	$[0.8,\infty)$	$[0.1530,\infty)$	$[0.1724,\infty)$
	[7]	1	$[-0.8,0.8]$	1	$[1.2,\infty)$	1	1	$[-0.8,0.8]$	$[1.2,\infty)$	2	$(0.2,1.8]$	$(0.2,1.8]$	$[2.4,\infty)$	$[0.8,\infty)$	$[0.1530,\infty)$	$[0.1724,\infty)$
w/ 5.1 DGP2	True	1	1.0432	-0.3448	0.3055	1	-2.0792	1.1369	1.8940	2	-1.0360	0.7921	2.1994	4.4433	-0.7357	0.1682
	CS	1	2.0198	-0.1766	1.8872	1	-0.7420	3.3497	3.4288	2	1.2778	3.1731	5.3160	2.8651	-0.2050	0.0470
	[1]	$\{-1,1\}$	$(-\infty,\infty)$	$(-\infty,0.8]$	$(-\infty,\infty)$	$\{-1,1\}$	$(-\infty,0.8]$	$(-\infty,\infty)$	$(-\infty,\infty)$	$\{-2,2\}$	$(-\infty,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$
	[2]	$\{-1,1\}$	$[0.2,\infty)$	$(-\infty,0.8]$	$[0.2,\infty)$	$\{-1,1\}$	$(-\infty,0.8]$	$[0.2,\infty)$	$[0.2,\infty)$	$\{-2,2\}$	$(-\infty,\infty)$	$(-\infty,\infty)$	$[0.4,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$
w/o 5.1 DGP2	True	1	1.0432	-1	0.8860	1	-1	1.1369	0.9109	2	0.0432	0.1369	1.7969	3.6167	-0.4340	-0.4052
	CS	1	2.0198	-1	10.6845	1	-1	3.3497	4.6213	2	1.0198	2.3497	15.3058	13.9364	5.7047	0.8688
	[3]	$\{-1,1\}$	$[0.5,\infty)$	$\{-1,1\}$	$[0.5,\infty)$	$\{-1,1\}$	$\{-1,1\}$	$[0.5,\infty)$	$[0.5,\infty)$	$\{-2,2\}$	$(-0.5,\infty)$	$[-0.5,\infty)$	$[1,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$
	[4]	$\{-1,1\}$	$[0.5,\infty)$	$\{-1,1\}$	$[0.5,\infty)$	$\{-1,1\}$	$\{-1,1\}$	$[0.5,\infty)$	$[0.5,\infty)$	$\{-2,2\}$	$(-0.5,\infty)$	$[-0.5,\infty)$	$[1,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$
	[5]	1	$[1.01,\infty)$	-1	$[1,\infty)$	1	-1	$[1.01,\infty)$	$[1,\infty)$	2	$[0.01,\infty)$	$[0.01,\infty)$	$[2,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$
	[6]	1	$[1.01,\infty)$	-1	$[1,\infty)$	1	-1	$[1.01,\infty)$	$[1,\infty)$	2	$[0.01,\infty)$	$[0.01,\infty)$	$[2,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$
	[7]	1	$[1.01,\infty)$	-1	$[1,\infty)$	1	-1	$[1.01,\infty)$	$[1,\infty)$	2	$[0.01,\infty)$	$[0.01,\infty)$	$[2,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$	$(-\infty,\infty)$
w/o 5.1 DGP3	True	1	0.4464	1	2.9478	1	1	0.1981	2.2425	2	1.4464	1.1981	5.1903	4.5458	1.5016	1.0546
	CS	1	0.0543	1	10.2726	1	1	0.0058	7.0958	2	1.0543	1.0058	17.3684	17.3083	6.7941	4.5653
	[3]	1	$(-\infty,0.8]$	$\{-1,1\}$	$[0.2,\infty)$	1	$\{-1,1\}$	$(-\infty,0.8]$	$[0.2,\infty)$	2	$(-\infty,1.8]$	$(-\infty,1.8]$	$[0.4,\infty)$	$[0.8,\infty)$	$[-0.8275,\infty)$	$[-0.8243,\infty)$
	[4]	1	$(-\infty,0.8]$	$\{-1,1\}$	$[0.2,\infty)$	1	$\{-1,1\}$	$(-\infty,0.8]$	$[0.2,\infty)$	2	$(-\infty,1.8]$	$(-\infty,1.8]$	$[0.4,\infty)$	$[0.8,\infty)$	$[-0.8275,\infty)$	$[-0.8243,\infty)$
	[5]	1	$[-0.8,0.8]$	1	$[1.2,\infty)$	1	1	$[-0.8,0.8]$	$[1.2,\infty)$	2	$[0.2,1.8]$	$[0.2,1.8]$	$[2.4,\infty)$	$[0.8,\infty)$	$[0.1960,\infty)$	$[0.2105,\infty)$
	[6]	1	$[-0.8,0.8]$	1	$[1.2,\infty)$	1	1	$[-0.8,0.8]$	$[1.2,\infty)$	2	$[0.2,1.8]$	$[0.2,1.8]$	$[2.4,\infty)$	$[0.8,\infty)$	$[0.1960,\infty)$	$[0.2105,\infty)$
	[7]	1	$[-0.8,0.8]$	1	$[1.2,\infty)$	1	1	$[-0.8,0.8]$	$[1.2,\infty)$	2	$[0.2,1.8]$	$[0.2,1.8]$	$[2.4,\infty)$	$[0.8,\infty)$	$[0.1960,\infty)$	$[0.2105,\infty)$

Table 2: Projections of the identified sets of U, V, Φ, $D_{22,11}(\Phi)$, $C_{21}(U)$, and $C_{21}(V)$ in the first simulation exercise.

	$r=3$			$r=4$			$r=5$		
Parameters	True & CS	[5]	[6]	True & CS	[5]	[6]	True & CS	[5]	[6]
U_{11}	1	1	1	1	1	1	1	1	1
U_{12}	0.2668	$[-0.8,0.8]$	$[-0.8,0.6]$	0.2774	$[-0.6,0.8]$	$[-0.6,0.8]$	0.3218	$[-0.7,0.4]$	$[-0.7,0.4]$
U_{13}	0.2775	$[-0.6,2.6]$	$[-0.6,0.8]$	0.2060	$[-0.8,0.6]$	$[-0.8,0.6]$	0.2354	$(-\infty,0.4]$	$(-\infty,-0.4]$
U_{14}				0.2292	$[-0.6,2.2]$	$[-0.6,2.2]$	0.2878	$[-0.7,3.8]$	$[-0.7,3.8]$
U_{15}							0.2878	$[-0.7,3.8]$	$[-0.7,3.8]$
U_{21}	1	1	1	1	1	1	1	$\{-1,1\}$	$\{-1,1\}$
U_{22}	3.8223	$[1.2,\infty)$	$[1.2,\infty)$	3.4317	$[1.2,\infty)$	$[1.2,\infty)$	6.8110	$[2.3,\infty]$	$[2.3,\infty]$
U_{23}	1.0129	$(-\infty,\infty)$	$(-\infty,\infty)$	1	$(-\infty,\infty)$	$(-\infty,\infty)$	0.3883	$(-\infty,\infty)$	$(-\infty,\infty)$
U_{24}				1.0695	$(-\infty,\infty)$	$(-\infty,\infty)$	1	$(-\infty,\infty)$	$(-\infty,\infty)$
U_{25}							1.3587	$(-\infty,\infty)$	$(-\infty,\infty)$
U_{31}	1	1	1	1	1	1	1	$\{-1,1\}$	$\{-1,1\}$
U_{32}	1.0204	$[1.2,\infty)$	$[1.2,\infty)$	1.0881	$[1.2,\infty)$	$[1.2,\infty)$	1.1246	$(-\infty,\infty)$	$(-\infty,\infty)$
U_{33}	5.3120	$[1.4,\infty)$	$[1.4,\infty)$	4.2173	$[1.4,\infty)$	$[1.4,\infty)$	3.7145	$[2.1,\infty)$	$[2.1,\infty)$
U_{34}				1	$(-\infty,\infty)$	$(-\infty,\infty)$	0.8665	$(-\infty,5.8]$	$(-\infty,5.8]$
U_{35}							0.8102	$(-\infty,\infty)$	$(-\infty,\infty)$
U_{41}				1	1	1	1	1	1
U_{42}				0.9144	$[-0.8,0.8]$	$[-0.8,0.8]$	1.1770	$(-\infty,\infty)$	$(-\infty,\infty)$
U_{43}				1.0617	$[0.2,\infty)$	$[0.2,\infty)$	1.1078	$(-\infty,5.9]$	$(-\infty,5.9]$
U_{44}				3.9006	$[0.4,\infty)$	$[0.6,\infty)$	4.5592	$[2,\infty)$	$[2,\infty)$
U_{45}							1.3727	$(-\infty,\infty)$	$(-\infty,\infty)$
U_{51}							1	1	1
U_{52}							0.8632	$[-0.7,1.6]$	$[-0.7,1.6]$
U_{53}							1.1269	$[0.3,\infty)$	$[0.3,\infty)$
U_{54}							1.0261	$(-\infty,5.9]$	$(-\infty,5.9]$
U_{55}							3.6316	$[0.6,\infty)$	$[0.6,\infty)$
V_{11}	1	1	1	1	1	1	1	1	1
V_{12}	1	1	1	1	1	1	1	$\{-1,1\}$	$\{-1,1\}$
V_{13}	1	1	1	1	1	1	1	$\{-1,1\}$	$\{-1,1\}$
V_{14}				1	1	1	1	1	1
V_{15}							1	1	1
V_{21}	0.2716	$[-0.6,0.8]$	$[-0.6,0.8]$	0.2107	$[-0.8,0.8]$	$[-0.8,0.8]$	0.1779	$[-0.7,0.4]$	$[-0.7,0.4]$
V_{22}	3.7589	$[1.2,\infty)$	$[1.2,\infty)$	2.9469	$[1.2,\infty)$	$[1.2,\infty)$	3.6739	$[2.6,\infty)$	$[2.6,\infty)$
V_{23}	1.2428	$[1.2,\infty)$	$[1.2,\infty)$	1.3043	$[1.2,\infty)$	$[1.2,\infty)$	0.7883	$(-\infty,\infty)$	$(-\infty,\infty)$
V_{24}				1.7284	$[1.2,\infty)$	$[1.2,\infty)$	0.8864	$(-\infty,4.9]$	$(-\infty,4.9]$
V_{25}							1.1024	$[0.3,\infty)$	$[0.3,\infty)$
V_{31}	0.2285	$[-0.8,0.6]$	$[-0.8,0.6]$	0.2635	$[-0.8,2.6]$	$[-0.8,1.4]$	0.2688	$[-0.7,1.6]$	$[-0.7,1.6]$
V_{32}	1.0389	$(-\infty,\infty)$	$(-\infty,\infty)$	1.1851	$(-\infty,\infty)$	$(-\infty,\infty)$	1.2342	$(-\infty,\infty)$	$(-\infty,\infty)$
V_{33}	5.3601	$[1.4,\infty)$	$[1.4,\infty)$	4.3093	$[1.4,\infty)$	$[1.4,\infty)$	4.8803	$[1.7,\infty)$	$[1.7,\infty)$
V_{34}				2.0316	$[1.4,\infty)$	$[1.4,\infty)$	1.1433	$[0.3,\infty)$	$[0.3,\infty)$
V_{35}							1.0688	$(-\infty,\infty)$	$(-\infty,\infty)$
V_{41}				0.1995	$[-1.2,0.8]$	$[-0.8,0.6]$	0.1942	$[-0.7,0.4]$	$[-0.7,0.4]$
V_{42}				0.9118	$(-\infty,\infty)$	$(-\infty,\infty)$	0.9721	$(-\infty,5.8]$	$(-\infty,5.8]$
V_{43}				1.3259	$(-\infty,\infty)$	$(-\infty,\infty)$	1.3373	$(-\infty,5.9]$	$(-\infty,5.9]$
V_{44}				9.1210	$[1.6,\infty)$	$[1.6,\infty)$	4.5940	$[2,\infty)$	$[2,\infty)$
V_{45}							1.3243	$[0.3,\infty)$	$[0.3,\infty)$
V_{51}							0.2329	$[-0.7,1.6]$	$[-07,1.6]$
V_{52}							0.8259	$(-\infty,\infty)$	$(-\infty,\infty)$
V_{53}							1.5688	$(-\infty,\infty)$	$(-\infty,\infty)$
V_{54}							1.1778	$(-\infty,\infty)$	$(-\infty,\infty)$
V_{55}							4.4209	$[1.7,\infty)$	$[1.7,\infty)$
Φ_{11}	2	2	2.	2	2	2	2	2	2
Φ_{12}	1.2668	$[0.2,1.8]$	$[0.2,1.6]$	1.2774	$[0.4,1.8]$	$[0.4,1.8]$	1.3218	$[-1.7,1.4]$	$[-1.7,1.4]$
Φ_{13}	1.2775	$[0.4,3.6]$	$[0.4,1.8]$	1.2060	$[0.2,1.6]$	$[0.2,1.6]$	1.2354	$(-\infty,1.4]$	$(-\infty,1.4]$
Φ_{14}				1.2292	$[0.4,3.2]$	$[0.4,2.4]$	1.2878	$[0.3,4.8]$	$[0.3,4.8]$
Φ_{15}							1.2878	$[0.3,4.8]$	$[0.3,4.8]$
Φ_{21}	1.2716	$[0.4,1.8]$	$[0.4,1.8]$	1.2107	$[0.2,1.8]$	$[0.2,1.6]$	1.1779	$[-1.7,1.4]$	$[-1.7,1.4]$
Φ_{22}	7.5812	$[2.4,\infty)$	$[2.4,\infty)$	6.3785	$[2.4,\infty)$	$[2.4,\infty)$	10.4849	$[4.9,\infty)$	$[4.9,\infty)$
Φ_{23}	2.2557	$(-\infty,\infty)$	$(-\infty,\infty)$	2.3043	$(-\infty,\infty)$	$(-\infty,\infty)$	1.1767	$(-\infty,\infty)$	$(-\infty,\infty)$
Φ_{24}				2.7979	$(-\infty,\infty)$	$(-\infty,\infty)$	1.8864	$(-\infty,\infty)$	$(-\infty,\infty)$
Φ_{25}							2.4611	$(-\infty,\infty)$	$(-\infty,\infty)$
Φ_{31}	1.2285	$[0.2,1.6]$	$[0.2,1.6]$	1.2635	$[0.2,3.6]$	$[0.2,2.4]$	1.2688	$[-1.7,2.6]$	$[-1.7,2.6]$
Φ_{32}	2.0593	$(-\infty,\infty)$	$(-\infty,\infty)$	2.2732	$(-\infty,\infty)$	$(-\infty,\infty)$	2.3588	$(-\infty,\infty)$	$(-\infty,\infty)$
Φ_{33}	10.7722	$[2.8,\infty)$	$[2.8,\infty)$	8.5266	$[2.8,\infty)$	$[2.8,\infty)$	8.5948	$[3.8,\infty)$	$[3.8,\infty)$
Φ_{34}				3.0316	$(-\infty,\infty)$	$(-\infty,\infty)$	2.0098	$(-\infty,\infty)$	$(-\infty,\infty)$
Φ_{35}							1.8790	$(-\infty,\infty)$	$(-\infty,\infty)$
Φ_{41}				1.1995	$[-0.2,1.8]$	$[-0.2,1.6]$	1.1942	$[0.3,1.4]$	$[0.3,1.4]$
Φ_{42}				1.8262	$(-\infty,\infty)$	$(-\infty,\infty)$	2.1491	$(-\infty,\infty)$	$(-\infty,\infty)$
Φ_{43}				2.3876	$(-\infty,\infty)$	$(-\infty,\infty)$	2.4451	$(-\infty,11.8]$	$(-\infty,11.8]$
Φ_{44}				13.0215	$[2,\infty)$	$[2.2,\infty)$	9.1532	$[4,\infty)$	$[4,\infty)$
Φ_{45}							2.6970	$(-\infty,\infty)$	$(-\infty,\infty)$
Φ_{51}							1.2329	$[0.3,2.6]$	$[0.3,2.6]$
Φ_{52}							1.6892	$(-\infty,\infty)$	$(-\infty,\infty)$
Φ_{53}							2.6957	$(-\infty,\infty)$	$(-\infty,\infty)$
Φ_{54}							2.2039	$(-\infty,\infty)$	$(-\infty,\infty)$
Φ_{55}							8.0525	$[2.2,\infty)$	$[2.2,\infty)$

Table 3: Projections of the identified sets of U, V, and Φ in the second simulation exercise.

	$r=3$			$r=4$			$r=5$		
Parameters	True & CS	[5]	[6]	True & CS	[5]	[6]	True & CS	[5]	[6]
$D_{22,11}(\Phi)$	7.0428	$[0.8,\infty)$	$[1,\infty)$	5.8904	$[0.8,\infty)$	$[1,\infty)$	9.9852	$[5.1,\infty)$	$[5.1,\infty)$
$D_{33,11}(\Phi)$	10.2662	$[-0.4,\infty)$	$[1.4,\infty)$	8.0571	$[-0.4,\infty)$	$[0.8,\infty)$	8.0906	$[2.9,\infty)$	$[2.9,\infty)$
$D_{44,11}(\Phi)$				12.5929	$[-1,\infty)$	$[0.2,\infty)$	8.6711	$[-0.2,\infty)$	$[-0.2,\infty)$
$D_{55,11}(\Phi)$							7.5318.	$[-3.1,\infty)$	$[-3.1,\infty)$
$D_{33,22}(\Phi)$	14.0383	$[0.8,\infty)$	$[0.8,\infty)$	10.3276	$[0.6,\infty)$	$[0.6,\infty)$	15.5443	$(-\infty,\infty)$	$(-\infty,\infty)$
$D_{44,22}(\Phi)$				14.7760	$(-\infty,\infty)$	$(-\infty,\infty)$	15.6027	$(-\infty,\infty)$	$(-\infty,\infty)$
$D_{55,22}(\Phi)$							14.3872	$(-\infty,\infty)$	$(-\infty,\infty)$
$D_{44,33}(\Phi)$				16.1290	$[0.4,\infty)$	$[0.4,\infty)$	13.2931	$(-\infty,\infty)$	$(-\infty,\infty)$
$D_{55,33}(\Phi)$							12.0726	$(-\infty,\infty)$	$(-\infty,\infty)$
$D_{55,44}(\Phi)$							12.3048	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{21}(U)$	0.9727	$[0.1188,\infty)$	$[0.2033,\infty)$	1.6252	$[-0.1308,\infty)$	$[-0.1232,\infty)$	3.1561	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{31}(U)$	2.9421	$[0.2659,\infty)$	$[0.3511,\infty)$	1.9579	$[0.0754,\infty)$	$[0.0830,\infty)$	1.6026	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{41}(U)$				1.9116	$[-0.5684,\infty)$	$[-0.4466,\infty)$	2.1100	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{51}(U)$							1.5932	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{32}(U)$	0.9694	$(-\infty,\infty)$	$(-\infty,\infty)$	0.3327	$(-\infty,\infty)$	$(-\infty,\infty)$	−1.5534	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{42}(U)$				0.2865	$(-\infty,\infty)$	$(-\infty,\infty)$	−1.0461	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{52}(U)$							−1.5629	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{43}(U)$				−0.0463	$(-\infty,\infty)$	$(-\infty,\infty)$	0.5073	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{53}(U)$							−0.0095	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{54}(U)$							−0.5168	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{21}(V)$	2.0310	$[0.2264,\infty)$	$[0.2264,\infty)$	1.4064	$[-0.0795,\infty)$	$[0.0076,\infty)$	1.6035	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{31}(V)$	3.0458	$[0.3581,\infty)$	$[0.3581,\infty)$	2.2766	$[0.3191,\infty)$	$[0.4062,\infty)$	2.4680	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{41}(V)$				4.9699	$[0.4239,\infty)$	$[0.5110,\infty)$	2.1398	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{51}(V)$							2.0755	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{32}(V)$	1.0148	$(-\infty,\infty)$	$(-\infty,\infty)$	0.8702	$(-\infty,\infty)$	$(-\infty,\infty)$	0.8645	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{42}(V)$				3.5634	$(-\infty,\infty)$	$(-\infty,\infty)$	0.5363	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{52}(V)$							0.4720	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{43}(V)$				2.6932	$(-\infty,\infty)$	$(-\infty,\infty)$	−0.3282	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{53}(V)$							−0.3925	$(-\infty,\infty)$	$(-\infty,\infty)$
$C_{54}(V)$							−0.0643	$(-\infty,\infty)$	$(-\infty,\infty)$

Table 4: Projections of the identified sets of $D(\Phi)$, $C(U)$, and $C(V)$ in the second simulation exercise.

6 Empirical application

In this section, we use our methodology to re-examine some of the questions considered in the empirical literature on the marriage market that have been answered by relying on the Multinomial Logit 1to1TU model.

An important question is whether educational sorting has changed over time as this can be key to understand the sources of inequality in intergenerational outcomes (see references in Section 1). Detecting changes in educational sorting is challenging because it requires us to disentangle the effect of changes in the marginal probability distribution of education categories from potential structural changes in the match surplus. In fact, men and especially women have become more educated over time. This implies that individuals with similar education levels are mechanically more likely to marry. We are thus interested in capturing the changes in educational sorting after having accounted for the variations naturally arising due to such distributional variations in education.

The literature offers two approaches to do this. The first consists of using indices of sorting which rely on comparing empirical matches to a random matching counterfactual (e.g., Eika, Mogstad, and Zafar, 2019 and other references in Section 1). The second consists of using a structural model of the marriage market to estimate individual preferences. For instance, one can take the 1to1TU model with $\mathcal{X} = \mathcal{Y} \equiv \{1, ..., r\}$ listing education categories. Recall the definition of the supermodular core of Φ, denoted as $D(\Phi)$, from (1) and (2). As per Definition 1 of Chiappori, Salanié, and Weiss (2017) and Definition 7 of Chiappori, Costa-Dias, and Meghir (2020), if each component of the vector $D(\Phi)$ is strictly positive, then there is positive educational sorting. Further, a given time period t displays more positive educational sorting than another time period t' if and only if $D(\Phi^t) > D(\Phi^{t'})$ componentwise. The second approach has been implemented by Chiappori, Salanié, and Weiss (2017), Chiappori, Costa-Dias, and Meghir (2020), and other authors cited in Section 1 within the Multinomial Logit structure of Choo and Siow (2006).

Both approaches lead to the conclusion that positive educational sorting has overall increased in the U.S. in the past decades, although there is some debate about this trend when we distinguish among education categories. For instance, Eika, Mogstad, and Zafar (2019) find that positive educational sorting has declined among the highly educated and increased among the less educated. Instead, Chiappori, Salanié, and Weiss (2017) find that positive educational sorting has increased in each education category, particularly at the top of the education distribution. Similarly, Chiappori, Costa-Dias, and Meghir (2020) find that positive educational sorting has increased in each education category, with the exception of high-school graduates. Note, however, that the conclusions achieved via the structural approach, while having the advantages of being directly related to the underlying force that determines matching, may be driven by the Multinomial Logit structure. Our objective is to use the methodology of Section 4 to investigate the robustness of those conclusions and, specifically, whether they are obtainable also under nonparametric distributional restrictions on the unobserved heterogeneity, such as those listed in Assumption 5.

In addition to studying the evolution of educational sorting, this section touches upon another important question in the empirical literature on the marriage market. In particular, as discussed by Chiappori, Salanié, and Weiss (2017), the increase in educational sorting makes a higher stock of human capital more valuable on the marriage market. Therefore, one should also expect an increase in the marital education premium, especially at the highest levels of education and for women. Based on the Multinomial Logit 1to1TU model, Chiappori, Salanié, and Weiss (2017) empirically confirm such a prediction for the U.S. As discussed in Section 4.2, our methodology does not allow us to compute the marital education premium. Nevertheless, it permits us to get bounds on quantities that contribute to the marital education premium, namely $C(U)$ and $C(V)$, as outlined in (3), (4), and (5). In turn, such bounds will help us make certain conclusions on the marital education premium without relying on parametric distributional restrictions.

The remainder of the section is organised as follows: in Section 6.1, we describe the data; in Section 6.2, we present and interpret our results.

6.1 Data

We focus on the U.S. marriage market and take our data from the American Community Survey, that is a representative extract of the census. To construct the final dataset, we follow the steps outlined in Section I.A and Appendix B of Chiappori, Salanié, and Weiss (2017). In particular, from the $21,583,529$ households in the 2008 to 2014 waves, we take all white adults who are out of school. We record the education level of each adult, by distinguishing four categories as in Chiappori, Costa-Dias, and Meghir (2020): high school dropouts (HSD, or "1"); high school graduates (HSG, or "2"); some college (SC, or "3"); four-year college graduates and graduate degrees (CG, or "4").[19] We treat individuals as married if they define themselves as such, without including cohabitation. We focus on first marriages and never-married singles. The final sample consists of $1,502,157$ couples and $136,052$ singles.

We define cohorts by using year of birth and take women to be one year younger. For instance, cohort 1940 includes all men born in year 1940 and all women born in year 1941. In turn, the sample analogue of $p^{1940}_{y|x}$ is computed by taking the ratio between the number of men of type $x \in \mathcal{X}$ who are born in year 1940 and marry a woman of type $y \in \mathcal{Y}_0$ born in any year, and the number of men of type $x \in \mathcal{X}$ who are born in year 1940. Similarly, the sample analogue of $p^{1940}_{x|y}$ is computed by taking the ratio between the number of women of type $y \in \mathcal{Y}$ who are born in year 1941 and marry a man of type $x \in \mathcal{X}_0$ born in any year, and the number of women of type $y \in \mathcal{Y}$ who are born in year 1941.[20] In what follows, we provide some descriptive analysis using

[19]For the white population, Chiappori, Salanié, and Weiss (2017) further distinguish between four-year college graduates and graduate degrees. We have run our analysis also under such additional distinction and obtained less informative bounds.

[20]We ignore the issue of cohort mixing in order to exactly mimic the data construction process of Chiappori,

all of the 27 cohorts between cohort 1940 and 1966, as in Chiappori, Salanié, and Weiss (2017). In our empirical application, we focus on a few representative cohorts, namely cohorts 1940, 1945, 1950, 1955, 1960, and 1966.

Figure 1 reveals that the proportion of college educated men increases until 1950, then drops, and finally reverses into an increase around 1960. Instead, the proportion of college educated women always increases. Moreover, the proportion of college educated women is lower than that of men 1940, while the opposite is true by 1966. These changes imply that the evolution of educational sorting cannot be inferred by simply comparing matching patters across cohorts.

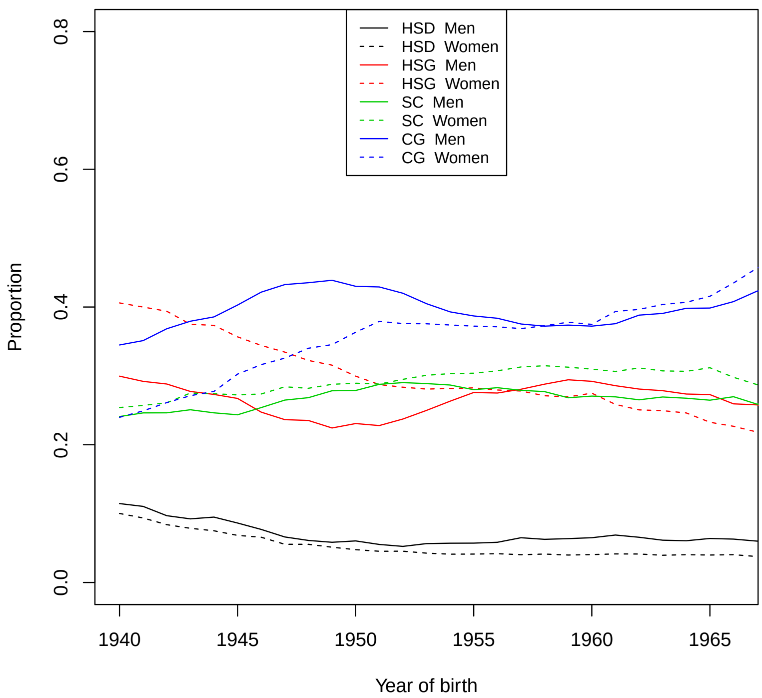
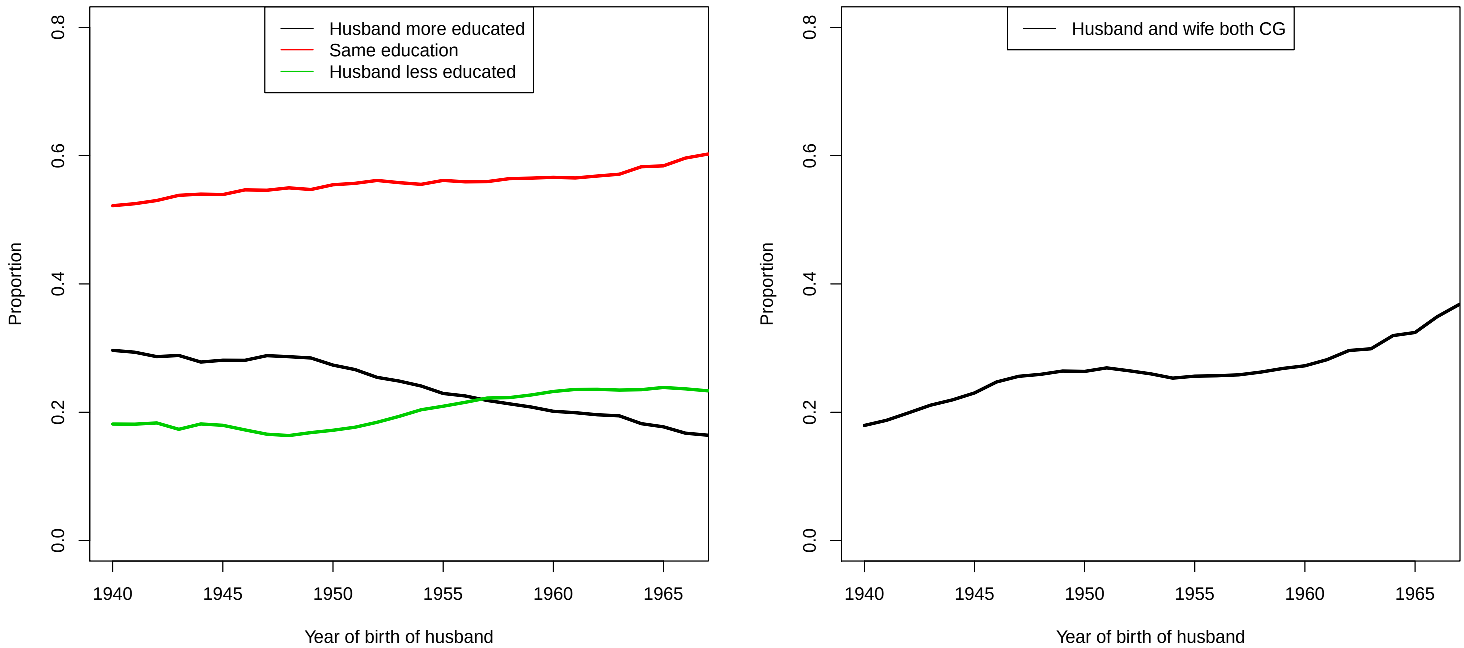


Figure 1: Education of men and women.

Figure 2 (a) shows an increase in the proportion of marriages of *like* with *like*. A substantial surge is also registered when focusing on the proportion of couples where both spouses have a college degree (Figure 2 (b)). However, as discussed earlier, these figures are not proof of an increase in educational sorting because they may be mechanically driven by changes in the proportions of individuals in each education category.

Salanié, and Weiss (2017) so as to make our conclusions as comparable as possible. In particular, given that the modal age difference within couples is one year in the data, Chiappori, Salanié, and Weiss (2017) concentrate their analysis on couples in which the age difference takes one year.



(a) (b)

Figure 2: Comparing spouses.

6.2 Results

For each of the six cohorts considered, we construct the identified sets of U, V, Φ, $D(\Phi)$, $C(U)$, and $C(V)$ under specifications [5], [6], and [7] of Table 1. Among the seven specifications of Table 1, we focus on specifications [5], [6], and [7] because they seem to deliver the tightest bounds according to our simulations in Section 5.

We start with discussing the results on Φ and $D(\Phi)$. Table 5 reports the projections of the estimated identified sets of Φ and $D(\Phi)$ under specification [5] of Table 1. Each row corresponds to a 2×2 sub-matrix of Φ that keeps same-education marriages in the main diagonal. Each column corresponds to a cohort. The estimated identified sets are obtained by implementing the methodology illustrated in Section 4.5 to the empirical distribution of the data. We adopt the scale normalisations (30) described in Section 5. These scale normalisations imply that the projection of the estimated identified set of Φ_{11} is a subset of $\{-2, 0, 2\}$. These scale normalisations also imply that the estimates of Chiappori, Salanié, and Weiss (2017) do not necessarily belong to our projections because we impose scale normalisations on U and V, whereas Chiappori, Salanié, and Weiss (2017) impose scale normalisations via the parameterisation of the taste shock distributions. As in Chiappori, Salanié, and Weiss (2017), we assume that the cohorts feature independent matching processes and apply the procedure separately to the sample of each cohort. However, our analysis is more robust in many ways. First, we allow the taste shock distributions to be arbitrarily different across cohorts, within the family of distributions outlined by specification [5] of Table 1. Second and importantly, we allow for potential endogeneity of education and conditional heteroskedasticity of the taste shocks. Third, the taste shocks can display any correlation structure. Instead, Chiappori, Salanié, and Weiss (2017) assume that, in each cohort, the taste shocks are i.i.d. standard Extreme Value Type I, independently distributed from education categories. They incorporate conditional heteroskedasticity in some final specification checks but encounter robustness issues possibly due to overparametrization. Finally, we allow Φ to be arbitrarily different across cohorts, whereas Chiappori, Salanié, and Weiss (2017) assume that Φ is identical across cohorts up to some drifts or linear/quadratic trends. In other words, Chiappori, Salanié, and Weiss (2017) restrict the evolution of Φ over time, while instead we remain agnostic about it so as to avoid misspecifications of the underlying dynamics.[21]

Overall, the conclusions from Table 5 can be summarised as follows. In most of the cases, the projections for Φ are quite similar or nested across cohorts. For instance, Figure 3 highlights that the projections for Φ in cohort 1966 (in blue) are often subsets of the projections for Φ in cohort 1940 (in black). Further, the ability of the model to recover the sign and magnitude of the Φ's components improves across cohorts. In fact, the number of projections for Φ that lie entirely on the negative or positive part of the real line and/or that are bounded increases across cohorts.

Nevertheless, a non-negligible number of projections for Φ remain unbounded at least on one side in each cohort. This implies that, when we take the algebraic sum of such projections in order to compute $D(\Phi)$, we obtain unbounded intervals. In particular, the projections for $D(\Phi)$ are completely uninformative up to cohort 1950. The projections for $D(\Phi)$ start to unambiguously reveal the presence of positive educational sorting in cohort 1955 within some education categories, by lying entirely on the positive part of the real line. In cohort 1966, the presence of positive educational sorting is confirmed within every education category. Further, in each of the cohorts 1955, 1960, and 1966, the interval of $D_{44,11}(\Phi)$ has the highest lower bound, indicating that positive educational sorting might be especially evident within the most (least) educated fraction of the population. In short, the 1to1TU model is progressively more informative over time about the direction of educational sorting.

However, nothing can be said about the evolution of positive educational sorting over time: given that all of the projections for $D(\Phi)$ are unbounded on the right-hand-side, positive educational sorting may increase or decrease across cohorts within each education category. This differs from the empirical findings of Chiappori, Salanié, and Weiss (2017) which, instead, reveal an increase in positive educational sorting over time within each education category, especially among the college educated people. Our results suggest that such a conclusion may be driven by the Multinomial Logit structure assigned to unobserved heterogeneity.

We now move to discuss the results on U, V, $C(U)$, and $C(V)$. Table 6 reports the projections of the estimated identified sets of U and $C(U)$ under specification [5] of Table 1. Every row

[21] More technically, Chiappori, Salanié, and Weiss (2017) *need* to assume that Φ is identical across cohorts up to some drifts or linear/quadratic trends in order to generate *over-identifying* restrictions which can be used, in turn, for testing changes in the structural parameters across cohorts (see also our discussion in Section 2). This is *not* needed here because we are in a partial identification framework.

corresponds to a 4×2 sub-matrix of U that spans over two education categories of men. Every column corresponds to a cohort. Table 7 reports the projections of the estimated identified sets of V and $C(V)$ under specification [5] of Table 1 and is structured analogously to Table 6.

As seen for Φ, the projections for U and V are quite similar or nested across cohorts. For instance, from Figures 4 and 5, we can notice that the projections for U and V in cohort 1966 (in blue) are often subsets of the projections for U and V in cohort 1940 (in black). Further, the projections for U in cohort 1966 are much tighter than the projections for U in cohort 1940. In fact, the identifying power of the model with respect to U improves across cohorts. For example, from Table 6, the proportion of bounded projections for U is $\frac{6}{16}$ in 1940, $\frac{8}{16}$ in 1945 and 1950, $\frac{9}{16}$ in 1950 and 1955, and $\frac{13}{16}$ in 1966. Table 6 does not reveal a similar trend for the women's side, with the proportion of bounded projections for V equal to $\frac{8}{16}$ in each cohort, except in cohort 1960 where it is $\frac{7}{16}$.

Besides, a non-trivial number of projections for U and V remain unbounded at least on one side. This implies that when we take the algebraic sum of such projections in order to compute $C(U)$ and $C(V)$, we obtain unbounded intervals in almost each cohort. In turn, the identified sets of the marital education premia will be unbounded as well (see Equation (3)). Hence, we find no evidence for the increase in those premia over time. This is particularly true for the women's side, where almost every component of $C(V)$ is unbounded on both sides in each cohort. Such conclusions differ from the empirical findings of Chiappori, Salanié, and Weiss (2017) which, instead, reveal an increase in the marital college premium over time, especially for women. As earlier, our results suggest that this may be driven by the Multinomial Logit assumption.

Tables 5, 6, and 7 are based on the estimated identified sets of the parameters. In Appendix G, we also report the projections of the corresponding 95% confidence regions to account for the sampling uncertainty (Tables G.1, G.2, and G.3). These confidence regions are constructed by using the bootstrap-based procedure of Bugni, Canay, and Shi (2017). In Appendix F, we discuss in detail how this procedure applies to our specific setting. The findings of Tables 5, 6, and 7 are overall confirmed.

Lastly, Tables 5, 6, and 7 are based on specification [5] of Table 1. We have also considered specifications [6] and [7], which are more restrictive. Specification [6] delivers numerically almost identical results. Specification [7] delivers slightly tighter bounds in certain cases without altering the above conclusions. We report the results under specification [7] in Appendix G (Tables G.4, G.5, and G.6).

Parameters	1940	1945	1950	1955	1960	1966
Φ_{11}	2	2	2	0	-2	-2
Φ_{12}	$[2.1111, \infty)$	$[0.1111, \infty)$	$[0.1111, \infty)$	$[-1.9091, -0.0909]$	$[-1.9091, -0.0909]$	$[-1.9091, -0.0909]$
Φ_{21}	$[-1.9091, -0.0707]$	$[-1.9091, -0.0707]$	$[-1.9091, -0.0707]$	$[-1.9091, -0.0707]$	$(-\infty, -2.0404]$	$(-\infty, -2.0404]$
Φ_{22}	$[2.2020, \infty)$	$[0.2828, \infty)$	$[0.2828, \infty)$	$[0.2828, \infty)$	$[0.2134, \infty)$	$[0.0297, \infty)$
$\boldsymbol{D_{22,11}(\Phi)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{[0.4444, \infty)}$	$\boldsymbol{[0.4141, \infty)}$	$\boldsymbol{[0.4141, \infty)}$
Φ_{11}	2	2	2	0	-2	-2
Φ_{13}	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, -1.2424]$	$(-\infty, -1.2424]$	$(-\infty, -1.2424]$
Φ_{31}	$(-\infty, -0.6869]$	$(-\infty, -0.6869]$	$(-\infty, -0.6869]$	$(-\infty, -0.6869]$	$(-\infty, -2.1515]$	$(-\infty, -2.1515]$
Φ_{33}	$(-\infty, \infty)$	$[-0.6667, \infty)$	$[-0.6667, \infty)$	$[-0.6667, \infty)$	$[-0.6667, \infty)$	$[-1, \infty)$
$\boldsymbol{D_{33,11}(\Phi)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{[1.2624, \infty)}$	$\boldsymbol{[0.7273, \infty)}$	$\boldsymbol{[0.3939, \infty)}$
Φ_{11}	2	2	2	0	-2	-2
Φ_{14}	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, -1.3939]$	$(-\infty, -1.3939]$	$(-\infty, -1.3939]$
Φ_{41}	$(-\infty, -1.1313]$	$(-\infty, -1.1313]$	$(-\infty, -1.1313]$	$(-\infty, -1.1313]$	$(-\infty, -2.3838]$	$(-\infty, -2.3838]$
Φ_{44}	$[-0.6869, \infty)$	$[-0.6869, \infty)$	$[-0.6869, \infty)$	$[-0.6869, \infty)$	$[-0.7879, \infty)$	$[-0.7879, \infty)$
$\boldsymbol{D_{44,11}(\Phi)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{[1.8384, \infty)}$	$\boldsymbol{[0.9899, \infty)}$	$\boldsymbol{[0.9899, \infty)}$
Φ_{22}	$[2.2020, \infty)$	$[0.2828, \infty)$	$[0.2828, \infty)$	$[0.2828, \infty)$	$[0.2134, \infty)$	$[0.0297, \infty)$
Φ_{23}	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$[-2.5354, 0.5960]$
Φ_{32}	$[-6.8485, \infty)$	$[-7.3535, \infty)$	$[-7.3535, \infty)$	$[-7.8586, \infty)$	$[-7.8586, \infty)$	$[-7.8586, 25.4747]$
Φ_{33}	$(-\infty, \infty)$	$[-0.6667, \infty)$	$[-0.6667, \infty)$	$[-0.6667, \infty)$	$[-0.6667, \infty)$	$[-1, \infty)$
$\boldsymbol{D_{33,22}(\Phi)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{[0.5556, \infty)}$
Φ_{22}	$[2.2020, \infty)$	$[0.2828, \infty)$	$[0.2828, \infty)$	$[0.2828, \infty)$	$[0.2134, \infty)$	$[0.0297, \infty)$
Φ_{24}	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$[-2.9394, 0.3232]$
Φ_{42}	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$
Φ_{44}	$[-0.6869, \infty)$	$[-0.6869, \infty)$	$[-0.6869, \infty)$	$[-0.6869, \infty)$	$[-0.7879, \infty)$	$[-0.7879, \infty)$
$\boldsymbol{D_{44,22}(\Phi)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$
Φ_{33}	$(-\infty, \infty)$	$[-0.6667, \infty)$	$[-0.6667, \infty)$	$[-0.6667, \infty)$	$[-0.6667, \infty)$	$[-1, \infty)$
Φ_{34}	$(-\infty, 16.3737]$	$[-17, 16.3737]$	$[-18.1919, 8.6263]$	$[-18.1919, 8.6263]$	$[-18.1919, 8.6263]$	$[-2.6970, 0.8788]$
Φ_{43}	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, 8.7273]$
Φ_{44}	$[-0.6869, \infty)$	$[-0.6869, \infty)$	$[-0.6869, \infty)$	$[-0.6869, \infty)$	$[-0.7879, \infty)$	$[-0.7879, \infty)$
$\boldsymbol{D_{44,33}(\Phi)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{(-\infty, \infty)}$	$\boldsymbol{[0.1717, \infty)}$

Table 5: Projections of the estimated identified sets of Φ and $D(\Phi)$ under specification [5] of Table 1. "1" denotes high school dropouts; "2" denotes high school graduates; "3" denotes some college; "4" denotes four-year college graduates and graduate degrees.

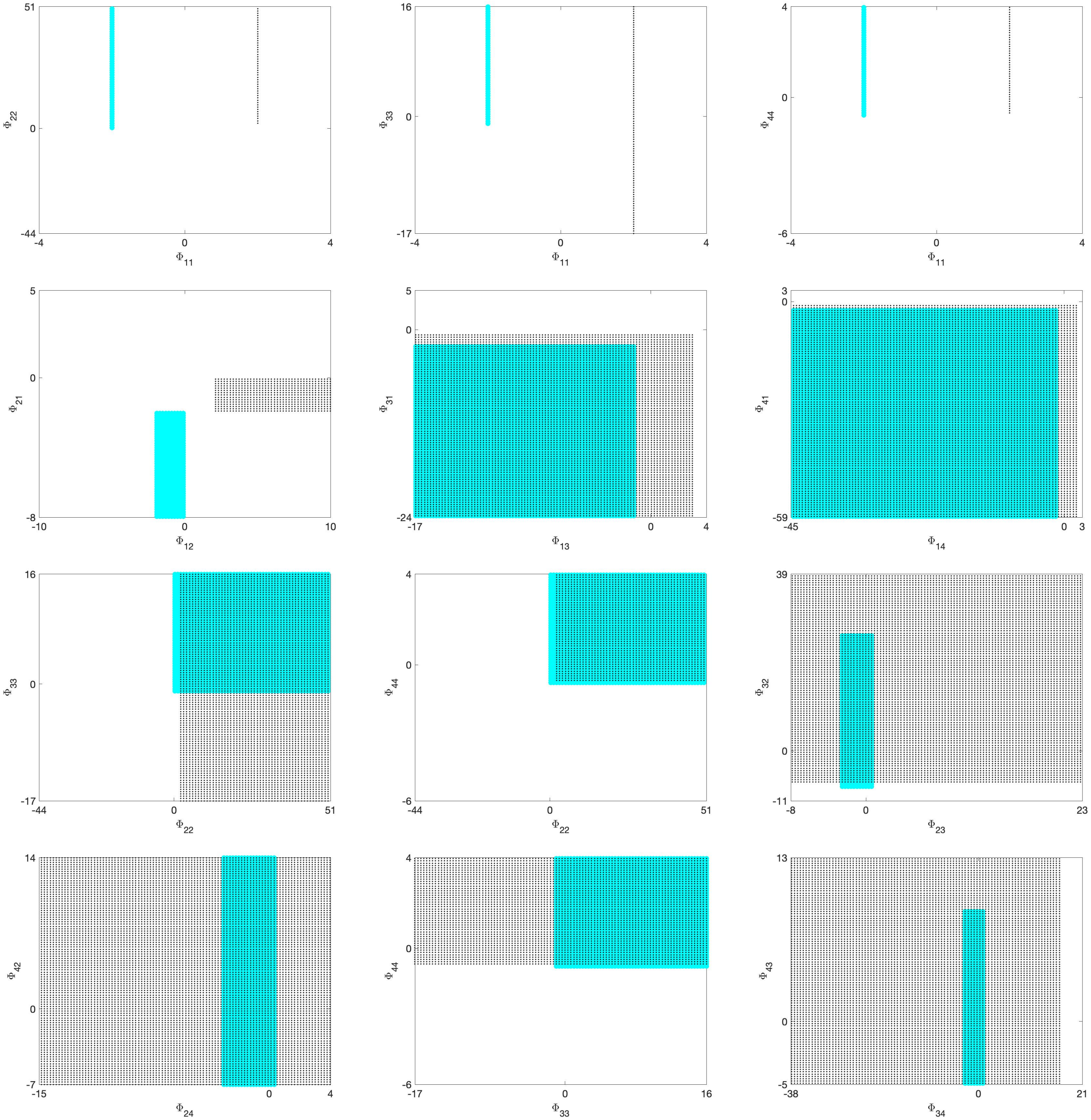


Figure 3: Projections of the estimated identified set of Φ under specification [5] of Table 1. The blue region refers to cohort 1966. The black region refers to cohort 1940. "1" denotes high school dropouts; "2" denotes high school graduates; "3" denotes some college; "4" denotes four-year college graduates and graduate degrees.

Parameters		1940		1945		1950		1955		1960		1966	
U_{11}	U_{12}	1	$[1.1111, \infty)$	1	$[1.1111, \infty)$	1	$[1.1111, \infty)$	-1	$[-0.9091, 0.9091]$	-1	$[-0.9091, 0.9091]$	-1	$[-0.9091, 0.9091]$
U_{13}	U_{14}	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, -0.2424]$	$(-\infty, -0.3939]$	$(-\infty, -0.2424]$	$(-\infty, -0.3939]$	$(-\infty, -0.2424]$	$(-\infty, -0.3939]$
U_{21}	U_{22}	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[-0.1117, 0.1414]$
U_{23}	U_{24}	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$[-1.7374, -0.1717]$	$[-1.9495, -0.2222]$
$C_{21}(U)$		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$[-0.5199, \infty)$**	
U_{11}	U_{12}	1	$[1.1111, \infty)$	1	$[1.1111, \infty)$	1	$[1.1111, \infty)$	-1	$[-0.9091, 0.9091]$	-1	$[-0.9091, 0.9091]$	-1	$[-0.9091, 0.9091]$
U_{13}	U_{14}	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, -0.2424]$	$(-\infty, -0.3939]$	$(-\infty, -0.2424]$	$(-\infty, -0.3939]$	$(-\infty, -0.2424]$	$(-\infty, -0.3939]$
U_{31}	U_{32}	-1	$[0.1111, \infty)$	-1	$[-0.3939, 15.7677]$	-1	$[-0.3939, 15.7677]$	-1	$[-0.8990, 0.6162]$	-1	$[-0.8990, 0.6162]$	-1	$[-0.8990, 0.1111]$
U_{33}	U_{34}	$(-\infty, \infty)$	$(-\infty, 15.6364]$	$[-0.3333, \infty)$	$[-16.5455, 15.6364]$	$[-0.3333, \infty)$	$[-17.7374, 7.8889]$	$[-0.3333, \infty)$	$[-17.7374, 7.8889]$	$[-0.3333, \infty)$	$[-17.7374, 7.8889]$	$[-0.6667, 0.6667]$	$[-2.2424, 0.1414]$
$C_{31}(U)$		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$(-0.5316, \infty)$**		**$[-0.5045, \infty)$**		**$[-0.5489, \infty)$**	
U_{11}	U_{12}	1	$[1.1111, \infty)$	1	$[1.1111, \infty)$	1	$[1.1111, \infty)$	-1	$[-0.9091, 0.9091]$	-1	$[-0.9091, 0.9091]$	-1	$[-0.9091, 0.9091]$
U_{13}	U_{14}	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, -0.2424]$	$(-\infty, -0.3939]$	$(-\infty, -0.2424]$	$(-\infty, -0.3939]$	$(-\infty, -0.2424]$	$(-\infty, -0.3939]$
U_{41}	U_{42}	-1	$[-0.4242, 3.8182]$	-1	$[-0.8485, 0.8485]$	-1	$[-0.8485, 0.8485]$	-1	$[-0.8485, 0.8485]$	-1	$[-0.8485, 0.6364]$	-1	$[-0.8485, 0.6364]$
U_{43}	U_{44}	$[-0.2727, 3.9091]$	$[-0.2424, \infty)$	$[-0.2727, 3.9091]$	$[-0.2424, \infty)$	$[-0.2727, 3.9091]$	$[-0.2424, \infty)$	$[-0.2727, 3.9091]$	$[-0.2424, \infty)$	$[-0.8182, 0.8182]$	$[-0.3434, \infty)$	$[-0.8182, 0.8182]$	$[-0.3434, \infty)$
$C_{41}(U)$		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$[-0.2643, \infty]$**		**$[-0.4050, \infty)$**		**$[-0.3389, \infty)$**	
U_{21}	U_{22}	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[-0.1117, 0.1414]$
U_{23}	U_{24}	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$[-1.7374, -0.1717]$	$[-1.9495, -0.2222]$
U_{31}	U_{32}	-1	$[0.1111, \infty)$	-1	$[-0.3939, 15.7677]$	-1	$[-0.3939, 15.7677]$	-1	$[-0.8990, 0.6162]$	-1	$[-0.8990, 0.6162]$	-1	$[-0.8990, 0.1111]$
U_{33}	U_{34}	$(-\infty, \infty)$	$(-\infty, 15.6364]$	$[-0.3333, \infty)$	$[-16.5455, 15.6364]$	$[-0.3333, \infty)$	$[-17.7374, 7.8889]$	$[-0.3333, \infty)$	$[-17.7374, 7.8889]$	$[-0.3333, \infty)$	$[-17.7374, 7.8889]$	$[-0.6667, 0.6667]$	$[-2.2424, 0.1414]$
$C_{32}(U)$		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$[-0.5188, 0.7789]$**	
U_{21}	U_{22}	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[-0.1117, 0.1414]$
U_{23}	U_{24}	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$[-1.7374, -0.1717]$	$[-1.9495, -0.2222]$
U_{41}	U_{42}	-1	$[-0.4242, 3.8182]$	-1	$[-0.8485, 0.8485]$	-1	$[-0.8485, 0.8485]$	-1	$[-0.8485, 0.8485]$	-1	$[-0.8485, 0.6364]$	-1	$[-0.8485, 0.6364]$
U_{43}	U_{44}	$[-0.2727, 3.9091]$	$[-0.2424, \infty)$	$[-0.2727, 3.9091]$	$[-0.2424, \infty)$	$[-0.2727, 3.9091]$	$[-0.2424, \infty)$	$[-0.2727, 3.9091]$	$[-0.2424, \infty)$	$[-0.8182, 0.8182]$	$[-0.3434, \infty)$	$[-0.8182, 0.8182]$	$[-0.3434, \infty)$
$C_{42}(U)$		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$[-0.3088, \infty)$**	
U_{31}	U_{32}	-1	$[0.1111, \infty)$	-1	$[-0.3939, 15.7677]$	-1	$[-0.3939, 15.7677]$	-1	$[-0.8990, 0.6162]$	-1	$[-0.8990, 0.6162]$	-1	$[-0.8990, 0.1111]$
U_{33}	U_{34}	$(-\infty, \infty)$	$(-\infty, 15.6364]$	$[-0.3333, \infty)$	$[-16.5455, 15.6364]$	$[-0.3333, \infty)$	$[-17.7374, 7.8889]$	$[-0.3333, \infty)$	$[-17.7374, 7.8889]$	$[-0.3333, \infty)$	$[-17.7374, 7.8889]$	$[-0.6667, 0.6667]$	$[-2.2424, 0.1414]$
U_{41}	U_{42}	-1	$[-0.4242, 3.8182]$	-1	$[-0.8485, 0.8485]$	-1	$[-0.8485, 0.8485]$	-1	$[-0.8485, 0.8485]$	-1	$[-0.8485, 0.6364]$	-1	$[-0.8485, 0.6364]$
U_{43}	U_{44}	$[-0.2727, 3.9091]$	$[-0.2424, \infty)$	$[-0.2727, 3.9091]$	$[-0.2424, \infty)$	$[-0.2727, 3.9091]$	$[-0.2424, \infty)$	$[-0.2727, 3.9091]$	$[-0.2424, \infty)$	$[-0.8182, 0.8182]$	$[-0.3434, \infty)$	$[-0.8182, 0.8182]$	$[-0.3434, \infty)$
$C_{43}(U)$		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$(-\infty, \infty)$**		**$[-0.5978, \infty)$**	

Table 6: Projections of the estimated identified sets of U and $C(U)$ under specification [5] of Table 1. “1” denotes high school dropouts; “2” denotes high school graduates; “3” denotes some college; “4” denotes four-year college graduates and graduate degrees.

Parameters		1940		1945		1950		1955		1960		1966	
V_{11}	V_{21}	1	$[-0.9091, 0.9293]$	1	$[-0.9091, 0.9293]$	1	$[-0.9091, 0.9293]$	1	$[-0.9091, 0.9293]$	-1	$(-\infty, -1.0404]$	-1	$(-\infty, -1.0404]$
V_{31}	V_{41}	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, -1.1551]$	$(-\infty, -1.3838]$	$(-\infty, -1.1551]$	$(-\infty, -1.3838]$
V_{12}	V_{22}	1	$[2.0606, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$
V_{32}	V_{42}	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, 25.3636]$	$(-\infty, \infty)$
$\boldsymbol{C_{21}(V)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$	
V_{11}	V_{21}	1	$[-0.9091, 0.9293]$	1	$[-0.9091, 0.9293]$	1	$[-0.9091, 0.9293]$	1	$[-0.9091, 0.9293]$	-1	$(-\infty, -1.0404]$	-1	$(-\infty, -1.0404]$
V_{31}	V_{41}	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, -1.1551]$	$(-\infty, -1.3838]$	$(-\infty, -1.1551]$	$(-\infty, -1.3838]$
V_{13}	V_{23}	-1	$[-0.4848, 14.2323]$	-1	$[-0.4848, 14.2323]$	-1	$[-0.4848, 14.2323]$	-1	$[-0.7980, 0.7677]$	-1	$[-0.4848, 14.2323]$	-1	$[-0.7980, 0.7677]$
V_{33}	V_{43}	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, 7.9091]$
$\boldsymbol{C_{31}(V)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$	
V_{11}	V_{21}	1	$[-0.9091, 0.9293]$	1	$[-0.9091, 0.9293]$	1	$[-0.9091, 0.9293]$	1	$[-0.9091, 0.9293]$	-1	$(-\infty, -1.0404]$	-1	$(-\infty, -1.0404]$
V_{31}	V_{41}	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, -1.1551]$	$(-\infty, -1.3838]$	$(-\infty, -1.1551]$	$(-\infty, -1.3838]$
V_{14}	V_{24}	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$
V_{34}	V_{44}	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$
$\boldsymbol{C_{41}(V)}$		$\boldsymbol{[-1.1268, \infty)}$		$\boldsymbol{[-1.1057, \infty)}$		$\boldsymbol{[-1.0301, \infty)}$		$\boldsymbol{[-0.9840, \infty)}$		$\boldsymbol{[0.3077, \infty)}$		$\boldsymbol{[0.2459, \infty)}$	
V_{12}	V_{22}	1	$[2.0606, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$
V_{32}	V_{42}	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, 25.3636]$	$(-\infty, \infty)$
V_{13}	V_{23}	-1	$[-0.4848, 14.2323]$	-1	$[-0.4848, 14.2323]$	-1	$[-0.4848, 14.2323]$	-1	$[-0.7980, 0.7677]$	-1	$[-0.4848, 14.2323]$	-1	$[-0.7980, 0.7677]$
V_{33}	V_{43}	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, 7.9091]$
$\boldsymbol{C_{32}(V)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$	
V_{12}	V_{22}	1	$[2.0606, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$
V_{32}	V_{42}	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, 25.3636]$	$(-\infty, \infty)$
V_{14}	V_{24}	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$
V_{34}	V_{44}	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$
$\boldsymbol{C_{42}(V)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$	
V_{13}	V_{23}	-1	$[-0.4848, 14.2323]$	-1	$[-0.4848, 14.2323]$	-1	$[-0.4848, 14.2323]$	-1	$[-0.7980, 0.7677]$	-1	$[-0.4848, 14.2323]$	-1	$[-0.7980, 0.7677]$
V_{33}	V_{43}	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, 7.9091]$
V_{14}	V_{24}	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$
V_{34}	V_{44}	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$
$\boldsymbol{C_{43}(V)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$	

Table 7: Projections of the estimated identified sets of V and $C(V)$ under specification [5] of Table 1. "1" denotes high school dropouts; "2" denotes high school graduates; "3" denotes some college; "4" denotes four-year college graduates and graduate degrees.

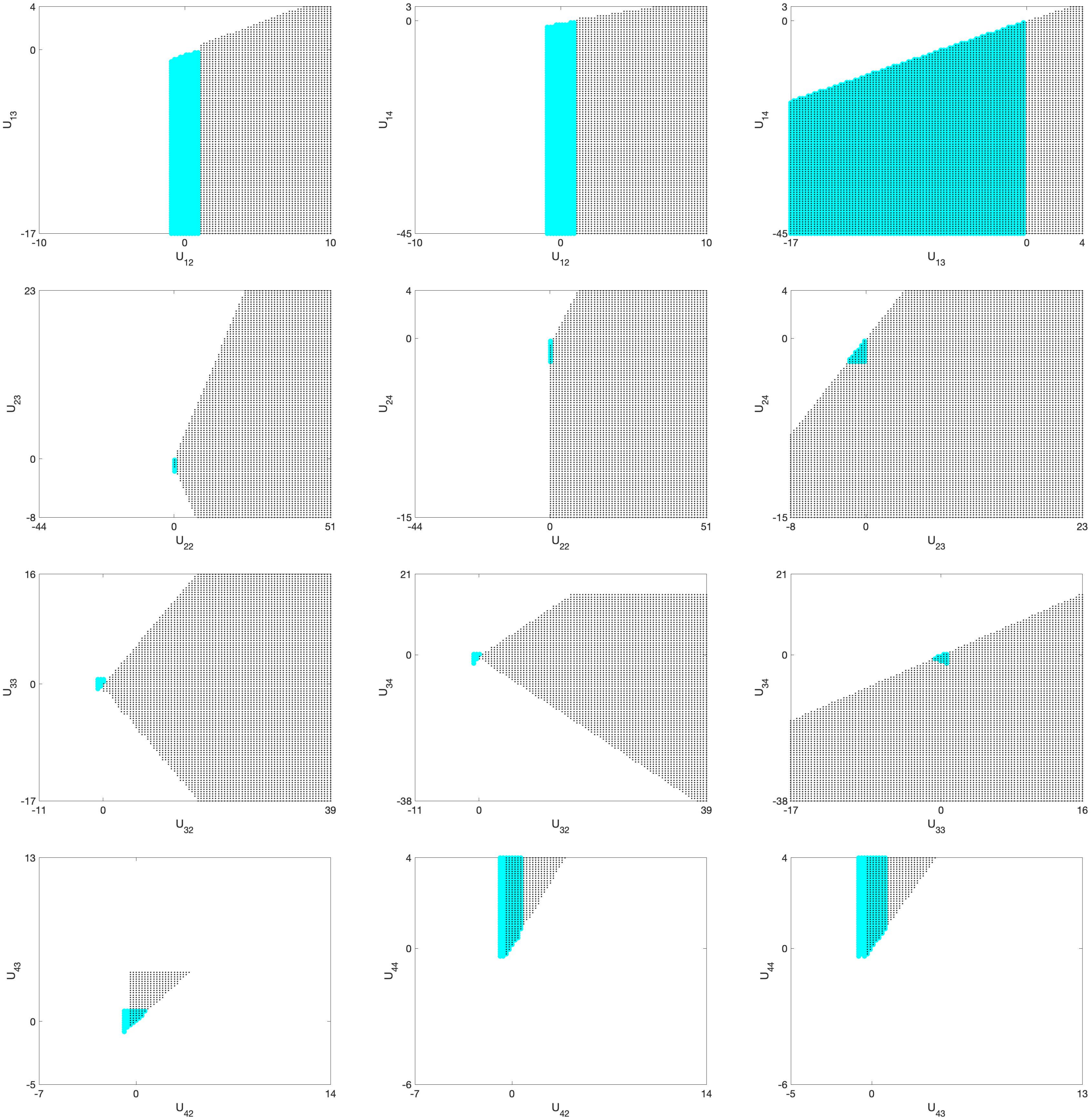


Figure 4: Projections of the estimated identified set of U under specification [5] of Table 1. The blue region refers to cohort 1966. The black region refers to cohort 1940. "1" denotes high school dropouts; "2" denotes high school graduates; "3" denotes some college; "4" denotes four-year college graduates and graduate degrees.

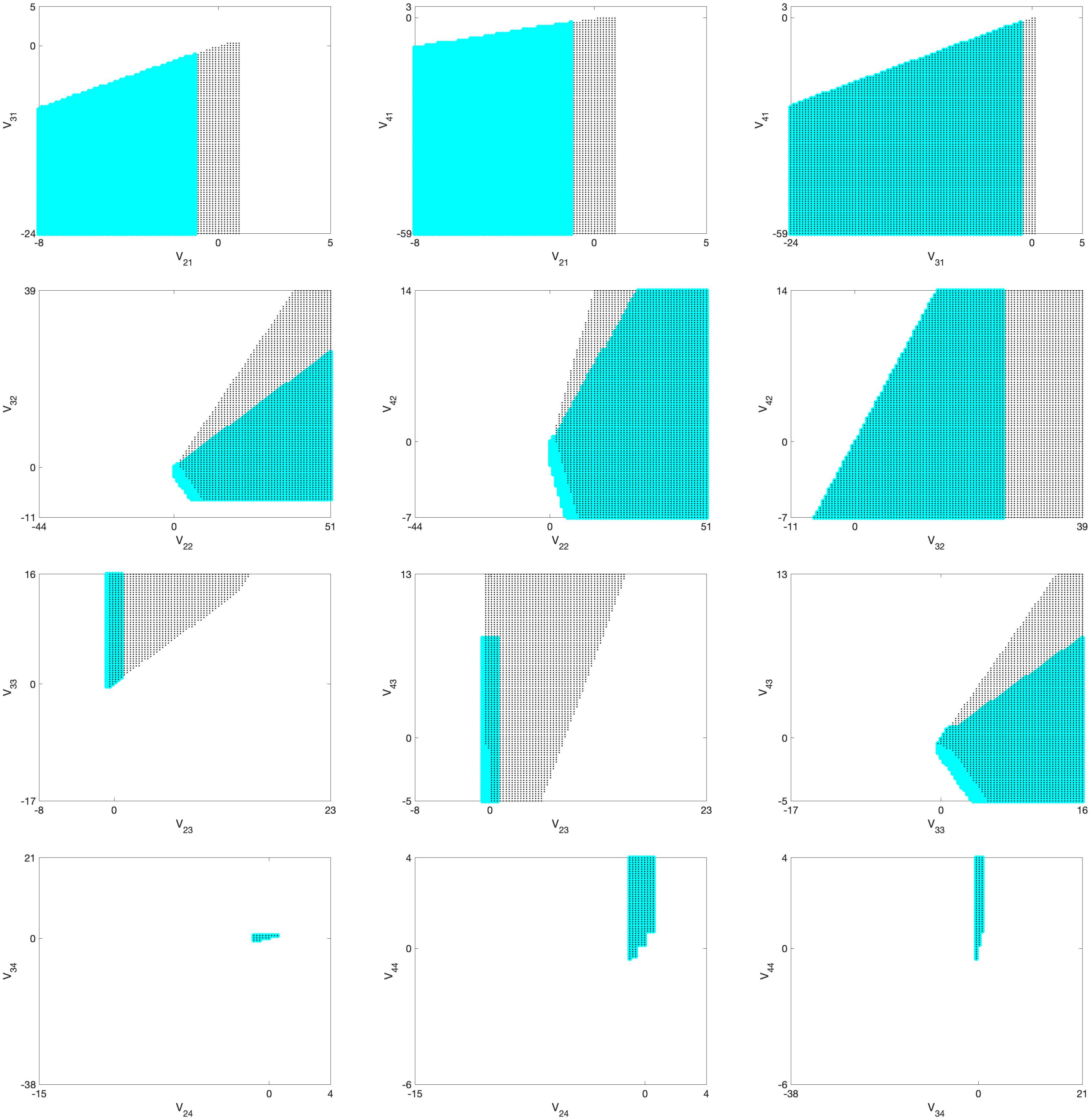


Figure 5: Projections of the estimated identified set of V under specification [5] of Table 1. The blue region refers to cohort 1966. The black region refers to cohort 1940. “1” denotes high school dropouts; “2” denotes high school graduates; “3” denotes some college; “4” denotes four-year college graduates and graduate degrees.

7 Conclusions

This paper investigates the identifying power of the 1to1TU model for the systematic match surplus and related policy-relevant quantities, when no parametric distributional assumption on the unobserved heterogeneity is imposed. We conclude our analysis by highlighting three main findings. First, some nonparametric distributional assumptions on the unobserved heterogeneity are needed for the 1to1TU model to retain some information about the systematic match surplus with data on one large market (even after accounting for scale and location normalisations). Second, we propose a computational approach for constructing the identified set of the systematic match surplus that is based on principles of linear programming and works under various classes of nonparametric distributional assumptions on the unobserved heterogeneity. Third, we use our methodology to re-examine some relevant questions in the empirical literature on the marriage market which have been studied under the Multinomial Logit assumption. Our results do not support the conclusions achieved under the Multinomial Logit assumption.

Lastly, we remark that we have been able to practically implement our methodology for up to five types on each side of the market, by parallelising our codes and using cluster facilities. Many empirical applications to the marriage market consider between three and five types, especially when focusing on education levels. Scaling up further might be computationally difficult because the linear programming problems to solve become large in the number of unknowns and constraints. Similar curses of dimensionality are faced by other nonparametric and partial identification procedures, but this need not stop researchers from finding methodologies that are more robust than fully parametric approaches.

References

Abbott, B., G. Gallipoli, C. Meghir, and G. L. Violante (2019): "Education Policy and Intergenerational Transfers in Equilibrium," *Journal of Political Economy*, 127(6), 2569–2624.

Akkus, O., J.A. Cookson, and A. Hortaçsu (2016): "The Determinants of Bank Mergers: A Revealed Preference Analysis," *Management Science*, 62(8), 2241–2258.

Andrews, D.W.K., and G. Soares (2010): "Inference for Parameters Defined by Moment Inequalities Using Generalized Moment Selection," *Econometrica*, 78(1), 119–157.

Baccara, M., A. İmrohoroğlu, A.J. Wilson, and L. Yariv (2012): "A Field Study on Matching with Network Externalities," *American Economic Review*, 102(5), 1773–1804.

Banal-Estañol, A., I. Macho-Stadler, D. Pérez-Castrillo (2018): "Endogenous Matching in University-Industry Collaboration: Theory and Empirical Evidence from the United Kingdom," *Management Science*, 64(4), 1591–1608.

Becker, G.S. (1973): "A Theory of Marriage: Part I," *Journal of Political Economy*, 81(4), 813–846.

Bertsimas, D., and J.N. Tsitsiklis (1997): *Introduction to Linear Optimisation*, Athena Scientific, Belmont, Massachusetts.

Bisin, A., and G. Tura (2020): "Marriage, Fertility, and Cultural Integration in Italy," NBER Working Paper No. 26303.

Bhattacharya, D. (2009): "Inferring Optimal Peer Assignment From Experimental Data," 104(486), 486–500.

Botticini, M, and A. Siow (2011): "Are There Increasing Returns to Scale in Marriage Markets?," Working Paper.

Brandt, L., A. Siow, and C. Vogel (2016): "Large Demographic Shocks and Small Changes in the Marriage Market," *Journal of the European Economic Association*, 14(6), 1437–1468.

Bruze, G., M. Svarer, and Y. Weiss (2015): "The Dynamics of Marriage and Divorce," *Journal of Labor Economics*, 33,(1) 123–170.

Bugni, F.A., I.A. Canay, and X. Shi (2017): " Inference for subvectors and other functions of partially identified parameters in moment inequality models," *Quantitative Economics*, 8(1), 1–38.

Chen, L. (2017): "Compensation, Moral Hazard, and Talent Misallocation in the Market for CEOs," SSRN Working Paper.

Chernozhukov, V., H. Hong, and E. Tamer (2007): "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica*, 75(5), 1243–1284.

Chiappori, P.-A. (2017): *Matching with Transfers: The Economics of Love and Marriage*, Princeton University Press.

Chiappori, P.-A., M. Costa-Dias, C. Meghir (2018): "The Marriage Market, Labor Supply, and Education Choice," *Journal of Political Economy*, 126(S1), S26-S72.

Chiappori, P.-A., M. Costa-Dias, S. Crossman, and C. Meghir (2020): "Changes in Assortative Matching and Inequality in Income: Evidence for the UK," *Fiscal Studies*, 41(1), 39–63.

Chiappori, P.-A., M. Costa-Dias, C. Meghir (2020): "Changes in Assortative Matching: Theory and Evidence for the US," IFS Working Paper W20/10.

Chiappori, P.-A., M. Iyigun, Y. Weiss (2009): "Investment in Schooling and the Marriage Market," *American Economic Review*, 99(5), 1689–1713.

Chiappori, P.-A., R.J. McCann, and L.P. Nesheim (2010): "Hedonic Price Equilibria, Stable Matching, and Optimal Transport: Equivalence, Topology, and Uniqueness," *Economic Theory*, 42(2), 317–354.

Chiappori, P.-A., R.J. McCann, and B. Pass (2020): "Multidimensional matching: theory and empirics," Working Paper.

Chiappori, P.-A., and B. Salanié (2016): "The Econometrics of Matching Models," *Journal of Economic Literature*, 54(3), 832–861.

Chiappori, P.-A., B. Salanié, and Y. Weiss (2017): "Partner Choice, Investment in Children, and the Marital College Premium," *American Economic Review*, 107(8), 2109–2167.

Chiappori, P.-A., B. Salanié, A. Tillman, and Y. Weiss (2008): "Assortative Matching on the Marriage Market: A Structural Investigation," slides available at `http://adres.ens.fr/IMG/pdf/09022009.pdf`.

Choo, E. (2015): "Dynamic Marriage Matching: An Empirical Framework," *Econometrica*, 83(4), 1373–1423.

Choo, E., and S. Seitz (2013): "The Collective Marriage Matching Model: Identification, Estimation, and Testing," *Structural Econometric Models (Advances in Econometrics)*, 31, 291–336.

Choo, E., and A. Siow (2006): "Who Marries Whom and Why," *Journal of Political Economy*, 114(1), 175–201.

Ciliberto, F., and E. Tamer (2009): "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, 77(6), 1791–1828.

Ciscato, E., A. Galichon, and M. Goussé (2020): "Like Attract Like? A Structural Comparison of Homogamy Across Same-Sex and Different-Sex Households," *Journal of Political Eocnomy*, 128(2), 740–781.

Ciscato, E., and S. Weber (2020) "The Role of Evolving Marital Preferences in Growing Income Inequality," *Journal of Population Economics*, 33(1), 307–347.

Christensen, T., and B. Connault (2019): "Counterfactual Sensitivity and Robustness," arXiv:1904.00989.

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

Dupuy, A., and S. Weber (2020): "Marital Patterns and Income Inequality," SSRN Working Paper.

Eika, L., M. Mogstad, and B. Zafar (2019): "Educational Assortative Mating and Household Income Inequality," *Journal of Political Economy*, 127(6), 2795–2835.

Fernández, R., N. Guner, and J. Knowles (2005): "Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality," *Quarterly Journal of Economics*, 120 (1), 273–344.

Fernández, R. and E. Rogerson (2001): "Sorting and Long-Run Inequality," *Quarterly Journal of Economics*, 116 (4), 1305–1341.

Fox, J.T. (2010): "Identification in Matching Games," *Quantitative Economics*, 1(2), 203–254.

Fox, J.T. (2018): "Estimating Matching Games with Transfers," *Quantitative Economics*, 9(1), 1–38.

Fox, J.T., C. Yang, and D.H. Hsu (2018): "Unobserved Heterogeneity in Matching Games," *Journal of Political Economy*, 126(4), 1339–1373.

Galichon, A., S.D. Kominers, and S. Weber (2019): "Costly Concessions: An Empirical Framework for Matching with Imperfectly Transferable Utility," *Journal of Political Economy*, 127(6), 2875–2925.

Galichon, A., and B. Salanié (2019): "IIA in Separable Matching Markets," Columbia University mimeo.

Galichon, A., and B. Salanié (2020): "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models," SSRN Working Paper.

Gayle, G.-L., and A. Shephard (2019): "Optimal Taxation, Marriage, Home Production, and Family Labor Supply," *Econometrica*, 87(1), 291–326.

Graham, B. (2011): "Econometric Methods for the Analysis of Assignment Problems in the Presence of Complementarity and Social Spillovers," in *Handbook of Social Economics*, ed. by J. Benhabib, M.O. Jackson, A. Bisin, 1B, 965–1052.

Graham, B. (2013a): "Comparative Static and Computational Methods for an Empirical One-To-One Transferable Utility Matching Model," *Advances in Econometrics: Structural Econometric Models*, 31(1), 151–179.

Graham, B. (2013b): "Errata in Econometric Methods for the Analysis of Assignment Problems in the Presence of Complementarity and Social Spillovers," Unpublished.

Graham, B. (2013c): "Efficient estimation of average match output under conditionally exogenous matching," Working Paper.

Graham, B., G. Imbens, and S. Ridder (2014): "Complementarity and aggregate implications of assortative matching: a nonparametric analysis," *Quantitative Economics*, 5(1), 29–66.

Greenwood, J., N. Guner, and J.A. Knowles (2003): "More on Marriage, Fertility, and the Distribution of Income," *International Economic Review*, 44(3), 827–862.

Greenwood, J., N. Guner, G. Kocharkov, and C. Santos (2014): "Marry your Like: Assortative Mating and Income Inequality," *American Economic Review: Papers and Proceedings*, 104(5), 348–353.

Heckman, J.J., and S. Mosso (2014): "The Economics of Human Development and Social Mobility," *Annual Review of Economics*, 6, 689–733.

Ho, K., and A.M. Rosen (2017): "Partial Identification in Applied Research: Benefits and Challenges," in *Advances in Economics and Econometrics: Eleventh World Congress of the Econometric Society (Econometric Society Monographs)*, ed. by B. Honore, A. Pakes, M. Piazzesi and L. Samuelson, 2, 307–359.

Kremer, M. (1997): "How Much Does Sorting Increase Inequality," *Quarterly Journal of Economics*, 112 (1), 115–139.

Liu, H. and J. Lu (2006) "Measuring the Degree of Assortative Mating," *Economics Letters*, 92(3), 317–322.

Mindruda, D. (2013): "Value Creation in University-Firm Research Collaborations: A Matching Approach," *Strategic management journal*, 34(6), 644–665.

Mindruda, D., M. Mohen, R. Agarwal (2016): "A Two-sided Matching Approach for Partner Selection and Assessing Complementarities in Inter-firm Alliances," *Strategic Management Journal*, 37(1), 206–231.

Molinari, F. (2020): "Microeconometrics with Partial Identification," forthcoming in the *Handbook of Econometrics*, Volume 7A.

Mourifié, I., and A. Siow (2020): "The Cobb Douglas Marriage Matching Function: Marriage Matching with Peer and Scale Effects," accepted at *Journal of Labor Economics*.

Romano, J. P., and A.M. Shaikh (2008): "Inference for Identifiable Parameters in Partially Identified Econometric Models," *Journal of Statistical Planning and Inference*, 138(9), 2786-2807.

Shapley, L., and M. Shubik (1972): "The Assignment Game I: The Core", *International Journal of Game Theory*, 1(1), 111–130.

Shen, J. (2019): "(Non-)Marital Assortative Mating and the Closing of the Gender Gap in Education," Working Paper.

Sinha, S. (2018): "Identification in One-to-One Matching Models with Nonparametric Unobservables," TSE Working Paper 18-897.

Siow, A. (2015): "Testing Becker's Theory of Positive Assortative Matching", *Journal of Labor Economics*, 33(2), 409-441.

Torgovitsky, A. (2019): "Partial Identification by Extending Subdistributions," *Quantitative Economics*, 10(1), 105–144.

Villani C., (2009): *Optimal Transport. Old and new.*, Grundlehren der Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences), 338. Springer.

A Proofs

A.1 Proof of Proposition 2

The proof is organised in the following steps. Fix any $\Phi \in \Theta$. Consider any $U \in \mathcal{U}$ and $V \in \mathcal{V}$ such that $U_{xy} + V_{xy} = \Phi_{xy}$ for each $(x, y) \in \mathcal{X} \times \mathcal{Y}$. In Steps 1-4, we show by construction that there exists $\{F_x\}_{x\in\mathcal{X}} \in \mathcal{F}$ such that $p_{y|x} = \kappa(U, F_x, y)$ for each $(x, y) \in \mathcal{X} \times \mathcal{Y}_0$. In Step 5, we replicate the same arguments for the women's side and show that there exists $\{G_y\}_{y\in\mathcal{Y}} \in \mathcal{G}$ such that $p_{x|y} = \kappa(V, G_y, x)$ for each $(x, y) \in \mathcal{X}_0 \times \mathcal{Y}$. In Step 6, we highlight that our conclusions do not change if we impose scale normalisations on U and V.

For simplicity of exposition, we provide the proof of Proposition 2 for the case $\mathcal{X}_0 \equiv \mathcal{Y}_0 \equiv \{0, 1, 2\}$. The proof for a generic case follows exactly the same steps, but becomes notationally more complicated.

Step 1 Start from the men's side. For every $(x, y) \in \mathcal{X} \times \mathcal{Y}_0$, let $\mathcal{R}^U_{y|x}$ be the set of taste shock realisations such that it is optimal for a man of type x to choose a woman of type y. More formally,

$$\begin{aligned}
\mathcal{R}^U_{0|x} &\equiv \{(e_0, e_1, e_2) \in \mathbb{R}^3 : e_1 - e_0 \leq -U_{x1}, e_2 - e_0 \leq -U_{x2}\},\\
\mathcal{R}^U_{1|x} &\equiv \{(e_0, e_1, e_2) \in \mathbb{R}^3 : e_1 - e_0 \geq -U_{x1}, e_1 - e_2 \geq U_{x2} - U_{x1}\},\\
\mathcal{R}^U_{2|x} &\equiv \{(e_0, e_1, e_2) \in \mathbb{R}^3 : e_2 - e_0 \geq -U_{x2}, e_1 - e_2 \leq U_{x2} - U_{x1}\},
\end{aligned}$$

for each $x \in \mathcal{X}$. Our objective is to show that there exists $\{F_x\}_{x\in\mathcal{X}} \in \mathcal{F}$ such that

$$p_{y|x} = \lambda_{F_x}(\mathcal{R}^U_{y|x}) \ \forall(x, y) \in \mathcal{X} \times \mathcal{Y}_0, \tag{A.1.1}$$

where λ_{F_x} is the measure associated with F_x.

Step 2 As usual in multinomial choice models, we can reduce the problem by one dimension. Hence, let $\zeta \equiv (\zeta_1, \zeta_2)$ be a 2-dimensional random vector with CDF conditional on $X_i = x$ denoted by H_x, for every $x \in \mathcal{X}$. H_x can be *any* continuous CDF with bounded support in $\mathbb{R}^2$. Further, let $H_{x,\mu,\sigma}$ be the CDF of $\sigma\zeta + \mu$ conditional on $X_i = x$, with $\mu \in \mathbb{R}^2$ and $\sigma \in \mathbb{R}_+$, for each $x \in \mathcal{X}$. Lastly, let

$$\begin{aligned}
\mathcal{Q}^U_{0|x} &\equiv \{(\tilde{e}_1, \tilde{e}_2) \in \mathbb{R}^2 : \tilde{e}_1 \leq -U_{x1}, \tilde{e}_2 \leq -U_{x2}\},\\
\mathcal{Q}^U_{1|x} &\equiv \{(\tilde{e}_1, \tilde{e}_2) \in \mathbb{R}^2 : \tilde{e}_1 \geq -U_{x1}, \tilde{e}_1 - \tilde{e}_2 \geq U_{x2} - U_{x1}\},\\
\mathcal{Q}^U_{2|x} &\equiv \{(\tilde{e}_1, \tilde{e}_2) \in \mathbb{R}^2 : \tilde{e}_2 \geq -U_{x2}, \tilde{e}_1 - \tilde{e}_2 \leq U_{x2} - U_{x1}\},
\end{aligned}$$

for each $x \in \mathcal{X}$.

Step 3 In this step we show that, for each $x \in \mathcal{X}$, there exists a location $\mu \in \mathbb{R}^2$ and a scale $\sigma \in \mathbb{R}_+$ (possibly different across x) such that

$$p_{y|x} = \lambda_{H_{x,\mu,\sigma}}(\mathcal{Q}^U_{y|x}) \ \forall y \in \mathcal{Y}_0,$$

where $\lambda_{H_{x,\mu,\sigma}}$ is the measure associated with $H_{x,\mu,\sigma}$.

Step 3.1 Fix $x \in \mathcal{X}$. Note that, by continuity of H_x, the mapping

$$(\mu, \sigma) \in \mathbb{R}^2 \times \mathbb{R}_+ \mapsto \lambda_{H_{x,\mu,\sigma}}(\mathcal{A}) \in [0, 1],$$

is continuous, for every measurable set $\mathcal{A} \subseteq \mathbb{R}^2$. Consequently, the mapping

$$h_x : (\mu, \sigma) \in \mathbb{R}^2 \times \mathbb{R}_+ \mapsto (\lambda_{H_{x,\mu,\sigma}}(\mathcal{Q}^U_{0|x}), \lambda_{H_{x,\mu,\sigma}}(\mathcal{Q}^U_{1|x}), \lambda_{H_{x,\mu,\sigma}}(\mathcal{Q}^U_{2|x})) \in [0, 1]^3,$$

is also continuous. Further, given that $\{\mathcal{Q}^U_{0|x}, \mathcal{Q}^U_{1|x}, \mathcal{Q}^U_{2|x}\}$ is a partition of $\mathbb{R}^2$, the image set of h_x is a subset of the unit simplex,

$$\mathcal{S} \equiv \{(t_1, t_2, t_3) \in \mathbb{R}^3 : t_1 + t_2 + t_3 = 1, t_1 \geq 0, t_2 \geq 0, t_3 \geq 0\}.$$

Step 3.2 We now show that the image set of h_x is the unit simplex. In fact, observe that:

1. Fix any point $z \in \mathcal{Q}^U_{0|x}$. By setting $\mu = z$ and shrinking σ, one can focus almost the entire image of h_x within $\mathcal{Q}^U_{0|x}$, i.e.,
$$\lim_{\sigma \to 0^+} h_x(z, \sigma) = (1, 0, 0).$$
Thus, the image set of h_x is arbitrarily close to the vertex $(1, 0, 0)$ of the unit simplex. Further, the image set of h_x contains such a vertex because the support of H_x is bounded. Repeat the same argument for $\mathcal{Q}^U_{1|x}$ and $\mathcal{Q}^U_{2|x}$. Conclude that the image set of h_x is arbitrarily close to (and include) all the vertices of the unit simplex.

2. Fix any two points, z, r, in the interiors of $\mathcal{Q}^U_{0|x}$ and $\mathcal{Q}^U_{1|x}$, respectively. Consider the continuous curve
$$t \in [a, b] \mapsto \gamma_{0,1}(t) \in \mathbb{R}^2$$
connecting z and r, while passing through the interior of $\mathcal{Q}^U_{0|x} \cup \mathcal{Q}^U_{1|x}$ and not through $\mathcal{Q}^U_{2|x}$. Note that such a continuous curve always exists because $\mathcal{Q}^U_{0|x}, \mathcal{Q}^U_{1|x}, \mathcal{Q}^U_{2|x}$ consists of three infinite angular sectors located at $(-U_{x1}, -U_{x2})$. Shrink σ arbitrarily close to zero and consider the continuous function
$$\iota_x : t \in [a, b] \mapsto h_x(\gamma_{0,1}(t), \sigma) \in \mathcal{S}.$$
Since σ is arbitrarily close to zero and $\gamma_{0,1}(t) \notin \mathcal{Q}^U_{2|x}$, it follows that the third component

of $\iota_x(t)$ is arbitrarily close to zero, for each $t \in [a, b]$. Further, observe that the endpoints of the image set of ι_x are arbitrarily close to the vertices $(1, 0, 0)$ and $(0, 1, 0)$ of the unit simplex. Therefore, the image set of ι_x is arbitrarily close to the edge of the unit simplex between the vertices $(1, 0, 0)$ and $(0, 1, 0)$. Further, by the same arguments of 1. of Step 3.2, the image set of ι_x contains such an edge because the support of H_x is bounded.

Repeat the same argument for the other two edges, by fixing any two points in the interiors of $\mathcal{Q}^U_{0|x}$ and $\mathcal{Q}^U_{2|x}$, and any two points in the interiors of $\mathcal{Q}^U_{1|x}$ and $\mathcal{Q}^U_{2|x}$.

Conclude that the image set of h_x is arbitrarily close to (and include) all the edges of the unit simplex.

The two facts above imply that the image set of h_x is the unit simplex.

Step 3.3 Since the image set of h_x is the unit simplex, there exists $\mu \in \mathbb{R}^2$ and $\sigma \in \mathbb{R}_+$ such that

$$p_{y|x} = \lambda_{H_{x,\mu,\sigma}}(\mathcal{Q}^U_{y|x}) \ \forall y \in \mathcal{Y}_0.$$

Step 4 In this step we go back to the original dimension of our problem. Consider any random vector $(\epsilon_{i0}, \tilde{\epsilon}_{i1}, \tilde{\epsilon}_{i2})$ such that $(\tilde{\epsilon}_{i1}, \tilde{\epsilon}_{i2})$ has CDF conditional on $X_i = x$ equal to $H_{x,\mu,\sigma}$, for every $x \in \mathcal{X}$. Take the random vector $\epsilon_i \equiv (\epsilon_{i0}, \tilde{\epsilon}_{i1} + \epsilon_{i0}, \tilde{\epsilon}_{i2} + \epsilon_{i0})$ and denote its CDF conditional on $X_i = x$ by F_x, for every $x \in \mathcal{X}$. By construction, $\{F_x\}_{x\in\mathcal{X}}$ solves (A.1.1).

Step 5 Repeat Steps 1-4 for the women's side to show that there exists $\{G_y\}_{y\in\mathcal{Y}} \in \mathcal{G}$ such that

$$p_{x|y} = \lambda_{G_y}(\mathcal{R}^V_{x|y}) \ \forall (y, x) \in \mathcal{Y} \times \mathcal{X}_0,$$

where λ_{G_y} is the measure associated with G_y. Hence, the result claimed by Proposition 2 follows.

Step 6 In Section 5 and Appendix E.1, we discuss some scale normalisations that one may want to impose on U and V.[22] In short, these scale normalisations amount to fixing to 1 the absolute value of some of the U and V's components. Since these are just specific values of U and V, the above steps still go through.

A.2 Proof of Proposition 3

The proof is organised in the following steps. In Step 0, we recall the notation introduced in Section 4.5.2 and introduce some new one. In Step 1, we present the notion of an equivalence class for every $U \in \mathcal{U}$ and prove that if $\tilde{U}, \hat{U} \in \mathcal{U}$ belong to the same equivalence class, then they induce the same feasible set of the linear programming problem of Section 4.5.1. In Step 2, we

[22]Note that appropriate location normalisations have been already imposed by Proposition 1, i.e., $U_{x0} = V_{0y} = 0$ for every $(x, y) \in \mathcal{X} \times \mathcal{Y}$.

show how such equivalence classes are related to the notion of π-ordering used in Proposition 3. In Step 3, we conclude. Remark A.1 explains how Proposition 3 is implemented in practice.

For simplicity of exposition, we provide the proof of Proposition 3 for the case $r = 2$ and Assumption 5.2. The proof for a generic case follows exactly the same steps, but becomes notationally more complicated. In the case considered, we have that

$$\begin{aligned}
&\mathcal{A}_{x,1,U} \equiv \{-U_{x1}, U_{x1}, 0, \infty, -\infty\},\\
&\mathcal{A}_{x,2,U} \equiv \{-U_{x2}, U_{x2}, 0, \infty, -\infty\},\\
&\mathcal{A}_{x,3,U} \equiv \{U_{x2} - U_{x1}, -U_{x2} + U_{x1}, 0, \infty, -\infty\},
\end{aligned}$$

for every $x \in \mathcal{X}$ and $U \in \mathcal{U}$. Therefore, for any given $U \in \mathcal{U}$ and by following Section 4.5.1,

$$\exists\ \{\Delta F_x\}_{x\in\mathcal{X}} \in \Delta\mathcal{F}^{\dagger} \text{ s.t. } p_{y|x} = \kappa(U, \Delta F_x, y)\ \forall(x, y) \in \mathcal{X} \times \mathcal{Y}_0,$$

if and only if, for every $x \in \mathcal{X}$, the following linear programming problem is feasible with respect to $\Delta\bar{F}^U_x : \mathcal{A}_{x,U} \to \mathbb{R}$,

$$p_{1|x} = 1 + \Delta\bar{F}^U_x(-U_{x1}, \infty, U_{x2} - U_{x1}) - \Delta\bar{F}^U_x(\infty, \infty, U_{x2} - U_{x1}) - \Delta\bar{F}^U_x(-U_{x1}, \infty, \infty), \tag{A.2.1}$$

$$p_{2|x} = \Delta\bar{F}^U_x(\infty, \infty, U_{x2} - U_{x1}) - \Delta\bar{F}^U_x(\infty, -U_{x2}, U_{x2} - U_{x1}), \tag{A.2.2}$$

$$p_{0|x} = \Delta\bar{F}^U_x(-U_{x1}, -U_{x2}, \infty), \tag{A.2.3}$$

$$\Delta\bar{F}^U_x(-U_{x1}, \infty, \infty) = 1 - \Delta\bar{F}^U_x(-U_{x1}, \infty, \infty), \tag{A.2.4}$$

$$\Delta\bar{F}^U_x(\infty, -U_{x2}, \infty) = 1 - \Delta\bar{F}^U_x(\infty, -U_{x2}, -\infty), \tag{A.2.5}$$

$$\Delta\bar{F}^U_x(\infty, \infty, U_{x2} - U_{x1}) = 1 - \Delta\bar{F}^U_x(\infty, \infty, -U_{x2} + U_{x1}), \tag{A.2.6}$$

$$\Delta\bar{F}^U_x(0, \infty, \infty) = 1/2, \tag{A.2.7}$$

$$\Delta\bar{F}^U_x(\infty, 0, \infty) = 1/2, \tag{A.2.8}$$

$$\Delta\bar{F}^U_x(\infty, \infty, 0) = 1/2, \tag{A.2.9}$$

$$\Delta\bar{F}^U_x(-\infty, t, q) = 0\ \forall(t, q) \in \mathcal{A}_{x,2,U} \times \mathcal{A}_{x,3,U}, \tag{A.2.10}$$

$$\Delta\bar{F}^U_x(t, -\infty, q) = 0\ \forall(t, q) \in \mathcal{A}_{x,1,U} \times \mathcal{A}_{x,3,U}, \tag{A.2.11}$$

$$\Delta\bar{F}^U_x(t, q, -\infty) = 0\ \forall(t, q) \in \mathcal{A}_{x,1,U} \times \mathcal{A}_{x,2,U}, \tag{A.2.12}$$

$$\Delta\bar{F}^U_x(\infty, \infty, \infty) = 1, \tag{A.2.13}$$

$$0 \le \Delta\bar{F}^U_x(t, q, r) \le 1\ \forall(t, q, r) \in \mathcal{A}_{x,U}, \tag{A.2.14}$$

$$\begin{matrix} [-\Delta\bar{F}^U_x(t,q,r) + \Delta\bar{F}^U_x(t',q,r) + \Delta\bar{F}^U_x(t,q',r) - \Delta\bar{F}^U_x(t',q',r) + & \forall(t,q,r),(t',q',r') \in \mathcal{A}_{x,U} \\ \Delta\bar{F}^U_x(t,q,r') - \Delta\bar{F}^U_x(t',q,r') - \Delta\bar{F}^U_x(t,q',r') + \Delta\bar{F}^U_x(t',q',r')] \ge 0 & \text{s.t. } (t,q,r) \le (t',q',r'). \end{matrix} \tag{A.2.15}$$

In the linear programming above: (A.2.1)-(A.2.3) match the predicted conditional choice probabilities with the empirical ones; (A.2.4)-(A.2.9) impose Assumption 5.2; (A.2.10)-(A.2.15) ensure that $\Delta\bar{F}^U_x$ can be extended to a proper conditional CDF.

Step 0 In this step, we recall the notation introduced in Section 4.5.2 and introduce some new ones.

Fix $U \in \mathcal{U}$ and $x \in \mathcal{X}$. In the setting considered, $\mathcal{A}_{x,l,U}$ has cardinality 5 for every $l \in \{1,2,3\}$. $\mathcal{A}_{x,U}$ has cardinality 5^3. The image set of $\Delta\bar{F}_x^U$, which we denote by $\Delta\bar{F}_x^U(\mathcal{A}_{x,U})$, has cardinality 5^3. In all such sets, possible repetitions of elements are kept.

For every $l \in \{1,2,3\}$, $\alpha_{x,l,U}$ is a 5×1 column vector listing the 5 elements of $\mathcal{A}_{x,l,U}$. $\alpha_{x,U}$ is a $5^3 \times 3$ matrix listing the 5^3 3-tuples of $\mathcal{A}_{x,U}$. Further, we reorder $\alpha_{x,U}$ lexicographically by row and call the reordered matrix $\alpha_{x,U}^L$.[23] $f_{x,U}$ is a $5^3 \times 1$ column vector listing the 5^3 elements of $\Delta\bar{F}_x^U(\mathcal{A}_{x,U})$. $\iota : \Delta\bar{F}_x^U(\mathcal{A}_{x,U}) \to \{1,2,...,5^3\}$, where $\iota(k)$ is the row index of scalar k in the vector $f_{x,U}$. $\tau : \mathcal{A}_{x,U} \to \{1,2,...,5^3\}$, where $\tau(k)$ is the row index of 3-tuple k in the matrix $\alpha_{x,U}^L$.

For every $l \in \{1,2,3\}$, Π_1 is the set of all possible permutations without repetition of $\{1,2,3,4,5\}$ and $\Pi_2 \equiv \{<,=\}^4$. $\pi : \bar{\mathbb{R}}^5 \to \Pi_1 \times \Pi_2$, where $\pi(\alpha) \equiv (\pi_1(\alpha), \pi_2(\alpha))$, $\forall \alpha \in \bar{\mathbb{R}}^5$. In particular, $\pi_1(\alpha)$ sorts the 5 elements of α from smallest to largest and reports their positions in the original vector. $\pi_2(\alpha)$ reports the relational operators, $<$ or $=$, among the sorted elements of α. When α contains multiple elements with the same value, then we adopt any convention on which element to sort first. We call $\pi(\alpha)$ the "π-ordering" of α. For instance, suppose $\alpha = (100, 99, 0, \infty, -\infty)$. Then, $\pi(\alpha) = \{(5,3,2,1,4),(<,<,<,<)\}$. Suppose $\alpha = (1,5,5,6,-\infty)$. Then, $\pi(\alpha) = \{(5,1,2,3,4),(<,<,=,<)\}$.

Step 1 In this step, we present the notion of an equivalence class for every $U \in \mathcal{U}$ and prove that if $\tilde{U}, \hat{U} \in \mathcal{U}$ belong to the same equivalence class, then they induce the same feasible set to the linear programming problem (A.2.1)-(A.2.15).

Let $x \in \mathcal{X}$ and $\tilde{U}, \hat{U} \in \mathcal{U}$. Define

$$\mathcal{C}_x(\tilde{U}) \equiv \Big\{\{(\tilde{t},\tilde{q},\tilde{r}),(\tilde{t}',\tilde{q}',\tilde{r}')\} : (\tilde{t},\tilde{q},\tilde{r}),(\tilde{t}',\tilde{q}',\tilde{r}') \in \mathcal{A}_x(\tilde{U}), (\tilde{t},\tilde{q},\tilde{r}) \leq (\tilde{t}',\tilde{q}',\tilde{r}')\Big\},$$

$$\mathcal{C}_x(\hat{U}) \equiv \Big\{\{(\hat{t},\hat{q},\hat{r}),(\hat{t}',\hat{q}',\hat{r}')\} : (\hat{t},\hat{q},\hat{r}),(\hat{t}',\hat{q}',\hat{r}') \in \mathcal{A}_x(\hat{U}), (\hat{t},\hat{q},\hat{r}) \leq (\hat{t}',\hat{q}',\hat{r}')\Big\}.$$

Definition A.1. (*Equivalence class*) Let $\tilde{U}, \hat{U} \in \mathcal{U}$. $\hat{U}$ is said to belong to the equivalence class of $\tilde{U}$ at $x \in \mathcal{X}$ if the following two conditions hold:

[23]For example, suppose that $\mathcal{A}_{x,U}$ contains three 3-tuples (instead of 5^3 3-tuples, for shortness of exposition): $(3,1,2),(2,3,4),(2,1,3)$. Then,

$$\alpha_{x,U}^L \equiv \begin{pmatrix} 2 & 1 & 3 \\ 2 & 3 & 4 \\ 3 & 1 & 2 \end{pmatrix}.$$

1. For every $\{(\tilde{t}, \tilde{q}, \tilde{r}), (\tilde{t}', \tilde{q}', \tilde{r}')\} \in \mathcal{C}_x(\tilde{U})$, there exists $\{(\hat{t}, \hat{q}, \hat{r}), (\hat{t}', \hat{q}', \hat{r}')\} \in \mathcal{C}_x(\hat{U})$ such that

$$\begin{aligned}
\iota(\Delta\bar{F}_x^{\tilde{U}}(\tilde{t}, \tilde{q}, \tilde{r})) &= \iota(\Delta\bar{F}_x^{\hat{U}}(\hat{t}, \hat{q}, \hat{r})), \\
\iota(\Delta\bar{F}_x^{\tilde{U}}(\tilde{t}', \tilde{q}, \tilde{r})) &= \iota(\Delta\bar{F}_x^{\hat{U}}(\hat{t}', \hat{q}, \hat{r})), \\
\iota(\Delta\bar{F}_x^{\tilde{U}}(\tilde{t}, \tilde{q}', \tilde{r})) &= \iota(\Delta\bar{F}_x^{\hat{U}}(\hat{t}, \hat{q}', \hat{r})), \\
\iota(\Delta\bar{F}_x^{\tilde{U}}(\tilde{t}', \tilde{q}', \tilde{r})) &= \iota(\Delta\bar{F}_x^{\hat{U}}(\hat{t}', \hat{q}', \hat{r})), \\
\iota(\Delta\bar{F}_x^{\tilde{U}}(\tilde{t}, \tilde{q}, \tilde{r}')) &= \iota(\Delta\bar{F}_x^{\hat{U}}(\hat{t}, \hat{q}, \hat{r}')), \\
\iota(\Delta\bar{F}_x^{\tilde{U}}(\tilde{t}', \tilde{q}, \tilde{r}')) &= \iota(\Delta\bar{F}_x^{\hat{U}}(\hat{t}', \hat{q}, \hat{r}')), \\
\iota(\Delta\bar{F}_x^{\tilde{U}}(\tilde{t}, \tilde{q}', \tilde{r}')) &= \iota(\Delta\bar{F}_x^{\hat{U}}(\hat{t}, \hat{q}', \hat{r}')), \\
\iota(\Delta\bar{F}_x^{\tilde{U}}(\tilde{t}', \tilde{q}', \tilde{r}')) &= \iota(\Delta\bar{F}_x^{\hat{U}}(\hat{t}', \hat{q}', \hat{r}')),
\end{aligned}$$

   and vice-versa.

2. $\pi_2(\alpha_{x,l,\tilde{U}}) = \pi_2(\alpha_{x,l,\hat{U}})$ for every $l \in \{1, 2, 3\}$.

Let $[\tilde{U}]_x$ denote the equivalence class of $\tilde{U}$ at $x \in \mathcal{X}$. ⋄

Lemma A.1. Let $x \in \mathcal{X}$ and $\tilde{U}, \hat{U} \in \mathcal{U}$. If $\hat{U} \in [\tilde{U}]_x$, then $\tilde{U}$ and $\hat{U}$ induce the same feasible set to the linear programming problem (A.2.1)-(A.2.15). ⋄

Proof. Let $x \in \mathcal{X}$ and $\tilde{U}, \hat{U} \in \mathcal{U}$. Suppose $\hat{U} \in [\tilde{U}]_x$. Take any $\{(\tilde{t}, \tilde{q}, \tilde{r}), (\tilde{t}', \tilde{q}', \tilde{r}')\} \in \mathcal{C}_x(\tilde{U})$ and a corresponding $\{(\hat{t}, \hat{q}, \hat{r}), (\hat{t}', \hat{q}', \hat{r}')\} \in \mathcal{C}_x(\hat{U})$ such that Condition 1 of Definition A.1 holds. Consider constraint (A.2.15) at $\{\tilde{U}, (\tilde{t}, \tilde{q}, \tilde{r}), (\tilde{t}', \tilde{q}', \tilde{r}')\}$, where the terms of the form $\Delta\bar{F}_x^{\tilde{U}}(\cdot)$ are unknowns to be determined. Relabel them as $\theta_{\iota(\Delta\bar{F}_x^{\tilde{U}}(\cdot))}$. Then, restate (A.2.15) as

$$\begin{aligned}
&-\theta_{\iota(\Delta\bar{F}_x^{\tilde{U}}(\tilde{t},\tilde{q},\tilde{r}))} + \theta_{\iota(\Delta\bar{F}_x^{\tilde{U}}(\tilde{t}',\tilde{q},\tilde{r}))} + \theta_{\iota(\Delta\bar{F}_x^{\tilde{U}}(\tilde{t},\tilde{q}',\tilde{r}))} - \theta_{\iota(\Delta\bar{F}_x^{\tilde{U}}(\tilde{t}',\tilde{q}',\tilde{r}))} \\
&+\theta_{\iota(\Delta\bar{F}_x^{\tilde{U}}(\tilde{t},\tilde{q},\tilde{r}'))} - \theta_{\iota(\Delta\bar{F}_x^{\tilde{U}}(\tilde{t}',\tilde{q},\tilde{r}'))} - \theta_{\iota(\Delta\bar{F}_x^{\tilde{U}}(\tilde{t},\tilde{q}',\tilde{r}'))} + \theta_{\iota(\Delta\bar{F}_x^{\tilde{U}}(\tilde{t}',\tilde{q}',\tilde{r}'))} \geq 0,
\end{aligned} \tag{A.2.16}$$

where θ is a $5^3 \times 1$ vector of unknowns and θ_h denotes the h-th element of θ. Similarly, consider the following relabelled constraint corresponding to $\hat{U}$,

$$\begin{aligned}
&-\theta_{\iota(\Delta\bar{F}_x^{\hat{U}}(\hat{t},\hat{q},\hat{r}))} + \theta_{\iota(\Delta\bar{F}_x^{\hat{U}}(\hat{t}',\hat{q},\hat{r}))} + \theta_{\iota(\Delta\bar{F}_x^{\hat{U}}(\hat{t},\hat{q}',\hat{r}))} - \theta_{\iota(\Delta\bar{F}_x^{\hat{U}}(\hat{t}',\hat{q}',\hat{r}))} \\
&+\theta_{\iota(\Delta\bar{F}_x^{\hat{U}}(\hat{t},\hat{q},\hat{r}'))} - \theta_{\iota(\Delta\bar{F}_x^{\hat{U}}(\hat{t}',\hat{q},\hat{r}'))} - \theta_{\iota(\Delta\bar{F}_x^{\hat{U}}(\hat{t},\hat{q}',\hat{r}'))} + \theta_{\iota(\Delta\bar{F}_x^{\hat{U}}(\hat{t}',\hat{q}',\hat{r}'))} \geq 0.
\end{aligned} \tag{A.2.17}$$

By Condition 1 of Definition A.1, the subscripts of θ in (A.2.16) and (A.2.17) are identical. Further, if some or all of the components of $(\tilde{t}, \tilde{q}, \tilde{r})$ are equal to $(\tilde{t}', \tilde{q}', \tilde{r}')$, then Condition 2 of Definition A.1 ensures that the same is true across $(\hat{t}, \hat{q}, \hat{r})$ and $(\hat{t}', \hat{q}', \hat{r}')$. Therefore, (A.2.16) and (A.2.17) are identical. Similar arguments can be repeated for every $\{(\tilde{t}, \tilde{q}, \tilde{r}), (\tilde{t}', \tilde{q}', \tilde{r}')\} \in \mathcal{C}_x(\tilde{U})$ and $\{(\hat{t}, \hat{q}, \hat{r}), (\hat{t}', \hat{q}', \hat{r}')\} \in \mathcal{C}_x(\hat{U})$ so that the constraints imposing 3-increasingness, (A.2.15), are identical under $\tilde{U}$ and $\hat{U}$.

If the constraints (A.2.15) are identical under $\tilde{U}$ and $\hat{U}$, then they induce the feasible set to the linear programming problem (A.2.1)-(A.2.15). This is because the only piece of (A.2.1)-(A.2.15) that can potentially generate different solutions for different values of U is the one requiring the function $\Delta\bar{F}_x^U$ to be 3-increasing. □

Step 2 In this step, we show how the equivalence classes of Step 1 are related to the notion of π-ordering used in Proposition 3.

Lemma A.2. Let $x \in \mathcal{X}$ and $\tilde{U}, \hat{U} \in \mathcal{U}$. If

i. $\pi_1(\alpha_{x,l,\tilde{U}}) = \pi_1(\alpha_{x,l,\hat{U}})$ for every $l \in \{1,2,3\}$,

ii. $\pi_2(\alpha_{x,l,\tilde{U}}) = \pi_2(\alpha_{x,l,\hat{U}})$ for every $l \in \{1,2,3\}$,

then $\hat{U} \in [\tilde{U}]_x$. ◇

Proof. Condition ii of Lemma A.2 coincides with Condition 2 of Definition A.1. Therefore, in what follows we show that Conditions i and ii of Lemma A.2 imply Condition 1 of Definition A.1.

Let $x \in \mathcal{X}$. Let $\tilde{U}, \hat{U} \in \mathcal{U}$ such that $\pi(\alpha_{x,l,\tilde{U}}) = \pi(\alpha_{x,l,\hat{U}})$ for every $l \in \{1,2,3\}$. Take any $\{(\tilde{t},\tilde{q},\tilde{r}),(\tilde{t}',\tilde{q}',\tilde{r}')\} \in \mathcal{C}_x(\tilde{U})$, i.e., a comparable pair of 3-tuples from $\mathcal{A}_x(\tilde{U})$. Pick $(\hat{t},\hat{q},\hat{r}),(\hat{t}',\hat{q}',\hat{r}') \in \mathcal{A}_x(\hat{U})$ such that $\tau((\hat{t},\hat{q},\hat{r})) = \tau((\tilde{t},\tilde{q},\tilde{r}))$ and $\tau((\hat{t}',\hat{q}',\hat{r}')) = \tau((\tilde{t}',\tilde{q}',\tilde{r}'))$. Given that $\pi_1(\alpha_{x,l,\tilde{U}}) = \pi_1(\alpha_{x,l,\hat{U}})$ for every $l \in \{1,2,3\}$, it should be that $\{(\hat{t},\hat{q},\hat{r}),(\hat{t}',\hat{q}',\hat{r}')\} \in \mathcal{C}_x(\hat{U})$. That is, by construction, $(\hat{t},\hat{q},\hat{r}),(\hat{t}',\hat{q}',\hat{r}')$ should be a comparable pair of 3-tuples from $\mathcal{A}_x(\hat{U})$. Moreover, given $\pi_1(\alpha_{x,l,\tilde{U}}) = \pi_1(\alpha_{x,l,\hat{U}})$ for every $l \in \{1,2,3\}$, it should be that

$$
\begin{aligned}
\tau((\tilde{t}',\tilde{q},\tilde{r})) &= \tau((\hat{t}',\hat{q},\hat{r})),\\
\tau((\tilde{t},\tilde{q}',\tilde{r})) &= \tau((\hat{t},\hat{q}',\hat{r})),\\
\tau((\tilde{t}',\tilde{q}',\tilde{r})) &= \tau((\hat{t}',\hat{q}',\hat{r})),\\
\tau((\tilde{t},\tilde{q},\tilde{r}')) &= \tau((\hat{t},\hat{q},\hat{r}')),\\
\tau((\tilde{t}',\tilde{q},\tilde{r}')) &= \tau((\hat{t}',\hat{q},\hat{r}')),\\
\tau((\tilde{t},\tilde{q}',\tilde{r}')) &= \tau((\hat{t},\hat{q}',\hat{r}')).
\end{aligned}
$$

Construct a new $5^3 \times 3$ matrix, where the first, second, and third columns are defined as

$$
\begin{aligned}
&\mathbf{1}_{5^0} \otimes \left[(\pi_1(\alpha_{x,1,\tilde{U}}))^T \otimes \mathbf{1}_{5^2}\right],\\
&\mathbf{1}_{5^1} \otimes \left[(\pi_1(\alpha_{x,2,\tilde{U}}))^T \otimes \mathbf{1}_{5^1}\right],\\
&\mathbf{1}_{5^2} \otimes \left[(\pi_1(\alpha_{x,3,\tilde{U}}))^T \otimes \mathbf{1}_{5^0}\right],
\end{aligned}
$$

respectively, with $\mathbf{1}_d$ denoting the d-dimensional vector of ones. Reorder this new matrix lexicographically by row and call the reordered matrix $b_{x,\tilde{U}}$. Observe that $b_{x,\tilde{U}}$ is a standardised relabelling of the matrix $\alpha^L_{x,\tilde{U}}$.[24] Construct the matrix $b_{x,\hat{U}}$ in a similar way.

Note that $\pi_2(\alpha_{x,l,\tilde{U}}) = \pi_2(\alpha_{x,l,\hat{U}})$ for every $l \in \{1,2,3\}$ ensures that $b_{x,\tilde{U}}$ and $b_{x,\hat{U}}$ are *logically* equivalent. Further, $\pi_1(\alpha_{x,l,\tilde{U}}) = \pi_1(\alpha_{x,l,\hat{U}})$ for every $l \in \{1,2,3\}$ implies $b_{x,\tilde{U}} = b_{x,\hat{U}}$. Therefore, it must be that Condition 1 of Definition A.1 is satisfied.

□

Step 3 In this step, we combine Steps 1 and 2 and conclude. Let $x \in \mathcal{X}$ and $\tilde{U}, \hat{U} \in \mathcal{U}$. Lemmas A.1 and A.2 imply that if $\alpha_{x,l,\tilde{U}}$ and $\alpha_{x,l,\hat{U}}$ have the same π-ordering for every $l \in \{1,2,3\}$ (or, equivalently, if $\tilde{U}$ and $\hat{U}$ induce the same π-ordering), then the linear programming problem (A.2.1)-(A.2.15) for $\tilde{U}$ has the same feasible set as the linear programming problem (A.2.1)-(A.2.15) for $\hat{U}$. Hence, the result claimed by Proposition 3 follows.

Remark A.1. (*Proposition 3 in practice*) In practice, we use Proposition 3 as follows. First, we generate a grid of points covering $\mathcal{U}$ as precisely as possible, depending on the available computational resources. Second, we find the π-ordering of each grid point for every $l \in \{1, ..., d\}$ and $x \in \mathcal{X}$. Third, we group the grid points producing the same π-ordering for every $l \in \{1, ..., d\}$ and $x \in \mathcal{X}$ into the same "partitioning set". In Matlab, steps 2 and 3 can be straightforwardly implemented using the function `sort`. Fourth, we take a representative grid point for each partitioning set. Fifth, for each of the selected grid points, we verify the feasibility of the linear programming problem of Section 4.5.1. If the linear programming problem is feasible, then the corresponding partitioning set belongs to the identified set (conditionally on the degeneracy condition being satisfied, as discussed in Section 4.5.3). Otherwise, the corresponding partitioning set does not belong to the identified set. The overall procedure can be easily parallelised.

Note that the number of partitioning sets increases with the granularity of the initial grid of points covering $\mathcal{U}$, with r, and with the number of nonparametric restrictions on $\{\Delta F_x\}_{x\in\mathcal{X}}$. However, providing a general formula for the number of partitioning sets does not seem viable. Note also that if $\Delta\mathcal{F}^\dagger = \Delta\mathcal{F}$, then there is only one partitioning set. ◇

A.3 Proof of Proposition 4

For simplicity of exposition, we provide the proof of Proposition 4 when $r = 2$. The proof for a generic r follows exactly the same steps, but becomes notationally more complicated.

Step 1 As highlighted in the main text, we should firstly observe that

$$\mathcal{B} \equiv \{(b_1, b_2, b_3) \in \mathbb{R}^3 : \ b_3 = b_1 - b_2\} = \{(b_1, b_2, b_3) \in \mathbb{R}^3 : \ b_1 = b_2 + b_3\}.$$

[24]The matrix $b_{x,U}$ abstracts from the magnitudes of the elements of $\alpha^L_{x,U}$ and captures their relative positions.

Accordingly, the relevant 3-dimensional boxes defined in Proposition 4 are

$$B_{1,2,3}(\hat{b},\tilde{b}) \equiv \{(x,y,z) \in \mathbb{R}^3 : \; x > \hat{b}+\tilde{b}, y \leq \hat{b}, z \leq \tilde{b}\},$$
$$Q_{1,2,3}(\hat{b},\tilde{b}) \equiv \{(x,y,z) \in \mathbb{R}^3 : \; x \leq \hat{b}+\tilde{b}, y > \hat{b}, z > \tilde{b}\},$$

for any $(\hat{b},\tilde{b}) \in \mathbb{Q}^2$.

Step 2 Let

$$A_1 \equiv \{(x,y,z) \in \mathbb{R}^3 : \; x > y+z\}.$$

We now show that

$$A_1 = \cup_{(\hat{b},\tilde{b})\in\mathbb{Q}^2} B_{1,2,3}(\hat{b},\tilde{b}).$$

It is clear that $\cup_{(\hat{b},\tilde{b})\in\mathbb{Q}^2} B_{1,2,3}(\hat{b},\tilde{b}) \subseteq A_1$. To prove the reverse, take any $(x,y,z) \in A_1$ and $\epsilon \equiv x-(y+z) > 0$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, there exists $p \in [y, y+\frac{\epsilon}{2}] \cap \mathbb{Q}$ and $q \in [z, z+\frac{\epsilon}{2}] \cap \mathbb{Q}$. Therefore, $x = y+z+\epsilon > p+q$ and, hence, $(x,y,z) \in B_{1,2,3}(p,q)$. Thus, $A_1 \subseteq \cup_{(\hat{b},\tilde{b})\in\mathbb{Q}^2} B_{1,2,3}(\hat{b},\tilde{b})$.

Step 3 Let

$$A_2 \equiv \{(x,y,z) \in \mathbb{R}^3 : \; x < y+z\}.$$

By following the same arguments of step 2, we can show that

$$A_2 = \cup_{(\hat{b},\tilde{b})\in\mathbb{Q}^2} Q_{1,2,3}(\hat{b},\tilde{b}).$$

Step 4 Assume $\lambda_{\Delta F_x}(B_{1,2,3}(\hat{b},\tilde{b})) = \lambda_{\Delta F_x}(Q_{1,2,3}(\hat{b},\tilde{b})) = 0 \; \forall (\hat{b},\tilde{b}) \in \mathbb{Q}^2$. Hence, by Step 3, A_1 and A_2 are disjoint and infinitely countable unions of zero probability measure sets. Note that $\mathcal{B}^c = A_1 \cup A_2$, where $\mathcal{B}^c$ denotes the complement of the region $\mathcal{B}$ in $\mathbb{R}^3$. Therefore, $\lambda_{\Delta F_x}(\mathcal{B}^c) = \lambda_{\Delta F_x}(A_1 \cup A_2) = 0$, which is equivalent to $\lambda_{\Delta F_x}(\mathcal{B}) = 1$.

Step 5 Conversely, assume $\lambda_{\Delta F_x}(\mathcal{B}) = 1$. This implies $\lambda_{\Delta F_x}(\cup_{(\hat{b},\tilde{b})\in\mathbb{Q}^2} B_{1,2,3}(\hat{b},\tilde{b})) = 0$ and $\lambda_{\Delta F_x}(\cup_{(\hat{b},\tilde{b})\in\mathbb{Q}^2} Q_{1,2,3}(\hat{b},\tilde{b})) = 0$. In turn, $\lambda_{\Delta F_x}(B_{1,2,3}(\hat{b},\tilde{b})) = \lambda_{\Delta F_x}(Q_{1,2,3}(\hat{b},\tilde{b})) = 0 \; \forall (\hat{b},\tilde{b}) \in \mathbb{Q}^2$.

B Theorem 1 in Torgovitsky (2019)

In this section we provide a formal statement of Theorem 1 in Torgovitsky (2019) within our framework. We start with outlining some useful assumptions. We refer the reader to Definitions 1-5, Lemmas 1-2, and Corollary 1 in Torgovitsky (2019), which are the other key results and definitions used by Theorem 1. In what follows, $\Delta F_x|_{\mathcal{C}}$ denotes the restriction of ΔF_x to a subset, $\mathcal{C}$, of its domain.

Assumption B.1. (*Assumption A, Torgovitsky, 2019*) $\Delta\mathcal{F}^\dagger$ satisfies the following properties: for each $\{\Delta F_x\}_{x\in\mathcal{X}} \in \Delta\mathcal{F}^\dagger$, it holds that

1. $\Delta F_x(a) = \Delta F_{\tilde{x}}(a)\ \forall x, \tilde{x} \in \{\mathcal{X} \cap \mathcal{X}^{\dagger}_{0,m_0}\}$, $\forall a \in \bar{\mathbb{R}}^d$, $\forall m_0 \in \{1, ..., M_0\}$, where each $\mathcal{X}^{\dagger}_{0,m_0}$ is a known (possibly empty) subset of $\mathcal{X}$.

2. $\Delta F_{x,l}(a) = \Delta F_{\tilde{x},l}(a)\ \forall x, \tilde{x} \in \{\mathcal{X} \cap \mathcal{X}^{\dagger}_{l,m_l}\}$, $\forall a \in \bar{\mathbb{R}}$, $\forall m_l \in \{1, ..., M_L\}$, $\forall l \in \{1, ..., d\}$, where each $\mathcal{X}^{\dagger}_{l,m_l}$ is a known (possibly empty) subset of $\mathcal{X}$.

3. $\{\Delta F_{x,l}\}_{x\in\mathcal{X}} \in \Delta\mathcal{F}^{\dagger}_l\ \forall l \in \{1, ..., d\}$, where $\Delta\mathcal{F}^{\dagger}_l$ is a known collection of families of one-dimensional conditional CDFs.

4. $\rho(U, \{\Delta F_x\}_{x\in\mathcal{X}}) \geq 0$ for some known vector-valued function ρ, where the inequality is interpreted component wise.

◇

In Assumption B.1, Conditions 1 and 2 are independence restrictions on $\{\Delta F_x\}_{x\in\mathcal{X}}$ and $\{\Delta F_{x,l}\}_{x\in\mathcal{X}}$, respectively. Condition 3 requires $\{\Delta F_{x,l}\}_{x\in\mathcal{X}}$ to be extendible in the sense described below in Proposition B.1. Condition 4 allows for miscellaneous restrictions, represented here by a function ρ chosen by the researcher. Any of the Conditions 1-4 can be made non-restrictive by using specific choices of $\mathcal{X}^{\dagger}_{0,m_0}$, $\mathcal{X}^{\dagger}_{l,m_l}$, $\Delta\mathcal{F}^{\dagger}_l$, or ρ. The restrictions listed in Assumption 5 of Section 4.5.1 can be written in terms of 1-4.

Assumption B.2. (*Condition U, Torgovitsky, 2019*) Suppose that $\Delta\mathcal{F}^{\dagger}$ satisfies Assumption B.1. A collection of sets, $\{\mathcal{A}_{x,U}\}_{x\in\mathcal{X}}$, satisfies Assumption B.2 if the following properties hold:

1. For every $x \in \mathcal{X}$, $\mathcal{A}_{x,U} \equiv \mathcal{A}_{x,1,U} \times \mathcal{A}_{x,2,U} \times ... \times \mathcal{A}_{x,d}(U)$, where $\mathcal{A}_{x,l,U} \subseteq \bar{\mathbb{R}}$ is closed and such that $\{\infty, -\infty\} \subseteq \mathcal{A}_{x,l,U}\ \forall l \in \{1, ..., d\}$.

2. There exists functions $\bar{\kappa}$ and $\bar{\rho}$ such that, $\forall \{\Delta F_x\}_{x\in\mathcal{X}} \in \Delta\mathcal{F}^{\dagger}$,

$$\kappa(U, \Delta F_x, y) = \bar{\kappa}(U, \Delta F_x|_{\mathcal{A}_{x,U}}, y)\ \forall (x, y) \in \mathcal{X} \times \mathcal{Y}_0,$$
$$\rho(U, \{\Delta F_x\}_{x\in\mathcal{X}}) = \bar{\rho}(U, \{\Delta F_x|_{\mathcal{A}_{x,U}}\}_{x\in\mathcal{X}}).$$

3. $\forall l \in \{1, ..., d\}$, there exists a collection of families of conditional subsdistributions, $\Delta\bar{\mathcal{F}}^{\dagger}_l$, such that

$$\Delta\bar{\mathcal{F}}^{\dagger}_l \text{ is extendible to } \Delta\mathcal{F}^{\dagger}_l,$$
$$\Delta\mathcal{F}^{\dagger}_l \text{ is reducible to } \Delta\bar{\mathcal{F}}^{\dagger}_l,$$
$$\forall \{\Delta\bar{F}^U_{x,l}\}_{x\in\mathcal{X}} \in \Delta\bar{\mathcal{F}}^{\dagger}_l, \text{ every } \Delta\bar{F}^U_{x,l} \text{ has common domain } \mathcal{A}_{x,l,U}.$$

4. $\mathcal{A}_{x,U} = \mathcal{A}_{\tilde{x},U}\ \forall x, \tilde{x} \in \{\mathcal{X} \cap \mathcal{X}^{\dagger}_{0,m_0}\}$ and $\forall m_0 \in \{1, ..., M_0\}$.

5. $\mathcal{A}_{x,l,U} = \mathcal{A}_{\tilde{x},l,U}\ \forall x, \tilde{x} \in \{\mathcal{X} \cap \mathcal{X}^{\dagger}_{l,m_l}\}$, $\forall m_l \in \{1, ..., M_L\}$, and $\forall l \in \{1, ..., d\}$.

◇

Proposition B.1. (*Theorem 1, Torgovitsky, 2019*) Suppose that $\Delta\mathcal{F}^\dagger$ can be represented as in Assumption B.1. Take any $U \in \mathcal{U}$. Let $\{\mathcal{A}_{x,U}\}_{x\in\mathcal{X}}$ be any collection of subsets of $\bar{\mathbb{R}}^d$ that satisfy Assumption B.2. Suppose that, $\forall x \in \mathcal{X}$, there exists $\Delta\bar{F}^U_x : \mathcal{A}_{x,U} \to \mathbb{R}$ such that:

$$
\begin{aligned}
&p_{y|x} = \bar{\kappa}(U, \Delta\bar{F}^U_x, y)\ \forall (x,y) \in \mathcal{X}\times\mathcal{Y}_0,\\
&\Delta\bar{F}^U_x \text{ is a } d\text{-dimensional subdistribution } \forall x \in \mathcal{X},\\
&\Delta\bar{F}^U_x(a) = \Delta\bar{F}^U_{\tilde{x}}(a)\ \forall x,\tilde{x} \in \{\mathcal{X}\cap\mathcal{X}^\dagger_{0,m_0}\},\ \forall a \in \mathcal{A}_{x,U},\ \forall m_0 \in \{1,...,M_0\},\\
&\Delta\bar{F}^U_{x,l}(a) = \Delta\bar{F}^U_{\tilde{x},l}(a)\ \forall x,\tilde{x} \in \{\mathcal{X}\cap\mathcal{X}^\dagger_{l,m_l}\},\ \forall a \in \mathcal{A}_{x,l,U},\ \forall m_l \in \{1,...,M_L\},\ \forall l \in \{1,...,d\},\\
&\{\Delta\bar{F}^U_{x,l}\}_{x\in\mathcal{X}} \in \Delta\mathcal{F}^\dagger_l\ \forall l \in \{1,...,d\},\\
&\bar{\rho}(U, \{\Delta\bar{F}^U_x\}_{x\in\mathcal{X}}) \geq 0.
\end{aligned}
$$

Then, $U \in \mathcal{U}^*$. ◇

C Other derivations

C.1 Derivation of (16)

For every $y \in \mathcal{Y}_0$, let $\mathcal{R}^U_{y|x}$ be the set of realisations of $\Delta\epsilon_i$ such that it is optimal for main i of type x to choose a woman of type y. More formally,

$$
\begin{aligned}
\mathcal{R}^U_{0|x} &\equiv \{(\Delta e_1, \Delta e_2, \Delta e_3) \in \mathbb{R}^3 : \Delta e_1 \geq -U_{x1}, \Delta e_3 \geq U_{x2} - U_{x1}\} = [-U_{x1}, \infty) \times (-\infty, \infty) \times [U_{x2} - U_{x1}, \infty),\\
\mathcal{R}^U_{1|x} &\equiv \{(\Delta e_1, \Delta e_2, \Delta e_3) \in \mathbb{R}^3 : \Delta e_2 \geq -U_{x2}, \Delta e_3 \leq U_{x2} - U_{x1}\} = (-\infty, \infty) \times [-U_{x2}, \infty) \times (-\infty, U_{x2} - U_{x1}],\\
\mathcal{R}^U_{2|x} &\equiv \{(\Delta e_1, \Delta e_2, \Delta e_3) \in \mathbb{R}^3 : \Delta e_1 \leq -U_{x1}, \Delta e_2 \leq U_{x2}\} = (-\infty, -U_{x1}] \times (-\infty, -U_{x2}] \times (-\infty, \infty),
\end{aligned}
$$

such that Δe_1, Δe_2, Δe_3 denote realisations of $\epsilon_{i1} - \epsilon_{i0}$, $\epsilon_{i2} - \epsilon_{i0}$, $\epsilon_{i1} - \epsilon_{i2}$, respectively.

Therefore, the restriction $p_{y|x} = \kappa(U, \Delta F_x, y)$ for every $y \in \mathcal{Y}_0$ can be more explicitly rewritten as

$$
\begin{aligned}
p_{0|x} &= \lambda_{\Delta F_x}([-U_{x1}, \infty) \times (-\infty, \infty) \times [U_{x2} - U_{x1}, \infty)),\\
p_{1|x} &= \lambda_{\Delta F_x}((-\infty, \infty) \times [-U_{x2}, \infty) \times (-\infty, U_{x2} - U_{x1}]),\\
p_{1|x} &= \lambda_{\Delta F_x}((-\infty, -U_{x1}] \times (-\infty, -U_{x2}] \times (-\infty, \infty)).
\end{aligned} \tag{C.1.1}
$$

Now, recall that the measure of a cube can be rewritten using the CDF. In particular, for every $x \in \mathcal{X}$ and $[a_1, b_1] \times [a_2, b_2] \times [a_3, b_3] \subseteq \mathbb{R}^3$, it holds that

$$
\begin{aligned}
&\lambda_{\Delta F_x}([a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]) =\\
&- \Delta F_x(a_1, a_2, a_3) + \Delta F_x(b_1, a_2, a_3) + \Delta F_x(a_1, b_2, a_3) - \Delta F_x(b_1, b_2, a_3)\\
&+ \Delta F_x(a_1, a_2, b_3) - \Delta F_x(b_1, a_2, b_3) - \Delta F_x(a_1, b_2, b_3) + \Delta F_x(b_1, b_2, b_3).
\end{aligned} \tag{C.1.2}
$$

By applying (C.1.2) to (C.1.1), we obtain (16).

C.2 Derivation of (29)

If $r = 2$, then

$$\mathcal{B} \equiv \{(b_1, b_2, b_3) \in \mathbb{R}^3 : \ b_3 = b_1 - b_2\} = \{(b_1, b_2, b_3) \in \mathbb{R}^3 : \ b_1 = b_2 + b_3\}.$$

Accordingly, the relevant 3-dimensional boxes defined in Proposition 4 are

$$\begin{aligned} B_{1,2,3}(\hat{b}, \tilde{b}) &\equiv \{(x, y, z) \in \mathbb{R}^3 : \ x > \hat{b} + \tilde{b}, y \leq \hat{b}, z \leq \tilde{b}\}, \\ Q_{1,2,3}(\hat{b}, \tilde{b}) &\equiv \{(x, y, z) \in \mathbb{R}^3 : \ x \leq \hat{b} + \tilde{b}, y > \hat{b}, z > \tilde{b}\}, \end{aligned}$$

for any $(\hat{b}, \tilde{b}) \in \mathbb{Q}^2$. Therefore,

$$\begin{aligned} \lambda_{\Delta F_x}(B_{t,p,q}(\hat{b}, \tilde{b})) &= \lambda_{\Delta F_x}((\hat{b} + \tilde{b}, \infty) \times (-\infty, \hat{b}] \times (-\infty, \tilde{b}]), \\ \lambda_{\Delta F_x}(Q_{t,p,q}(\hat{b}, \tilde{b})) &= \lambda_{\Delta F_x}((-\infty, \hat{b} + \tilde{b}] \times (\hat{b}, \infty) \times (\tilde{b}, \infty)). \end{aligned} \tag{C.2.1}$$

By applying (C.1.2) to (C.2.1), we obtain (29).

C.3 Matrix $\boldsymbol{\alpha_{x,U}}$ used in Section 4.5.2

$$\alpha_{x,U} \equiv \begin{pmatrix} -U_{x1} & -U_{x2} & U_{x2} - U_{x1} \\ -U_{x1} & -U_{x2} & \infty \\ -U_{x1} & -U_{x2} & -\infty \\ -U_{x1} & -\infty & U_{x2} - U_{x1} \\ -U_{x1} & -\infty & \infty \\ -U_{x1} & -\infty & -\infty \\ -U_{x1} & \infty & U_{x2} - U_{x1} \\ -U_{x1} & \infty & \infty \\ -U_{x1} & \infty & -\infty \\ -U_{x2} & -U_{x2} & U_{x2} - U_{x1} \\ -U_{x2} & -U_{x2} & \infty \\ -U_{x2} & -U_{x2} & -\infty \\ -U_{x2} & -\infty & U_{x2} - U_{x1} \\ -U_{x2} & -\infty & \infty \\ -U_{x2} & -\infty & -\infty \\ -U_{x2} & \infty & U_{x2} - U_{x1} \\ -U_{x2} & \infty & \infty \\ -U_{x2} & \infty & -\infty \\ U_{x2} - U_{x1} & -U_{x2} & U_{x2} - U_{x1} \\ U_{x2} - U_{x1} & -U_{x2} & \infty \\ U_{x2} - U_{x1} & -U_{x2} & -\infty \\ U_{x2} - U_{x1} & -\infty & U_{x2} - U_{x1} \\ U_{x2} - U_{x1} & -\infty & \infty \\ U_{x2} - U_{x1} & -\infty & -\infty \\ U_{x2} - U_{x1} & \infty & U_{x2} - U_{x1} \\ U_{x2} - U_{x1} & \infty & \infty \\ U_{x2} - U_{x1} & \infty & -\infty \end{pmatrix}$$

D Example of a linear programming problem

Fix $U \in \mathcal{U}$. Impose Assumption 5.2. Select, for simplicity, two 2-tuples from $\mathbb{Q}^2$, $(\hat{b}^1, \tilde{b}^1), (\hat{b}^2, \tilde{b}^2)$ to approximate (27). Therefore,

$$\begin{aligned}
&\mathcal{A}_{x,1}(U) \equiv \{-U_{x1}, -U_{x1}, 0, \hat{b}^1 + \tilde{b}^1, -\hat{b}^1 - \tilde{b}^1, \hat{b}^2 + \tilde{b}^2, -\hat{b}^2 - \tilde{b}^2, \infty, -\infty\},\\
&\mathcal{A}_{x,2}(U) \equiv \{-U_{x2}, -U_{x2}, 0, \hat{b}^1, -\hat{b}^1, \hat{b}^2, -\hat{b}^2, \infty, -\infty\},\\
&\mathcal{A}_{x,3}(U) \equiv \{U_{x2} - U_{x1}, -U_{x2} + U_{x1}, 0, \tilde{b}^1, -\tilde{b}^1, \tilde{b}^2, -\tilde{b}^2 \infty, -\infty\},
\end{aligned}$$

for every $x \in \mathcal{X}$. By the arguments of Sections 4.5.1 and 4.5.3, if $U \in \mathcal{U}^*$, then the following linear programming problem is feasible with respect to $\Delta \bar{F}_x^U : \mathcal{A}_{x,U} \to \mathbb{R}$ for every $x \in \mathcal{X}$:

$$p_{1|x} = 1 + \Delta \bar{F}_x^U(-U_{x1}, \infty, U_{x2} - U_{x1}) - \Delta \bar{F}_x^U(\infty, \infty, U_{x2} - U_{x1}) - \Delta \bar{F}_x^U(-U_{x1}, \infty, \infty), \tag{D.1}$$

$$p_{2|x} = \Delta \bar{F}_x^U(\infty, \infty, U_{x2} - U_{x1}) - \Delta \bar{F}_x^U(\infty, -U_{x2}, U_{x2} - U_{x1}), \tag{D.2}$$

$$p_{0|x} = \Delta \bar{F}_x^U(-U_{x1}, -U_{x2}, \infty), \tag{D.3}$$

$$\Delta \bar{F}_x^U(-U_{x1}, \infty, \infty) = 1 - \Delta \bar{F}_x^U(-U_{x1}, \infty, \infty), \tag{D.4}$$

$$\Delta \bar{F}_x^U(\infty, -U_{x2}, \infty) = 1 - \Delta \bar{F}_x^U(\infty, -U_{x2}, -\infty), \tag{D.5}$$

$$\Delta \bar{F}_x^U(\infty, \infty, U_{x2} - U_{x1}) = 1 - \Delta \bar{F}_x^U(\infty, \infty, -U_{x2} + U_{x1}), \tag{D.6}$$

$$\Delta \bar{F}_x^U(0, \infty, \infty) = 1/2, \tag{D.7}$$

$$\Delta \bar{F}_x^U(\infty, 0, \infty) = 1/2, \tag{D.8}$$

$$\Delta \bar{F}_x^U(\infty, \infty, 0) = 1/2, \tag{D.9}$$

$$\Delta \bar{F}_x^U(\hat{b}^1 + \tilde{b}^1, \infty, \infty) = 1 - \Delta \bar{F}_x^U(-\hat{b}^1 - \tilde{b}^1, \infty, \infty), \tag{D.10}$$

$$\Delta \bar{F}_x^U(\hat{b}^2 + \tilde{b}^2, \infty, \infty) = 1 - \Delta \bar{F}_x^U(-\hat{b}^2 - \tilde{b}^2, \infty, \infty), \tag{D.11}$$

$$\Delta \bar{F}_x^U(\infty, \hat{b}_1, \infty) = 1 - \Delta \bar{F}_x^U(\infty, -\hat{b}_1, -\infty), \tag{D.12}$$

$$\Delta \bar{F}_x^U(\infty, \hat{b}_2, \infty) = 1 - \Delta \bar{F}_x^U(\infty, -\hat{b}_2, -\infty), \tag{D.13}$$

$$\Delta \bar{F}_x^U(\infty, \infty, \tilde{b}_1) = 1 - \Delta \bar{F}_x^U(\infty, \infty, -\tilde{b}_1), \tag{D.14}$$

$$\Delta \bar{F}_x^U(\infty, \infty, \tilde{b}_2) = 1 - \Delta \bar{F}_x^U(\infty, \infty, -\tilde{b}_2), \tag{D.15}$$

$$\Delta \bar{F}_x^U(\infty, \hat{b}^1, \tilde{b}^1) - \Delta \bar{F}_x^U(\hat{b}^1 + \tilde{b}^1, \hat{b}^1, \tilde{b}^1) = 0, \tag{D.16}$$

$$\Delta \bar{F}_x^U(\hat{b}^1 + \tilde{b}^1, \infty, \infty) - \Delta \bar{F}_x^U(\hat{b}^1 + \tilde{b}^1, \hat{b}^1, \infty) - \Delta \bar{F}_x^U(\hat{b}^1 + \tilde{b}^1, \infty, \tilde{b}^1) + \Delta \bar{F}_x^U(\hat{b}^1 + \tilde{b}^1, \hat{b}^1, \tilde{b}^1) = 0, \tag{D.17}$$

$$\Delta \bar{F}_x^U(\infty, \hat{b}^2, \tilde{b}^2) - \Delta \bar{F}_x^U(\hat{b}^2 + \tilde{b}^2, \hat{b}^2, \tilde{b}^2) = 0, \tag{D.18}$$

$$\Delta \bar{F}_x^U(\hat{b}^2 + \tilde{b}^2, \infty, \infty) - \Delta \bar{F}_x^U(\hat{b}^2 + \tilde{b}^2, \hat{b}^2, \infty) - \Delta \bar{F}_x^U(\hat{b}^2 + \tilde{b}^2, \infty, \tilde{b}^2) + \Delta \bar{F}_x^U(\hat{b}^2 + \tilde{b}^2, \hat{b}^2, \tilde{b}^2) = 0, \tag{D.19}$$

$$\Delta \bar{F}_x^U(-\infty, t, q) = 0 \ \forall (t, q) \in \mathcal{A}_{x,2,U} \times \mathcal{A}_{x,3,U}, \tag{D.20}$$

$$\Delta \bar{F}_x^U(t, -\infty, q) = 0 \ \forall (t, q) \in \mathcal{A}_{x,1,U} \times \mathcal{A}_{x,3,U}, \tag{D.21}$$

$$\Delta \bar{F}_x^U(t, q, -\infty) = 0 \ \forall (t, q) \in \mathcal{A}_{x,1,U} \times \mathcal{A}_{x,2,U}, \tag{D.22}$$

$$\Delta \bar{F}_x^U(\infty, \infty, \infty) = 1, \tag{D.23}$$

$$0 \leq \Delta \bar{F}_x^U(t, q, r) \leq 1 \ \forall (t, q, r) \in \mathcal{A}_{x,U}, \tag{D.24}$$

$$\begin{matrix} [-\Delta \bar{F}_x^U(t, q, r) + \Delta \bar{F}_x^U(t', q, r) + \Delta \bar{F}_x^U(t, q', r) - \Delta \bar{F}_x^U(t', q', r) + \\ \Delta \bar{F}_x^U(t, q, r') - \Delta \bar{F}_x^U(t', q, r') - \Delta \bar{F}_x^U(t, q', r') + \Delta \bar{F}_x^U(t', q', r')] \geq 0 \end{matrix} \quad \begin{matrix} \forall (t, q, r), (t', q', r') \in \mathcal{A}_{x,U} \\ \text{s.t. } (t, q, r) \leq (t', q', r'). \end{matrix} \tag{D.25}$$

In the linear programming above: (D.1)-(D.3) match the predicted conditional choice probabilities with the empirical ones; (D.4)-(D.15) impose Assumption 5.2; (D.16)-(D.19) approximate the degeneracy condition as discussed in Section 4.5.3; (D.20)-(D.25) ensure that $\Delta \bar{F}^U_x$ can be extended to a proper conditional CDF.

E Additional details on the simulations

In this section, we report additional details on the simulations discussed in Section 5.

E.1 Normalisations

For shortness of exposition, we focus on the men's side. Analogous arguments hold for the women's side. Appropriate location normalisations have been already imposed by Proposition 1, i.e., $U_{x0} = 0$ for every $x \in \mathcal{X}$. We also want to impose some scale normalisations. Here, we distinguish our approach depending on whether Assumption 5.1 (independence between $\Delta\epsilon_i$ and X_i) is imposed. Specifically, when Assumption 5.1 is not imposed, we define

$$\mathcal{U} \equiv \Big\{(U_{xy}\ \forall (x,y) \in \mathcal{X} \times \mathcal{Y}_0) \in \mathbb{R}^{|\mathcal{X}\times\mathcal{Y}_0|}: \\ \text{[location normalisation]} \quad U_{x0} = 0\ \forall x \in \mathcal{X} \qquad \text{(E.1.1)} \\ \text{[scale normalisation]} \quad U_{x1} \in \{-1,1\}\ \forall x \in \mathcal{X}\Big\},$$

where the first condition is a location normalisation and the second condition is a scale normalisation.[25] Instead, when Assumption 5.1 is imposed, we define

$$\mathcal{U} \equiv \Big\{(U_{xy}\ \forall (x,y) \in \mathcal{X} \times \mathcal{Y}_0) \in \mathbb{R}^{|\mathcal{X}\times\mathcal{Y}_0|}: \\ \text{[location normalisation]} \quad U_{x0} = 0\ \forall x \in \mathcal{X} \qquad \text{(E.1.2)} \\ \text{[scale normalisation]} \quad U_{11} \in \{-1,1\}\Big\},$$

where the first condition is as before and the second condition is a different scale normalisation.[26] Note that, when Assumption 5.1 is not imposed, we include one scale normalisation for each $x \in \mathcal{X}$. In fact, if U belongs to $\mathcal{U}^*$ for some $\{\Delta F_x\}_{x\in\mathcal{X}}$, then any rescaled version of $\{\Delta F_x\}_{x\in\mathcal{X}}$

[25] For instance, if $r = 2$, we consider vectors like

$$\Big(\frac{U_{10}-U_{10}}{|U_{11}|}, \frac{U_{11}-U_{10}}{|U_{11}|}, \frac{U_{12}-U_{10}}{|U_{11}|}, \frac{U_{20}-U_{20}}{|U_{21}|}, \frac{U_{21}-U_{20}}{|U_{21}|}, \frac{U_{22}-U_{20}}{|U_{21}|}\Big).$$

[26] For instance, if $r = 2$, we consider vectors like

$$\Big(\frac{U_{10}-U_{10}}{|U_{11}|}, \frac{U_{11}-U_{10}}{|U_{11}|}, \frac{U_{12}-U_{10}}{|U_{11}|}, \frac{U_{20}-U_{20}}{|U_{11}|}, \frac{U_{21}-U_{20}}{|U_{11}|}, \frac{U_{22}-U_{20}}{|U_{11}|}\Big).$$

induces some scalar multiples of U to also belong to $\mathcal{U}^*$. Hence, the number of scale normalisations to impose on $\mathcal{U}$ is equal to the number of conditional CDFs to recover, that is $|\mathcal{X}|$. Instead, when Assumption 5.1 is imposed, we include just one scale normalisation. This is because determining whether U belongs to $\mathcal{U}^*$ requires recovering only one admissible conditional CDF, ΔF, where $\Delta F \equiv \Delta F_x \; \forall x \in \mathcal{X}$. Hence, when Assumption 5.1 is not imposed, we include *more* scale normalisations than when Assumption 5.1 is imposed.

E.2 DGPs considered in the first simulation exercise

In DGPs 2 and 3, for both sides of the market, we assume that every mixture component has mean $\mu \equiv (0, 0)$. In DGP 2, the two mixture components have variance-covariance matrices

$$\Sigma = \begin{pmatrix} 1 & 0.99 & 0.99 \\ 0.99 & 1 & 0.99 \\ 0.99 & 0.99 & 1 \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} 100 & -20 & -20 \\ -20 & 100 & -20 \\ -20 & -20 & 100 \end{pmatrix}.$$

In DGP 3, conditional on $X_i = 1$ (or, $Y_i = 1$), the two mixture components have variance-covariance matrices

$$\Sigma = \begin{pmatrix} 1 & 0.99 & 0.99 \\ 0.99 & 1 & 0.99 \\ 0.99 & 0.99 & 1 \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} 100 & -20 & -20 \\ -20 & 100 & -20 \\ -20 & -20 & 100 \end{pmatrix}.$$

Still in DGP 3, conditional on $X_i = 2$ (or, $Y_i = 2$), the two mixture components have variance-covariance matrices

$$\Sigma = \begin{pmatrix} 1 & 0.99 & 0.99 \\ 0.99 & 1 & 0.99 \\ 0.99 & 0.99 & 1 \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} 40 & 0 & 0 \\ 0 & 40 & 0 \\ 0 & 0 & 40 \end{pmatrix}.$$

E.3 Computation of U^{true} and V^{true}

Under DGP1, one can compute U^{true} and V^{true} by using the Multinomial Logit formulas of Choo and Siow (2006), i.e.,

$$U^{\text{true}}_{xy} = \log\left(\frac{p_{y|x}}{p_{0|x}}\right) \quad \text{and} \quad V^{\text{true}}_{xy} = \log\left(\frac{p_{x|y}}{p_{0|y}}\right),$$

for every $(x, y) \in \mathcal{X} \times \mathcal{Y}$. Under DGP2 and DGP3, one can compute U^{true} and V^{true} by using Proposition 2 in Galichon and Salanié (2020). This proposition claims that

$$U^{\text{true}}_{xy} = \frac{\partial F^*_x(\{p_{y|x}\}_{y \in \mathcal{Y}_0})}{\partial p_{y|x}} \quad \text{and} \quad V^{\text{true}}_{xy} = \frac{\partial G^*_y(\{p_{x|y}\}_{x \in \mathcal{X}_0})}{\partial p_{y|x}},$$

for every $(x,y) \in \mathcal{X} \times \mathcal{Y}$, where $F_x^*(\{p_{y|x}\}_{y\in\mathcal{Y}_0})$ is the Legendre-Fenchel transform of F_x evaluated at $\{p_{y|x}\}_{y\in\mathcal{Y}_0}$ and $G_y^*(\{p_{x|y}\}_{x\in\mathcal{X}_0})$ is the Legendre-Fenchel transform of G_y evaluated at $\{p_{x|y}\}_{x\in\mathcal{X}_0}$. Under DGP2 and DGP3 $F_x^*(\{p_{y|x}\}_{y\in\mathcal{Y}_0})$ and $G_y^*(\{p_{x|y}\}_{x\in\mathcal{X}_0})$ cannot be computed in closed form and, hence, we obtain them by simulation. Finally, we calculate their numerical derivatives.

F Inference

Section 4 studies identification by relying on the assumption that the empirical conditional choice probabilities, $\{p_{y|x}\}_{(x,y)\in\mathcal{X}\times\mathcal{Y}_0}$ and $\{p_{x|y}\}_{(x,y)\in\mathcal{X}_0\times\mathcal{Y}}$, are known by the analyst. When conducting inference, the analyst should replace them with consistent sample analogues, resulting from sampling at random from the market at the individual level or at the household level. This section illustrates how to construct a confidence region for the vector of parameters (U,V). In particular, we reformulate our problem using unconditional moment equalities and apply the bootstrap-based procedure of Bugni, Canay, and Shi (2017) (hereafter, BCS).

We assume, for simplicity, that our sample is composed by the same number, n, of men and women. Given $\alpha \in (0,1)$, a uniformly asymptotically valid $(1-\alpha)\%$ confidence region for (U,V) is obtained by inverting a test for the hypothesis

$$\mathrm{H}_0 : (U,V) = (U^0,V^0) \quad \text{vs.} \quad \mathrm{H}_1 : (U,V) \neq (U^0,V^0), \tag{F.1}$$

given hypothetical values (U^0,V^0). The test rejects H_0 if $TS_n(U^0,V^0) > \hat{c}_{n,1-\alpha}(U^0,V^0)$, where $TS_n(U^0,V^0)$ is a test statistic and $\hat{c}_{n,1-\alpha}(U^0,V^0)$ is a critical value. The remainder of the section explains how to compute $TS_n(U^0,V^0)$ and $\hat{c}_{n,1-\alpha}(U^0,V^0)$.

In order to define the test statistic, we start by introducing some useful notation. First, fix $U \in \mathcal{U}$ and $V \in \mathcal{V}$ and rewrite the identifying restrictions,

$$\begin{aligned} p_{y|x} &= \kappa(U, \Delta\bar{F}_x^U, y) \ \forall(x,y) \in \mathcal{X} \times \mathcal{Y}_0, \\ p_{x|y} &= \kappa(V, \Delta\bar{G}_y^V, x) \ \forall(x,y) \in \mathcal{X}_0 \times \mathcal{Y}, \end{aligned} \tag{F.2}$$

as unconditional moment equalities. To do so, define

$$\begin{aligned} m_i(U, \Delta\bar{F}_x^U, y) &\equiv \mathbb{1}\{X_i = x, Y_{i'} = y\} - \kappa(U, \Delta\bar{F}_x^U, y)\mathbb{1}\{X_i = x\} \ \forall(x,y) \in \mathcal{X} \times \mathcal{Y}_0, \\ m_j(V, \Delta\bar{G}_y^V, x) &\equiv \mathbb{1}\{Y_j = y, X_{j'} = x\} - \kappa(V, \Delta\bar{G}_y^V, x)\mathbb{1}\{Y_j = y\} \ \forall(x,y) \in \mathcal{X}_0 \times \mathcal{Y}, \end{aligned}$$

where i' and j' denote the labels of the i and j's spouse, respectively. Therefore, (F.2) is equivalent to

$$\begin{aligned} \mathbb{E}(m_i(U, \Delta\bar{F}_x^U, y)) &= 0 \ \forall(x,y) \in \mathcal{X} \times \mathcal{Y}_0, \\ \mathbb{E}(m_j(V, \Delta\bar{G}_y^V, x)) &= 0 \ \forall(x,y) \in \mathcal{X}_0 \times \mathcal{Y}. \end{aligned} \tag{F.3}$$

Second, given $(x,y) \in \mathcal{X} \times \mathcal{Y}$, let $\mathcal{S}_{U,V,x,y}$ be the collection of functions $\{\Delta\bar{F}_x^U\}_{x\in\mathcal{X}}$ and $\{\Delta\bar{G}_y^V\}_{y\in\mathcal{Y}}$ which are extendable to proper families of conditional CDFs in $\Delta\mathcal{F}^\dagger$ and $\Delta\mathcal{G}^\dagger$, respectively, with support contained in the region $\mathcal{B}$.

Third, define the sample analogue of (F.3) as

$$\begin{aligned}\bar{m}_n(U, \Delta\bar{F}_x^U, y) &\equiv \frac{1}{n}\sum_{i=1}^{n} m_i(U, \Delta\bar{F}_x^U, y) \ \forall(x,y) \in \mathcal{X}\times\mathcal{Y}_0, \\ \bar{m}_n(V, \Delta\bar{G}_y^V, x) &\equiv \frac{1}{n}\sum_{j=1}^{n} m_j(V, \Delta\bar{G}_y^V, x) \ \forall(x,y) \in \mathcal{X}_0\times\mathcal{Y}.\end{aligned} \tag{F.4}$$

Thus, the analyst can use the following profiled test statistic:

$$TS_n(U,V) \equiv \inf_{\left\{\{\Delta\bar{F}_x^U\}_{x\in\mathcal{X}}, \{\Delta\bar{G}_y^V\}_{y\in\mathcal{Y}}\right\} \in \mathcal{S}_{U,V,x,y}} Q_n(U,V), \tag{F.5}$$

where

$$Q_n(U,V) \equiv \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}_0} \sqrt{n}|\bar{m}_n(U,\Delta\bar{F}_x^U, y)| + \sum_{(x,y)\in\mathcal{X}_0\times\mathcal{Y}} \sqrt{n}|\bar{m}_n(V,\Delta\bar{G}_y^V, x)|$$

Intuitively, $TS_n(U,V)$ is built by imposing a penalty for each sample moment in (F.4) that is different from zero, hence violating the population counterparts in (F.3).

In order to compute the critical value, BCS suggest two approximations to the distribution of the profiled test statistic. In particular, as the first approximation, compute

$$TS_n^{DR}(U,V) \equiv \inf_{\left\{\{\Delta\bar{F}_x^U\}_{x\in\mathcal{X}}, \{\Delta\bar{G}_y^V\}_{y\in\mathcal{Y}}\right\} \in \hat{\mathcal{S}}_{U,V,x,y}} \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}_0} |\nu_n^*(U,\Delta\bar{F}_x^U, y)| + \sum_{(x,y)\in\mathcal{X}_0\times\mathcal{Y}} \sqrt{n}|\nu_n^*(V,\Delta\bar{G}_y^V, x)|, \tag{F.6}$$

where $\nu_n^*(U,\Delta\bar{F}_x^U, y)$ and $\nu_n^*(V,\Delta\bar{G}_y^V, x)$ are stochastic processes defined as

$$\begin{aligned}\nu_n^*(U,\Delta\bar{F}_x^U, y) &\equiv \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(m_i(U,\Delta\bar{F}_x^U, y) - \bar{m}_n(U,\Delta\bar{F}_x^U, y))\zeta_i, \\ \nu_n^*(V,\Delta\bar{G}_y^V, x) &\equiv \frac{1}{\sqrt{n}}\sum_{j=1}^{n}(m_j(V,\Delta\bar{G}_y^V, x) - \bar{m}_n(V,\Delta\bar{G}_y^V, x))\zeta_j,\end{aligned}$$

and $\{\zeta_i\}_{i=1}^n, \{\zeta_j\}_{j=1}^n$ are i.i.d. standard normals, independent of the data. Further, $\hat{\mathcal{S}}_{U,V,x,y}$ is the set of minimisers of $Q_n(U,V)$, i.e.,

$$\hat{\mathcal{S}}_{U,V,x,y} \equiv \left\{\left\{\{\Delta\bar{F}_x^U\}_{x\in\mathcal{X}}, \{\Delta\bar{G}_y^V\}_{y\in\mathcal{Y}}\right\} \in \mathcal{S}_{U,V,x,y} : Q_n(U,V) \leq TS_n(U,V)\right\}.$$

As the second approximation, compute

$$TS_n^{PR}(U,V) \equiv \inf_{\left\{\{\Delta\bar{F}_x^U\}_{x\in\mathcal{X}},\{\Delta\bar{G}_y^V\}_{y\in\mathcal{Y}}\right\}\in S_{U,V,x,y}} \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}_0} |\nu_n^*(U,\Delta\bar{F}_x^U,y)+l_n(U,\Delta\bar{F}_x^U,y)| + \sum_{(x,y)\in\mathcal{X}_0\times\mathcal{Y}} \sqrt{n}|\nu_n^*(V,\Delta\bar{G}_y^V,x)+l_n(V,\Delta\bar{G}_y^V,x)|, \tag{F.7}$$

where $l_n(U,\Delta\bar{F}_x^U,y)$ and $l_n(V,\Delta\bar{G}_y^V,x)$ are slackness functions defined as

$$l_n(U,\Delta\bar{F}_x^U,y) \equiv \kappa_n^{-1}\sqrt{n}\bar{m}_n(U,\Delta\bar{F}_x^U,y),$$
$$l_n(V,\Delta\bar{G}_y^V,x) \equiv \kappa_n^{-1}\sqrt{n}\bar{m}_n(V,\Delta\bar{G}_y^V,x),$$

and κ_n is a tuning paramater that satisfies $\lim_{n\to\infty}\kappa_n=\infty$ and $\lim_{n\to\infty}\frac{\kappa_n}{\sqrt{n}}=0$. For instance, $\kappa_n\equiv\sqrt{\log(n)}$.
Given (F.6) and (F.7), the critical value under H_0, $\hat{c}_{n,1-\alpha}(U^0,V^0)$, is the $1-\alpha$ quantile of

$$TS_n^{MR}(U^0,V^0)\equiv\min\{TS_n^{DR}(U^0,V^0),TS_n^{PR}(U^0,V^0)\},$$

and can be computed via bootstrapping.

Once the 95% confidence region for (U,V) is obtained by following the above procedure, we project it along various dimensions to get the confidence regions for Φ, $D(\Phi)$, $C(U)$ and $C(V)$.

Remark F.1. (*Computational simplifications for inference*) We outline a few important simplifications that help us to reduce the computational burden of the above procedure. First, note that the minimisation problems in (F.5) and (F.7) can be rewritten as linear minimisation problems. For instance, take (F.5). As shown for instance by Bertsimas and Tsitsiklis (1997) (Section 1, p.18), (F.5) is equivalent to

$$\inf_{\substack{\left\{\{\Delta\bar{F}_x^U\}_{x\in\mathcal{X}},\{\Delta\bar{G}_y^V\}_{y\in\mathcal{Y}}\right\}\in S_{U,V,x,y}\\ \tau_{x,y}^+,\tau_{x,y}^-\in\mathbb{R}\ \forall(x,y)\in\mathcal{X}\times\mathcal{Y}_0\\ \tau_{y,x}^+,\tau_{y,x}^-\in\mathbb{R}\ \forall(x,y)\in\mathcal{X}_0\times\mathcal{Y}}} \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}_0}\sqrt{n}(\tau_{x,y}^+ + \tau_{x,y}^-) + \sum_{(x,y)\in\mathcal{X}_0\times\mathcal{Y}}\sqrt{n}(\tau_{y,x}^+ + \tau_{y,x}^-),$$
$$\text{s.t.}$$
$$\tau_{x,y}^+,\tau_{x,y}^-\ge 0\ \forall(x,y)\in\mathcal{X}\times\mathcal{Y}_0, \tag{F.8}$$
$$\tau_{y,x}^+,\tau_{y,x}^-\ge 0\ \forall(x,y)\in\mathcal{X}_0\times\mathcal{Y},$$
$$\tau_{x,y}^+ + \tau_{x,y}^- = \bar{m}_n(U,\Delta\bar{F}_x^U,y)\ \forall(x,y)\in\mathcal{X}\times\mathcal{Y}_0,$$
$$\tau_{y,x}^+ + \tau_{y,x}^- = \bar{m}_n(V,\Delta\bar{G}_y^V,x)\ \forall(x,y)\in\mathcal{X}_0\times\mathcal{Y}.$$

Given our identification arguments, (F.8) is a linear minimisation problem. Similar considerations can be applied to (F.7).

Second, to construct a confidence region, the analyst should run the test (F.1) for every $(U,V)\in\mathcal{U}\times\mathcal{V}$, which can be computationally burdensome. However, this can be simplified by exploiting Proposition 3. In fact, take any $(U, V), (\tilde{U}, \tilde{V}) \in \mathcal{U} \times \mathcal{V}$ such that $\pi(\alpha_{x,l,U}) = \pi(\alpha_{x,l,\tilde{U}})$ and $\pi(\alpha_{y,l,V}) = \pi(\alpha_{y,l,\tilde{V}})$ for every $l \in \{1, ..., d\}$, $x \in \mathcal{X}$, and $y \in \mathcal{Y}$. Then, by Proposition 3, $TS_n(U, V) = TS_n(\tilde{U}, \tilde{V})$, $TS_n^{DR}(U, V) = TS_n^{DR}(\tilde{U}, \tilde{V})$, and $TS_n^{PR}(U, V) = TS_n^{PR}(\tilde{U}, \tilde{V})$. Therefore, one can run the test (F.1) only for a handful of $(U^0, V^0) \in \mathcal{U} \times \mathcal{V}$.

Third, for the BCS's method to work, it is enough for the set $\hat{\mathcal{S}}_{U,V,x,y}$ used in (F.6) to be a good approximation of the set of minimisers of $Q_n(U, V)$. For instance, BCS's results hold as long as $\hat{\mathcal{S}}_{U,V,x,y}$ contains one minimiser of $Q_n(U, V)$. Therefore, the analyst can consider the minimiser of $Q_n(U, V)$ found when computing $TS_n(U, V)$.

Fourth, note that the minimisation problems in (F.5), (F.6), and (F.7) can be made separable across the two sides of the market if the two sides of the market are assumed to be stochastically independent. In this case, the analyst can decompose each minimisation problem into two smaller minimisation problems on both sides.

Fifth, if Assumption 5.1 is not imposed, then the men's minimisation problem is further separable across each type of men. Similarly, the women's minimisation problem is further separable across each type of women. Therefore, the analyst can decompose the men's minimisation problem into $|\mathcal{X}|$ smaller minimisation problems. Similarly, the analyst can decompose the women's minimisation problem into $|\mathcal{Y}|$ smaller minimisation problems.

Lastly, to preserve linearity we have decided not to rescale the moment equalities by their standard deviations. The cost is losing the scale invariance property of the test statistic. Some of the earlier papers on inference for partially identified parameters, such as Chernozhukov, Hong, and Tamer (2007), Romano and Shaikh (2008), and Ciliberto and Tamer (2009), consider modified methods of moments estimators that are not scale invariant. However, Andrews and Soares (2010) discuss that this may lead to poor power performances.

◇

G Additional details on the empirical application

In this section, we present additional details on the empirical application discussed in Section 6. Tables 5, 6, and 7 are based on the estimated identified sets of the parameters. To account for the sampling uncertainty, below we report the projections of the corresponding 95% confidence regions (Tables G.1, G.2, and G.3). These confidence regions are constructed by using the bootstrap-based procedure of Bugni, Canay, and Shi (2017) as discussed in Appendix F. The intervals in Tables G.1, G.2, and G.3 are equal to or contain the intervals in Tables 5, 6, and 7, respectively. The fact that some intervals in Tables G.1, G.2, and G.3 are equal to the intervals in Tables 5, 6, and 7, respectively, could be due to the fact that we constructed the estimated identified sets and confidence regions by evaluating a coarse grid of parameter values.

Tables G.4, G.5, and G.6 report the projections of the estimated identified sets of U, V, Φ, $D(\Phi)$, $C(U)$, and $C(V)$ under specification [7] of Table 1. We have highlighted in blue the numbers that differ from Tables 5, 6, and 7.

Parameters	1940		1945		1950		1955		1960		1966	
Φ_{11} Φ_{12}	2	$[0.1111, \infty)$	$\{0, 2\}$	$(-\infty, \infty)$	$\{0, 2\}$	$[-1.3030, \infty)$	0	$[-1.9091, -0.0909]$	$\{-2, 0\}$	$[-1.9091, -0.0909]$	$\{-2, 0\}$	$[-1.9091, 6.7778]$
Φ_{21} Φ_{22}	$[-1.9091, -0.0707]$	$(-\infty, \infty)$	$[-1.9091, 1.9293]$	$[0.2828, \infty)$	$[-1.9091, -0.0707]$	$[0.2828, \infty)$	$(-\infty, -0.0707]$	$[0.2828, \infty)$	$(-\infty, \infty)$	$[0.2134, \infty)$	$(-\infty, -0.0707]$	$[0.0297, \infty)$
$\boldsymbol{D_{22,11}(\Phi)}$	$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{[0.4444, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{[-6.4242, \infty)}$	
Φ_{11} Φ_{13}	2	$(-\infty, \infty)$	$\{0, 2\}$	$(-\infty, \infty)$	$\{0, 2\}$	$(-\infty, \infty)$	0	$(-\infty, -1.2424]$	$\{-2, 0\}$	$(-\infty, -1.2424]$	$\{-2, 0\}$	$(-\infty, -1.2424]$
Φ_{31} Φ_{33}	$(-\infty, -0.1010]$	$(-\infty, \infty)$	$(-\infty, -0.6869]$	$[-0.6667, \infty)$	$(-\infty, -0.6869]$	$[-0.6667, \infty)$	$(-\infty, -0.6869]$	$[-0.6667, \infty)$	$(-\infty, -1.2727]$	$[-0.6667, \infty)$	$(-\infty, -1.2727]$	$[-1, \infty)$
$\boldsymbol{D_{33,11}(\Phi)}$	$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{[1.2624, \infty)}$		$\boldsymbol{[-0.1515, \infty)}$		$\boldsymbol{[-0.4848, \infty)}$	
Φ_{11} Φ_{14}	2	$(-\infty, \infty)$	$\{0, 2\}$	$(-\infty, \infty)$	$\{0, 2\}$	$(-\infty, \infty)$	0	$(-\infty, -1.3939]$	$\{-2, 0\}$	$(-\infty, -1.3939]$	$\{-2, 0\}$	$(-\infty, -1.3939]$
Φ_{41} Φ_{44}	$(-\infty, 0.8687]$	$[-0.6869, \infty)$	$(-\infty, 0.8687]$	$[-0.9899, \infty)$	$(-\infty, -1.1313]$	$[-0.6869, \infty)$	$(-\infty, -1.1313]$	$[-0.7879, \infty)$	$(-\infty, -1.7576]$	$[-0.7879, \infty)$	$(-\infty, -1.7576]$	$[-0.7879, \infty)$
$\boldsymbol{D_{44,11}(\Phi)}$	$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{[1.7374, \infty)}$		$\boldsymbol{[0.9190, \infty)}$		$\boldsymbol{[2.3636, \infty)}$	
Φ_{22} Φ_{23}	$(-\infty, \infty)$	$(-\infty, \infty)$	$[0.2828, \infty)$	$(-\infty, \infty)$	$[0.2828, \infty)$	$(-\infty, \infty)$	$[0.2828, \infty)$	$(-\infty, \infty)$	$[0.2134, \infty)$	$(-\infty, \infty)$	$[0.0297, \infty)$	$[-2.5354, 15.6263]$
Φ_{32} Φ_{33}	$[-6.8485, \infty)$	$(-\infty, \infty)$	$[-7.3535, \infty)$	$[-0.6667, \infty)$	$[-7.3535, \infty)$	$[-0.6667, \infty)$	$[-7.8586, \infty)$	$[-0.6667, \infty)$	$[-7.8586, \infty)$	$[-0.6667, \infty)$	$[-7.8586, \infty)$	$[-1, \infty)$
$\boldsymbol{D_{33,22}(\Phi)}$	$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{[0.4141, \infty)}$	
Φ_{22} Φ_{24}	$(-\infty, \infty)$	$(-\infty, \infty)$	$[0.2828, \infty)$	$(-\infty, \infty)$	$[0.2828, \infty)$	$(-\infty, \infty)$	$[0.2828, \infty)$	$(-\infty, \infty)$	$[0.2134, \infty)$	$(-\infty, \infty)$	$[0.0297, \infty)$	$[-2.9394, 0.3232]$
Φ_{42} Φ_{44}	$(-\infty, \infty)$	$[-0.6869, \infty)$	$(-\infty, \infty)$	$[-0.9899, \infty)$	$(-\infty, \infty)$	$[-0.6869, \infty)$	$(-\infty, \infty)$	$[-0.7879, \infty)$	$(-\infty, \infty)$	$[-0.7879, \infty)$	$(-\infty, \infty)$	$[-0.7879, \infty)$
$\boldsymbol{D_{44,22}(\Phi)}$	$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$	
Φ_{33} Φ_{34}	$(-\infty, \infty)$	$(-\infty, 16.3737]$	$[-0.6667, \infty)$	$[-17, 16.3737]$	$[-0.6667, \infty)$	$[-18.1919, 16.3737]$	$[-0.6667, \infty)$	$[-18.1919, 8.6263]$	$[-0.6667, \infty)$	$[-18.1919, 8.6263]$	$[-1, \infty)$	$[-18.1919, 8.6263]$
Φ_{43} Φ_{44}	$(-\infty, \infty)$	$[-0.6869, \infty)$	$(-\infty, \infty)$	$[-0.9899, \infty)$	$(-\infty, \infty)$	$[-0.6869, \infty)$	$(-\infty, \infty)$	$[-0.7879, \infty)$	$(-\infty, \infty)$	$[-0.7879, \infty)$	$(-\infty, \infty)$	$[-0.7879, \infty)$
$\boldsymbol{D_{44,33}(\Phi)}$	$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$	

Table G.1: Projections of the 95% confidence regions for Φ and $D(\Phi)$ under specification [5] of Table 1. “1” denotes high school dropouts; “2” denotes high school graduates; “3” denotes some college; “4” denotes four-year college graduates and graduate degrees.

Parameters	1940		1945		1950		1955		1960		1966	
U_{11} U_{12}	1	$[1.1111, \infty)$	$\{-1, 1\}$	$(-\infty, \infty)$	$\{-1, 1\}$	$[-0.3030, \infty)$	-1	$[-0.9091, 0.9091]$	-1	$[-0.9091, 0.9091]$	$\{-1, 1\}$	$[-0.9091, 7.7778]$
U_{13} U_{14}	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, -0.2424]$	$(-\infty, -0.3939]$	$(-\infty, -0.2424]$	$(-\infty, -0.3939]$	$(-\infty, -0.2424]$	$(-\infty, -0.3939]$
U_{21} U_{22}	-1	$[0.1414, \infty)$	$\{-1, 1\}$	$(-\infty, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[-0.1117, 0.1414]$
U_{23} U_{24}	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$[-1.7374, -0.1717]$	$[-1.9495, -0.2222]$
$\boldsymbol{C_{21}(U)}$	$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{[-1.7015, \infty)}$	
U_{11} U_{12}	1	$[1.1111, \infty)$	$\{-1, 1\}$	$(-\infty, \infty)$	$\{-1, 1\}$	$[-0.3030, \infty)$	-1	$[-0.9091, 0.9091]$	-1	$[-0.9091, 0.9091]$	$\{-1, 1\}$	$[-0.9091, 7.7778]$
U_{13} U_{14}	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, -0.2424]$	$(-\infty, -0.3939]$	$(-\infty, -0.2424]$	$(-\infty, -0.3939]$	$(-\infty, -0.2424]$	$(-\infty, -0.3939]$
U_{31} U_{32}	-1	$[0.1111, \infty)$	-1	$[-0.3939, 15.7677]$	-1	$[-0.3939, 15.7677]$	-1	$[-0.8990, 0.6162]$	-1	$[-0.8990, 0.6162]$	-1	$[-0.8990, 0.6162]$
U_{33} U_{34}	$(-\infty, \infty)$	$(-\infty, 15.6364]$	$[-0.3333, \infty)$	$[-16.5455, 15.6364]$	$[-0.3333, \infty)$	$[-17.7374, 15.6364]$	$[-0.3333, \infty)$	$[-17.7374, 7.8889]$	$[-0.3333, \infty)$	$[-17.7374, 7.8889]$	$[-0.6667, \infty)$	$[-17.7374, 7.8889]$
$\boldsymbol{C_{31}(U)}$	$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-0.5316, \infty)}$		$\boldsymbol{[-0.5045, \infty)}$		$\boldsymbol{[-1.7305, \infty)}$	
U_{11} U_{12}	1	$[1.1111, \infty)$	$\{-1, 1\}$	$(-\infty, \infty)$	$\{-1, 1\}$	$[-0.3030, \infty)$	-1	$[-0.9091, 0.9091]$	-1	$[-0.9091, 0.9091]$	$\{-1, 1\}$	$[-0.9091, 7.7778]$
U_{13} U_{14}	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, -0.2424]$	$(-\infty, -0.3939]$	$(-\infty, -0.2424]$	$(-\infty, -0.3939]$	$(-\infty, -0.2424]$	$(-\infty, -0.3939]$
U_{41} U_{42}	$\{-1, 1\}$	$(-\infty, \infty)$	$\{-1, 1\}$	$(-\infty, 11.8788]$	-1	$[-0.8485, 0.8485]$	-1	$[-0.8485, 0.8485]$	-1	$[-0.8485, 0.6364]$	-1	$[-0.8485, 0.6364]$
U_{43} U_{44}	$(-\infty, \infty)$	$[-0.2424, \infty)$	$(-\infty, \infty)$	$[-0.5455, \infty)$	$[-0.2727, 3.9091]$	$[-0.2424, \infty)$	$[-0.8182, 3.9091]$	$[-0.3424, \infty)$	$[-0.8182, 0.8182]$	$[-0.3434, \infty)$	$[-0.8182, 0.8182]$	$[-0.3434, \infty)$
$\boldsymbol{C_{41}(U)}$	$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{[-0.4244, \infty]}$		$\boldsymbol{[-0.4050, \infty)}$		$\boldsymbol{[-1.5205, \infty)}$	
U_{21} U_{22}	-1	$[0.1414, \infty)$	$\{-1, 1\}$	$(-\infty, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[-0.1117, 0.1414]$
U_{23} U_{24}	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$[-1.7374, -0.1717]$	$[-1.9495, -0.2222]$
U_{31} U_{32}	-1	$[0.1111, \infty)$	-1	$[-0.3939, 15.7677]$	-1	$[-0.3939, 15.7677]$	-1	$[-0.8990, 0.6162]$	-1	$[-0.8990, 0.6162]$	-1	$[-0.8990, 0.6162]$
U_{33} U_{34}	$(-\infty, \infty)$	$(-\infty, 15.6364]$	$[-0.3333, \infty)$	$[-16.5455, 15.6364]$	$[-0.3333, \infty)$	$[-17.7374, 15.6364]$	$[-0.3333, \infty)$	$[-17.7374, 7.8889]$	$[-0.3333, \infty)$	$[-17.7374, 7.8889]$	$[-0.6667, \infty)$	$[-17.7374, 7.8889]$
$\boldsymbol{C_{32}(U)}$	$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{[-0.5188, \infty)}$	
U_{21} U_{22}	-1	$[0.1414, \infty)$	$\{-1, 1\}$	$(-\infty, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[-0.1117, 0.1414]$
U_{23} U_{24}	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$[-1.7374, -0.1717]$	$[-1.9495, -0.2222]$
U_{41} U_{42}	$\{-1, 1\}$	$(-\infty, \infty)$	$\{-1, 1\}$	$(-\infty, 11.8788]$	-1	$[-0.8485, 0.8485]$	-1	$[-0.8485, 0.8485]$	-1	$[-0.8485, 0.6364]$	-1	$[-0.8485, 0.6364]$
U_{43} U_{44}	$(-\infty, \infty)$	$[-0.2424, \infty)$	$(-\infty, \infty)$	$[-0.5455, \infty)$	$[-0.2727, 3.9091]$	$[-0.2424, \infty)$	$[-0.8182, 3.9091]$	$[-0.3424, \infty)$	$[-0.8182, 0.8182]$	$[-0.3434, \infty)$	$[-0.8182, 0.8182]$	$[-0.3434, \infty)$
$\boldsymbol{C_{42}(U)}$	$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{[-0.3088, \infty)}$	
U_{31} U_{32}	-1	$[0.1111, \infty)$	-1	$[-0.3939, 15.7677]$	-1	$[-0.3939, 15.7677]$	-1	$[-0.8990, 0.6162]$	-1	$[-0.8990, 0.6162]$	-1	$[-0.8990, 0.6162]$
U_{33} U_{34}	$(-\infty, \infty)$	$(-\infty, 15.6364]$	$[-0.3333, \infty)$	$[-16.5455, 15.6364]$	$[-0.3333, \infty)$	$[-17.7374, 15.6364]$	$[-0.3333, \infty)$	$[-17.7374, 7.8889]$	$[-0.3333, \infty)$	$[-17.7374, 7.8889]$	$[-0.6667, \infty)$	$[-17.7374, 7.8889]$
U_{41} U_{42}	$\{-1, 1\}$	$(-\infty, \infty)$	$\{-1, 1\}$	$(-\infty, 11.8788]$	-1	$[-0.8485, 0.8485]$	-1	$[-0.8485, 0.8485]$	-1	$[-0.8485, 0.6364]$	-1	$[-0.8485, 0.6364]$
U_{43} U_{44}	$(-\infty, \infty)$	$[-0.2424, \infty)$	$(-\infty, \infty)$	$[-0.5455, \infty)$	$[-0.2727, 3.9091]$	$[-0.2424, \infty)$	$[-0.8182, 3.9091]$	$[-0.3424, \infty)$	$[-0.8182, 0.8182]$	$[-0.3434, \infty)$	$[-0.8182, 0.8182]$	$[-0.3434, \infty)$
$\boldsymbol{C_{43}(U)}$	$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$	

Table G.2: Projections of the 95% confidence regions for U and $C(U)$ under specification [5] of Table 1. “1” denotes high school dropouts; “2” denotes high school graduates; “3” denotes some college; “4” denotes four-year college graduates and graduate degrees.

Parameters	1940		1945		1950		1955		1960		1966	
V_{11} V_{21}	1	$[-0.9091, 0.9293]$	1	$[-0.9091, 0.9293]$	1	$[-0.9091, 0.9293]$	1	$(-\infty, 0.9293]$	$\{-1, 1\}$	$(-\infty, \infty)$	-1	$(-\infty, 0.9293]$
V_{31} V_{41}	$(-\infty, 0.8990]$	$(-\infty, -0.1313]$	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, -0.2727]$	$(-\infty, -0.7576]$	$(-\infty, -0.2727]$	$(-\infty, -0.7576]$
V_{12} V_{22}	$\{-1, 1\}$	$(-\infty, \infty)$	$\{-1, 1\}$	$(-\infty, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$
V_{32} V_{42}	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$
$\boldsymbol{C_{21}(V)}$	$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$	
V_{11} V_{21}	1	$[-0.9091, 0.9293]$	1	$[-0.9091, 0.9293]$	1	$[-0.9091, 0.9293]$	1	$(-\infty, 0.9293]$	$\{-1, 1\}$	$(-\infty, \infty)$	-1	$(-\infty, 0.9293]$
V_{31} V_{41}	$(-\infty, 0.8990]$	$(-\infty, -0.1313]$	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, -0.2727]$	$(-\infty, -0.7576]$	$(-\infty, -0.2727]$	$(-\infty, -0.7576]$
V_{13} V_{23}	-1	$[-0.4848, 14.2323]$	-1	$[-0.4848, 14.2323]$	-1	$[-0.4848, 14.2323]$	-1	$[-0.7980, 14.2323]$	-1	$[-0.7980, 14.2323]$	-1	$[-0.7980, 15.7980]$
V_{33} V_{43}	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$
$\boldsymbol{C_{31}(V)}$	$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$	
V_{11} V_{21}	1	$[-0.9091, 0.9293]$	1	$[-0.9091, 0.9293]$	1	$[-0.9091, 0.9293]$	1	$(-\infty, 0.9293]$	$\{-1, 1\}$	$(-\infty, \infty)$	-1	$(-\infty, 0.9293]$
V_{31} V_{41}	$(-\infty, 0.8990]$	$(-\infty, -0.1313]$	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, 0.3131]$	$(-\infty, -0.1313]$	$(-\infty, -0.2727]$	$(-\infty, -0.7576]$	$(-\infty, -0.2727]$	$(-\infty, -0.7576]$
V_{14} V_{24}	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$
V_{34} V_{44}	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$
$\boldsymbol{C_{41}(V)}$	$\boldsymbol{[-1.1268, \infty)}$		$\boldsymbol{[-1.1057, \infty)}$		$\boldsymbol{[-1.0301, \infty)}$		$\boldsymbol{[-0.9840, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{[-0.3166, \infty)}$	
V_{12} V_{22}	$\{-1, 1\}$	$(-\infty, \infty)$	$\{-1, 1\}$	$(-\infty, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$
V_{32} V_{42}	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$
V_{13} V_{23}	-1	$[-0.4848, 14.2323]$	-1	$[-0.4848, 14.2323]$	-1	$[-0.4848, 14.2323]$	-1	$[-0.7980, 14.2323]$	-1	$[-0.7980, 14.2323]$	-1	$[-0.7980, 15.7980]$
V_{33} V_{43}	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$
$\boldsymbol{C_{32}(V)}$	$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$	
V_{12} V_{22}	$\{-1, 1\}$	$(-\infty, \infty)$	$\{-1, 1\}$	$(-\infty, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$	-1	$[0.1414, \infty)$
V_{32} V_{42}	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$	$[-6.9596, \infty)$	$(-\infty, \infty)$
V_{14} V_{24}	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$
V_{34} V_{44}	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$
$\boldsymbol{C_{42}(V)}$	$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$	
V_{13} V_{23}	-1	$[-0.4848, 14.2323]$	-1	$[-0.4848, 14.2323]$	-1	$[-0.4848, 14.2323]$	-1	$[-0.7980, 14.2323]$	-1	$[-0.7980, 14.2323]$	-1	$[-0.7980, 15.7980]$
V_{33} V_{43}	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$	$[-0.3333, \infty)$	$(-\infty, \infty)$
V_{14} V_{24}	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$	-1	$[-0.9899, 0.5455]$
V_{34} V_{44}	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$	$[-0.4545, 0.7374]$	$[-0.4444, \infty)$
$\boldsymbol{C_{43}(V)}$	$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$		$\boldsymbol{(-\infty, \infty)}$	

Table G.3: Projections of the 95% confidence regions for V and $C(V)$ under specification [5] of Table 1. "1" denotes high school dropouts; "2" denotes high school graduates; "3" denotes some college; "4" denotes four-year college graduates and graduate degrees.

Parameters	1940		1945		1950		1955		1960		1966	
Φ_{11} Φ_{12}	2	[2.1111, ∞)	2	[0.1111, ∞)	2	[0.1111, ∞)	0	[−1.9091, −0.0909]	−2	[−1.9091, −0.0909]	−2	[−1.9091, −0.0909]
Φ_{21} Φ_{22}	[−1.9091, −0.0707]	[2.2020, ∞)	[−1.9091, −0.0707]	[0.2828, ∞)	[−1.9091, −0.0707]	[0.2828, ∞)	[−1.9091, −0.0707]	[0.2828, ∞)	(−∞, −2.0404]	[0.2134, ∞)	(−∞, −2.0404]	[0.0297, ∞)
$\boldsymbol{D_{22,11}(\Phi)}$	**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**[0.4444, ∞)**		**[0.4141, ∞)**		**[0.4141, ∞)**	
Φ_{11} Φ_{13}	2	(−∞, ∞)	2	(−∞, ∞)	2	(−∞, ∞)	0	(−∞, −1.2424]	−2	(−∞, −1.2424]	−2	(−∞, −1.2424]
Φ_{31} Φ_{33}	(−∞, −0.6869]	(−∞, ∞)	(−∞, −0.6869]	[−0.6667, ∞)	(−∞, −0.6869]	[−0.6667, ∞)	(−∞, −0.6869]	[−0.6667, ∞)	(−∞, −2.1515]	[−0.6667, ∞)	(−∞, −2.1515]	[−1, ∞)
$\boldsymbol{D_{33,11}(\Phi)}$	**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**[1.2624, ∞)**		**[0.7273, ∞)**		**[0.3939, ∞)**	
Φ_{11} Φ_{14}	2	(−∞, ∞)	2	(−∞, ∞)	2	(−∞, ∞)	0	(−∞, −1.3939]	−2	(−∞, −1.3939]	−2	(−∞, −1.3939]
Φ_{41} Φ_{44}	(−∞, −1.1313]	[−0.1818, ∞)	(−∞, −1.1313]	[0.1212, ∞)	(−∞, −1.1313]	[0.1212, ∞)	(−∞, −1.1313]	[0.1212, ∞)	(−∞, −2.3838]	[0.1212, ∞)	(−∞, −2.3838]	[−0.7879, ∞)
$\boldsymbol{D_{44,11}(\Phi)}$	**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**[2.6465, ∞)**		**[1.8990, ∞)**		**[2.3838, ∞)**	
Φ_{22} Φ_{23}	[2.2020, ∞)	(−∞, ∞)	[0.2828, ∞)	(−∞, ∞)	[0.2828, ∞)	(−∞, ∞)	[0.2828, ∞)	(−∞, ∞)	[0.2134, ∞)	(−∞, ∞)	[0.0297, ∞)	[−2.5354, 0.5960]
Φ_{32} Φ_{33}	[−6.8485, ∞)	(−∞, ∞)	[−7.3535, ∞)	[−0.6667, ∞)	[−7.3535, ∞)	[−0.6667, ∞)	[−7.8586, ∞)	[−0.6667, ∞)	[−7.8586, ∞)	[−0.6667, ∞)	[−7.8586, 25.4747]	[−1, ∞)
$\boldsymbol{D_{33,22}(\Phi)}$	**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**[0.5556, ∞)**	
Φ_{22} Φ_{24}	[2.2020, ∞)	(−∞, ∞)	[0.2828, ∞)	(−∞, ∞)	[0.2828, ∞)	(−∞, ∞)	[0.2828, ∞)	(−∞, ∞)	[0.2134, ∞)	(−∞, ∞)	[0.0297, ∞)	[−2.9394, 0.3232]
Φ_{42} Φ_{44}	(−∞, ∞)	[−0.1818, ∞)	(−∞, ∞)	[0.1212, ∞)	(−∞, ∞)	[0.1212, ∞)	(−∞, ∞)	[0.1212, ∞)	(−∞, ∞)	[0.1212, ∞)	(−∞, ∞)	[−0.7879, ∞)
$\boldsymbol{D_{44,22}(\Phi)}$	**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**	
Φ_{33} Φ_{34}	(−∞, ∞)	(−∞, 16.3737]	[−0.6667, ∞)	[−17, 16.3737]	[−0.6667, ∞)	[−18.1919, 8.6263]	[−0.6667, ∞)	[−18.1919, 8.6263]	[−0.6667, ∞)	[−18.1919, 8.6263]	[−1, ∞)	[−2.6970, 0.8788]
Φ_{43} Φ_{44}	(−∞, ∞)	[−0.1818, ∞)	(−∞, ∞)	[0.1212, ∞)	(−∞, ∞)	[0.1212, ∞)	(−∞, ∞)	[0.1212, ∞)	(−∞, ∞)	[0.1212, ∞)	(−∞, 8.7273]	[−0.7879, ∞)
$\boldsymbol{D_{44,33}(\Phi)}$	**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**[0.1818, ∞)**	

Table G.4: Projections of the estimated identified sets of Φ and $D(\Phi)$ under specification [7] of Table 1. "1" denotes high school dropouts; "2" denotes high school graduates; "3" denotes some college; "4" denotes four-year college graduates and graduate degrees.

Parameters		1940		1945		1950		1955		1960		1966	
U_{11}	U_{12}	1	[1.1111, ∞)	1	[1.1111, ∞)	1	[1.1111, ∞)	−1	[−0.9091, 0.9091]	−1	[−0.9091, 0.9091]	−1	[−0.9091, −0.1010]
U_{13}	U_{14}	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, −0.2424]	(−∞, −0.3939]	(−∞, −0.2424]	(−∞, −0.3939]	(−∞, −0.2424]	(−∞, −0.8788]
U_{21}	U_{22}	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[−0.1117, 0.1414]
U_{23}	U_{24}	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	[−1.7374, −0.1717]	[−1.9495, −0.2222]
$\boldsymbol{C_{21}(U)}$		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**[−0.2606, ∞)**	
U_{11}	U_{12}	1	[1.1111, ∞)	1	[1.1111, ∞)	1	[1.1111, ∞)	−1	[−0.9091, 0.9091]	−1	[−0.9091, 0.9091]	−1	[−0.9091, −0.1010]
U_{13}	U_{14}	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, −0.2424]	(−∞, −0.3939]	(−∞, −0.2424]	(−∞, −0.3939]	(−∞, −0.2424]	(−∞, −0.8788]
U_{31}	U_{32}	−1	[0.1111, ∞)	−1	[−0.3939, 15.7677]	−1	[−0.3939, 15.7677]	−1	[−0.8990, 0.6162]	−1	[−0.8990, 0.6162]	−1	[−0.8990, 0.1111]
U_{33}	U_{34}	(−∞, ∞)	(−∞, 15.6364]	[−0.3333, ∞)	[−16.5455, 15.6364]	[−0.3333, ∞)	[−17.7374, 7.8889]	[−0.3333, ∞)	[−17.7374, 7.8889]	[−0.3333, ∞)	[−17.7374, 7.8889]	[−0.6667, 0.6667]	[−2.2424, 0.1414]
$\boldsymbol{C_{31}(U)}$		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−0.5316, ∞)**		**[−0.5045, ∞)**		**[−0.2826, ∞)**	
U_{11}	U_{12}	1	[1.1111, ∞)	1	[1.1111, ∞)	1	[1.1111, ∞)	−1	[−0.9091, 0.9091]	−1	[−0.9091, 0.9091]	−1	[−0.9091, −0.1010]
U_{13}	U_{14}	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, −0.2424]	(−∞, −0.3939]	(−∞, −0.2424]	(−∞, −0.3939]	(−∞, −0.2424]	(−∞, −0.8788]
U_{41}	U_{42}	−1	[−0.4242, 3.8182]	−1	[−0.8485, 0.8485]	−1	[−0.8485, 0.8485]	−1	[−0.8485, 0.8485]	−1	[−0.8485, 0.6364]	−1	[−0.8485, 0.6364]
U_{43}	U_{44}	[−0.2727, 3.9091]	[−0.2424, ∞)	[−0.2727, 3.9091]	[0.0606, ∞)	[−0.2727, 3.9091]	[0.0606, ∞)	[−0.2727, 3.9091]	[0.0606, ∞)	[−0.8182, 0.8182]	[0.0606, ∞)	[−0.8182, 0.8182]	[0.0606, ∞)
$\boldsymbol{C_{41}(U)}$		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−0.1026, ∞]**		**[−0.1828, ∞)**		**[0.1618, ∞)**	
U_{21}	U_{22}	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[−0.1117, 0.1414]
U_{23}	U_{24}	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	[−1.7374, −0.1717]	[−1.9495, −0.2222]
U_{31}	U_{32}	−1	[0.1111, ∞)	−1	[−0.3939, 15.7677]	−1	[−0.3939, 15.7677]	−1	[−0.8990, 0.6162]	−1	[−0.8990, 0.6162]	−1	[−0.8990, 0.1111]
U_{33}	U_{34}	(−∞, ∞)	(−∞, 15.6364]	[−0.3333, ∞)	[−16.5455, 15.6364]	[−0.3333, ∞)	[−17.7374, 7.8889]	[−0.3333, ∞)	[−17.7374, 7.8889]	[−0.3333, ∞)	[−17.7374, 7.8889]	[−0.6667, 0.6667]	[−2.2424, 0.1414]
$\boldsymbol{C_{32}(U)}$		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**[−0.5188, 0.7789]**	
U_{21}	U_{22}	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[−0.1117, 0.1414]
U_{23}	U_{24}	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	(−∞, ∞)	[−1.7374, −0.1717]	[−1.9495, −0.2222]
U_{41}	U_{42}	−1	[−0.4242, 3.8182]	−1	[−0.8485, 0.8485]	−1	[−0.8485, 0.8485]	−1	[−0.8485, 0.8485]	−1	[−0.8485, 0.6364]	−1	[−0.8485, 0.6364]
U_{43}	U_{44}	[−0.2727, 3.9091]	[−0.2424, ∞)	[−0.2727, 3.9091]	[0.0606, ∞)	[−0.2727, 3.9091]	[0.0606, ∞)	[−0.2727, 3.9091]	[0.0606, ∞)	[−0.8182, 0.8182]	[0.0606, ∞)	[−0.8182, 0.8182]	[0.0606, ∞)
$\boldsymbol{C_{42}(U)}$		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**[−0.0674, ∞)**	
U_{31}	U_{32}	−1	[0.1111, ∞)	−1	[−0.3939, 15.7677]	−1	[−0.3939, 15.7677]	−1	[−0.8990, 0.6162]	−1	[−0.8990, 0.6162]	−1	[−0.8990, 0.1111]
U_{33}	U_{34}	(−∞, ∞)	(−∞, 15.6364]	[−0.3333, ∞)	[−16.5455, 15.6364]	[−0.3333, ∞)	[−17.7374, 7.8889]	[−0.3333, ∞)	[−17.7374, 7.8889]	[−0.3333, ∞)	[−17.7374, 7.8889]	[−0.6667, 0.6667]	[−2.2424, 0.1414]
U_{41}	U_{42}	−1	[−0.4242, 3.8182]	−1	[−0.8485, 0.8485]	−1	[−0.8485, 0.8485]	−1	[−0.8485, 0.8485]	−1	[−0.8485, 0.6364]	−1	[−0.8485, 0.6364]
U_{43}	U_{44}	[−0.2727, 3.9091]	[−0.2424, ∞)	[−0.2727, 3.9091]	[0.0606, ∞)	[−0.2727, 3.9091]	[0.0606, ∞)	[−0.2727, 3.9091]	[0.0606, ∞)	[−0.8182, 0.8182]	[0.0606, ∞)	[−0.8182, 0.8182]	[0.0606, ∞)
$\boldsymbol{C_{43}(U)}$		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**[−0.3565, ∞)**	

Table G.5: Projections of the estimated identified sets of U and $C(U)$ under specification [7] of Table 1. “1” denotes high school dropouts; “2” denotes high school graduates; “3” denotes some college; “4” denotes four-year college graduates and graduate degrees.

Parameters	1940		1945		1950		1955		1960		1966	
V_{11} V_{21}	1	[−0.9091, 0.9293]	1	[−0.9091, 0.9293]	1	[−0.9091, 0.9293]	1	[−0.9091, 0.9293]	−1	(−∞, −1.0404]	−1	(−∞, −1.0404]
V_{31} V_{41}	(−∞, 0.3131]	(−∞, −0.1313]	(−∞, 0.3131]	(−∞, −0.1313]	(−∞, 0.3131]	(−∞, −0.1313]	(−∞, 0.3131]	(−∞, −0.1313]	(−∞, −1.1551]	(−∞, −1.3838]	(−∞, −1.1551]	(−∞, −1.3838]
V_{12} V_{22}	1	[2.0606, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)
V_{32} V_{42}	[−6.9596, ∞)	(−∞, ∞)	[−6.9596, ∞)	(−∞, ∞)	[−6.9596, ∞)	(−∞, ∞)	[−6.9596, ∞)	(−∞, ∞)	[−6.9596, ∞)	(−∞, ∞)	[−6.9596, 25.3636]	(−∞, ∞)
$\boldsymbol{C_{21}(V)}$	**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**	
V_{11} V_{21}	1	[−0.9091, 0.9293]	1	[−0.9091, 0.9293]	1	[−0.9091, 0.9293]	1	[−0.9091, 0.9293]	−1	(−∞, −1.0404]	−1	(−∞, −1.0404]
V_{31} V_{41}	(−∞, 0.3131]	(−∞, −0.1313]	(−∞, 0.3131]	(−∞, −0.1313]	(−∞, 0.3131]	(−∞, −0.1313]	(−∞, 0.3131]	(−∞, −0.1313]	(−∞, −1.1551]	(−∞, −1.3838]	(−∞, −1.1551]	(−∞, −1.3838]
V_{13} V_{23}	−1	[−0.4848, 14.2323]	−1	[−0.4848, 14.2323]	−1	[−0.4848, 14.2323]	−1	[−0.7980, 0.7677]	−1	[−0.4848, 14.2323]	−1	[−0.7980, 0.7677]
V_{33} V_{43}	[−0.3333, ∞)	(−∞, ∞)	[−0.3333, ∞)	(−∞, ∞)	[−0.3333, ∞)	(−∞, ∞)	[−0.3333, ∞)	(−∞, ∞)	[−0.3333, ∞)	(−∞, ∞)	[−0.3333, ∞)	(−∞, 7.9091]
$\boldsymbol{C_{31}(V)}$	**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**	
V_{11} V_{21}	1	[−0.9091, 0.9293]	1	[−0.9091, 0.9293]	1	[−0.9091, 0.9293]	1	[−0.9091, 0.9293]	−1	(−∞, −1.0404]	−1	(−∞, −1.0404]
V_{31} V_{41}	(−∞, 0.3131]	(−∞, −0.1313]	(−∞, 0.3131]	(−∞, −0.1313]	(−∞, 0.3131]	(−∞, −0.1313]	(−∞, 0.3131]	(−∞, −0.1313]	(−∞, −1.1551]	(−∞, −1.3838]	(−∞, −1.1551]	(−∞, −1.3838]
V_{14} V_{24}	−1	[−0.9899, 0.5455]	−1	[−0.9899, 0.5455]	−1	[−0.9899, 0.5455]	−1	[−0.9899, 0.5455]	−1	[−0.9899, 0.5455]	−1	[−0.9899, 0.5455]
V_{34} V_{44}	[−0.4545, 0.7374]	[0.0606, ∞)	[−0.4545, 0.7374]	[0.0606, ∞)	[−0.4545, 0.7374]	[0.0606, ∞)	[−0.4545, 0.7374]	[0.0606, ∞)	[−0.4545, 0.7374]	[0.0606, ∞)	[−0.4545, 0.7374]	[0.0606, ∞)
$\boldsymbol{C_{41}(V)}$	**[−0.7883, ∞)**		**[−0.7691, ∞)**		**[−0.7212, ∞)**		**[−0.6957, ∞)**		**[0.5931, ∞)**		**[0.5433, ∞)**	
V_{12} V_{22}	1	[2.0606, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)
V_{32} V_{42}	[−6.9596, ∞)	(−∞, ∞)	[−6.9596, ∞)	(−∞, ∞)	[−6.9596, ∞)	(−∞, ∞)	[−6.9596, ∞)	(−∞, ∞)	[−6.9596, ∞)	(−∞, ∞)	[−6.9596, 25.3636]	(−∞, ∞)
V_{13} V_{23}	−1	[−0.4848, 14.2323]	−1	[−0.4848, 14.2323]	−1	[−0.4848, 14.2323]	−1	[−0.7980, 0.7677]	−1	[−0.4848, 14.2323]	−1	[−0.7980, 0.7677]
V_{33} V_{43}	[−0.3333, ∞)	(−∞, ∞)	[−0.3333, ∞)	(−∞, ∞)	[−0.3333, ∞)	(−∞, ∞)	[−0.3333, ∞)	(−∞, ∞)	[−0.3333, ∞)	(−∞, ∞)	[−0.3333, ∞)	(−∞, 7.9091]
$\boldsymbol{C_{32}(V)}$	**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**	
V_{12} V_{22}	1	[2.0606, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)	−1	[0.1414, ∞)
V_{32} V_{42}	[−6.9596, ∞)	(−∞, ∞)	[−6.9596, ∞)	(−∞, ∞)	[−6.9596, ∞)	(−∞, ∞)	[−6.9596, ∞)	(−∞, ∞)	[−6.9596, ∞)	(−∞, ∞)	[−6.9596, 25.3636]	(−∞, ∞)
V_{14} V_{24}	−1	[−0.9899, 0.5455]	−1	[−0.9899, 0.5455]	−1	[−0.9899, 0.5455]	−1	[−0.9899, 0.5455]	−1	[−0.9899, 0.5455]	−1	[−0.9899, 0.5455]
V_{34} V_{44}	[−0.4545, 0.7374]	[0.0606, ∞)	[−0.4545, 0.7374]	[0.0606, ∞)	[−0.4545, 0.7374]	[0.0606, ∞)	[−0.4545, 0.7374]	[0.0606, ∞)	[−0.4545, 0.7374]	[0.0606, ∞)	[−0.4545, 0.7374]	[0.0606, ∞)
$\boldsymbol{C_{42}(V)}$	**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**	
V_{13} V_{23}	−1	[−0.4848, 14.2323]	−1	[−0.4848, 14.2323]	−1	[−0.4848, 14.2323]	−1	[−0.7980, 0.7677]	−1	[−0.4848, 14.2323]	−1	[−0.7980, 0.7677]
V_{33} V_{43}	[−0.3333, ∞)	(−∞, ∞)	[−0.3333, ∞)	(−∞, ∞)	[−0.3333, ∞)	(−∞, ∞)	[−0.3333, ∞)	(−∞, ∞)	[−0.3333, ∞)	(−∞, ∞)	[−0.3333, ∞)	(−∞, 7.9091]
V_{14} V_{24}	−1	[−0.9899, 0.5455]	−1	[−0.9899, 0.5455]	−1	[−0.9899, 0.5455]	−1	[−0.9899, 0.5455]	−1	[−0.9899, 0.5455]	−1	[−0.9899, 0.5455]
V_{34} V_{44}	[−0.4545, 0.7374]	[0.0606, ∞)	[−0.4545, 0.7374]	[0.0606, ∞)	[−0.4545, 0.7374]	[0.0606, ∞)	[−0.4545, 0.7374]	[0.0606∞)	[−0.4545, 0.7374]	[0.0606, ∞)	[−0.4545, 0.7374]	[0.0606, ∞)
$\boldsymbol{C_{43}(V)}$	**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**		**(−∞, ∞)**	

Table G.6: Projections of the estimated identified sets of V and $C(V)$ under specification [7] of Table 1. "1" denotes high school dropouts; "2" denotes high school graduates; "3" denotes some college; "4" denotes four-year college graduates and graduate degrees.