Econometrica, Vol. 92, No. 3 (May, 2024), 749–774

IDENTIFICATION AND ESTIMATION IN MANY-TO-ONE TWO-SIDED MATCHING WITHOUT TRANSFERS

YINGHUA HE
Department of Economics, Rice University

SHRUTI SINHA
Toulouse School of Economics, University of Toulouse Capitole

XIAOTING SUN
Department of Economics, Simon Fraser University

In a setting of many-to-one two-sided matching with nontransferable utilities, for example, college admissions, we study conditions under which preferences of both sides are identified with data on one single market. Regardless of whether the market is centralized or decentralized, assuming that the observed matching is stable, we show nonparametric identification of preferences of both sides under certain exclusion restrictions. To take our results to the data, we use Monte Carlo simulations to evaluate different estimators, including the ones that are directly constructed from the identification. We find that a parametric Bayesian approach with a Gibbs sampler works well in realistically sized problems. Finally, we illustrate our methodology in decentralized admissions to public and private schools in Chile and conduct a counterfactual analysis of an affirmative action policy.

KEYWORDS: Many-to-one two-sided matching, nontransferable utility, nonparametric identification, college admissions, school choice.

1. INTRODUCTION

IN A MANY-TO-ONE TWO-SIDED MATCHING MARKET, agents are categorized into two sides; everyone on one side has preferences over those on the other side; an agent on only one of the two sides can have multiple match partners from the other side. Many real-life markets fit this description, for example, the medical resident match (Roth (1984), Agarwal (2015)) in the U.S., school admissions in Chile (Gazmuri (2017)) and Hungary

YingHua He: yinghua.he@rice.edu
Shruti Sinha: shruti.sinha@tse-fr.eu
Xiaoting Sun: xiaoting_sun@sfu.ca
This paper subsumes Sun's job market paper "Identification and Estimation of Many-to-One Matching With an Application to the U.S. College Admissions" (first draft: 2018) and He and Sinha's previously circulated paper "Identification and Estimation in Many-to-One Two-Sided Matching Without Transfers" (first draft: 2020). We thank two anonymous referees for their insightful comments and also Nikhil Agarwal, Cristina Gualdani, Thierry Magnac, Rosa Matzkin, Ismael Mourifie, Shuyang Sheng, and Xun Tang for their useful comments. We would also like to thank the seminar and conference participants at 2018 California Econometrics Conference, 2018 Midwest Econometrics Group Conference, 2019 Canadian Econometric Study Group Meetings, 2019 Network Econometrics Junior's conference at Northwestern University, 2019 Seattle-Vancouver Econometrics Conference, Auburn University, Econometric Society World Congress 2020, EEA-ESEM 2019, ICEF Moscow, Imperial College, Jinan University, Johns Hopkins University, National University of Singapore, Optimal Transport and Econometrics Workshop at University of Washington, Peking University, Rice University, Simon Fraser University, Toulouse School of Economics, UCLA, UC Riverside, University of Florida, University of Toronto, UNSW Sydney, and Western University. We thank the Ministry of Education and Agencia Educacion in Chile for providing the data for this paper. Thanks also to Alípio Ferreira Cantisani for excellent research assistance. Sinha acknowledges funding from the French National Research Agency (ANR) under the Investments for the Future (Investissements d'Avenir) program, grant ANR-17-EURE-0010.

(Aue, Klein, and Ortega (2020)), college admissions in the U.S., and graduate program admissions in France (He and Magnac (2022)). Such markets often exclude personalized transfers, even though limited monetary exchanges may exist. Hence, the literature defines it as matching without transfers or matching with nontransferable utility.

While the literature has extensively studied this type of matching theoretically (see, e.g., Roth and Sotomayor (1992), Azevedo and Leshno (2016)), its econometrics is less explored. Our paper aims to make a contribution by answering the following questions: Are the preferences of both sides identified from data on who matches with whom? If so, how can the preferences be estimated?

To fix ideas, we proceed in the language of college admissions. We derive a set of sufficient conditions under which both student and college preferences are nonparametrically identified. Our results are obtained from a single market in which there are a continuum of students and a fixed number of colleges. We use that to approximate a single large market. Further, we provide an estimation procedure that is practical even in settings with many agents, allowing for rich observed and unobserved heterogeneity. Understanding agent preferences is often crucial for policymaking, and one may analyze a wide range of counterfactual policies with estimated preferences. Potentially, our results open a new avenue of research on such matching markets.

The main challenge in identifying student preferences is that each student's *actual* choice set is unobservable to the researcher. For student i to be able to enroll at college c, college c needs to accept i. The same difficulty exists in the identification of college preferences. Moreover, each student's and each college's choice sets are endogenously determined in equilibrium without market-clearing prices.

In our continuum setting, we assume that an observed matching is stable. That is, no college prefers to reject any of its currently matched students to vacate a seat, and no student prefers to leave her current match to become unmatched or matched with a college that is willing to accept her and, if necessary, reject one of its currently matched students. Stability is often imposed in the study of various matching markets (see, for a survey, Chiappori and Salanié (2016)) and is satisfied in equilibrium in our setting in certain game-theoretical models (Artemov, Che, and He (2023), Fack, Grenet, and He (2019)).

Importantly, there is generically a unique stable matching that is characterized by the colleges' admission cutoffs (Azevedo and Leshno (2016)). When college preferences over individual students are represented by utility functions, a college's cutoff is the lowest utility level among its matched students. Cutoffs further define a student's actual choice set in equilibrium, called *feasible set*. A college is in a student's feasible set if the college's utility of being matched with her is higher than its cutoff. Stability implies that a student is matched with her most-preferred feasible college, similar to a discrete choice problem, except that feasible sets are unobservable and heterogeneous.

A simple equation, called the i-c match probability, is the key to understanding our identification result. Specifically, the conditional probability of student i being matched with college c is the sum of conditional probabilities of i choosing c from a given feasible set L weighted by the conditional probability of facing L:

$$\begin{aligned}
&\mathbb{P}(\text{student } i \text{ is matched with college } c|x_i) \\
&= \sum_{\text{all possible feasible sets, } L} \underbrace{\mathbb{P}\big(L \text{ is } i\text{'s feasible set } |x_i\big)}_{\equiv A(\text{college preferences})} \\
&\quad \cdot \underbrace{\mathbb{P}\big(c \text{ is } i\text{'s most-preferred college in } L|L, x_i\big)}_{\equiv B(\text{student preferences})},
\end{aligned}$$

where x_i consists of all observed characteristics of student i (e.g., pair-specific characteristics like distance to colleges). The equation provides a decomposition of the preferences of the two sides: for each given L, piece A only depends on the preferences of *all* colleges given cutoffs, while B only depends on i's preferences over all colleges.

We then detail a set of exclusion restrictions, among other regularity conditions, such that the excluded variables act as "demand shifters" and "feasible-set shifters" (or supply shifters). Sufficient variation in these excluded variables identifies the preferences of colleges and students using the i-c match probability described above.

Here are some intuitions. For a college d, an (i, d)-specific demand shifter traces out how i's preferences for d affects the i-c match probability. Similarly, an (i, d)-specific feasible-set shifter traces out how d's preference for i affects the i-c match probability. A nonexcluded variable affects the i-c match probability through preferences on both sides for all colleges. By taking derivatives of the i-c match probability with respect to (w.r.t.) all the variables, excluded and nonexcluded, we derive systems of linear equations that link the effects of variations in demand and supply. Hence, the identification problem reduces to setting up systems of linear equations and ensuring the existence of a unique solution.

The second objective of our paper is to provide practical methods that can be used to analyze real-life markets. We achieve this by deriving theoretical guidelines and showcasing a practical estimation method.

When taken to the data, the requirement of a large number of excluded variables may be difficult to meet. To address this, we theoretically characterize the tradeoff between exclusion restrictions and the degree of identifiable preference heterogeneity (Proposition 3.7). The researcher can use this result as a guideline for empirical studies when having insufficient excluded variables.

We also need a practical estimation method to take these identification results to the data. In fact, our identification arguments are constructive, leading to nonparametric and semiparametric estimators. Monte Carlo simulations suggest that estimating the matrices of partial derivatives in the linear systems using the average derivative estimators of Powell, Stock, and Stoker (1989) performs well in finite samples only when the curse of dimensionality is not severe. In a reasonably sized problem, we resort to a parametric Bayesian approach with a Gibbs sampler (Rossi, Allenby, and McCulloch (2012)), resembling applications such as Logan, Hoff, and Newton (2008) for one-to-one two-sided matching and Abdulkadiroğlu, Agarwal, and Pathak (2017) for a one-sided problem. We demonstrate its good performance in Monte Carlo simulations with high dimensionality.

As an empirical application, we consider the decentralized admissions to secondary schools in Chile. To the best of our knowledge, this is one of the first attempts to estimate the preferences of both sides in a *decentralized* market of many-to-one two-sided matching without transfers. There is no clearinghouse, and students do not submit rank-order lists of schools. By allowing flexible preference heterogeneity in the Bayesian approach, we estimate the preferences of students and schools. We also consider a counterfactual policy in which students from low-income families are prioritized for admissions to all schools. Segregation in terms of ability and income decreases, albeit slightly. We find that simply giving low-income students access to schools may not significantly change matching outcomes due to student preferences.

Related Literature. This paper is related to the literature on the identification of matching models; Table I provides an incomplete summary.

TABLE I

IDENTIFICATION RESULTS OF MATCHING MODELS.

	Transferable Utility (TU)	Nontransferable Utility (NTU)
One-to-one	The match surplus is identified (see, e.g., Choo and Siow (2006), Fox (2010), Galichon and Salanie (2022)).	The match surplus is identified (see, e.g., Dagsvik (2000), Menzel (2015)).
Many-to-one	The utility function and the distribution of the unobservables of both sides are identified in a homogeneous setting (see, e.g., Diamond and Agarwal (2017)).	The utility function and the distribution of the unobservables of both sides are identified in a homogeneous setting (see, e.g., Diamond and Agarwal (2017)). *Our paper*: The utility functions of both sides with heterogeneity and the distribution of the unobservables are identified.
Many-to-many	The match surplus and/or the distribution of the unobservables are identified (see, e.g., Fox (2010), Fox, Yang, and Hsu (2018)).	The match surplus is identified (see, e.g., Menzel (2022)).

This literature is split into several strands depending on the preference structures of the agents—transferable utility (TU) and nontransferable utilities (NTU);[1] and the maximum number of links an agent is permitted to form across sides—one-to-one, many-to-one, and many-to-many (see Chiappori and Salanié (2016), for a survey).

There is also a close relationship between the one-to-one TU matching model (Choo and Siow (2006), Diamond and Agarwal (2017), Fox (2010), Galichon and Salanie (2022), Gualdani and Sinha (2023), Sinha (2015)) and the many-to-one NTU matching model considered here. Market-clearing college cutoffs in our setting play the role of market-clearing shadow prices, although the endogenous cutoffs do not determine how the surplus is split among the agents.

Most of the work on identification within the NTU framework focuses on one-to-one markets (Dagsvik (2000), Menzel (2015)). Allowing for infinitely many agents on *both* sides of a many-to-one matching market, Agarwal (2015) and Diamond and Agarwal (2017) prove identification under a homogeneity restriction on the preferences,[2] and Ederer (2022) shows identification by relaxing the homogeneity assumption but still restricting preferences. Agarwal and Somaini (2023) provide a recent survey on empirical models of NTU matching.

Many-to-one NTU matching has been empirically studied in the context of secondary school admissions in Hungary (Aue, Klein, and Ortega (2020)) and graduate program admissions in France (He and Magnac (2022)). Their data include information on the preferences of both sides that is reported to a centralized mechanism. Therefore, they can independently identify and estimate the preferences of each side, essentially reducing the two-sided matching to two separate one-sided problems.

[1]See Galichon, Kominers, and Weber (2019) for an example of imperfectly transferable utility (ITU) models.

[2]When studying the medical resident match, Agarwal (2015) discusses some intuitions of using exclusion restrictions to identify heterogeneous preferences on each side of the market.

Centralized many-to-one NTU matching in the context of school choice has been studied extensively, both theoretically since Abdulkadiroğlu and Sönmez (2003) and empirically (e.g., Abdulkadiroğlu, Agarwal, and Pathak (2017), Agarwal and Somaini (2018), Calsamiglia, Fu, and Güell (2020), Fack, Grenet, and He (2019), He (2017)). In this literature, school preferences are (assumed to be) known because schools rank students according to certain pre-specified rules. The problem then reduces to identifying and estimating student preferences.

Feasible sets in our setting resemble endogenous consideration sets that arise in one-sided decision problems. In this sense, our paper relates to the growing strand of literature that studies the econometrics of decision problems under consideration set formation and unobserved choice set heterogeneity (see Crawford, Griffith, and Iaria (2021), for a survey). Our contribution here is that we provide a structural two-sided setting where the consideration probabilities in a student's decision problem are entirely determined by the college (supply side) preferences. Along similar lines to ours, Agarwal and Somaini (2022) study consumer choice models with latent choice-set constraints. Their identification conditions and ours are nonnested (see Section 3).

The remaining paper is organized as follows: Section 2 describes the model and data generating process; Section 3 discusses the identification of preferences of the agents on both sides of the market; Section 4 illustrates an empirical analysis of the match between students and secondary schools in Chile; and Section 5 concludes.

2. MODEL

For the sake of exposition, our model is set up as a college admissions problem. Consider a single market with a continuum of students and finitely many colleges. The set of all students is $\mathbf{I}$, with a probability measure Q defined over it,[3] and the set of all colleges is $\mathbf{C} = \{1, 2, \dots, C\}$. College $c \in \mathbf{C}$ has a capacity $q_c \in (0, 1)$.

For $i \in \mathbf{I}$, the utility of being matched with college c is u_{ic}; for $c \in \mathbf{C}$, the utility of being matched with i is v_{ci}. To prepare for our identification results in Section 3, we assume additive separability and excluded variables in the utility functions,[4] although the rest of the current section applies to more general models.

The utilities u_{ic} and v_{ci} depend on the vector (z_i, y_i, w_i), which is observable to the researcher, and $\epsilon_i = (\epsilon_{i1} \dots, \epsilon_{iC})$ and $\eta_i = (\eta_{1i} \dots, \eta_{Ci})$, which are unobservable to the researcher. Specifically, $z_i \in \mathcal{Z} \subseteq \mathbb{R}^{d_z}$, $y_i = (y_{i1} \dots, y_{iC}) \in \mathcal{Y} = \mathcal{Y}_1 \times \cdots \times \mathcal{Y}_C \subseteq \mathbb{R}^C$, and $w_i = (w_{1i} \dots, w_{Ci}) \in \mathcal{W} \subseteq \mathbb{R}^C$. Further, for any $c \in \mathbf{C}$, ϵ_{ic} and η_{ci} are scalar random variables. (ϵ_i, η_i) are independent and identically distributed (i.i.d.) draws from a joint distribution F. We make no restrictions on this joint distribution, allowing for arbitrary correlations within (ϵ_i, η_i).

ASSUMPTION 2.1: *Let $u^c : \mathcal{Z} \to \mathbb{R}$, $r^c : \mathcal{Y}_c \to \mathbb{R}$, and $v^c : \mathcal{Z} \to \mathbb{R}$ be nonparametric functions, such that*

$$u_{ic} = \tau_{ic} + \epsilon_{ic} \quad \textit{and} \quad v_{ci} = \iota_{ci} + \eta_{ci}, \forall c \in \mathbf{C}, \tag{1}$$

where $\tau_{ic} = u^c(z_i) + r^c(y_{ic})$ and $\iota_{ci} = v^c(z_i) + w_{ci}$.

[3]Our probability space is $(\mathbf{I}, \mathcal{B}(\mathbf{I}), Q)$ where $\mathcal{B}(\mathbf{I})$ is the Borel set of $\mathbf{I}$, $Q : \mathcal{B}(\mathbf{I}) \to [0, 1]$, $Q(\mathbf{I}) = 1$.

[4]See Supplemental Appendix A (He, Sinha, and Sun (2024)) for our identification results in a more general nonseparable utility specification.

Scalar y_{ic} is a demand shifter and scalar w_{ci} is a supply shifter: y_{ic} enters only u_{ic} and is excluded from all other utility functions, and w_{ci} enters only v_{ci}.

We impose scale normalization on each side. For students, there exists a known value $\overline{y}_c$ in the interior of its support such that $\frac{\partial r^c(\overline{y}_c)}{\partial y_{ic}} = 1$.[5] This holds trivially if $r^c(y_{ic}) = y_{ic}$. For colleges, we assume w_{ci} enters v_{ci} linearly with a coefficient normalized to one. As detailed below, this linearity assumption allows us to vary w_i to construct a sufficient number of equations without increasing the number of unknowns.[6] The assumptions on y_i and w_i can be switched, that is, having nonlinearity in w_i and linearity in y_i.

Students can remain unmatched, or equivalently, be matched with an outside option denoted by "0." The utility of the outside option is normalized to 0, $u_{i0} = 0\ \forall i \in \mathbf{I}$. We assume that colleges have responsive preferences.[7] This implies that the total utility of a college from being matched with a subset of students (up to its capacity) is increasing in its utility from each student; for example, the total utility is the sum of the utility from each of its matched students. College c has an acceptability threshold, T_c, and finds student i unacceptable if $v_{ci} < T_c$.

With the data from one such continuum market on $\{(z_i, y_i, w_i)\}_i$, $\{q_c\}_c$, and who matches with whom, we aim to identify student and college preferences by identifying $\{u^c, r^c, v^c, T_c\}_c$ and F, although as we shall see, $\{T_c\}_c$ are not always point identified. Note that (z_i, y_i, w_i) does not include college-specific variables that are constant across students, as these will be absorbed by the college-specific utility functions, (u^c, r^c, v^c).

We use the continuum market to approximate a data generating process in a large finite market as follows: $(z_i, y_i, w_i, \epsilon_i, \eta_i)$ is an i.i.d. draw from its joint distribution, college c's capacity is a q_c-fraction of the total number of students, and $\{u_{ic}, v_{ci}, T_c\}_{i,c}$ determines both sides' preferences. This approximation is close to the matching outcomes when we use the equilibrium concept that will be introduced in Section 2.1.[8]

REMARK 2.2: The location normalization in the functions is worth highlighting because the joint distribution of (ϵ_i, η_i), F, is fully nonparametric. In student preferences, we already impose the normalization, $u_{i0} = 0$, but we need another location normalization for each c on either $u^c(z_i) + r^c(y_{ic})$ or ϵ_{ic} to separately identify the two. Similarly, for each college c's preferences, we need to location-normalize two of the three model primitives $(v^c(z_i), \eta_{ci}, T_c)$ to pin down the third.

[5]If y_{ic} is known to have a negative effect on student preferences, the partial is normalized to -1. The same applies to w_i.

[6] The linearity in w_{ci} can be relaxed. We can allow w_{ci} to enter v_{ci} through a nonlinear function, which is known for some colleges. Namely, assume $v_{ci} = v^c(z_i) + s^c(w_{ci}) + \eta_{ci}$, where $s^c\colon \mathcal{W}_c \to \mathbb{R}$ is a nonparametric function such that (i) for each $c \in \overline{\mathbf{C}} \subset \mathbf{C}$, s^c is *unknown* with $\frac{\partial s^c(\overline{w}_c)}{\partial w_{ci}} = 1$ for some known value $\overline{w}_c$; and (ii) for each $c \in \mathbf{C} \setminus \overline{\mathbf{C}}$, s^c is *known*. In this specification, our identification relies on varying w_{ci} for $c \in \mathbf{C} \setminus \overline{\mathbf{C}}$ instead of all $c \in \mathbf{C}$.

[7]For $\varepsilon > 0$, let $N_\varepsilon(i)$ and $N_\varepsilon(i')$ be a neighborhood of students around v_{ci} and $v_{ci'}$, respectively, such that $Q(N_\varepsilon(i)) = Q(N_\varepsilon(i'))$. Responsive preferences imply that, for any $\mathbf{I}^c \subset \mathbf{I}$ with $Q(\mathbf{I}^c) \leq q_c - Q(N_\varepsilon(i))$, $N_\varepsilon(i) \subset \mathbf{I} \setminus \mathbf{I}^c$, and $N_\varepsilon(i') \subset \mathbf{I} \setminus \mathbf{I}^c$, college c prefers $\mathbf{I}^c \cup N_\varepsilon(i)$ to $\mathbf{I}^c \cup N_\varepsilon(i')$ if and only if $v_{ci} > v_{ci'}$. See Roth and Sotomayor (1992) for a definition in a case with discrete students.

[8]In a typical identification argument, the number of observations is taken to infinity while the "game" is kept constant. Our setting contains one single matching game that changes with the market size. Proposition 3 of Azevedo and Leshno (2016), Proposition 4 of Fack, Grenet, and He (2019), and Corollary 2 of Artemov, Che, and He (2023) imply that, under certain conditions, the equilibrium outcome in the continuum approximates well an equilibrium outcome in a large finite market. Such an approximation is also used in the network literature (e.g., Menzel (2022)).

In Section 3.1, we identify the derivatives of (u^c, r^c, v^c), so this additional location normalization is not needed. However, it is necessary for identifying T_c and F in Section 3.2, where we location-normalize $u^c(z_i) + r^c(y_{ic})$, $v^c(z_i)$, and η_{ci}.

2.1. *Matching and Stable Matching*

We define a matching function, or simply, a matching, $\mu : \mathbf{I} \to \mathbf{C} \cup \{0\}$, such that (i) $\mu(i) = c \iff i \in \mu^{-1}(c)$, and (ii) $\forall c \in \mathbf{C}$, $\mu^{-1}(c) \subseteq \mathbf{I}$, where $0 \leq Q(\mu^{-1}(c)) \leq q_c$.

The following concepts are important for our analysis: individual rationality, blocking pairs, and stability. For notational reasons, we define them in the case with discrete students, corresponding to our empirical application. The precise definitions for a model with a continuum of students can be found in Azevedo and Leshno (2016), with measure-zero sets of students appropriately dealt with.

A matching μ is *individually rational* if $u_{i\mu(i)} \geq u_{i0}$ and $v_{\mu(i)i} \geq T_{\mu(i)}$ for all $i \in \mathbf{I}$. A student-college pair $(i, c) \in \mathbf{I} \times \mathbf{C}$ *blocks* a matching μ if (i) student i strictly prefers college c to her current match $\mu(i)$, $u_{ic} > u_{i\mu(i)}$; and (ii) either college c has excess capacity, $Q(\mu^{-1}(c)) < q_c$, or college c prefers i to one of its matched students, $\exists i' \in \mu^{-1}(c)$, s.t., $v_{ci} > v_{ci'}$. Finally, a matching is *stable* if it is individually rational and not blocked by any pair $(i, c) \in \mathbf{I} \times \mathbf{C}$.

We assume that the matching in the data is stable.[9] A stable matching exists and is generically unique (Azevedo and Leshno (2016)).[10] Moreover, a stable matching is characterized by college cutoffs. College c's cutoff is determined by its least-preferred matched student when its capacity constraint is binding; otherwise, it coincides with the acceptability threshold. Let δ_c be college c's cutoff. Then $\forall c \in \mathbf{C}$,

$$\delta_c = \inf_{j \in \mu^{-1}(c)} v_{cj} \quad \text{if } Q\big(\mu^{-1}(c)\big) = q_c; \qquad \delta_c = T_c \quad \text{if } Q\big(\mu^{-1}(c)\big) < q_c. \tag{2}$$

By definition, $\delta_c \geq T_c$. Under the assumption of responsive college preferences, to determine if student i can be accepted by c, we just need to compare v_{ci} and δ_c. With nonresponsive preferences, how c ranks i and j would depend on who else c accepts, and δ_c alone would not be sufficient to determine if i could have been accepted by c.

In a stable matching of the continuum market, $\{u^c, r^c, v^c, T_c\}_{c \in \mathbf{C}}$, and F imply a unique vector of cutoffs, $\{\delta_c\}_c$. Therefore, $\{\delta_c\}_c$ is merely a shorthand notation for the expression in equation (2) rather than additional parameters.[11]

[9]Stability can be achieved in certain equilibrium if students apply to all acceptable colleges and if a stable mechanism, for example, the deferred acceptance (Gale and Shapley (1962)), is used to find the matching. Theoretically, provided that students know what criteria colleges use to rank them, stability can still be satisfied in equilibrium, even if students choose not to apply to all acceptable colleges due to application costs (Fack, Grenet, and He (2019)) or if students make certain application mistakes (Artemov, Che, and He (2023)). Importantly, achieving stability does not require the market to be centralized, as shown in laboratory experiments (see, e.g., Pais, Pintér, and Veszteg, 2020), and steps in mechanisms such as the deferred acceptance can be implemented in a decentralized fashion (Grenet, He, and Kübler (2022)).

[10]We need the regularity condition that the set $\{i \in \mathbf{I} : \mu(i)$ is strictly less preferred than $c\}$ is open $\forall c \in \mathbf{C}$. This condition implies that a stable matching always allows an extra measure zero set of students into a college when this can be done without compromising stability.

[11]As mentioned earlier, especially, in footnote 8, the continuum approximates a large finite market. The literature cited therein shows that equilibrium cutoffs, hence matching outcomes, in the large market can be close to $\{\delta_c\}_c$ with agents being practically "cutoff-takers." That is, each agent's realized preferences have a negligible effect on cutoffs in large markets.

2.2. *Two-Sided Discrete Choice Problem in a Stable Matching*

College c is said to be *feasible* for student i if and only if $v_{ci} \geq \delta_c$. A student can "choose" to match with any of her feasible colleges, but not any infeasible college. We call the set of all feasible colleges of a student her *feasible set*. Let $\mathcal{L}$ be the collection of the 2^C possible feasible sets, $\mathcal{L} \equiv \{L : 0 \in L, L \setminus \{0\} \subseteq \mathbf{C}\}$. By construction, the outside option always belongs to every feasible set. A matching is stable if and only if every student is matched with her most-preferred feasible college (Fack, Grenet, and He (2019)). A classic issue in two-sided matching is that students' feasible sets are determined endogenously, unobserved by the researcher, and heterogeneous across students.

For any given matching μ, the probability that $L \in \mathcal{L}$ is student i's feasible set conditional on (z_i, w_i) is

$$\begin{aligned}\mathbb{P}(\text{feasible set is } L|z_i, w_i) &= \mathbb{P}(v_{ci} \geq \delta_c \ \forall c \in L; v_{di} < \delta_d \ \forall d \notin L|z_i, w_i) \\ &\equiv \lambda_L(\iota_{1i}, \dots, \iota_{Ci}). \end{aligned} \tag{3}$$

If each student's feasible set was observed, λ_L could be identified from the data, and recovering student and college preferences would follow from standard arguments in the discrete choice literature. However, we do not observe the feasible sets.

Similarly, for students, $\mathbb{P}(c = \arg\max_{d\in L} u_{id}|L, z_i, y_i)$ is the probability that utility-maximizing students with observables (z_i, y_i) "choose" c from feasible set L. Conditional on L, this probability only depends on student preferences. We define

$$g_{c,L}(\tau_{ic}; \tau_{id}, d \neq c) \equiv \mathbb{P}\Big(c = \arg\max_{d\in L} u_{id}|L, z_i, y_i\Big), \quad \forall c \in \mathbf{C}, \tag{4}$$

where the first argument of $g_{c,L}$ is always τ_{ic}. If $d \notin L$, $g_{c,L}$ does not vary with τ_{id}.

3. NONPARAMETRIC IDENTIFICATION

We now turn to nonparametrically identifying the distribution of student and college preferences given a stable matching μ and covariates (z_i, y_i, w_i) in *one* market.

A matching is stable if and only if every student is matched with the most-preferred college in her feasible set. Thus, stability implies, for $c \in \mathbf{C} \cup \{0\}$,

$$\begin{aligned}\sigma_c(z_i, y_i, w_i) &\equiv \mathbb{P}\big(\mu(i) = c|z_i, y_i, w_i\big) \\ &= \sum_{L\in\mathcal{L}} \mathbb{P}(\text{feasible set is } L|z_i, w_i) \cdot \mathbb{P}\Big(c = \arg\max_{d\in L} u_{id}|L, z_i, y_i\Big) \\ &= \sum_{L\in\mathcal{L}} \lambda_L(\iota_{1i}, \dots, \iota_{Ci}) \cdot g_{c,L}(\tau_{ic}; \tau_{id}, d \neq c). \end{aligned} \tag{5}$$

Hence, the conditional match probability, $\sigma_c(z_i, y_i, w_i)$, which is the fraction of students with (z_i, y_i, w_i) matched with c and known in the population data, is linked to the model through student and college preferences.

Below, we first study the conditions under which the functions $\{u^c, r^c, v^c\}_c$ are nonparametrically identified; we then identify the joint distribution of (ϵ_i, η_i), F. Later in Section 3.3, we present results that impose fewer requirements on the data.

3.1. *Identifying the Derivatives of the Utility Functions*

We nonparametrically identify the derivatives of the functions $\{u^c, r^c, v^c\}_c$ w.r.t. the observables. With these derivatives identified, the functions are identified up to a constant, provided that z_i and y_i have full support. The idea is to use the variation in the excluded variables to trace out how each argument in equation (5) affects the conditional match probabilities. Specifically, the excluded variables in student preferences (y_i) only shift demand, while the excluded variables in college preferences (w_i) shift supply, or feasible sets. The effect of other variables (z_i) that enter both demand and supply can be written as a combination of the effects of y_i and w_i. This leads to a system of linear equations in the derivatives of the conditional match probabilities, whose solution is our parameters of interest.

3.1.1. *A Simple Example With One College*

We describe the intuition for identification in a one-college example, $\mathbf{C} = \{1\}$. Student utility functions are $u_{i1} = u^1(z_i) + r^1(y_{i1}) + \epsilon_{i1}$ for college 1 and $u_{i0} = 0$ for the outside option. Here, z_i is a scalar and $\frac{\partial r^1(\overline{y}_1)}{\partial y_1} = 1$ for a known value, $\overline{y}_1$. College 1's utility function is $v_{1i} = v^1(z_i) + w_{1i} + \eta_{1i}$, and the (unobserved) cutoff is δ_1.

To identify $\frac{\partial u^1}{\partial z_i}$ and $\frac{\partial v^1}{\partial z_i}$, we fix $y_{i1} = \overline{y}_1$ and consider any value (z, w_1) in the interior of its support. Figure 1(a) shows that the space of $(\epsilon_{i1}, \eta_{1i})$ is partitioned into four parts based on the feasibility of college 1 and student i's preferences (i.e., the acceptability of college 1 to i). Moreover, $\mu(i) = 1$ if and only if $\epsilon_{i1} > -u^1(z) - r^1(\overline{y}_1)$ (college 1 is acceptable to i) and $\eta_{1i} > \delta_1 - v^1(z) - w_1$ (college 1 is feasible to i).

Figures 1(b)–(d) depict how the marginal effect of z_i on match probability is linked to the marginal effects of the excluded variables, y_{i1} and w_{1i}. Panel (b) describes the marginal effect of y_{i1}. Specifically, decreasing y_{i1} from $\overline{y}_1$ to $\overline{y}_1 - \Delta y$ makes college 1 less attractive to student i, and the region in which $\mu(i) = 1$ shrinks along the horizontal ϵ_{i1}-axis. The area I_1 depicts the set of students whose match differs when y_{i1} decreases. The induced change in the match probability, $\sigma_1 = \mathbb{P}(\mu(i) = 1 | (z_i, y_{i1}, w_{1i}) = (z, \overline{y}_1, w_1)) \equiv \mathbb{P}(\mu(i) = 1 | z, \overline{y}_1, w_1)$, is the mass that the density of $(\epsilon_{i1}, \eta_{1i})$ puts on I_1, or for $(z_i, y_{i1}, w_{1i}) = (z, \overline{y}_1, w_1)$,

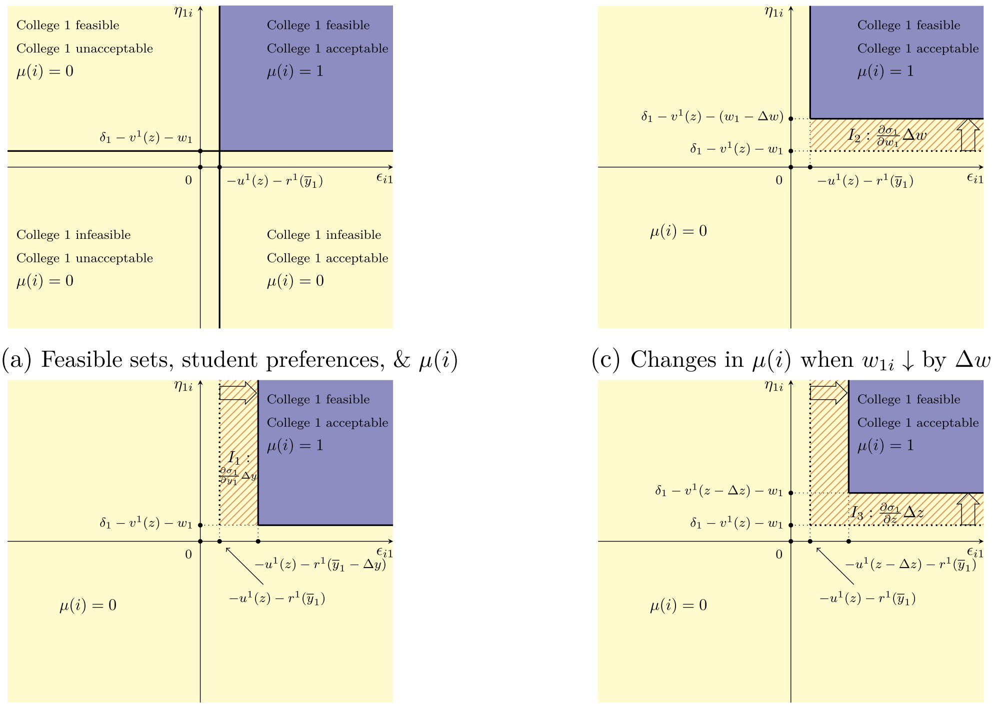
$$\frac{\partial \sigma_1(z, \overline{y}_1, w_1)}{\partial y_{i1}} = \frac{\partial r^1(\overline{y}_1)}{\partial y_{i1}} \cdot \sum_{L \in \mathcal{L}} \lambda_L(\iota_1) \cdot \frac{\partial g_{1,L}(\tau_1)}{\partial \tau_{i1}} = \sum_{L \in \mathcal{L}} \lambda_L(\iota_1) \cdot \frac{\partial g_{1,L}(\tau_1)}{\partial \tau_{i1}}, \tag{6}$$

where $\tau_1 \equiv u^1(z) + r^1(\overline{y}_1)$ and $\iota_1 \equiv v^1(z) + w_1$. By scale normalization, $\frac{\partial r^1(\overline{y}_1)}{\partial y_1} = 1$.

Panel (c) shows a similar graph in which decreasing w_{1i} from w_1 to $w_1 - \Delta w$ makes college 1 less likely to be feasible to student i. Hence, the region $\mu(i) = 1$ shrinks along the vertical η_{1i}-axis. The change in the match probability induced by the decrease in w_{1i} is the mass that the density of $(\epsilon_{i1}, \eta_{1i})$ puts on the area I_2, or

$$\frac{\partial \sigma_1(z, \overline{y}_1, w_1)}{\partial w_{1i}} = \sum_{L \in \mathcal{L}} \frac{\partial \lambda_L(\iota_1)}{\partial \iota_{1i}} \cdot g_{1,L}(\tau_1). \tag{7}$$

In panel (d), the decrease in z_i reduces the attractiveness and feasibility of college 1 for student i, because z_i enters both student and college preferences. Besides, how z_i changes



(a) Feasible sets, student preferences, & $\mu(i)$

(c) Changes in $\mu(i)$ when $w_{1i} \downarrow$ by Δw

(b) Changes in $\mu(i)$ when $y_{i1} \downarrow$ by Δy

(d) Changes in $\mu(i)$ when $z_i \downarrow$ by Δz

FIGURE 1.—Partitioning the Space of Unobservables in the One-College Case. *Notes*: Panel (a) describes the partition of the $(\epsilon_{i1}, \eta_{1i})$ space given $(z_i, y_{i1}, w_{1i}) = (z, \overline{y}_1, w_1)$ by student i's feasible set and preferences. The other panels show the changes in $\mu(i)$ when y_i decreases by Δy and affects only student preferences (panel b), when w_{1i} decreases by Δw and affects only feasible set (panel c), and when z_i decreases by Δz (panel d).

the region of $\mu(i) = 1$ relies on the shape of the functions u^1 and v^1. The change in the match probability induced by the change in z_i corresponds to the area I_3, or

$$\frac{\partial \sigma_1(z, \overline{y}_1, w_1)}{\partial z_i} = \frac{\partial u^1(z)}{\partial z_i} \cdot \sum_{L\in\mathcal{L}} \lambda_L(\iota_1) \cdot \frac{\partial g_{1,L}(\tau_1)}{\partial \tau_{i1}} + \frac{\partial v^1(z)}{\partial z_i} \cdot \sum_{L\in\mathcal{L}} \frac{\partial \lambda_L(\iota_1)}{\partial \iota_{1i}} \cdot g_{1,L}(\tau_1). \tag{8}$$

Our identification result relies on the changes caused by z_i, y_{i1}, and w_{1i}. Plugging equations (6) and (7) into equation (8), we have, for $(z_i, y_{i1}, w_{1i}) = (z, \overline{y}_1, w_1)$,

$$\frac{\partial \sigma_1(z, \overline{y}_1, w_1)}{\partial z_i} = \frac{\partial u^1(z)}{\partial z_i} \cdot \frac{\partial \sigma_1(z, \overline{y}_1, w_1)}{\partial y_{i1}} + \frac{\partial v^1(z)}{\partial z_i} \cdot \frac{\partial \sigma_1(z, \overline{y}_1, w_1)}{\partial w_{1i}}. \tag{9}$$

This equation reflects the chain rule: the effect of z_i on the match probability, $\frac{\partial \sigma_1}{\partial z_i}$, is realized through its effects on utilities u_{i1} and v_{1i}, captured by $\frac{\partial u^1}{\partial z_i}$ and $\frac{\partial v^1}{\partial z_i}$, and the effects of the utilities on the match probability, captured by $\frac{\partial \sigma_1}{\partial y_{i1}}$ and $\frac{\partial \sigma_1}{\partial w_{1i}}$. In equation (9), the derivatives of the match probability can be recovered from the population data, and the two unknowns, $\frac{\partial u^1}{\partial z_i}$ and $\frac{\partial v^1}{\partial z_i}$ are the parameters of interest.

Importantly, when w_{1i} varies, the conditional match probability changes, but the two unknowns remain constant, a consequence of w_{1i} entering v_{1i} in a known way. If, for any z, $(\epsilon_{i1}, \eta_{1i})$ has enough variation such that two distinct values of w_{1i} produce two linearly independent equations that have a unique solution, we identify the unknowns. This requirement is formalized as Condition 3.3 later, which imposes a mild restriction on the distribution of $(\epsilon_{i1}, \eta_{1i})$ as detailed in Supplemental Appendix C.1.

We then identify $\frac{\partial r^1}{\partial y_{i1}}$, which is not one when $y_{i1} \neq \overline{y}_1$. For any value (z, y_1, w_1), plugging equations (6) and (7) into equation (8), we rearrange and obtain

$$\left(\frac{\partial \sigma_1(z, y_1, w_1)}{\partial z_i} - \frac{\partial v^1(z)}{\partial z_i} \cdot \frac{\partial \sigma_1(z, y_1, w_1)}{\partial w_{1i}}\right) \cdot \frac{\partial r^1(y_1)}{\partial y_{i1}} = \frac{\partial u^1(z)}{\partial z_i} \cdot \frac{\partial \sigma_1(z, y_1, w_1)}{\partial y_{i1}}, \quad (10)$$

where all the terms except for $\frac{\partial r^1}{\partial y_{i1}}$ are either identified or known. For any y_1, if $\frac{\partial \sigma_1}{\partial z_i} - \frac{\partial v^1}{\partial z_i} \cdot \frac{\partial \sigma_1}{\partial w_{1i}} \neq 0$ for some value (z, w_1), $\frac{\partial r^1}{\partial y_{i1}}$ is identified; otherwise, equation (10) implies that $\frac{\partial \sigma_1}{\partial y_{i1}} = 0$ for all (z, w_1), and thus $\frac{\partial r^1}{\partial y_{i1}}$ is also identified and equal to zero.

Below, we extend this example to the case with multiple colleges. We derive equation (9) for each college, in which the marginal effect of z_i on the probability of being matched with each college is the sum of its marginal effects on (u^c, v^c) for all c. By varying $\{w_{ci}\}_c$, we form a system of equations in $\{\frac{\partial u^c}{\partial z_i}, \frac{\partial v^c}{\partial z_i}\}_c$. The identification of $\frac{\partial r^c}{\partial y_{ic}}$ is the same as above and relies on a generalized version of equation (10) because we can identify $\frac{\partial r^c}{\partial y_{ic}}$ for each c separately by holding $\frac{\partial r^d}{\partial y_{id}} = 1$ for $d \neq c$.

3.1.2. *Formal Identification Results*

Our nonparametric identification of $\{\frac{\partial u^c}{\partial z_i}, \frac{\partial r^c}{\partial y_{ic}}, \frac{\partial v^c}{\partial z_i}\}_c$ extends Matzkin (2019) who uses excluded variables to identify nonparametric nonseparable discrete choice models.

ASSUMPTION 3.1: (*i*) *z_i, y_i, and w_i are continuously distributed*; (*ii*) *for each $c \in \mathbf{C}$, the functions*, (u^c, r^c, v^c), *are continuously differentiable*; *and* (*iii*) *F is continuously differentiable*.

Part (i) of Assumption 3.1 requires that all covariates are continuous (but not necessarily have full support), which is relaxed in Section 3.3.

ASSUMPTION 3.2: (ϵ_i, η_i) *is distributed independently of* (z_i, y_i, w_i).

This exogeneity assumption is made for simplicity. One way to relax this assumption is to adopt a control function approach. See Supplemental Appendix B for a discussion.

Additionally, we need a condition on the derivatives of match probabilities w.r.t. the excluded variables. Let $\Pi_y(z_i, y_i, w_i)$ be a $C \times C$ Jacobian matrix of the match probabilities w.r.t. the excluded variable y_i, whose (c, d) element is $\frac{\partial \sigma_c(z_i, y_i, w_i)}{\partial y_{id}}$. Similarly, let $\Pi_w(z_i, y_i, w_i)$ be a $C \times C$ Jacobian matrix w.r.t. the excluded variable w_i. Fix $y_i = \overline{y}$, where $\overline{y} = (\overline{y}_1, \dots, \overline{y}_C)$. We then consider a pair of distinct values of w_i, $\widehat{w}$ and $\widetilde{w}$, and define a

$2C \times 2C$ matrix evaluated at $(z_i, y_i) = (z, \overline{y})$,

$$\Pi(z, \overline{y}, \widehat{w}, \widetilde{w}) \equiv \begin{pmatrix} \Pi_y(z, \overline{y}, \widehat{w}) & \Pi_w(z, \overline{y}, \widehat{w}) \\ \Pi_y(z, \overline{y}, \widetilde{w}) & \Pi_w(z, \overline{y}, \widetilde{w}) \end{pmatrix}.$$

We impose the following testable condition on $\Pi(z, \overline{y}, \widehat{w}, \widetilde{w})$.[12]

CONDITION 3.3: *For a given value z in the interior of $\mathcal{Z}$, there exist two values of w_i, $\widehat{w}$ and $\widetilde{w}$, in w_i's support conditional on $(z_i, y_i) = (z, \overline{y})$ such that $\Pi(z, \overline{y}, \widehat{w}, \widetilde{w})$ has rank $2C$.*

Note that, for any value of z_i, Condition 3.3 only needs *two* values of w_i at which $\Pi(z, \overline{y}, \widehat{w}, \widetilde{w})$ is full-rank. In other words, Condition 3.3 can hold even when there are *infinitely many* values of w_i at which $\Pi(z, \overline{y}, \widehat{w}, \widetilde{w})$ is *not* full-rank.

Nevertheless, the condition may fail in some reasonable scenarios. For example, if there is a college that students with $z_i = z$ consider unacceptable for a neighborhood around $\overline{y}$, Condition 3.3 is violated.[13] Similarly, the condition fails if there is a college that always considers students with $z_i = z$ unacceptable for all values of w_i.

For a set of sufficient and testable conditions for Condition 3.3, consider $\widehat{w}$ that is large enough such that all colleges are feasible when $w_i = \widehat{w}$. In this case, the upper half of $\Pi(z, \overline{y}, \widehat{w}, \widetilde{w})$ corresponds to a one-sided discrete choice model, where the feasible-set shifter w_i has no impact on the matching probabilities, and thus $\Pi_w(z, \overline{y}, \widehat{w}) = \mathbf{0}_{C\times C}$. This makes $\Pi(z, \overline{y}, \widehat{w}, \widetilde{w})$ a lower triangular block matrix. Then Condition 3.3 holds if and only if both $\Pi_y(z, \overline{y}, \widehat{w})$ and $\Pi_w(z, \overline{y}, \widetilde{w})$ are invertible, which can be achieved under the connected substitutes conditions (Berry, Gandhi, and Haile (2013), Theorem 2).[14, 15] This set of sufficient conditions suggests that Condition 3.3 holds under common logit or probit models provided that w_i has a reasonably large support. It is worth noting that the large support of w_i above is used to facilitate a clear connection with the connected substitutes conditions, but may be stronger than necessary. In Supplemental Appendix C.2, we

[12]For a statistical test of Condition 3.3 given a value of z_i and a pair of values of w_i, one may use the method proposed by Chen and Fang (2019). Testing H_0: $\mathrm{rank}(\Pi(z, \overline{y}, \widehat{w}, \widetilde{w})) \leq 2C - 1$ against H_1: $\mathrm{rank}(\Pi(z, \overline{y}, \widehat{w}, \widetilde{w})) > 2C - 1$ is a special case of setup (1) in Chen and Fang (2019, p. 1788). In practice, one needs to find two values of w_i satisfying Condition 3.3, which can be achieved by the following procedure: (i) choose m pairs of w_i values, (ii) for each pair, apply this test and calculate the p values, and (iii) use the Holm–Bonferroni method to control the overall size of this multiple hypothesis testing problem and then determine which null hypothesis, if any, is rejected. The pair of values of w_i associated with any rejected hypothesis satisfies Condition 3.3.

[13]Specifically, when the match probability with college c is zero for a neighborhood around $\overline{y}$ and for all values of w_i, its partial derivatives with college c w.r.t. y_i (evaluated at $\overline{y}$) and w_i are always zero. Hence, $\Pi(z, \overline{y}, \widehat{w}, \widehat{w})$ contains two rows of zeros (i.e., the c^{th} and $(c + C)^{th}$ rows). Moreover, the derivatives of all match probabilities w.r.t. y_{ic} (evaluated at $\overline{y}$) and w_{ci} are always zero. Hence, $\Pi(z, \overline{y}, \widehat{w}, \widehat{w})$ also contains two columns of zeros (i.e., the c^{th} and $(c + C)^{th}$ columns).

[14]The invertibility of $\Pi_y(z, \overline{y}, \widehat{w})$ holds under the following two assumptions: (i) for each $d \in \mathbf{C} \cup \{0\}$, $\frac{\partial \sigma_d(z_i, y_i, w_i)}{\partial y_{ic}} \leq 0$ for all $c \in \mathbf{C}\backslash\{d\}$; (ii) for any nonempty $\overline{\mathbf{C}} \subseteq \mathbf{C}$, there exists $c \in \overline{\mathbf{C}}$ and $d \notin \overline{\mathbf{C}}$ such that $\frac{\partial \sigma_d(z, \overline{y}, \widehat{w})}{\partial y_{ic}} < 0$. The invertibility of $\Pi_w(z, \overline{y}, \widetilde{w})$ holds under similar assumptions.

[15]This sufficient condition for Condition 3.3 provides us with a clear comparison with a related study, Agarwal and Somaini (2022). Their conditions for the identification of a similar model and our conditions are not nested. Specifically, they assume large support on w_i, impose a substitution condition on $\Pi_y(z, y, w)$ stronger than the one in Berry, Gandhi, and Haile (2013), and require the substitution condition to hold for *all* but a finite set of y_i values. In contrast, in this sufficient condition for Condition 3.3, with large support of w_i, we need a substitution condition on both $\Pi_y(z, \overline{y}, w)$ and $\Pi_w(z, \overline{y}, w)$ à la Berry, Gandhi, and Haile (2013), but only for $y_i = \overline{y}$.

analyze common logit and probit models with two or more colleges. The results suggest that Condition 3.3 is satisfied for a wide range of values of $(\widehat{w}, \widetilde{w})$; in fact, the *failure* of Condition 3.3 imposes strict restrictions on the supply, or the conditional probabilities of different feasible sets.[16]

PROPOSITION 3.4: *For given (z, y) in the interior of $\mathcal{Z} \times \mathcal{Y}$, under Assumptions 2.1, 3.1, 3.2, and Condition 3.3, $\{\frac{\partial u^c(z)}{\partial z_i^k}, \frac{\partial v^c(z)}{\partial z_i^k}, \frac{\partial r^c(y_c)}{\partial y_{ic}}\}_{c \in \mathbf{C}}$ are identified for $k = 1, \dots, d_z$.*

Our identification arguments proceed in two steps. First, to identify $\frac{\partial u^c(z)}{\partial z_i^k}$ and $\frac{\partial v^c(z)}{\partial z_i^k}$, we use the variation in the excluded variables (y_i, w_i) to derive a system of linear equations that generalizes equation (9). Fixing $y_i = \overline{y}$, for any value (z, w) in the interior of $\mathcal{Z} \times \mathcal{W}$, for each college $d \in \mathbf{C}$, and for the kth component of z_i, z_i^k,

$$\frac{\partial \sigma_d(z, \overline{y}, w)}{\partial z_i^k} = \sum_{c \in \mathbf{C}} \frac{\partial \sigma_d(z, \overline{y}, w)}{\partial y_{ic}} \cdot \frac{\partial u^c(z)}{\partial z_i^k} + \sum_{c \in \mathbf{C}} \frac{\partial \sigma_d(z, \overline{y}, w)}{\partial w_{ci}} \cdot \frac{\partial v^c(z)}{\partial z_i^k}, \tag{11}$$

which gives C linear equations of $2C$ unknowns. Note that equation (11) is for one value of w_i. By evaluating equation (11) at $d = 1, \dots, C$ and two different values of w_i, $\widehat{w}$, and $\widetilde{w}$, and stacking them together, we have

$$\begin{pmatrix} \dfrac{\partial \boldsymbol{\sigma}(z, \overline{y}, \widehat{w})}{\partial z_i^k} \\ \dfrac{\partial \boldsymbol{\sigma}(z, \overline{y}, \widetilde{w})}{\partial z_i^k} \end{pmatrix} = \Pi(z, \overline{y}, \widehat{w}, \widetilde{w}) \cdot \begin{pmatrix} \dfrac{\partial \boldsymbol{u}(z)}{\partial z_i^k} \\ \dfrac{\partial \mathbf{v}(z)}{\partial z_i^k} \end{pmatrix}, \tag{12}$$

where $\boldsymbol{\sigma} \equiv (\sigma_1, \dots, \sigma_C)'$, $\boldsymbol{u} \equiv (u^1, \dots, u^C)'$, and $\boldsymbol{v} \equiv (v^1, \dots, v^C)'$. Both $\Pi(z, \overline{y}, \widehat{w}, \widetilde{w})$ and the left-hand side are known from the population data. Hence, the invertibility of $\Pi(z, \overline{y}, \widehat{w}, \widetilde{w})$ in Condition 3.3 guarantees the existence of a unique solution to this system, leading to the identification of $\frac{\partial u^c(z)}{\partial z_i^k}$ and $\frac{\partial v^c(z)}{\partial z_i^k}$. Importantly, this only requires a pair of distinct values of w_i to satisfy Condition 3.3.

Second, we identify $\frac{\partial r^c(y_c)}{\partial y_{ic}}$ for each $c \in \mathbf{C}$ by generalizing equation (10) for the one-college example. See the Appendix for detailed proof.

In certain empirical applications, identifying the above derivatives is sufficient, in which case a large-support assumption on excluded variables is not needed. When the functions (u^c, v^c, r^c) must be identified, a full-support condition on their arguments, (z_i, y_i), is often imposed. This is the case for our next result of identifying the joint distribution F and cutoffs δ_c. Additionally, a full-support assumption on w_i is also required. Hence, we will assume all observables, (z_i, y_i, w_i), have full support.

3.2. *Identifying the Cutoffs and Joint Distribution of Unobservables*

We now formalize the assumptions that are needed for the identification of the cutoffs, $\{\delta_c\}_c$, and the joint distribution of unobservables, F.

[16]Also, see Supplemental Appendix C.1 for a one-college example, where we show that Condition 3.3 holds for all distribution of η_{1i} except for the exponential distribution.

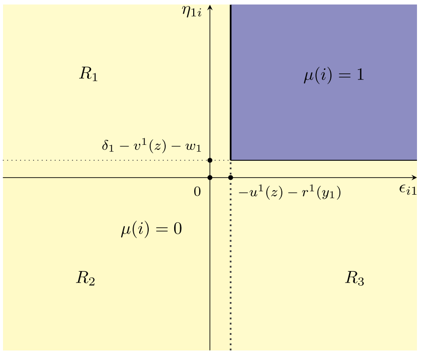


FIGURE 2.—Partitioning the Space of Unobservables in the One-College Case. *Notes*: This figure shows the partition of the space of the unobservables $(\epsilon_{i1}, \eta_{1i})$ by the matching outcome ($\mu(i) = 0$ or 1) in a one-college setting given $(z_i, y_{i1}, w_{1i}) = (z, y_1, w_1)$.

ASSUMPTION 3.5: *For each $c \in \mathbf{C}$, (i) the functions $u^c + r^c$ and v^c are identified;*[17] *(ii) y_{ic} and w_{ci} possess an everywhere positive Lebesgue density conditional on z_i; (iii) the range of the function r^c is the whole real line, and $\mathcal{W} = \mathbb{R}^C$; and (iv) the ρ_c-quantile of the marginal distribution of η_{ci} is 0, that is, $Quantile_{\eta_{ci}}(\rho_c) \equiv \inf\{\eta_c : F_{\eta_{ci}}(\eta_c) \geq \rho_c\} = 0$, for an arbitrary $\rho_c \in (0, 1)$.*

Part (iv) is a location normalization on college preferences, as mentioned in Remark 2.2, which can be done college-by-college. Alternatively, one may replace part (iv) by normalizing cutoffs $\{\delta_c\}_c$ to zero.

Before presenting our formal results, we give some intuitions. To identify cutoff δ_c, under the full-support assumption on student preferences (parts (ii) and (iii) of Assumption 3.5), we consider a mass of students to whom all colleges except for c are unacceptable. The probability of these students matching with c is $1 - F_{\eta_{ci}}(\delta_c - v^c(z_i) - w_{ci})$. Given the location normalization of $F_{\eta_{ci}}$ (part (iv) of Assumption 3.5), finding the maximum value of $v^c(z_i) + w_{ci}$ that sets this probability to $1 - \rho_c$ identifies δ_c.

To identify F, we use the conditional probability of being *unmatched*. Let us illustrate the intuition in the same one-college example as in Section 3.1.1. For any given value (z, y_1, w_1), Figure 2 shows the partition of the space of $(\epsilon_{i1}, \eta_{1i})$ by matching outcome, $\mu(i) = 0$ or 1. The corresponding area of $\mu(i) = 0$ can be decomposed into three parts, R_1, R_2, and R_3. Moreover,

$$\begin{aligned}
&\mathbb{P}\big(\mu(i) = 0 \big| (z_i, y_{i1}, w_{1i}) = (z, y_1, w_1)\big) \\
&\quad = \mathbb{P}\big(\epsilon_{i1} < -u^1(z) - r^1(y_1) \text{ or } \eta_{1i} < \delta_1 - v^1(z) - w_1 | z, y_1, w_1\big) \\
&\quad = \mathbb{P}(R_1 \cup R_2 | z, y_1, w_1) + \mathbb{P}(R_3 \cup R_2 | z, y_1, w_1) - \mathbb{P}(R_2 | z, y_1, w_1),
\end{aligned} \tag{13}$$

where $\mathbb{P}(R_2 | z, y_1, w_1)$ is the joint CDF of $(\epsilon_{i1}, \eta_{1i})$, $F(-u^1(z) - r^1(y_1), \delta_1 - v^1(z) - w_1)$, our parameter of interest.

Further, $\mathbb{P}(R_1 \cup R_2 | z, y_1, w_1) = F_{\epsilon_{i1}}(-u^1(z) - r^1(y_1))$ is the marginal CDF of ϵ_{i1}. It can be identified by considering the subset of students whose value of w_{1i} is high enough so

[17] This requires Proposition 3.4, a full support assumption on (z_i, y_i), and location normalization on $u^c + r^c$ and v^c for each c. See Remark 2.2 for a discussion on location normalization.

that the college will be feasible no matter what value η_{1i} takes. That is, we can identify $\mathbb{P}(R_1 \cup R_2|z, y_1, w_1)$ by "shutting down" the effects of college preferences. Similarly, $\mathbb{P}(R_3 \cup R_2|z, y_1, w_1)$ can be identified by focusing on the subset of students whose value of $r^1(y_{i1})$ is large enough so that those students find college 1 acceptable no matter what value ϵ_{i1} takes. As $\mathbb{P}(\mu(i) = 0|z, y_1, w_1)$ is known from the population data, once $\mathbb{P}(R_1 \cup R_2|z, y_1, w_1)$ and $\mathbb{P}(R_3 \cup R_2|z, y_1, w_1)$ are identified, equation (13) implies that $\mathbb{P}(R_2|z, y_1, w_1)$, and thus F are identified.

PROPOSITION 3.6: *Under Assumptions 2.1, 3.2, and 3.5, (i) cutoffs $\{\delta_c\}_c$ are identified; (ii) the joint distribution of (ϵ_i, η_i), F, is identified; and (iii) acceptability threshold T_c for any college c with vacancies is identified, but T_c for c with a binding capacity constraint is only partially identified, $T_c \leq \delta_c$.*

To show part (ii) in a many-college setting, with the full, large support (parts (ii) and (iii) of Assumptions 3.5), we apply the same argument as in the one-college example to each college c sequentially, using extreme values of $r^c(y_{ic})$ or w_{ci} to "shut down" the effects of students' preference for college c or c's preference for students.

Part (iii) is a consequence of part (i) and the definition of cutoffs (equation (2)). If c reaches its capacity, T_c can be any value below δ_c and result in the same stable matching. We can identify T_c by imposing additional assumptions, for example, a full-capacity college having the same acceptability threshold as some college with vacancies.

3.3. *Practical Issues*

When our identification results are taken to the data, there can be practical issues. For example, vector z_i may include discrete variables such as gender, and the researcher may not have sufficient excluded variables. Below, we address these issues.

Discrete Random Variables. Our results can be extended to the case where z_i contains discrete variables. Suppose that $z_i = (z_i^1, z_i^2)$, where z_i^1 is a vector of discrete variables and z_i^2 is a vector of continuous variables. Let the support of z_i^1 be a finite set of points $\{z^{1,1}, z^{1,2}, \dots, z^{1,J}\}$. For $z_i^1 = z^{1,j}$, we define functions $(u^{c,j} + r^c, v^{c,j})$. Conditional on $z_i^1 = z^{1,1}$, we apply the results in Sections 3.1 and 3.2 to identify $\{u^{c,1} + r^c, v^{c,1}, \delta_c\}_c$ and F, requiring y_i and w_i being continuous, Assumptions 3.1(ii) and (iii), 3.2, 3.5(ii)–(iv), Condition 3.3, location normalization on $u^{c,1} + r^c$ and $v^{c,1}$ for each c, and (y_i, z_i^2) having full support. For $j \neq 1$, conditional on $z_i^1 = z^{1,j}$, using the conditional match probability of a mass of students for whom c is the only acceptable college, we can use F to identify $v^{c,j}$ for each c; similarly, using the conditional match probability of a mass of students whose feasible set is $\{0, c\}$, we can identify the function $u^{c,j} + r^c$.

Insufficient Excluded Variables. Allowing for college-level heterogeneity implies that we need to recover $3C$ preference parameters, $\{u^c, r^c, v^c\}_c$. As seen in Section 3.1, we need $2C$ excluded variables for identification of their derivatives.

The lack of sufficient excluded variables leads to a loss of identification, but not all is lost. We show that there is a trade-off between the identifiable degree of heterogeneity and the number of excluded variables. Consider $\mathbf{C}_1, \mathbf{C}_2 \subseteq \mathbf{C}$ with cardinality κ_1 and κ_2, respectively. For all $c \in \mathbf{C}_1$, there is a single excluded variable in student preferences, $y_{ic} = y_{i*}$; for all $c \in \mathbf{C}_2$, there is a single excluded variable in college preferences, $w_{ci} = w_{*i}$. We have the following identification result.

PROPOSITION 3.7: *Suppose the preference heterogeneity is reduced*: $u^c = u^*$ *and* $r^c = r^*$ *for all* $c \in \mathbf{C}_1$, *and* $v^c = v^*$ *for all* $c \in \mathbf{C}_2$. *For any* $c \in \mathbf{C}$, *the derivatives of* (u^c, r^c, v^c) *are identified if*: (i) *Assumptions* 2.1, 3.1, *and* 3.2 *hold*; *and* (ii) *the following rank conditions hold*: (a) *If* $C \leq \kappa_1 + \kappa_2 - 2$, *for any* z *in the interior of* $\mathcal{Z}$, *there exists a value of* w_i, $\widehat{w}$, *in* w_i*'s support conditional on* $(z_i, y_i) = (z, \overline{y})$ *such that* $(\Pi_y(z, \overline{y}, \widehat{w}), \Pi_w(z, \overline{y}, \widehat{w}))$ *has column rank at least* $2C - \kappa_1 - \kappa_2 + 2$; *or* (b) *if* $C > \kappa_1 + \kappa_2 - 2$, *for any* z *in the interior of* $\mathcal{Z}$, *there exist two values of* w_i, $\widehat{w}$ *and* $\widetilde{w}$, *in* w_i*'s support conditional on* $(z_i, y_i) = (z, \overline{y})$ *such that* $\Pi(z, \overline{y}, \widehat{w}, \widetilde{w})$ *has column rank at least* $2C - \kappa_1 - \kappa_2 + 2$.

In other words, we need only $C - \kappa_1 + 1$ excluded variables on the student side and $C - \kappa_2 + 1$ on the college side to identify the less heterogeneous model.[18] Importantly, this still allows for coefficient heterogeneity across colleges in $\mathbf{C}_1$ and $\mathbf{C}_2$. For example, it can be incorporated parametrically by interacting college-specific observables with z_i. For $c \in \mathbf{C}_1$, let $u^c(z_i) = \beta p_c z_i$, where p_c is a college-specific observable. This amounts to z_i having a college-specific parameter $\beta_c = \beta p_c$.

Since our identification approach is constructive, it directly implies a nonparameteric estimator. In the Monte Carlo simulations in Supplemental Appendix D, we apply this estimator in a semiparametric setting to avoid the well-known curse of dimensionality; yet the curse remains. Instead, we find that a parametric model based on a Bayesian approach works well. Below, we apply it to a real-life setting.

4. SECONDARY SCHOOL ADMISSIONS IN CHILE

Guided by our identification results, we study the admissions to public and private secondary schools (grades 9–12) in Chile in 2007. The market is organized similar to college admissions in the U.S. It is decentralized, both sides have preferences, and students do not submit rank-order lists of schools. We use a parametric Bayesian approach for preference estimation and then conduct counterfactual analysis.

4.1. *Institutional Background and Data*

Since 2003, secondary education has been compulsory for all Chileans up to 21 years of age. In principle, a public school must accept any student who is willing to enroll; a private school can be subsidized by the government or nonsubsidized, but in either case, it can select students based on its preferences.

We focus on a relatively independent market, *Market Valparaiso*, that includes three municipalities (Valparaiso, Viña del Mar, and Concon).[19] Our data includes the municipality of a student's residence and the geographical coordinates of each school, which identifies every agent in the market. A student is defined to be in Market Valparaiso if in 2008 she resided in a municipality within the boundary of the market. A secondary school is in Market Valparaiso if it is located in and admits students from the market. In total, there are 9314 students and 117 schools. This reasonably large size makes it plausible that the continuum market in Section 3 is a good approximation.

We use the SIMCE data set, provided by the Agency for the Quality of Education (Ministry of Education of Chile (n.d.)), on all 10th graders in 2008 to identify who started

[18]Note that the rank condition in Proposition 3.7 is weaker than Condition 3.3.

[19]In 2007, only 6.0% of 9th graders in Market Valparaiso schools resided outside of the market, while merely 1.6% of 9th graders residing in Market Valparaiso attended schools outside the area.

secondary school in 2007. The SIMCE is Chile's standardized testing program and tracks students' math and language performance. The data includes students' parental income, parental schooling, and other characteristics from a parental questionnaire sent home with students. Most school attributes are calculated from student characteristics of *the 10th graders in a school in 2006*, and thus are pre-determined in the 2007 admissions that we study. As tuition fees are largely fixed within each school and not completely flexible at the student level, we consider the problem as matching without transfers. See Supplemental Appendix E for more details on data construction and summary statistics of student characteristics and school attributes.

4.2. *Empirical Model*

We allow student preferences to be school-type-specific. For student i, the utility of attending school c of type $t \in \{\text{public, private nonsubsidized, private subsidized}\}$ is

$$u_{ict} = \alpha_{ft} \times \mathit{female}_i + \alpha_{mt} \times \mathit{male}_i + X'_{ic}\beta_t + \zeta_t \epsilon_{ic}, \tag{14}$$

where α_{ft} is a school-type fixed effect for female students; female_i is a dummy variable for female; α_{mt} and male_i are similarly defined; ϵ_{ic} is i.i.d. standard normal; ζ_t (> 0) allows type-specific variances; and X_{ic} are student-school-specific variables, including (i) the distance between i's residence and school c; (ii) 6 school attributes, and (iii) 3 interactions between school attributes and student characteristics.

Each student has an outside option, $u_{i0} = \epsilon_{i0}$, with ϵ_{i0} being standard normal. We impose the usual scale normalization through $\zeta_t = 1$ for public school,[20] while the location normalization is imposed by setting the deterministic part of u_{i0} to zero.

School preferences are also type-specific, but public schools do not have a utility function because they cannot select students. For private school c of type t (subsidized or nonsubsidized), its acceptability threshold is $T_c = 0$, and its utility function is

$$v_{cit} = \theta_t + Z'_{ci}\gamma_t + \eta_{ci}, \tag{15}$$

where θ_t is a type-specific intercept; η_{ci} is i.i.d. standard normal; and the vector Z_{ci} includes (i) 5 student characteristics and (ii) 3 interactions between student characteristics and school attributes.

In school preferences, the variance of η_{ci} being one is the scale normalization and $T_c = 0$ is the location normalization. Allowing the type-specific intercept θ_t and assuming $T_c = 0$ imply that schools of the same type have the same acceptability threshold. Because each type has some schools with vacancies, we can separately identify δ_c and T_c. Otherwise, we would lose its identification (Proposition 3.6).

The above specification uses distance as an (i, c)*-pair-specific* excluded variable in student preferences. In school preferences, math and language scores are *student-specific* excluded variables, implying that there are insufficient excluded variables to estimate school-specific utility functions. Using the results in Proposition 3.7, we limit preference heterogeneity and assume type-specific utility functions for schools. Our scale and location normalization is guided by Proposition 3.6 and Remark 2.2, while certain normalization is imposed via F.

[20]The variance of ϵ_{i0} is also normalized to be one because there is insufficient variation to estimate the variance of u_{i0}. Only 144 students (out of 9314) choose an outside option (see Table E.III).

4.3. *Estimation and Results*

We use a Bayesian approach with a Gibbs sampler for the estimation, which is first illustrated in Monte Carlo simulations (Supplemental Appendix D.3). We provide the details on the updating of the Markov Chain in Supplemental Appendices D.3 and F.

The estimation results are summarized in Table II. A caveat is in order when we interpret the results. Because we do not deal with endogeneity issues that may arise due to the correlation between preference shocks and school attributes or student characteristics, the estimates may not have a causal interpretation.

Panel A shows the estimates of student preferences. Most coefficients are of an expected sign. Interestingly, the coefficient on tuition is positive in the utility function for public schools, and parental income pushes it toward negative. This may reflect the fact that tuition at public schools is generally low (see Table E.IV) and may be correlated with unobserved school quality. There are also a few coefficients with an unexpected sign in school preferences. Panel B shows that nonsubsidized schools negatively value a student's parental income and mother's education. This may be because, on average, students at

TABLE II

ESTIMATION RESULTS: STUDENT AND SCHOOL PREFERENCES.

			Private Schools			
	Public Schools		Subsidized		Nonsubsidized	
	coef.	s.e.	coef.	s.e.	coef.	s.e.
Panel A. Student Preferences						
Female	−0.175	(0.836)	8.763	(0.581)	10.762	(5.937)
Male	−0.183	(0.835)	8.720	(0.576)	9.915	(5.941)
Distance	−0.175	(0.004)	−0.140	(0.008)	−0.673	(0.075)
log(tuition)	0.458	(0.095)	−0.446	(0.082)	−16.858	(1.680)
log(tuition) × log(income)	−0.019	(0.008)	0.047	(0.006)	0.958	(0.093)
log(median income)	−0.186	(0.072)	−0.524	(0.063)	2.704	(0.613)
Teacher experience	0.000	(0.001)	0.003	(0.001)	0.026	(0.017)
Fraction of female students	−0.111	(0.034)	−0.380	(0.052)	−1.818	(0.833)
Average composite score	−1.687	(0.123)	−2.451	(0.142)	−18.576	(2.753)
Average composite score × Composite score	5.930	(0.219)	5.663	(0.216)	19.642	(1.716)
Average mother's education	0.065	(0.022)	−0.269	(0.023)	−1.745	(0.337)
Average mother's education × Mother's education	0.003	(0.001)	0.008	(0.001)	0.056	(0.007)
Standard deviation of the utility shock	Normalized to 1		0.986	(0.062)	9.222	(0.920)
Panel B. School Preferences						
Constant			3.960	(0.744)	16.947	(3.014)
Female			−2.926	(0.193)	−1.264	(0.273)
Female × Fraction of female			6.134	(0.353)	1.718	(0.426)
Math score			−2.467	(0.473)	−6.447	(0.852)
Average math score × Math score			6.953	(0.909)	6.120	(0.908)
Language score			−1.026	(0.391)	−4.293	(0.813)
Average language score × Language score			3.105	(0.749)	4.655	(1.178)
Mother's education			−0.021	(0.014)	−0.101	(0.041)
log(income)			−0.302	(0.063)	−0.878	(0.190)

Note: This table presents the posterior mean and standard deviation of each coefficient in student and school utility functions (equations (14) and (15)). The Bayesian approach goes through a Markov Chain 1.75 million times, and the last 0.75 million iterations are used to calculate these statistics. Our Monte Carlo simulations suggest that the posterior standard deviation provides a useful estimate of the standard deviation of our estimators, albeit with a slight tendency to underestimate. See Table D.II for more details.

those schools often have a high income (above 1.4 million CLP; see Table E.III) and a highly educated mother (above 17 years).

Recall that each school's acceptability threshold is normalized to zero, so we can use the estimated school preferences to calculate if a student is acceptable to a private school. On average, a subsidized school finds 68% of the students acceptable, while a nonsubsidized school finds 82% acceptable. The higher acceptability rate at nonsubsidized schools does not imply that students are more often matched with them because their high tuition lowers their desirability to many students, especially those with a low parental income (see Table II).

Before conducting counterfactual analysis, we evaluate model fit in Supplemental Appendix F. It shows that our model fits the data reasonably well when we compare the observed matching with the one predicted based on our model.

4.4. *Counterfactual: Prioritizing Low-Income Students*

We consider a counterfactual policy in which students from low-income families are prioritized for admissions to all schools. A student is of low income if her parental income is among the lowest 40%.[21] Each private school's preferences over students are made lexicographical: low-income students are above others, and within each group of students, a school ranks them as in the current regime; all low-income students are acceptable, while others' acceptability is the same as in the current regime. We do not change anything in student or school preferences beyond the admission priority, which may mitigate the potential bias in our estimation due to aforementioned unaddressed endogeneity issues.

To simulate the counterfactual outcome, we keep 1500 draws of the utilities in the Markov Chain in the Bayesian estimation.[22] We run the Gale–Shapley deferred acceptance with each draw and obtain 1500 sets of counterfactual stable matchings. We then report the average of these counterfactual matchings; recall that we only have one matching outcome under the current regime—the observed one.

Table III presents the results. There are several noticeable patterns when we move from the current regime to the counterfactual. First, low-income students are in schools that have higher-ability and higher-income students in the same cohort, while the opposite is true for nonlow-income students. Second, some low-income students leave public schools for private subsidized schools, while crowding out some other students to public schools. Lastly, the policy benefits low-income students and hurts others. On average, low-income students' welfare gain is equivalent to decreasing travel distance (to a public school) by 0.433 km. This gain is concentrated among 16.4% of the low-income students, while others are not affected. Correspondingly, 9.1% of the nonlow-income students are worse off, and none is better off.

These results indicate that low-income students dislike private nonsubsidized schools. We explore why it is the case. With the 1500 draws of student preferences, we examine low-income students' favorite school of each type. We find that low-income students on average value their favorite public school at 2.25 and private subsidized school at 2.18, while their favorite nonsubsidized school is only valued at −23.39, in general unaccept-

[21]This resembles a policy adopted in 2008 in Chile as documented by Gazmuri (2017). It benefits 44% of elementary school students in 2012 in terms of admission priorities and a tuition waiver. Supplemental Appendix Table F.V shows summary statistics of the students by income status.

[22]Specifically, there are 15 blocks of 100 draws. The blocks are equally spaced in the 0.75 million iterations in the Markov chain that are used to calculate the posterior means and standard deviations.

TABLE III

SORTING AND STUDENT WELFARE IN THE CURRENT AND COUNTERFACTUAL REGIMES.

	Low-Income Students		Nonlow-Income Students	
	Current (1)	Counterfactual (2)	Current (3)	Counterfactual (4)
Average composite score (same cohort) at matched school	0.374	0.386	0.586	0.577
Average parental income (same cohort) at matched school	216,456	224,098	595,031	589,388
Fraction enrolled at each school type:				
Public	0.672	0.568	0.230	0.266
Private subsidized	0.310	0.417	0.528	0.490
Private nonsubsidized	0.002	0.002	0.226	0.228
Outside option	0.016	0.013	0.015	0.016
Welfare effects of moving from current to counterfactual:				
Average utility change (reduction in distance, km)	0.433		−0.267	
Winners (fraction)	0.164		0.000	
Losers (fraction)	0.000		0.091	
Indifferent (fraction)	0.836		0.909	

Note: The outcome in the current regime is the one observed in the data. To simulate the counterfactual outcome, we keep 1500 draws of the utilities in the Markov Chain in the Bayesian estimation and obtain 1500 sets of stable matching outcomes. The statistics for the counterfactual regime are averages across the 1500 outcomes. The average utility change is measured in terms of willingness to travel to a public school in kilometers.

able to them. We calculate the contribution of different variables to these differences by shutting down their effect in the utility functions. The results show that high tuition at nonsubsidized schools is the main contributor. It is worth noting that tuition could be correlated with unobserved school quality. Hence, low-income students may dislike a nonsubsidized school because of its high costs and/or their tastes. Additionally, mother's education, both a student's own and a school's average, is also an important factor. Student ability, measured by their composite score, and distance to each school do not appear to be as important.

In sum, giving low-income students access to schools fails to significantly change matching outcomes due to their own preferences. Low-income students are deterred from private nonsubsidized schools by their high tuition. These findings are in line with the preference heterogeneity documented in public school choice (see, e.g., Abdulkadiroğlu, Agarwal, and Pathak (2017)), although tuition plays no role there.

5. CONCLUDING REMARKS

We study nonparametric identification of agent preferences in many-to-one two-sided matching without transfers. We derive a set of sufficient conditions for identification and provide guidance for empirical studies. For example, our results clarify the data requirement for the identification of various degrees of preference heterogeneity. To take our results to the data of a reasonably-sized market, we propose a Bayesian approach with a Gibbs sampler whose performance is illustrated in Monte Carlo simulations. Our model encompasses many real-life matching markets, such as college admissions and school choice in many countries, in which our identification results and empirical method can be applied. Hence, this paper opens a new avenue for empirical research.

We illustrate our method in the context of secondary school admissions in Chile. As an example of the usefulness of the estimates, we consider a counterfactual policy in which students from low-income families are prioritized for admissions to all schools. Although the policy benefits low-income students, its effects are small. Such insights are difficult to obtain without estimating the preferences of both sides. In this sense, our method can help provide an ex ante evaluation of a range of alternative policies.

APPENDIX

Throughout this Appendix, for notational simplicity, we let $x_i = (z_i, y_i, w_i)$.

PROOF OF PROPOSITION 3.4: Equation (5) can be rewritten as

$$\sigma_c(x_i) = \Lambda_c(\tau_{i1}, \dots, \tau_{iC}; \iota_{1i}, \dots, \iota_{Ci}) = \Lambda_c(\tau_i, \iota_i),$$

where $\tau_i = (\tau_{i1}, \dots, \tau_{iC})$, $\iota_i = (\iota_{1i}, \dots, \iota_{Ci})$, and Λ_c denotes some unknown function. Recall that $\tau_{ic} = u^c(z_i) + r^c(y_{ic})$ and $\iota_{ci} = v^c(z_i) + w_{ci}$. Under Assumption 3.1, Λ_c, u^c, r^c, and v^c are continuously differentiable and the observables are all continuously distributed.

For colleges $c, d \in \mathbf{C}$, taking derivatives of $\sigma_d(x_i)$ w.r.t. y_{ic}, w_{ci}, and z_i^k (the kth component of z_i), respectively, one obtains

$$\frac{\partial \sigma_d(x_i)}{\partial y_{ic}} = \frac{\partial \Lambda_d(\tau_i, \iota_i)}{\partial \tau_{ic}} \frac{\partial r^c(y_{ic})}{\partial y_{ic}}, \qquad \frac{\partial \sigma_d(x_i)}{\partial w_{ci}} = \frac{\partial \Lambda_d(\tau_i, \iota_i)}{\partial \iota_{ci}}, \tag{16}$$

$$\frac{\partial \sigma_d(x_i)}{\partial z_i^k} = \sum_{c \in \mathbf{C}} \frac{\partial \Lambda_d(\tau_i, \iota_i)}{\partial \tau_{ic}} \frac{\partial u^c(z_i)}{\partial z_i^k} + \sum_{c \in \mathbf{C}} \frac{\partial \Lambda_d(\tau_i, \iota_i)}{\partial \iota_{ci}} \frac{\partial v^c(z_i)}{\partial z_i^k}. \tag{17}$$

First, we show the identification of the derivatives of the functions u^c and v^c. We fix $y_i = \overline{y}$ and consider (z, w) in the interior of $\mathcal{Z} \times \mathcal{W}$. Recall that $\frac{\partial r^c(\overline{y}_c)}{\partial y_{ic}} = 1$ due to the scale normalization. Substituting equation (16) into equation (17), we get

$$\frac{\partial \sigma_d(z, \overline{y}, w)}{\partial z_i^k} = \sum_{c \in \mathbf{C}} \frac{\partial \sigma_d(z, \overline{y}, w)}{\partial y_{ic}} \frac{\partial u^c(z)}{\partial z_i^k} + \sum_{c \in \mathbf{C}} \frac{\partial \sigma_d(z, \overline{y}, w)}{\partial w_{ci}} \frac{\partial v^c(z)}{\partial z_i^k}. \tag{18}$$

Suppose that two different values of the C-dimensional vector of excluded regressors w_i, $\widehat{w}$ and $\widetilde{w}$, satisfy Condition 3.3. We define $\widehat{x} = (z, \overline{y}, \widehat{w})$ and $\widetilde{x} = (z, \overline{y}, \widetilde{w})$. Further, let $\widehat{\sigma}_c = \mathbb{P}(\mu(i) = c|\widehat{x})$ and $\widetilde{\sigma}_c = \mathbb{P}(\mu(i) = c|\widetilde{x})$. By evaluating equation (18) at $d = 1, \dots, C$

and $x_i = \widehat{x}, \widetilde{x}$, and stacking them together, we have

$$\underbrace{\begin{pmatrix} \frac{\partial \widehat{\sigma}_1}{\partial z_i^k} \\ \vdots \\ \frac{\partial \widehat{\sigma}_C}{\partial z_i^k} \\ \frac{\partial \widetilde{\sigma}_1}{\partial z_i^k} \\ \vdots \\ \frac{\partial \widetilde{\sigma}_C}{\partial z_i^k} \end{pmatrix}}_{\Sigma_{z_i^k}(z,\overline{y},\widehat{w},\widetilde{w})} = \underbrace{\begin{pmatrix} \frac{\partial \widehat{\sigma}_1}{\partial y_{i1}} & \cdots & \frac{\partial \widehat{\sigma}_1}{\partial y_{iC}} & \frac{\partial \widehat{\sigma}_1}{\partial w_{1i}} & \cdots & \frac{\partial \widehat{\sigma}_1}{\partial w_{Ci}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \widehat{\sigma}_C}{\partial y_{i1}} & \cdots & \frac{\partial \widehat{\sigma}_C}{\partial y_{iC}} & \frac{\partial \widehat{\sigma}_C}{\partial w_{1i}} & \cdots & \frac{\partial \widehat{\sigma}_C}{\partial w_{Ci}} \\ \frac{\partial \widetilde{\sigma}_1}{\partial y_{i1}} & \cdots & \frac{\partial \widetilde{\sigma}_1}{\partial y_{iC}} & \frac{\partial \widetilde{\sigma}_1}{\partial w_{1i}} & \cdots & \frac{\partial \widetilde{\sigma}_1}{\partial w_{Ci}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \widetilde{\sigma}_C}{\partial y_{i1}} & \cdots & \frac{\partial \widetilde{\sigma}_C}{\partial y_{iC}} & \frac{\partial \widetilde{\sigma}_C}{\partial w_{1i}} & \cdots & \frac{\partial \widetilde{\sigma}_C}{\partial w_{Ci}} \end{pmatrix}}_{\Pi(z,\overline{y},\widehat{w},\widetilde{w})} \times \begin{pmatrix} \frac{\partial u^1(z)}{\partial z_i^k} \\ \vdots \\ \frac{\partial u^C(z)}{\partial z_i^k} \\ \frac{\partial v^1(z)}{\partial z_i^k} \\ \vdots \\ \frac{\partial v^C(z)}{\partial z_i^k} \end{pmatrix}. \tag{19}$$

The first C rows of the matrix $\Pi(z, \overline{y}, \widehat{w}, \widetilde{w})$ contain derivatives of the conditional match probabilities w.r.t. the excluded regressors, evaluated at $x_i = \widehat{x}$. The second C rows of $\Pi(z, \overline{y}, \widehat{w}, \widetilde{w})$ are constructed similarly, evaluating the derivatives at $x_i = \widetilde{x}$.

Under Condition 3.3, there always exist $\widehat{w}$ and $\widetilde{w}$ such that $\Pi(z, \overline{y}, \widehat{w}, \widetilde{w})$ in equation (19) is invertible. We then identify the derivatives of u^c and v^c for all c by solving the system of linear equations. Formally, let $\Pi^c_{z_i^k}(z, \overline{y}, \widehat{w}, \widetilde{w})$ be the matrix formed by replacing the c^{th} column of matrix $\Pi(z, \overline{y}, \widehat{w}, \widetilde{w})$ by the vector $\Sigma_{z_i^k}(z, \overline{y}, \widehat{w}, \widetilde{w})$ (defined in equation (19)). By the Cramer's rule, for any $c \in \mathbf{C}$,

$$\frac{\partial u^c(z)}{\partial z_i^k} = \frac{\left|\Pi^c_{z_i^k}(z, \overline{y}, \widehat{w}, \widetilde{w})\right|}{\left|\Pi(z, \overline{y}, \widehat{w}, \widetilde{w})\right|}, \qquad \frac{\partial v^c(z)}{\partial z_i^k} = \frac{\left|\Pi^{c+C}_{z_i^k}(z, \overline{y}, \widehat{w}, \widetilde{w})\right|}{\left|\Pi(z, \overline{y}, \widehat{w}, \widetilde{w})\right|}.$$

So far, we have only considered z_i that shows up in the utility functions for all colleges and for both sides. For any element of z_i that is excluded from certain utility functions, the identification is a special case of the above proof by noting that some derivatives of the utility functions are zero.

Second, we identify the derivative of the function r^c for all c. We start with r^1 and fix $y_{ic} = \overline{y}_c$ for $c \in \mathbf{C}\backslash\{1\}$; under the scale normalization, $\frac{\partial r^c(\overline{y}_c)}{\partial y_{ic}} = 1$ for $c \in \mathbf{C}\backslash\{1\}$. For any (y_1, z, w) in the interior of $\mathcal{Y}_1 \times \mathcal{Z} \times \mathcal{W}$, substituting equation (16) into equation (17), for each $d \in \mathbf{C}$, we obtain

$$\left(\frac{\partial \sigma_d}{\partial z_i^k} - \sum_{c\in\mathbf{C}} \frac{\partial \sigma_d}{\partial w_{ci}} \frac{\partial v^c}{\partial z_i^k} - \sum_{c\in\mathbf{C}\backslash\{1\}} \frac{\partial \sigma_d}{\partial y_{ic}} \frac{\partial u^c}{\partial z_i^k}\right) \frac{\partial r^1}{\partial y_{i1}} = \frac{\partial \sigma_d}{\partial y_{i1}} \frac{\partial u^1}{\partial z_i^k}, \tag{20}$$

where all the terms except for $\frac{\partial r^1(y_1)}{\partial y_{i1}}$ are known or already identified. If $\frac{\partial \sigma_d}{\partial z_i^k} - \sum_{c\in\mathbf{C}} \frac{\partial \sigma_d}{\partial w_{ci}} \frac{\partial v^c}{\partial z_i^k} - \sum_{c\in\mathbf{C}\backslash\{1\}} \frac{\partial \sigma_d}{\partial y_{ic}} \frac{\partial u^c}{\partial z_i^k} \neq 0$ for certain (z, w) and some $d \in \mathbf{C}$, $\frac{\partial r^1(y_1)}{\partial y_{i1}}$ is identified from equation (20); otherwise, equation (20) implies that $\frac{\partial \sigma_d}{\partial y_{i1}} = 0$ for all values (z, w) and all $d \in \mathbf{C}$, and thus $\frac{\partial r^1(y_1)}{\partial y_{i1}}$ is also identified and equal to zero. The derivative of the function r^c for $c \in \mathbf{C} \setminus \{1\}$ can be identified in the same manner. *Q.E.D.*

PROOF OF PROPOSITION 3.6: We start with the identification of the cutoffs $\{\delta_c\}_c$, that is, part (i). Because of the large support assumption on r^c (parts (ii) and (iii) of Assumption 3.5) for each $c \in \mathbf{C}$, there exists $\mathcal{J}_c \subseteq \mathcal{X}$ such that for any $x_i \in \mathcal{J}_c$, c is the only acceptable college with probability one, and that $Q(i \in \mathbf{I} : x_i \in \mathcal{J}_c) > 0$. Then

$$\mathbb{P}\big(\mu(i) = 0 | x_i \in \mathcal{J}_c\big) = \mathbb{P}\big(v^c(z_i) + w_{ci} + \eta_{ci} < \delta_c | x_i \in \mathcal{J}_c\big) = F_{\eta_{ci}}(\delta_c - \iota_{ci}),$$

where the last equality is due to $\iota_{ci} = v^c(z_i) + w_{ci}$ and the independence between η_{ci} and x_i (Assumption 3.2). There exists a unique ι_c^* such that $\delta_c - \iota_c^* = \text{Quantile}_{\eta_{ci}}(\rho_c) = \inf\{(\delta_c - \iota_{ci}) : F_{\eta_{ci}}(\delta_c - \iota_{ci}) \geq \rho_c\}$. By Assumption 3.5(iv), $\delta_c - \iota_c^* = 0$, which identifies δ_c. To show part (iii), we use the definition of cutoffs, equation (2).

To prove part (ii), the identification of the distribution F, following equation (5) for $c = 0$, the conditional probability of being unmatched can be rewritten as

$$\mathbb{P}\big(\mu(i) = 0 | x_i\big) = \Lambda_0(\tau_{i1}, \dots, \tau_{iC}; \iota_{1i}, \dots, \iota_{Ci}) = \Lambda_0(\tau_i, \iota_i). \tag{21}$$

Given that the functions $\{u^c + r^c, v^c\}_c$ are identified (part (i) of Assumption 3.5), the arguments in equation (21), $\{\tau_{ic}, \iota_{ci}\}_c$, are known. Because $\mathbb{P}(\mu(i) = 0 | x_i)$ is observed from the population data, the function Λ_0 is identified.

For each $c \in \mathbf{C}$, define $A_{ic} \equiv \{u_{ic} < 0\} = \{\tau_{ic} + \epsilon_{ic} < 0\} = \{\epsilon_{ic} < -\tau_{ic}\}$, and $B_{ci} \equiv \{v_{ci} < \delta_c\} = \{\iota_{ci} + \eta_{ci} < \delta_c\} = \{\eta_{ci} < \delta_c - \iota_{ci}\}$. By the independence between (ϵ_i, η_i) and x_i (Assumption 3.2), the parameter of interest can be written as

$$F(-\tau_{i1}, \dots, -\tau_{iC}, \delta_1 - \iota_{1i}, \dots, \delta_C - \iota_{Ci}) = \mathbb{P}\big(\cap_{c=1}^C (A_{ic} \cap B_{ci}) | x_i\big).$$

Further, the conditional probability of being unmatched is

$$\begin{aligned}
\Lambda_0(\tau_i, \iota_i) &= \mathbb{P}\left(\bigcap_{c=1}^{C} (A_{ic} \cup B_{ci}) | x_i\right) \\
&= \mathbb{P}\left(\bigcap_{c=1}^{C-1} (A_{ic} \cup B_{ci}) \cap (A_{iC} \cup B_{Ci}) | x_i\right) \\
&= \mathbb{P}\left(\left(\bigcap_{c=1}^{C-1} (A_{ic} \cup B_{ci}) \cap A_{iC}\right) \cup \left(\bigcap_{c=1}^{C-1} (A_{ic} \cup B_{ci}) \cap B_{Ci}\right) \Big| x_i\right) \\
&= \mathbb{P}\left(\bigcap_{c=1}^{C-1} (A_{ic} \cup B_{ci}) \cap A_{iC} | x_i\right) + \mathbb{P}\left(\bigcap_{c=1}^{C-1} (A_{ic} \cup B_{ci}) \cap B_{Ci} | x_i\right) \\
&\quad - \mathbb{P}\left(\left(\bigcap_{c=1}^{C-1} (A_{ic} \cup B_{ci}) \cap A_{iC} \cap B_{Ci}\right) \Big| x_i\right) \\
&= \Lambda_0(\tau_{i1}, \dots, \tau_{iC}; \iota_{1i}, \dots, \iota_{(C-1)i}, \infty) + \Lambda_0(\tau_{i1}, \dots, \tau_{i(C-1)}, \infty; \iota_{1i}, \dots, \iota_{Ci}) \\
&\quad - \mathbb{P}\left(\left(\bigcap_{c=1}^{C-1} (A_{ic} \cup B_{ci}) \cap A_{iC} \cap B_{Ci}\right) \Big| x_i\right).
\end{aligned} \tag{22}$$

Let $H_C(\tau_i, \iota_i) \equiv \mathbb{P}((\bigcap_{c=1}^{C-1}(A_{ic} \cup B_{ci}) \cap A_{iC} \cap B_{Ci})|x_i)$. It is identified from equation (22) because Λ_0 is identified. Moreover, similar derivations yield

$$
\begin{aligned}
& H_C(\tau_i, \iota_i) \\
&\quad = \mathbb{P}\left(\bigcap_{c=1}^{C-2}(A_{ic} \cup B_{ci}) \cap A_{i(C-1)} \cap A_{iC} \cap B_{Ci}|x_i\right) \\
&\qquad + \mathbb{P}\left(\bigcap_{c=1}^{C-2}(A_{ic} \cup B_{ci}) \cap B_{(C-1)i} \cap A_{iC} \cap B_{Ci}|x_i\right) \\
&\qquad - \mathbb{P}\left(\left(\bigcap_{c=1}^{C-2}(A_{ic} \cup B_{ci}) \cap A_{i(C-1)} \cap B_{(C-1)i} \cap A_{iC} \cap B_{Ci}\right)\Big|x_i\right) \\
&\quad = H_C(\tau_{i1}, \dots, \tau_{iC}; \iota_{1i}, \dots \iota_{(C-2)i}, \infty, \iota_{Ci}) + H_C(\tau_{i1}, \dots, \tau_{i(C-2)}, \infty, \tau_{iC}; \iota_{1i}, \dots, \iota_{Ci}) \\
&\qquad - \mathbb{P}\left(\left(\bigcap_{c=1}^{C-2}(A_{ic} \cup B_{ci}) \cap A_{i(C-1)} \cap B_{(C-1)i} \cap A_{iC} \cap B_{Ci}\right)\Big|x_i\right),
\end{aligned} \tag{23}
$$

where in the penultimate line, the effects of $\iota_{(C-1)i}$ and $\tau_{i(C-1)}$ are "shut down" in the two terms, respectively. Equation (23) then identifies $H_{C-1}(\tau_i, \iota_i) \equiv \mathbb{P}((\bigcap_{c=1}^{C-2}(A_{ic} \cup B_{ci}) \cap A_{i(C-1)} \cap B_{(C-1)i} \cap A_{iC} \cap B_{Ci})|x_i)$.

Repeat the above argument and define a sequence of functions recursively until $H_2(\tau_i, \iota_i) = H_2(\tau_{i1}, \dots, \tau_{iC}; \infty, \iota_{2i}, \dots, \iota_{Ci}) + H_2(\infty, \tau_{i2}, \dots, \tau_{iC}; \iota_{1i}, \dots, \iota_{Ci}) - H_1(\tau_i, \iota_i)$, where on the RHS, the effects of ι_{1i} and τ_{i1} are "shut down" in the first two terms, respectively. Every function in the sequence is identified.

It then follows that $F(-\tau_{i1}, \dots, -\tau_{iC}, \delta_1 - \iota_{1i}, \dots, \delta_C - \iota_{Ci}) = \mathbb{P}(\bigcap_{c=1}^{C}(A_{ic} \cap B_{ci})|x_i) \equiv H_1(\tau_i, \iota_i)$ is identified. *Q.E.D.*

PROOF OF PROPOSITION 3.7: We use the same argument as in the proof of Proposition 3.4, except that the matrix in equation (19) reduces in dimension due to the additional homogeneity restrictions. Specifically, with some abuse of notation, suppose that vector $\boldsymbol{u}$ consists of u^* and u^c $\forall c \in \mathbf{C} \setminus \mathbf{C}_1$, and that $\boldsymbol{v}$ consist of v^* and v^c $\forall c \in \mathbf{C} \setminus \mathbf{C}_2$. The vectors y_i and w_i are defined similarly, and suppose that $(\Pi_y(z, \overline{y}, \widehat{w}), \Pi_w(z, \overline{y}, \widehat{w}))$ has eliminated the duplicated elements accordingly. Fix $(y_i, w_i) = (\overline{y}, \widehat{w})$ and consider z in the interior of $\mathcal{Z}$. We can rewrite equation (18) as

$$
\frac{\partial \sigma_d}{\partial z_i^k} = \frac{\partial \sigma_d}{\partial y_{i*}} \frac{\partial u^*}{\partial z_i^k} + \sum_{c \in \mathbf{C} \setminus \mathbf{C}_1} \frac{\partial \sigma_d}{\partial y_{ic}} \frac{\partial u^c}{\partial z_i^k} + \frac{\partial \sigma_d}{\partial w_{*i}} \frac{\partial v^*}{\partial z_i^k} + \sum_{c \in \mathbf{C} \setminus \mathbf{C}_2} \frac{\partial \sigma_d}{\partial w_{ci}} \frac{\partial v^c}{\partial z_i^k}. \tag{24}
$$

Stacking equation (24) for all $d \in \mathbf{C}$, we obtain

$$
\frac{\partial \boldsymbol{\sigma}(z, \overline{y}, \widehat{w})}{\partial z_i^k} = \left(\Pi_y(z, \overline{y}, \widehat{w}), \Pi_w(z, \overline{y}, \widehat{w})\right) \times \begin{pmatrix} \dfrac{\partial \boldsymbol{u}(z)}{\partial z_i^k} \\ \dfrac{\partial \mathbf{v}(z)}{\partial z_i^k} \end{pmatrix}. \tag{25}
$$

When the number of parameters $2C - \kappa_1 - \kappa_2 + 2$ is at most C, equation (25) implies identification; otherwise, we can identify the parameters by considering a pair of distinct values of w_i, as in the proof of Proposition 3.4. *Q.E.D.*

REFERENCES

ABDULKADIROĞLU, ATILA, AND TAYFUN SÖNMEZ (2003): "School Choice: A Mechanism Design Approach," *American Economic Review*, 93 (3), 729–747. [753]

ABDULKADIROĞLU, ATILA, NIKHIL AGARWAL, AND PARAG A. PATHAK (2017): "The Welfare Effects of Coordinated Assignment: Evidence From the New York City High School Match," *American Economic Review*, 107 (12), 3635–3689. [751,753,768]

AGARWAL, NIKHIL (2015): "An Empirical Model of the Medical Match," *American Economic Review*, 105 (7), 1939–1978. [749,752]

AGARWAL, NIKHIL, AND PAULO SOMAINI (2018): "Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism," *Econometrica*, 86 (2), 391–444. [753]

——— (2022): "Demand Analysis Under Latent Choice Constraints," NBER Working Paper No. w29993. [753,760]

——— (2023): "Empirical Models of Non-Transferable Utility Matching," *Online and Matching-Based Market Design*, 530–551, Cambridge University Press. [752]

ARTEMOV, GEORGY, YEON-KOO CHE, AND YINGHUA HE (2023): "Stable Matching With Mistaken Agents," *Journal of Political Economy Microeconomics*, 1 (2), 270–320. [750,754,755]

AUE, ROBERT, THILO KLEIN, AND JOSUÉ ORTEGA (2020): "What Happens When Separate and Unequal School Districts Merge?" ZEW Discussion Paper, (20-032). [750,752]

AZEVEDO, EDUARDO M., AND JACOB D. LESHNO (2016): "A Supply and Demand Framework for Two-Sided Matching Markets," *Journal of Political Economy*, 124 (5), 1235–1268. [750,754,755]

BERRY, STEVEN, AMIT GANDHI, AND PHILIP HAILE (2013): "Connected Substitutes and Invertibility of Demand," *Econometrica*, 81 (5), 2087–2111. [760]

CALSAMIGLIA, CATERINA, CHAO FU, AND MAIA GÜELL (2020): "Structural Estimation of a Model of School Choices: The Boston Mechanism versus Its Alternatives," *Journal of Political Economy*, 128 (2), 642–680. [753]

CHEN, QIHUI, AND ZHENG FANG (2019): "Improved Inference on the Rank of a Matrix," *Quantitative Economics*, 10 (4), 1787–1824. [760]

CHIAPPORI, PIERRE-ANDRÉ, AND BERNARD SALANIÉ (2016): "The Econometrics of Matching Models," *Journal of Economic Literature*, 54 (3), 832–861. [750,752]

CHOO, EUGENE, AND ALOYSIUS SIOW (2006): "Who Marries Whom and Why," *Journal of Political Economy*, 114 (1), 175–201. [752]

CRAWFORD, GREGORY S., RACHEL GRIFFITH, AND ALESSANDRO IARIA (2021): "A Survey of Preference Estimation With Unobserved Choice Set Heterogeneity," *Journal of Econometrics*, 222 (1), 4–43. [753]

DAGSVIK, JOHN K. (2000): "Aggregation in Matching Markets," *International Economic Review*, 41 (1), 27–58. [752]

DIAMOND, WILLIAM, AND NIKHIL AGARWAL (2017): "Latent Indices in Assortative Matching Models," *Quantitative Economics*, 8 (3), 685–728. [752]

EDERER, TIM (2022): "Two-Sided Matching Without Transfers: A Unifying Empirical Framework," Working Paper, Toulouse School of Economics. [752]

FACK, GABRIELLE, JULIEN GRENET, AND YINGHUA HE (2019): "Beyond Truth-Telling: Preference Estimation With Centralized School Choice and College Admissions," *American Economic Review*, 109 (4), 1486–1529. [750,753-756]

FOX, JEREMY T. (2010): "Identification in Matching Games," *Quantitative Economics*, 1 (2), 203–254. [752]

FOX, JEREMY T., CHENYU YANG, AND DAVID H. HSU (2018): "Unobserved Heterogeneity in Matching Games," *Journal of Political Economy*. [752]

GALE, DAVID, AND LLOYD SHAPLEY (1962): "College Admissions and the Stability of Marriage," *The American Mathematical Monthly*, 69 (1), 9–15. [755]

GALICHON, ALFRED, AND BERNARD SALANIE (2022): "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models," *The Review of Economic Studies*, 89 (5), 2600–2629. [752]

GALICHON, ALFRED, SCOTT DUKE KOMINERS, AND SIMON WEBER (2019): "Costly Concessions: An Empirical Framework for Matching With Imperfectly Transferable Utility," *Journal of Political Economy*, 127 (6), 2875–2925. [752]

GAZMURI, ANA (2017): "School Segregation in the Presence of Student Sorting and Cream-Skimming," Working paper, Toulouse School of Economics. [749,767]

GRENET, JULIEN, YINGHUA HE, AND DOROTHEA KÜBLER (2022): "Preference Discovery in University Admissions: The Case for Dynamic Multioffer Mechanisms," *Journal of Political Economy*, 130 (6), 1427–1476. [755]

GUALDANI, CRISTINA, AND SHRUTI SINHA (2023): "Partial Identification in Matching Models for the Marriage Market," *Journal of Political Economy*, 131 (5), 1109–1171. [752]

HE, YINGHUA (2017): "Gaming the Boston School Choice Mechanism in Beijing," Working Paper, Rice University and Toulouse School of Economics. [753]

HE, YINGHUA, AND THIERRY MAGNAC (2022): "Application Costs and Congestion in Matching Markets," *The Economic Journal*, 132 (648), 2918–2950. [750,752]

HE, YINGHUA, SHRUTI SINHA, AND XIAOTING SUN (2024): "Supplement to 'Identification and Estimation in Many-to-One Two-Sided Matching Without Transfers'," *Econometrica Supplemental Material*, 92, https://doi.org/10.3982/ECTA19636. [753]

LOGAN, JOHN ALLEN, PETER D. HOFF, AND MICHAEL A. NEWTON (2008): "Two-Sided Estimation of Mate Preferences for Similarities in Age, Education, and Religion," *Journal of the American Statistical Association*, 103 (482), 559–569. [751]

MATZKIN, ROSA L. (2019): "Constructive Identification in Some Nonseparable Discrete Choice Models," *Journal of Econometrics*, 211 (1), 83–103. [759]

MENZEL, KONRAD (2015): "Large Matching Markets as Two-Sided Demand Systems," *Econometrica*, 83 (3), 897–941. [752]

——— (2022): "Strategic Network Formation With Many Agents," Working Paper, NYU. [752,754]

MINISTRY OF EDUCATION OF CHILE (n.d.): "Base de Datos de la Agencia de Calidad de la Educación," Santiago, Chile. [764]

PAIS, JOANA, ÁGNES PINTÉR, AND RÓBERT F. VESZTEG (2020): "Decentralized Matching Markets With(out) Frictions: A Laboratory Experiment," *Experimental Economics*, 23 (1), 212–239. [755]

POWELL, JAMES L., JAMES H. STOCK, AND THOMAS M. STOKER (1989): "Semiparametric Estimation of Index Coefficients," *Econometrica*, 57 (6), 1403–1430. [751]

ROSSI, PETER E., GREG M. ALLENBY, AND ROB MCCULLOCH (2012): *Bayesian Statistics and Marketing*. John Wiley & Sons. [751]

ROTH, ALVIN E. (1984): "The Evolution of the Labor Market for Medical Interns and Residents: A Case Study in Game Theory," *Journal of political Economy*, 92 (6), 991–1016. [749]

ROTH, ALVIN E., AND MARILDA SOTOMAYOR (1992): "Two-Sided Matching," *Handbook of game theory with economic applications*, 1, 485–541. [750,754]

SINHA, SHRUTI (2015): "Identification and Estimation in One-to-One Matching Models With Nonparametric Unobservables," Working Paper, Toulouse School of Economics. [752]

Co-editor Guido W. Imbens handled this manuscript.

Manuscript received 4 April, 2021; final version accepted 21 February, 2024; available online 23 February, 2024.

The replication package for this paper is available at https://doi.org/10.5281/zenodo.10667328. The authors were granted an exemption to publish parts of their data because either access to these data is restricted or the authors do not have the right to republish them. However, the authors included in the package, on top of the codes and the parts of the data that are not subject to the exemption, a simulated or synthetic dataset that allows running the codes. The Journal checked the data and the codes for their ability to generate all tables and figures in the paper and approved online appendices. Whenever the available data allowed, the Journal also checked for their ability to reproduce the results. However, the synthetic/simulated data are not designed to produce the same results. Given the highly demanding nature of the algorithms, the reproducibility checks were run on a simplified version of the code, which is also available in the replication package.