Taylor & Francis
Taylor & Francis Group

Journal of Business & Economic Statistics

ISSN: 0735-0015 (Print) 1537-2707 (Online) Journal homepage: https://www.tandfonline.com/loi/ubes20

Identification and Efficiency Bounds for the Average Match Function Under Conditionally Exogenous Matching

Bryan S. Graham, Guido W. Imbens & Geert Ridder

To cite this article: Bryan S. Graham, Guido W. Imbens & Geert Ridder (2020) Identification and Efficiency Bounds for the Average Match Function Under Conditionally Exogenous Matching, Journal of Business & Economic Statistics, 38:2, 303-316, DOI: 10.1080/07350015.2018.1497509

To link to this article: https://doi.org/10.1080/07350015.2018.1497509

Accepted author version posted online: 19 Jul 2018.
Published online: 23 Oct 2018.

Identification and Efficiency Bounds for the Average Match Function Under Conditionally Exogenous Matching

Bryan S. Graham
Department of Economics, University of California - Berkeley, Berkeley, CA 94720-3888, and National Bureau of Economic Research, Cambridge, MA 02138 (*bgraham@econ.berkeley.edu*)

Guido W. Imbens
Graduate School of Business, Stanford University, Stanford, CA 94305-5015, and National Bureau of Economic Research, Cambridge, MA 02138 (*imbens@stanford.edu*)

Geert Ridder
Department of Economics and USC INET, University of Southern California, Los Angeles, CA 90089 (*ridder@usc.edu*)

Consider two heterogenous populations of agents who, when matched, jointly produce an output, Y. For example, teachers and classrooms of students together produce achievement, parents raise children, whose life outcomes vary in adulthood, assembly plant managers and workers produce a certain number of cars per month, and lieutenants and their platoons vary in unit effectiveness. Let $W \in \mathbb{W} = \{w_1, \ldots, w_J\}$ and $X \in \mathbb{X} = \{x_1, \ldots, x_K\}$ denote agent types in the two populations. Consider the following matching mechanism: take a random draw from the $W = w_j$ subgroup of the first population and match her with an independent random draw from the $X = x_k$ subgroup of the second population. Let $\beta(w_j, x_k)$, the *average match function* (AMF), denote the expected output associated with this match. We show that (i) the AMF is identified when matching is conditionally exogenous, (ii) conditionally exogenous matching is compatible with a pairwise stable aggregate matching equilibrium under specific informational assumptions, and (iii) we calculate the AMF's semiparametric efficiency bound.

KEY WORDS: Average match function; Causal inference; Choo–Siow model; Matching; Semiparametric efficiency.

There are two populations, say teachers and classrooms of students. (We maintain this running example for much of what follows for expository reasons, but our results are not restricted to this case. See Boyd et al. (2013) for empirical context and Graham (2011a) for other empirical examples and references.) Let $W \in \mathbb{W} = \{w_1, \ldots, w_J\}$ denote the observable type of a teacher and $U \in \mathbb{U}$ unobserved teacher attributes.[1] The dimension of U is unrestricted. The J support points of W may encode, for example, different unique combinations of years of teaching experience, levels of education, race, and gender; U corresponds to unobserved dimensions of teacher quality. Teachers are heterogenous. Let R denote a vector of observable "proxies" for U; R may have both discrete and continuous components. We clarify the properties of R further below. All diversity in the population of teachers is captured by the triple (W, R, U). We index a random draw from this population by the subscript i, such that (W_i, R_i, U_i) corresponds to the ith random draw (teacher). A generic random draw is denoted by (W, R, U) (i.e., subscripts omitted).

Let $X \in \mathbb{X} = \{x_1, \ldots, x_K\}$ be the observable type of a classroom and $V \in \mathbb{V}$ unobserved classroom attributes. The dimension of V is unrestricted. The K types of classroom could enumerate different unique combinations of classroom size and/or student gender/ethnicity mixes. Let S denote an observed vector of proxies for V. We index a random draw from the population of classrooms by the superscript h, such that (X^h, S^h, V^h) equals measured and unmeasured characteristics of the hth random draw (classroom). The sub- and superscript notation emphasizes the two-population aspect of our setup.

Teachers and classrooms of students are matched (i.e., paired with one another) through some process. In this article, we only consider one-to-one matching. Restrictions on this process will be imposed in Section 3. Once paired they jointly produce the output, $Y \in \mathbb{Y} \subset \mathbb{R}$, say, student achievement.

[1] In what follows random variables are denoted by capital Roman letters, specific realizations by lower case Roman letters, and their support by blackboard bold Roman letters. That is, Y, y, and $\mathbb{Y}$, respectively, denote a generic random draw of, a specific value of, and the support of, Y.

Journal of Business & Economic Statistics
April 2020, Vol. 38, No. 2
DOI: 10.1080/07350015.2018.1497509

Associated with each teacher–classroom pair is a *potential* or *conjectural* output (Holland 1986; Manski 2007; Imbens and Rubin 2015). Let $Y_h(i)$ denote the potential output when teacher i matches with classroom h; in production function form,

$$Y_h(i) = g\left(W_i, X^h, U_i, V^h\right). \tag{1}$$

Now consider two teachers, i and i', both of type $W_i = W_{i'} = w$. A key feature of our setup is that there is no representation of the potential outcome for *classroom h* in terms of its assigned teacher's type, w, *alone*. This follows because, in general, $U_i \neq U_{i'}$ and hence

$$Y_h(i) = g(w, X^h, U_i, V^h) \neq g(w, X^h, U_{i'}, V^h) = Y_h(i'),$$

due to heterogeneity in unobserved teacher quality. Consequently, we cannot write $Y_h(i) = Y_h(w)$. Although teachers i and i' may be of the same observed type, $W_i = W_{i'} = w$, they would typically differ in terms of their unobserved "quality," $U_i \neq U_{i'}$.

Equation (1) is a production function with two *heterogenous* inputs, W and X. This contrasts with the standard single agent production function, where output across heterogenous firms varies with the level of a *homogenous* input (e.g., Chamberlain 1984; Griliches and Mairesse 1996; Olley and Pakes 1996). If we set $U_i = \bar{u}$ for all i, we recover this familiar single agent problem. Such a restriction would ensure that, conditional on their type, teachers are a homogenous input. We could then write a classroom's conjectural output as a function of its assigned teacher type alone:

$$Y_h(i) = Y_h(w) = \bar{g}(w, X^h, V^h) = g(w, X^h, \bar{u}, V^h). \tag{2}$$

In (2) achievement (output) across heterogenous classrooms (firms) is a function of the level of the homogenous input, teacher type (capital), $W_i = w$. In this article, we instead consider the nonstandard case, where observed output is generated according to (1). Loosely speaking, *both* the "firm" (classroom) and the "input" (teacher) are heterogenous in our setup. This raises new issues.[2]

Let $h = m(i)$ equal the classroom assigned to teacher i under the status quo (i.e., observed) matching (so that $m^{-1}(h) = i$). For simplicity, we assume that (i) the populations of teachers and classrooms are the same size and (ii) that all classrooms are assigned a teacher in the status quo matching. Observed output is therefore given by

$$Y_i = g(W_i, X^{m(i)}, U_i, V^{m(i)}). \tag{3}$$

In what follows we write $Y_i = Y_{m(i)}(i)$, $X_i = X^{m(i)}$, and $V_i = V^{m(i)}$ to simplify the notation. Put differently, the i subscript will be used to index both teachers and teacher–classroom matches (the latter in the status quo assignment only). Let $\{Z_i\}_{i=1}^N$ denote a random sample of size N, from the status quo distribution of matches, of $Z_i = (X_i, W_i, R_i', S_i', Y_i)'$.

The econometrician seeks to use this random sample to make inferences about average output across different *counterfactual* reallocations of teachers to classrooms. Specifically, we consider the following thought experiment. A social planner takes a random draw from the subpopulation of type $W_i = w$ teachers. She then takes an independent random draw from the subpopulation of type $X^h = x$ classrooms. The expected outcome associated with pairing together these two draws is

$$\beta(w, x) \stackrel{\text{def}}{\equiv} \iint g(w, x, u, v) f_{U|W}(u|w) f_{V|X}(v|x) du dv. \tag{4}$$

We call (4) the *average match function* (AMF; see Graham 2011a). The AMF is a building block for conducting inference on counterfactual reallocations. Observe that we make no presumption of independence between W and U or X and V. The distribution of teacher ability, U, may vary systematically with observed years of teacher experience, W. Because reallocations leave the joint distributions of (W_i, U_i) and (X^h, V^h) unchanged, $F_{U|W}$ and $F_{V|X}$ are the correct distributions to integrate over in (4).

In contrast, dependence between U and V, given $(W = w, X = x)$, generates a wedge between observed average output under the status quo matching, $\mathbb{E}[Y|W = w, X = x]$, and the AMF, $\beta(w, x)$, due to matching on unobservables. Say we wish to learn about the average match output when experienced teachers, $W_i = w$, are assigned to classrooms with low prior achievement, $X^h = x$. If, in the status quo matching, among experienced teachers, those with high ability, U_i, are matched to high ability, V^h, classrooms (among those with the low prior achievement), then there will be dependence between U and V given $(W = w, X = x)$. Therefore, the observed average outcome does not equal the causal match effect $\beta(w, x)$, but an upwardly biased estimate of it. This bias is induced by the (positive) correlation of U and V conditional on W and X. This matching bias occurs if conditional exogeneity fails.

The AMF is defined with reference to a hypothetical matching scheme which rules out such dependence by construction. This is analogous to the conceptual role played by random assignment in the program evaluation literature.

1. MAIN CONTRIBUTIONS

In this article, we present three results. First, we show that $\beta(w, x)$ is identified under a conditionally exogenous matching assumption. Our assumption is a multi-agent generalization of the "selection on observables" or "unconfoundedness" assumption familiar from the program evaluation literature (e.g., Heckman, Smith, and Clements 1997; Imbens 2004). Second, we show that, under certain assumptions about agents' information sets, our conditionally exogenous matching assumption is consistent with pairwise stability in an aggregate transferable utility (TU) matching market of the type introduced by Choo and Siow (2006a,b) and recently extended by a number of authors (e.g., Graham 2013; Dupuy and Galichon 2014; Chiappori, Salanié, and Weiss 2015; Galichon and Salanié 2015). (See Dagsvik 2000; Galichon and Hsieh 2015; Menzel

[2]If we consider classroom h's "treatment" to be the assignment to a *specific* teacher, then the fact that classroom h has a different potential outcome when assigned teacher i vs. teacher i' is not a violation of SUTVA (see Imbens and Rubin 2015). However, there is a violation of SUTVA if, instead, we consider the *type* of i as the treatment (e.g., assignment to an inexperienced vs. experienced teacher). This follows, as explained in the main text, because teachers of the same observed type may vary in terms of unobserved, output-effecting, attributes. Another violation of SUTVA implicit in our setup is that of no treatment interference. Interference in our setting arises because if classroom h is assigned to teacher i, then classroom h' cannot be assigned to teacher i; matching is one-to-one and rivalrous. Graham (2011a, Section 5) anticipated some of the discussion which follows.

2015 for related contributions to nontransferable utility (NTU) matching problems.) This result provides guidance to practitioners interested in applying our results outside of quasi-experimental settings. In particular, it suggests what types of variables should be included in the proxy vectors R and S. Third, we characterize the semiparametric efficiency bound for $\beta(w, x)$. The bound is complex, involving several integral equations, but nevertheless provides insights useful for efficient estimation.

In more specific forms than (4), the average match function (AMF) was introduced by Graham, Imbens and Ridder (2007, 2014) and Graham (2011a). The first two articles consider estimation of the AMF under various forms of random assignment/matching and more restrictive unobserved heterogeneity structures than allowed for here. The last of these articles does consider covariate adjustment and also provides a preliminary identification result (Graham 2011a, Proposition 3.2). This result relies on a "conditional inclusive definition of types" assumption (p. 980). Under doubly randomized assignments—as defined below—the AMF of Graham (2011a) and the one presented here coincide. Under conditional double randomization, the two identification results diverge. Evidently Assumption 3.2 of Graham (2011a) is consequential; we fully dispense with it here. The key technical tool for doing so is Lemma 1. These differences/improvements reflect our own improved understanding of matching models over time.[3]

This article is also related to Graham, Imbens, and Ridder (2014), which presented explicit estimators for various reallocation effects. In that article (i) teacher and classroom types were assumed continuously valued, (ii) methods for covariate adjustment were not presented (limiting the included formal results to experimental settings), and (iii) the *heterogenous* two-agent aspect of the problem was not explicitly developed. Finally, no analysis of semiparametric efficiency was undertaken. Our results are also related to the very large literature on efficient covariate adjustment in program evaluation problems (see Imbens and Wooldridge 2009 for a recent review). While covariate adjustment is a well-studied problem in single agent models, going back at least to the work of Yule (1897) on the causes of pauperism in late 19th century England, we are aware of no prior research on covariate adjustment for multi-agent models beyond that discussed above.

2. NOTATION

In what follows, random variables are denoted by capital Roman letters, specific realizations by lower case Roman letters, and their support by blackboard bold Roman letters. That is, Y, y, and $\mathbb{Y}$, respectively, denote a generic random draw of, a specific value of, and the support of, Y. The joint density of X and Y is denoted by $f_{X,Y}(x, y)$ or $f(x, y)$. The latter representation is only used when doing so causes no ambiguity. Similar conventions are followed for conditional densities.

3. IDENTIFICATION

We assume that the econometrician is able to collect a random sample of output measurements and agent observables from a status quo population of matches.

Assumption 1 (Random Sampling). Let $Z_i = (X_i, W_i, R_i', S_i', Y_i)$; $\{Z_i\}_{i=1}^N$ is a random sequence drawn from the status quo population of matches with distribution function F.

Our key identifying assumption restricts the structure of the status quo matching.

Assumption 2 (Conditionally Exogenous Matching).

$$(X, V, S) \perp U \,|\, W = w, R = r, \quad (W, U, R) \perp V \,|\, X = x, S = s$$

for all $(w, r) \in \mathbb{W} \times \mathbb{R}$ and all $(x, s) \in \mathbb{X} \times \mathbb{S}$.

Assumption 2 implies that, conditional on teacher observed attributes (W, R), her unobserved quality, U, has no predictive power for classroom characteristics. Likewise, conditional on observed classroom attributes (X, S), unobserved classroom attributes, V, have no predictive power for teacher characteristics. In Section 4, we show that Assumption 2 is consistent with a Choo–Siow aggregate matching market equilibrium under specific assumptions about agents' prematch information sets.

To better understand Assumption 2, we first prove the following factorization lemma. This lemma features in the proof of our main identification result, Proposition 1.

Lemma 1 (Factorization). Under Assumption 2

$$f_{U,V|W,X,R,S}(u, v|\, w, x, r, s) = f_{U|W,R}(u|\, w, r)\, f_{V|X,S}(v|\, x, s).$$

Proof. The first part of Assumption 2 gives the joint density factorization

$$\begin{aligned} &f_{W,X,R,S,U,V}(w, x, r, s, u, v) \\ &\quad = f_{X,V,S|W,R}(x, v, s|\, w, r)\, f_{U|W,R}(u|\, w, r)\, f_{W,R}(w, r), \end{aligned}$$

while the second part gives

$$\begin{aligned} &f_{W,X,R,S,U,V}(w, x, r, s, u, v) \\ &\quad = f_{W,U,R|X,S}(w, u, r|\, x, s)\, f_{V|X,S}(v|\, x, s)\, f_{X,S}(x, s). \end{aligned}$$

Conditioning on all observables therefore gives the pair of equalities

$$\begin{aligned} &f_{U,V|W,X,R,S}(u, v|\, w, x, r, s) \\ &\quad = f_{V|W,X,R,S}(v|\, w, x, r, s)\, f_{U|W,R}(u|\, w, r) \\ &\quad = f_{V|X,S}(v|\, x, s)\, f_{U|W,X,R,S}(u|\, w, x, r, s). \end{aligned}$$

Integrating over u then gives

$$f_{V|W,X,R,S}(v|\, w, x, r, s) = f_{V|X,S}(v|\, x, s), \qquad (5)$$

[3] The "inclusive definition of type" assumption imposes independence of, in the current notation, W_i and U_i and also of X^h and V^h. This assumption, which also features in Graham, Imbens, and Ridder (2010), is not imposed here. This difference is consequential when controlling for additional covariates. This can be seen by comparing the form of the identification result in Graham (2011a, Proposition 3.2) with the one outlined below. Here, identification involves averaging over the product of two conditional distributions, not two marginals as in Graham (2011a). Analog estimators based on the two results will numerically differ. Graham (2011a) did not derive a semiparametric efficiency bound for his estimand, but it too will differ from the one outlined here. We prefer the set of assumptions maintained here (which are weaker).

which after substitution gives

$$f_{U,V|W,X,R,S}(u,v|w,x,r,s) = f_{U|W,R}(u|w,r)\, f_{V|X,S}(v|x,s)$$

as claimed. □

Equation (5), in the proof to Lemma 1, highlights a key implication of Assumption 2: conditional on a classroom's observed attributes, X and S, the observed attributes of their assigned teacher, W and R, do not predict unobserved classroom attributes, V. Conversely, conditional on W and R, classroom characteristics, X and S, do not predict unobserved teacher attributes, U. Assumption 2 implies that within $W = w, R = r$ by $X = x, S = s$ cells, there is no matching on unobservables between teachers and classrooms. (Graham 2011a, equation (8)) instead studies assignments where $f_{U,V|W,X,R,S}(u,v|w,x,r,s) = f_{U|R}(u|r)f_{V|S}(v|s)$, which is stronger than maintained here.)

Note that

$$\begin{aligned} f_{U,V|W,X}(u,v|w,x) &= \iint f_{U|W,R}(u|w,r)\, f_{V|X,S}(v|x,s) \\ &\quad \times f_{R,S|W,X}(r,s|w,x)\, \mathrm{d}r\mathrm{d}s \\ &\neq f_{U|W}(u|w)\, f_{V|X}(v|x), \end{aligned} \tag{6}$$

so that Assumption 2 *does* allow for matching on unobservables within the *coarser* $W = w$ by $X = x$ cells. However within $W = w, R = r$ by $X = x, S = s$ cells matching is "as if" random. We call this *conditionally exogenous matching.*

Assumption 2 may hold for two reasons. First, it can hold by design. In that case, the researcher chooses a feasible joint distribution for (W, X, R, S), but forms a $(W, R) = (w, r)$ to $(X, S) = (x, s)$ match by taking a random draw from the subpopulation of teachers homogenous in $(W, R) = (w, r)$ and matching her with an independent random draw from the subpopulation of classrooms homogenous in $(X, S) = (x, s)$. This is a doubly randomized assignment scheme (see Graham 2008, 2011a). Note, as indicated by (6), this scheme does allow for sorting on unobservables within $W = w$ by $X = x$ cells. As shown below, the presence of the proxies R and S allows the researcher to "undo" this sorting to recover the AMF.

Second, Assumption 2 is also an equilibrium property of a Choo and Siow (2006a,b) type aggregate matching market (under certain restrictions on agents' information sets). We develop this result in Section 4.

Identification of $\beta(w, x)$ also requires a support condition.

Assumption 3 (Support).
(i) If $f_{S|X}(s|x)f_{R|W}(r|w) > 0$, then $f_{R,S|W,X}(r,s|w,x) > 0$,
(ii) $\pi_{xw} = \Pr(W = w, X = x) > 0$.

Note that, under part (ii) of Assumption 3, the reverse of the implication stated in part (i) holds as well. Assumption 3 therefore requires that the support of $f_{R,S|W,X}(r,s|w,x)$ equals the product of the supports of $f_{R|W}(r|w)$ and $f_{S|X}(s|x)$. Observe that the set

$$\mathbb{S}_{RS}^{\text{feasible}}(w,x) = \{r, s : f_{R|W}(r|w) f_{S|X}(s|x) > 0\}$$

equals the *feasible* joint support of R and S across the set of $W = w$ to $X = x$ matches. This set contains all logically possible combinations of $R = r$ and $S = s$ that *might be observed* in a $W = w$ to $X = x$ match. Identification requires that the *actual* support

$$\mathbb{S}_{RS}^{\text{actual}}(w,x) = \{r, s : f_{R,S|W,X}(r,s|w,x) > 0\}$$

and the feasible one overlap.

It is useful to connect this assumption to the familiar overlap condition found in the program evaluation literature. Doing so also allows us to introduce some notation that will be used in the efficiency bound calculation presented in Section 5. Under Assumption 3, Bayes' law gives

$$f_{R|W}(r|w) = \frac{p_w(r) f_R(r)}{\rho_w},\quad f_{S|X}(s|x) = \frac{p_x(s) f_S(s)}{\lambda_x},$$
$$f_{R,S|W,X}(r,s|w,x) = \frac{p_{wx}(r,s) f_{R,S}(r,s)}{\pi_{wx}},$$

where we define the conditional probabilities

$$\begin{aligned} p_w(r) &= \Pr(W = w | R = r) \\ p_x(s) &= \Pr(X = x | S = s) \\ p_{wx}(r,s) &= \Pr(W = w, X = x | R = r, S = s), \end{aligned}$$

and also the unconditional probabilities $\rho_w = \Pr(W = w)$, $\lambda_x = \Pr(X = x)$, and $\pi_{wx} = \Pr(W = w, X = x)$. (In certain instances, we will also use the notation $p_j(r) = \Pr(W = w_j | R = r)$, $p_k(s) = \Pr(X = x_k | S = s)$, and $p_{jk}(r,s) = \Pr(W = w_j, X = x_k | R = r, S = s)$ for $j = 1, \ldots, J$ and $k = 1, \ldots, K$.)

Under Assumption 3, we have

$$\begin{aligned} \mathbb{S}_{RS}^{\text{feasible}}(w,x) &= \{r, s : p_w(r) p_x(s) > 0\}, \\ \mathbb{S}_{RS}^{\text{actual}}(w,x) &= \{r, s : p_{wx}(r,s) > 0\}. \end{aligned} \tag{7}$$

The equalities in (7) suggest the following reformulation of Assumption 3:

Assumption 4 (Strong Overlap). $p_{wx}(r,s) \geq \kappa > 0$ for all (r, s) such that $p_w(r)p_x(s) > 0$.

As suggested by its label, Assumption 4, is related to the overlap assumption made in the program evaluation literature (e.g., Hahn 1998; Imbens 2004). It ensures that all logically possible combinations of R and S that could be observed in a $W = w$ to $X = x$ match are in fact observed in the set of status quo w-to-x matches. It would be useful to develop tests, heuristic or formal, for assessing Assumption 4 in practice.

Consider the mean regression of Y given W, X, R, S

$$\mathbb{E}[Y | W = w, X = x, R = r, S = s] \overset{\text{def}}{\equiv} q(w,x,r,s). \tag{8}$$

Note that $q(w, x, r, s)$ is a structural object under (1) and Assumptions 1, 2, and 3. Specifically, the difference

$$q(w, x', r, s') - q(w, x, r, s),$$

gives the expected change in output when a teacher with characteristics $(W, R) = (w, r)$ is assigned to a classroom with characteristics $(X, S) = (x, s)$ instead of one with characteristics $(X, S) = (x', s')$.

Our main identification result is

Proposition 1. (Identification) Under (1) and Assumptions 1, 2, and 3

$$\beta(w,x) = \frac{1}{\rho_w \lambda_x} \iint q(w,x,r,s)\, p_w(r)\, p_x(s)\, f_R(r)\, f_S(s)\, drds \quad (9)$$

$$= \int_s \int_r q(w,x,r,s)\, f_{R|W}(r|\, w)\, f_{S|X}(s|\, x)\, drds. \quad (10)$$

Proof. First, observe that under Assumption 2, we have, invoking Lemma 1,

$$\begin{aligned} q(w,x,r,s) &= \int_v \int_u g(w,x,u,v)\, f_{U,V|W,X,R,S}(u,v|\, w,x,r,s)\, dudv \\ &= \int_v \int_u g(w,x,u,v)\, f_{U|W,R}(u|\, w,r)\, f_{V|X,S}(v|\, x,s)\, dudv. \end{aligned}$$

Second, from Bayes' rule $f_{R|W}(r|w) = \frac{p_w(r) f_R(r)}{\rho_w}$ and $f_{S|X}(s|x) = \frac{p_x(s) f_S(s)}{\lambda_x}$. This and the second equality above yields

$$\begin{aligned} &\frac{1}{\rho_w \lambda_x} \int_s \int_r q(w,x,r,s)\, p_w(r)\, p_x(s)\, f_R(r)\, f_S(s)\, drds \\ &= \int_s \int_r q(w,x,r,s)\, f_{R|W}(r|\, w)\, f_{S|X}(s|\, x)\, drds \\ &= \int_s \int_r \left[\int_v \int_u g(w,x,u,v)\, f_{U|W,R}(u|\, w,r)\, f_{V|X,S}(v|\, x,s)\, dudv \right] \\ &\quad \times f_{R|W}(r|\, w)\, f_{S|X}(s|\, x)\, drds \\ &= \int_s \int_r \int_v \int_u g(w,x,u,v)\, f_{U,R|W}(u,r|\, w)\, f_{V,S|X}(v,s|\, x)\, dudrdvds \\ &= \int_v \int_u g(w,x,u,v)\, f_{U|W}(u|\, w)\, f_{V|X}(v|\, x)\, dudv \\ &= \beta(w,x). \end{aligned}$$

Note that for (9) and (10) to be well-defined, we require Assumption 3. Since all the components to the right of the equalities in (9) and (10) are asymptotically revealed under random sampling (Assumption 1), the result follows. □

We discuss some implications of Proposition 1 for estimation after presenting its semiparametric efficiency bound in Section 5. A key implication of Proposition 1, useful for both estimation and efficiency bound analysis, is the moment restriction

$$\mathbb{E}\left[\frac{1}{\rho_w \lambda_x} \frac{f(R)\, f(S)}{f(R,S)} \frac{p_w(R)\, p_x(S)}{p_{wx}(R,S)} T_{wx} Y - \beta(w,x) \right] = 0,$$

for $T_{wx} = 1$ when a random draw from the status quo distribution of matches is of type $W_i = w$ and $X_i = x$ and zero otherwise.

4. PAIRWISE STABILITY AND EXOGENEITY

In this section, we relate our exogenous matching condition (Assumption 2) to the notion of pairwise stability in transferable utility (TU) one-to-one matching problems (e.g., Shapley and Shubik 1971; Becker 1973). Our point of departure is the aggregate matching setup introduced by Choo and Siow (2006a,b). In this framework, the econometrician observes the match frequencies $\pi_{jk} = \Pr(W_i = w_j, X^{m(i)} = x_k)$ for $j = 1, \ldots, J$ and $k = 1, \ldots, K$ and, from this joint distribution of match types and equilibrium restrictions, seeks to recover (features of) the distribution of unobserved agent preferences. Match surplus and transfers are unobserved.

We add to this setup the observable match output Y_i. Match output will generally covary with the match surplus agents' actually care about, but it need not be coincident with it. For example, the surplus associated with a specific marriage may vary with (expected) child outcomes, but would generally not be coincident with them. Our question is: under what restrictions on agents' preferences and information sets will the observed matching be both (i) pairwise stable and (ii) satisfy Assumption 2? Our conclusion is that Assumption 2 can hold in settings where agents purposively choose match partners. More constructively, our analysis provides guidance as to what types of measures to include in the proxy variable vectors R_i and S^h. Proposition 2 formalizes and generalizes the stylized analysis for the $K = J = 2$ special case sketched in Graham, Imbens, and Ridder (2014, sec. 3).

For what follows, maintaining the assumption of one-to-one matching, it is pedagogically convenient to think of the first population as consisting of firms (teachers) and the second of workers (classrooms). The match output associated with the pairing of firm i and worker h is now restricted to equal

$$g(W_i, X^h, U_i, V^h) = \beta(W_i, X^h) + U_i(X^h) + V^h(W_i), \quad (11)$$

where $\beta(w,x)$ is an unrestricted function, $U_i(x) = \sum_{k=1}^{K} \mathbf{1}(x = x_k) U_{ki}$ and $V^h(w) = \sum_{j=1}^{J} \mathbf{1}(w = w_j) V_j^h$ with $V^h = (V_1^h, \ldots, V_J^h)'$ and $U_i = (U_{i1}, \ldots, U_{iK})'$. We normalize $U_i(x)$ to be conditionally mean zero (i.e., $\mathbb{E}[U_i(x)|W_i] = 0$ for all $x \in \mathbb{X}$) and impose the analogous restriction on $V^h(w)$. These normalizations imply that the average match function (AMF) equals $\beta(w,x)$. Note that average output across observed $W = w$ to $X = x$ matches may not equal the AMF. Matching bias is possible.

Equation (11) is restrictive, ruling out interactive effects in the unobservable productivity vectors $U_i = (U_{1i}, \ldots, U_{Ki})'$ and $V^h = (V_1^h, \ldots, V_J^h)'$. This type of separability restriction plays an essential role in the empirical structural matching literature (see Assumption 2 of Galichon and Salanié 2015). To better understand the content of (11), consider two firms, i and i', of the same type, say w, and two workers, h and h', also of the same type, say x. Under (11) the aggregate output associated with the i-to-h and i'-to-h' matching equals

$$2\beta(w,x) + U_i(x) + V^h(w) + U_{i'}(x) + V^{h'}(w),$$

which exactly coincides with that of the alternative i-to-h' and i'-to-h matching. Any rearrangement of matches *within* a $W = w$ and $X = x$ cell leaves aggregate output unchanged (although individual match output may, of course, change).

We now turn to firm and worker *preferences*. The surplus firm i gets from matching with worker h equals

$$\Pi_i(X^h) = \beta(W_i, X^h) - \tau(W_i, X^h) + U_i(X^h) + \tilde{\varepsilon}_i(X^h),$$

where $\tau(w_j, x_k)$ equals the *equilibrium* transfer a type $W_i = w_j$ firm "pays" a type $X^h = x_k$ worker (transfers may be negative), and $\tilde{\varepsilon}_i(x) = \sum_{k=1}^{K} \mathbf{1}(x = x_k)\tilde{\varepsilon}_{ki}$ is an additional source of unobserved firm-specific heterogeneity. We introduce this term to allow for a divergence between the net match output of interest to the econometrician, and the net match surplus agents actually care about. When these two objects coincide $\tilde{\varepsilon}_i(x)$ will equal zero for all $x \in \mathbb{X}$. (The development in this section employs a variant of the notation used in Graham (2013), we gloss over several interesting subtleties of the CS framework, referring the reader to, for example, Galichon and Salanié (2015) for a rigorous and comprehensive exposition.)

The surplus worker h gets from matching with firm i equals

$$\Xi^h(W_i) = \tau(W_i, X^h) + V^h(W_i) + \tilde{\upsilon}^h(W_i),$$

with $\tilde{\upsilon}^h(w) = \sum_{j=1}^{J} \mathbf{1}(w = w_j)\tilde{\upsilon}_j^h$ introduced for the same reason as $\tilde{\varepsilon}_i(x)$.

We impose the following informational structure: prior to matching firms and workers observe their own and candidate partners' types, know the form of $\beta(w, x)$, and also observe transfers. While agents also observe the vectors $\tilde{\varepsilon}_i = (\tilde{\varepsilon}_{1i}, \ldots, \tilde{\varepsilon}_{Ki})'$ and $\tilde{\upsilon}^h = (\tilde{\upsilon}_1^h, \ldots, \tilde{\upsilon}_J^h)'$, they do not observe $U_i = (U_{1i}, \ldots, U_{Ki})'$ and $V^h = (V_1^h, \ldots, V_J^h)'$. This means that the ex post utility (and output) associated with any given match is imperfectly known to agents ex ante. While U_i and V^h are unobserved, agents have at their disposal the signals R_i and S^h. We assume that these signals satisfy:

Assumption 5 (Signals). (i) Prior to matching firms and workers observe $(W_i, X^h, R_i, S^h, \tilde{\varepsilon}_i, \tilde{\upsilon}^h)$, (ii)

$$V^h \perp (W_i, R_i, \tilde{\varepsilon}_i) | X^h, S^h, \tilde{\upsilon}^h \tag{12}$$

$$U_i \perp \left(X^h, S^h, \tilde{\upsilon}^h\right) \Big| W_i, R_i, \tilde{\varepsilon}_i, \tag{13}$$

and (iii)

$$U_i \perp \tilde{\varepsilon}_i | W_i, R_i \tag{14}$$

$$V^h \perp \tilde{\upsilon}^h \Big| X^h, S^h. \tag{15}$$

Part (i) of Assumption 5 defines what is observed by agents i and h when they match (but prior to realizing match output). Parts (ii) and (iii) restrict the relationship between what is known and unknown by agents at the time of matching. Consider conditions (12) and (14), which restrict the predictability of worker and firm productivity, respectively, $V^h = (V_1^h, \ldots, V_J^h)$ and $U_i = (U_{i1}, \ldots, U_{iK})'$. We omit a discussion of conditions (13) and (15), as they are analogous to (12) and (14).

Condition (12) implies that firm i's own attributes—W_i, R_i, $\tilde{\varepsilon}_i$—have no predictive power for worker $h's$ unobserved productivity, V^h, conditional on her attributes—X^h, S^h, $\tilde{\upsilon}^h$. In words, conditional on what the two agents know about the worker, what is additionally known about the firm cannot be used to predict worker productivity. This appears to be a natural assumption in our context. Note that (12) alone does not impose restrictions on the joint distribution of $(W, X, R, S, \tilde{\varepsilon}, \tilde{\upsilon})$. This implies, for example, that agents could assortatively match on $\tilde{\varepsilon}_k$ and $\tilde{\upsilon}_j$ within a $W = w_j, R = r$ by $X = x_k, S = s$ match cell if they so wanted. Recall, further, that $\tilde{\varepsilon}_k$ and $\tilde{\upsilon}_j$ are unobserved by the econometrician.

Some further reification/simplification may be helpful. Let $X^h \in \{0, 1\}$ denote whether a worker is "skilled" and $W_i \in \{0, 1\}$ whether a firm is "high tech." We can think of $\tilde{\varepsilon}_i$ as a vector of firm-specific *cost* (or taste) shocks. If $\tilde{\varepsilon}_i(0)$ is low, then it is especially costly for firm i to hire a low skilled, $X^h = 0$, worker. The firm knows this, and also $\tilde{\varepsilon}_i(1)$, prior to choosing the type of worker they hire. In contrast, think of U_i as a vector of idiosyncratic firm-specific *productivity* shocks. If $U_i(1)$ is high, then firm i will produce more output when they hire a skilled, $X^h = 1$, worker. However, the productivity shocks $U_i(0)$ and $U_i(1)$ are only imperfectly observed, and hence acted upon, by the firm at the time of hiring. The firm uses the productivity signal R_i to forecast $U_i = (U_i(0), U_i(1))'$. A second key difference between $\tilde{\varepsilon}_i$ and U_i is that while the realization of the latter is directly reflected in observed match output, the influence of the former on output is only indirect—via the choice of the type of worker hired.

In this example, condition (14) implies that firm-specific cost and productivity shocks are independent conditional on own type and the vector of productivity signals, R_i. Among, say, high tech, $W_i = 1$, firms with identical signals, R_i, variation in the firm-specific costs (or tastes) of hiring the two types of works is independent of the corresponding firm-specific benefits.

Condition (14) *is* restrictive. It implies that W and R contain all variables that simultaneously predict U *and* $\tilde{\varepsilon}$ or, equivalently, all variables that predict match surplus, and hence choice, *and also* affect match output. This condition is analogous to the "selection on observables" assumption familiar from the program evaluation literature. In that context, treatment exogeneity requires that the set of pretreatment conditioning variables used by the econometrician include all *joint* predictors of the treatment and outcome. The appropriateness of condition (14) is context specific. It will be violated, for example, if there exists a component of $\tilde{\varepsilon}$, which is part of the firm's information set that covaries with productivity, U, conditional of those parts of the information sets that are observed by the econometrician (i.e., W and R). If R plausibly approximates those components of a firm's prematch information set that are also likely to predict productivity, U, then invoking Assumption 5 is reasonable.

Under Assumption 5 we prove the following Lemma.

Lemma 2 (Factorization with Signals). Under Assumption 5

$$\begin{aligned} & f_{U,V|W,X,R,S,\tilde{\varepsilon},\tilde{\upsilon}}\left(u, v | w, x, r, s, \tilde{\varepsilon}, \tilde{\upsilon}\right) \\ & \quad = f_{U|W,R}\left(u | w, r\right) f_{V|W,R}\left(v | x, s\right). \end{aligned} \tag{16}$$

Proof. See the Appendix. The argument is similar to that used to show Lemma 1. □

Agents directly act on their knowledge of $(W_i, X^h, R_i, S^h, \tilde{\varepsilon}_i, \tilde{\upsilon}^h)$ when matching, inducing a specific equilibrium match density $f_{W,X,R,S,\tilde{\varepsilon},\tilde{\upsilon}}(w, x, r, s, \tilde{\varepsilon}, \tilde{\upsilon})$ in the process. Observe that Assumption 5 alone does not restrict this match density (beyond the requirements of feasibility). As noted above, sorting on $\tilde{\varepsilon}$ and $\tilde{\upsilon}$, for example, is allowed. However Lemma 2 shows that an implication of Assumption 5 is that any such sorting does not induce sorting on U and V conditional on $W, X, R, S, \tilde{\varepsilon}$, and $\tilde{\upsilon}$.

Using Lemma 2, we compute firm i's expected utility from matching with worker h as

$$\begin{aligned} & \mathbb{E}\left[\Pi_i\left(X^h\right) \Big| W_i, X^h, R_i, S^h, \tilde{\varepsilon}_i, \tilde{\upsilon}^h\right] \\ & \quad = \beta\left(W_i, X^h\right) - \tau\left(W_i, X^h\right) + \tilde{\varepsilon}_i\left(X^h\right), \end{aligned} \tag{17}$$

where $\bar{\varepsilon}_i(X^h)$ is the firm's forecast of $U_i(X^h) + \tilde{\varepsilon}_i(X^h)$

$$\bar{\varepsilon}_i\left(X^h\right) = \sum_{k=1}^{K} \mathbf{1}\left(X^h = x_k\right) \mathbb{E}\left[U_{ki}|\, W_i, R_i\right] + \tilde{\varepsilon}_i(X^h). \quad (18)$$

Note that the utility firm i expects to receive when matching with worker h depends on worker h's type alone. Although the firm also observes the worker attributes S^h and $\tilde{\upsilon}^h$, its expected utility is invariant to them. This result is an implication of the separable form of the CS utility function, something we inherit from the structural matching literature, as well as our assumption about agent information sets.

Similarly, we compute worker h's expected utility from matching with firm i as

$$\mathbb{E}[\Xi^h(W_i)|W_i, X^h, R_i, S^h, \tilde{\varepsilon}_i, \tilde{\upsilon}^h] = \tau(W_i, X^h) + \bar{\upsilon}^h(W_i), \quad (19)$$

with the corresponding forecast of $V^h(W_i) + \tilde{\upsilon}^h(W_i)$ for worker h equal to

$$\bar{\upsilon}^h(W_i) = \sum_{j=1}^{J} \mathbf{1}\left(W_i = w_j\right) \mathbb{E}\left[V_j^h \middle| X^h, S^h\right] + \tilde{\upsilon}^h(W_i). \quad (20)$$

Worker h's expected utility from matching with firm i depends on firm i's type alone. Although the worker also observes the firm attributes R_i and $\tilde{\varepsilon}_i$, her expected utility is invariant to them.

Under (17) and (19) firm and worker partner choice, respectively, satisfies

$$\pi_{jk}^D = \Pr\left(k = \arg\max_{x \in \{x_1,\ldots,x_K\}} \left[\beta\left(w_j, x\right) - \tau\left(w_j, x\right) + \bar{\varepsilon}(x)\right]\right) \quad (21)$$

and

$$\pi_{jk}^S, = \Pr\left(j = \arg\max_{w \in \{w_1,\ldots,w_J\}} \left[\tau\left(w, x_k\right) + \bar{\upsilon}(w)\right]\right), \quad (22)$$

which coincide with the choice rules of the generalized CS model. The transfers, $\tau(w, x)$, adjust so as to ensure that $\pi_{jk}^D = \pi_{jk}^S$ for all j, k in equilibrium. This ensures that the "demand" for type $X^h = x$ workers by type $W_i = w$ firms coincides with the "supply" of type $X^h = x$ workers to type $W_i = w$ firms (e.g., Graham 2013). Galichon and Salanié (2015, Theorems 1 and 2) showed that equilibrium exists and is unique as long as $\bar{\varepsilon}_i(x)$ for all $x \in \mathbb{X}$ and $\bar{\upsilon}^h(w)$ for all $w \in \mathbb{W}$ have sufficiently large support (which we assume here).

The CS equilibrium induces a particular type of sorting. Consider the subpopulation of type $W_i = w$ firms. Among these firms the subset that matches with type $X^h = x$ workers will differ from the subset that matches with type $X^h = x'$ workers. Specifically, from the choice rule (21), the distribution of $\bar{\varepsilon}(x)$ and $\bar{\varepsilon}(x')$ will differ between the two groups. Because R_i covaries with $\bar{\varepsilon}_i$—see (18) above—the distribution of R_i may differ across the two groups as well. Finally, because R_i covaries with U_i, the distribution of U_i may differ across the two groups. Consequently average output across $W = w$ to $X = x$ matches will not generally coincide with the AMF. This is because the distribution of firm productivity in this cell may differ from that across the entire subpopulation of $W = w$ firms in a CS equilibrium. A similar reasoning can be used to describe how, among workers of the same type, the distribution of worker ability will vary with the chosen type of the matched firm.

Under Assumption 5, the availability of R and S is sufficient to "undo" any biases caused by CS matching. Together (11) and Lemma 2 imply (multiplying both sides of (16) by $f_{\tilde{\varepsilon},\tilde{\upsilon}|W,X,R,S}(\tilde{\varepsilon}, \tilde{\upsilon}|w, x, r, s)$ and integrating over $\tilde{\varepsilon}$ and $\tilde{\upsilon}$ gives

$$f_{U,V|W,X,R,S}\left(u, v|\, w, x, r, s\right) = f_{U|W,R}\left(u|\, w, r\right) f_{V|W,R}\left(v|\, x, s\right))$$

that the proxy variable regression function (8) equals

$$q(w_j, x_k, r, s) = \beta(w_j, x_k) + \mathbb{E}\left[U_k|\, W = w, R = r\right] + \mathbb{E}[V_j|X = x, S = s] \quad (23)$$

for all combinations of $j = 1, \ldots, J$ and $k = 1, \ldots, K$. Plugging (23) into the right-hand side of (9) or (10) and evaluating then gives

$$\int_s \int_r q(w_j, x_k, r, s) f(r|w_j) f(s|x_k) dr ds = \beta(w_j, x_k).$$

Lemma 2 and (23) thus give.

Proposition 2 (CS Equilibrium and Exogeneity). When match surplus and output takes the form described above, and agents' prematch information sets satisfy Assumption 5, agents will (i) match according to (21) and (22), (ii) transfers adjust to clear the market, and (iii) the equilibrium matching will satisfy the conditionally exogenous matching condition (Assumption 2).

For an empirical researcher contemplating invoking Assumption 2 in a setting where agents *choose* match partners in a decentralized way (with transferable utility), assessing the plausibility of Assumption 5 is key. This condition is analogous to conditions for input exogeneity in single-agent models (e.g., Chamberlain 1984; Olley and Pakes 1996). (In Chamberlain's (1984) example, the farmer knows land quality (unobserved by the econometrician) when choosing her input level, but is unable to forecast weather. Weather influences farm output, but only after input choices are made.) Consider the teacher-to-classroom matching problem introduced in the introduction. In that example R_i should include attributes that correlate with teacher productivity, U_i. Likewise S^h should include student characteristics that are associated with high levels of achievement, V^h. It may be that there are additional (unobserved) teacher and student attributes that influence the matching process, for example, some teachers may especially prefer to work close to where they live. In that case, condition (14) would require that conditional on R_i, a teacher's commuting tastes do not help to predict her unobserved productivity.

To be clear our conclusion is not that Assumption 2 is suitable for routine use in all observational settings, rather it is that it (i) *can* be appropriate in certain well-defined settings and (ii) it is possible to reason about such settings in a ways familiar from the single agent observational context (e.g., Heckman, Smith, and Clements 1997; Imbens 2004). As in the single agent context, articulating the relationship between the agents' and the econometrician's information sets is central.

Research designs based on conditional exogeneity assumptions (i.e., "selection on observables," "unconfoundedness," etc.) have proved to be a very durable, albeit controversial, part of the researcher's toolkit (e.g., Chamberlain 1984; Olley and Pakes 1996; Griliches and Mairesse 1998). Our view, shaped by the observation that Assumption 2 is compatible with the leading empirical model of one-to-one matching under certain informational assumptions, is that covariate adjustment can play a similar role in multi-agent production problems.

5. SEMIPARAMETRIC EFFICIENCY BOUND

Our final result, Theorem 1, characterizes the semiparametric efficiency bound for $\beta(w,x)$ under (1) and Assumptions 1, 2, and 3. As in Graham (2011b), a multinomial approximation (not reported) was used to conjecture the form of the bound, with the formal result following from a pathwise derivative calculation, as in Newey (1990). We have also verified our derivation of the efficient influence function using the method of Newey (1994b) and the moment condition $\mathbb{E}[\frac{1}{\rho_w\lambda_x}\frac{f(R)f(S)}{f(R,S)}\frac{p_w(R)p_x(S)}{p_{wx}(R,S)}T_{wx}Y-\beta(w,x)]=0$.

Let $D_w(W)=D_w=1$ if $W=w$ and zero otherwise. Let $E_x(X)=E_x=1$ if $X=x$ and zero otherwise. Let $T_{wx}(W,X)=T_{wx}=1$ if $W=w$ and $X=x$ and zero otherwise. Let $\beta_{wx}=\beta(w,x)$ and define the candidate efficient influence function

$$\phi_0(Z,\beta_{wx},h(Z))=\psi_0(Z,\beta_{wx},h(Z))+\psi_R(Z,\beta_{wx},h(Z))+\psi_S(Z,\beta_{wx},h(Z)),\tag{24}$$

where

$$h(Z)=(f(R,S),f(R|W),f(S|X),\rho_\omega,\lambda_x,p_{wx}(R,S),q(w,x,R,S),e_S(w,x,R),e_R(w,x,S))'$$

and

$$\psi_0(Z,\beta_{wx},h(Z))=\frac{f(R|W)f(S|X)}{f(R,S)}\frac{T_{wx}}{p_{wx}(R,S)}\times(Y-q(w,x,R,S))$$

$$\psi_R(Z,\beta_{wx},h(Z))=\frac{D_w}{\rho_w}(e_S(w,x,R)-\beta_{wx})$$

$$\psi_S(Z,\beta_{wx},h(Z))=\frac{E_x}{\lambda_x}(e_R(w,x,S)-\beta_{wx})$$

with

$$e_S(w,x,r)=\int q(w,x,r,s)f(s|x)ds$$

$$e_R(w,x,s)=\int q(w,x,r,s)f(r|w)dr.$$

Define the candidate variance bound

$$\begin{aligned}\mathcal{I}_0(\beta_{wx})^{-1}&=\mathbb{E}\left[\left\{\frac{f(R|W=w)f(S|X=x)}{f(R,S)}\right\}^2\frac{\sigma^2_{wx}(R,S)}{p_{wx}(R,S)}\right]\\&+\frac{1}{\rho_w}\mathbb{E}\left[(e_S(w,x,R)-\beta_{wx})^2\middle|W=w\right]\\&+\frac{1}{\lambda_x}\mathbb{E}\left[(e_R(w,x,S)-\beta_{wx})^2\middle|X=x\right]\\&+2\frac{\pi_{wx}}{\rho_w\lambda_x}\mathbb{E}[(e_S(w,x,R)-\beta_{wx})(e_R(w,x,S)-\beta_{wx})|\\&W=w,X=x]\end{aligned}\tag{25}$$

with

$$\sigma^2_{wx}(r,s)=\mathbb{V}(Y|W=w,X=x,R=r,S=s).$$

Theorem 1. The semiparametric efficiency bound for β_{wx} in the problem defined by (1) and Assumptions 1, 2, and 3 is equal to $\mathcal{I}_0(\beta_{wx})$ with an efficient influence function of $\phi_0(Z,\beta_{wx},h(Z))$.

Proof. See the Appendix. □

Both the efficient influence function and the variance bound have straightforward interpretations. Consider first the influence function. Its first term, $\psi_0(Z,\beta_{wx},h(Z))$, reflects the asymptotic penalty associated with not knowing conditional distribution of Y given (W,X,R,S). The second and third terms, $\psi_R(Z,\beta_{wx},h(Z))$ and $\psi_S(Z,\beta_{wx},h(Z))$, reflect the contributions of uncertainty about, respectively, the conditional distributions of R given W and S given X. The interpretation of $\mathcal{I}_0(\beta_{wx})^{-1}$ is analogous, with its last term arising from covariance between $\psi_R(Z,\beta_{wx},h(Z))$ and $\psi_S(Z,\beta_{wx},h(Z))$.

One implication of Theorem 1 likely to be of direct interest to empirical researchers is its implications for the relationship between the quality of overlap and feasible precision. Under conditional homoscedasticity, the firm term in $\mathcal{I}_0(\beta_{wx})^{-1}$ is proportional to

$$\mathbb{E}\left[\left\{\frac{f(R|W=w)f(S|X=x)}{f(R,S)}\right\}^2\frac{1}{p_{wx}(R,S)}\right],$$

which will be large when the conditional probability of a $W_i=w$ to $X^h=x$ match is very low for enough combinations of $R_i=r$ and $S^h=s$ appearing in $\mathbb{S}^{\text{feasible}}_{RS}(w,x)$. Developing simple diagnostics for assessing our support condition (Assumption 4), similar to those available in the program evaluation setting, would be useful.

Although we do not formally present an AMF estimator that attains the bound of Theorem 1, in the next section we discuss some procedures which are likely to do so under appropriate regularity conditions. An analog estimate of the efficient influence function given in Equation (24) could be used to construct a consistent variance estimate for these AMF estimators.

6. FURTHER RESEARCH DIRECTIONS

In this article, we have characterized a method of covariate adjustment appropriate for two-agent models. (The extension to settings with more than two agents appears to be straightforward. Graham, Imbens, and Ridder (2010) provided one motivating example for such an extension.) When matching is conditionally exogenous our approach to covariate adjustment recovers a well-defined causal object: the *average match function* (AMF). Although, as in other areas of applied social science research, the econometrician may be interested in "controlling for" observed covariate differences even if Assumption 2 does not hold (exactly) (see Keiding and Clayton 2014). Our efficiency bound calculation characterizes the maximum asymptotic precision possible when undertaking such covariate adjustment. The bound is valid for the estimand defined by the right-hand side of (10) irrespective of whether it also coincides with the AMF.

Recovering structural objects via covariate adjustment can be controversial in some settings (see Freedman 1997). Proposition 2 relates our conditionally exogenous matching assumption to the structural TU matching model of Choo and Siow (2006a,b). This model has been an object of intense study, development, and application in recent years (see Chiappori and Salanié 2016 for a survey). Proposition 2 shows that a status quo matching can both satisfy our key identifying assumption (Assumption 2) *and* be consistent with a TU matching equilibrium. This result requires maintaining certain assumptions about agents' information sets (Assumption 5) and provides guidance regarding which types of measures should be included in the teacher and classroom proxy variables, respectively, R and S.

We have not presented an estimator for the AMF, instead we leave this exercise to future research. However, the structure of the efficient influence function in Theorem 1 suggest several possibilities. Perhaps the most obvious is the following "double average" estimator

$$\hat{\beta}_{\mathrm{DA}}(w,x) = \frac{N^{-2}\sum_{i=1}^{N}\sum_{j=1}^{N} 1(W_i=w)\,1(X_j=x)\,\hat{q}(w,x,R_i,S_j)}{N^{-2}\sum_{i=1}^{N}\sum_{j=1}^{N} 1(W_i=w)\,1(X_j=x)}, \quad (26)$$

where $\hat{q}(w,x,R_i,S_j)$ is a preliminary nonparametric estimate. This estimator is similar to the partial mean estimator introduced by Newey (1994a), but instead requires "integration" with respect to a product of two marginal distributions (as opposed to integrating with respect to a single joint distribution). This feature is reflected in the V-Statistic structure of (26). Statistically, (26) corresponds to the random matching estimator introduced in Graham, Imbens, and Ridder (2014) for the special case where $W_i = w$ for all i and $X_j = x$ for all j (i.e., when there is only one type of teacher and only one type of classroom). (Graham, Imbens, and Ridder (2014) estimated $\hat{q}(w,x,r,s)$ by a particular kernel regression estimator designed to deal with boundary bias.) In that case the efficient influence function given in Theorem 1 also corresponds to the influence function derived (by brute force) in Graham, Imbens, and Ridder (2014). This suggests that an efficient estimator for the AMF could be constructed by adapting the regularity conditions and specific estimation procedures presented there (likewise it implies semiparametric efficiency of the random matching estimator). Empirical researchers might consider using a flexible parametric estimate of $\hat{q}(w,x,r,s)$ in practice.

The form of the efficient influence function also suggests an inverse probability weighting type (IPW) estimator. In particular, under (1) and Assumptions 1, 2, and 4, we have

$$\beta(w,x) = \mathbb{E}\left[\frac{1}{\rho_w\lambda_x}\frac{f(R)f(S)}{f(R,S)}\frac{p_w(R)\,p_x(S)}{p_{wx}(R,S)}T_{wx}Y\right].$$

This suggests an estimator, akin the one studied by Hirano, Imbens, and Ridder (2003) for the single agent case, of

$$\hat{\beta}_{\mathrm{IPW}}(w,x) = \frac{1}{N}\frac{1}{\hat{\rho}_w\hat{\lambda}_x}\sum_{i=1}^{N}\frac{\hat{f}(R_i)\hat{f}(S_i)}{\hat{f}(R_i,S_i)}\frac{\hat{p}_w(R_i)\,\hat{p}_x(S_i)}{\hat{p}_{wx}(R_i,S_i)}T_{wx,i}Y_i. \quad (27)$$

It would also be of interest to construct locally efficient, doubly robust, estimators, as has been done in the program evaluation context (see Graham, Pinto, and Egel 2012, Graham, Pinto, and Egel 2016, and the references cited therein).

The AMF provides information on how match output varies across different types of agent pairings. We close our article by briefly outlining how to integrate the AMF into an explicit social planning problem. We assume the social planner knows $\beta(w,x)$ for all $(w,x)\in\mathbb{W}\times\mathbb{X}$ (perhaps up to sampling uncertainty). She also knows the marginal distributions of teacher and classroom types, respectively, $\rho=(\rho_1,\ldots,\rho_J)'$ for $\rho_j=\Pr(W_i=w_j)$ and $\lambda=(\lambda_1,\ldots,\lambda_K)'$ for $\lambda_k=\Pr(X^h=x_k)$ (again perhaps up to sampling uncertainty). She does not observe $(R_i',U_i')'$ or $(S^h,V^h)'$ or is unable/unwilling to act on this knowledge if she does. Put differently, the planner is constrained to consider only doubly randomized reallocations (Graham 2008, 2011a).

Recall that $\pi_{jk}=\Pr(W=w_j, X=x_k)$ for $j=1,\ldots J$ and $k=1,\ldots,K$. The planner's problem is to choose a $\pi=(\pi_{11},\ldots,\pi_{1K},\ldots,\pi_{J1},\ldots,\pi_{JK})'$ that maximizes expected output

$$\theta(\pi)=\sum_{j=1}^{J}\sum_{k=1}^{K}\beta(w_j,x_k)\,\pi_{jk} \quad (28)$$

subject to the $J+K$ feasibility constraints:

$$\sum_{k=1}^{K}\pi_{jk}=\rho_j,\ \ j=1,\ldots,J \quad (29)$$

$$\sum_{j=1}^{J}\pi_{jk}=\lambda_k,\ \ k=1,\ldots,K.$$

See Graham, Imbens, and Ridder (2007).

Since $\sum_{j=1}^{J}\sum_{k=1}^{K}\pi_{jk}=1$, one constraint is redundant. Table 1 depicts the structure of a feasible assignment. By substituting out the feasibility constraints, an assignment can be represented in terms of $(J-1)(K-1)$ probabilities.

Graham (2011a) showed that the difference between two doubly randomized allocations, π' and π is given by

$$\theta(\pi')-\theta(\pi)=\sum_{j=1}^{J-1}\sum_{k=1}^{K-1}(\pi'_{jk}-\pi_{jk})\left(\beta(w_J,x_K)-\beta(w_J,x_k)-\left[\beta(w_j,x_K)-\beta(w_j,x_k)\right]\right). \quad (30)$$

Equation (30) indicates that the average outcome properties of an allocation depend critically on the complementarity properties of the average match function (AMF). Of particular interest is the difference between a candidate assignment π and the completely random matching $\pi_{jk}^{\mathrm{rdm}}=\rho_j\lambda_k$ for all $j=1,\ldots,J$ and $k=1,\ldots,K$:

$$\theta(\pi')-\theta(\pi^{\mathrm{rdm}})=\sum_{j=1}^{J-1}\sum_{k=1}^{K-1}(\pi'_{jk}-\rho_j\lambda_k)\left(\beta(w_J,x_K)-\beta(w_J,x_k)-\left[\beta(w_j,x_K)-\beta(w_j,x_k)\right]\right). \quad (31)$$

Equation (31) suggests that outcome-maximizing assignments will tend to be *assortative* $(\pi'_{jk}>\rho_j\lambda_k)$ in regions of *complementarity* $(\beta(w_J,x_K)-\beta(w_J,x_k)-[\beta(w_j,x_K)-\beta(w_j,x_k)]>0)$ and *anti-assortative* $(\pi'_{jk}<\rho_j\lambda_k)$ in regions of *substitutability* $(\beta(w_J,x_K)-\beta(w_J,x_k)-[\beta(w_j,x_K)-\beta(w_j,x_k)]<0)$.

The semiparametric efficiency bound for $\theta(\pi)$, for a given *fixed* assignment, π, should follow relatively easily from

Table 1. The structure of feasible assignments

Teachers/Classrooms	x_1	$\cdots$	x_{K-1}	x_K	$f_W(w)$
w_1	π_{11}	$\cdots$	π_{1K-1}	$\rho_1 - \sum_{k=1}^{K-1} \pi_{1k}$	ρ_1
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
w_{J-1}	π_{J-11}	$\cdots$	π_{J-1K-1}	$\rho_{J-1} - \sum_{k=1}^{K-1} \pi_{J-1k}$	ρ_{J-1}
w_J	$\lambda_1 - \sum_{j=1}^{J-1} \pi_{j1}$	$\cdots$	$\lambda_{K-1} - \sum_{j=1}^{J-1} \pi_{jK-1}$	$1 - \sum_{j=1}^{J-1} \rho_j - \sum_{k=1}^{K-1} \lambda_k + \sum_{j=1}^{J-1} \sum_{k=1}^{K-1} \pi_{jk}$	ρ_J
$f_X(x)$	λ_1		λ_{K-1}	λ_K	

NOTES: The $j = 1, \ldots, J$ types of teachers are enumerated in the first column, with the marginal frequency of each type given in the last column. The $k = 1, \ldots, K$ types of classrooms are enumerated in the first row, with the marginal frequency of each type given in the last row. The joint distribution of teachers and classrooms is characterized by the interior probabilities. The feasibility constraints are used to reduce the parameterization of an assignment to $(J-1)(K-1)$ probabilities.

Theorem 1. Likewise an efficient estimate, $\hat{\theta}(\pi)$, should be straightforward to construct, given the availability on an efficient estimate of the AMF at all points in $(w, x) \in \mathbb{W} \times \mathbb{X}$. There remain interesting decision theoretical questions regarding how to implement an optimal assignment on the basis of sample information alone.

APPENDIX: PROOFS

Proof of Lemma 2. To economize on the notation we drop subscripts from densities in what follows. Recall the notation $\tilde{\varepsilon}_i = (\tilde{\varepsilon}_{1i}, \ldots, \tilde{\varepsilon}_{Ki})'$ and $\tilde{\upsilon}^h = (\tilde{\upsilon}_1^h, \ldots, \tilde{\upsilon}_J^h)'$. We begin by factoring the joint density of all firm and worker attributes as

$$f(u, \upsilon, w, x, r, s, \tilde{\varepsilon}, \tilde{\upsilon}) = f(\upsilon, x, s, \tilde{\upsilon} | u, w, r, \tilde{\varepsilon}) f(u, \tilde{\varepsilon} | w, r) f(w, r)$$
$$= f(\upsilon, x, s, \tilde{\upsilon} | u, w, r, \tilde{\varepsilon}) f(\tilde{\varepsilon} | w, r) \times f(u | w, r) f(w, r), \quad \text{(A.1)}$$

where the second equality follows from (14) in the main text. An analogous calculation gives the parallel factorization

$$f(u, \upsilon, w, x, r, s, \tilde{\varepsilon}, \tilde{\upsilon}) = f(u, w, r, \tilde{\varepsilon} | \upsilon, x, s, \tilde{\upsilon}) f(\tilde{\upsilon} | x, s) f(\upsilon | x, s) f(x, s). \quad \text{(A.2)}$$

Dividing (A.1) by $f(w, x, r, s, \tilde{\varepsilon}, \tilde{\upsilon})$ yields

$$f(u, \upsilon | w, x, r, s, \tilde{\varepsilon}, \tilde{\upsilon}) = \frac{f(\upsilon, x, s, \tilde{\upsilon} | u, w, r, \tilde{\varepsilon}) f(\tilde{\varepsilon} | w, r) f(u | w, r) f(w, r)}{f(x, s, \tilde{\upsilon} | w, r, \tilde{\varepsilon}) f(w, r, \tilde{\varepsilon})}$$
$$= \frac{f(\upsilon, x, s, \tilde{\upsilon} | u, w, r, \tilde{\varepsilon}) f(u | w, r)}{f(x, s, \tilde{\upsilon} | u, w, r, \tilde{\varepsilon})}$$
$$= f(\upsilon | w, x, r, s, u, \tilde{\varepsilon}, \tilde{\upsilon}) f(u | w, r), \quad \text{(A.3)}$$

where the second equality follows from (13) of the main text.

Dividing (A.2) by $f(w, x, r, s, \tilde{\varepsilon}, \tilde{\upsilon})$ and invoking (12) yields the parallel result

$$f(u, \upsilon | w, x, r, s, \tilde{\varepsilon}, \tilde{\upsilon}) = f(u | w, x, r, s, \upsilon, \tilde{\varepsilon}, \tilde{\upsilon}) f(\upsilon | x, s). \quad \text{(A.4)}$$

Integrating (A.4) with respect to u gives

$$f(\upsilon | w, x, r, s, u, \tilde{\varepsilon}, \tilde{\upsilon}) = f(\upsilon | x, s). \quad \text{(A.5)}$$

Substituting (A.5) into (A.3) yields the density factorization

$$f(u, \upsilon | w, x, r, s, \tilde{\varepsilon}, \tilde{\upsilon}) = f(u | w, r) f(\upsilon | x, s)$$

as claimed.

Proof of Theorem 1. In calculating the semiparametric efficiency bound for the model defined by (1) and Assumptions 1–4 we follow the general approach of Bickel et al. (1993) and, especially, Newey (1990, sec. 3). First, we characterize the nuisance tangent space. Second, we demonstrate pathwise differentiability of the average match function $\beta_{jk} = \beta(w_j, x_k)$. The efficient influence function is the projection of the pathwise derivative onto the nuisance tangent space. In the present case, the pathwise derivative is an element of the tangent space and therefore coincides with the required projection (i.e., β_{jk} is a parameter of an unrestricted distribution and hence the pathwise derivative is unique; see Newey 1994b). The main result then follows from Theorem 3.1 of Newey (1990, p. 106).

Step 1: Characterization of tangent space

The joint density function of $Z = (W, X, Y, R', S')'$, recalling that

$$p_{jk}(r, s) = \Pr\left(W = w_j, X = x_k \middle| R = r, S = s\right),$$

$\rho_j = \Pr(W = w_j)$ and $\lambda_k = \Pr(X = x_k)$, is conveniently factorized as follows:

$$f(y, w, x, r, s) = \prod_{j=1}^{J} \prod_{k=1}^{K} f\left(y | w_j, x_k, r, s\right)^{d_j e_k} \times f(r, s | w_j, x_k)^{d_j e_k} \Pr\left(W = w_j, X = x_k\right)^{d_j e_k}$$
$$= \prod_{j=1}^{J} \prod_{k=1}^{K} f\left(y | w_j, x_k, r, s\right)^{d_j e_k} \times \left[\frac{f\left(w_j, x_k, r, s\right)}{f\left(w_j, r\right) f(x_k, s)} f\left(r | w_j\right) f(s | x_k) \rho_j \lambda_k\right]^{d_j e_k}$$
$$= \prod_{j=1}^{J} \prod_{k=1}^{K} f\left(y | w_j, x_k, r, s\right)^{d_j e_k} \times \left[\frac{p_{jk}(r, s)}{p_j(r) p_k(s)} \frac{f(r, s)}{f(r) f(s)} f\left(r | w_j\right) f(s | x_k) \rho_j \lambda_k\right]^{d_j e_k},$$

where we suppress the functional dependence of d_j on w and e_k on x. (That is, $D_j = D_j(W) = 1$ if $W = w_j$ and zero otherwise and $E_k = E_k(X) = 1$ if $X = x_k$ and zero otherwise.) Recall also that $p_j(r) = \Pr(W = w_j | R = r)$ and $p_k(s) = \Pr(X = x_k | S = s)$.

Consider a regular parametric submodel with $f(y, w, x, r, s; \eta) = f(y, w, x, r, s)$ at $\eta = \eta_0$. The submodel joint density is given by

$$f(y, w, x, r, s; \eta) = \prod_{j=1}^{J} \prod_{k=1}^{K} f\left(y | w_j, x_k, r, s; \eta\right)^{d_j e_k} \times \left[\frac{p_{jk}(r, s; \eta)}{p_j(r; \eta) p_k(s; \eta)} \frac{f(r, s; \eta)}{f(r; \eta) f(s; \eta)} \times f\left(r | w_j; \eta\right) f(s | x_k; \eta) \rho_j(\eta) \lambda_k(\eta)\right]^{d_j e_k}.$$

The submodel log-likelihood is

$$
\begin{aligned}
&\ln f(y, w, x, r, s; \eta) \\
&= \sum_{j=1}^{J}\sum_{k=1}^{K} d_j e_k \ln f\left(y \middle| w_j, x_k, r, s; \eta\right) \\
&\quad + \sum_{j=1}^{J}\sum_{k=1}^{K} d_j e_k \ln\left\{\frac{p_{jk}(r, s; \eta)}{p_j(r; \eta)\, p_k(s; \eta)} \frac{f(r, s; \eta)}{f(r; \eta) f(s; \eta)}\right\} \\
&\quad + \sum_{j=1}^{J}\sum_{k=1}^{K} d_j e_k \ln f\left(r \middle| w_j; \eta\right) + \sum_{j=1}^{J}\sum_{k=1}^{K} d_j e_k \ln f(s | x_k; \eta) \\
&\quad + \sum_{j=1}^{J}\sum_{k=1}^{K} d_j e_k \ln \rho_j(\eta) + \sum_{j=1}^{J}\sum_{k=1}^{K} d_j e_k \ln \lambda_k(\eta) \\
&= \sum_{j=1}^{J}\sum_{k=1}^{K} d_j e_k \ln f\left(y \middle| w_j, x_k, r, s; \eta\right) \\
&\quad + \sum_{j=1}^{J}\sum_{k=1}^{K} d_j e_k \ln\left\{\frac{p_{jk}(r, s; \eta)}{p_j(r; \eta)\, p_k(s; \eta)} \frac{f(r, s; \eta)}{f(r; \eta) f(s; \eta)}\right\} \\
&\quad + \sum_{j=1}^{J} d_j \ln f\left(r \middle| w_j; \eta\right) + \sum_{k=1}^{K} e_k \ln f(s | x_k; \eta) \\
&\quad + \sum_{j=1}^{J} d_j \ln \rho_j(\eta) + \sum_{k=1}^{K} e_k \ln \lambda_k(\eta),
\end{aligned}
$$

with an associated score vector of

$$
\begin{aligned}
s_\eta(y, w, x, r, s; \eta) &= \sum_{j=1}^{J}\sum_{k=1}^{K} d_j e_k s_\eta\left(y \middle| w_j, x_k, r, s; \eta\right) \\
&\quad + \sum_{j=1}^{J}\sum_{k=1}^{K} d_j e_k k_\eta\left(w_j, x_k, r, s; \eta\right) \\
&\quad + \sum_{j=1}^{J} d_j t_\eta\left(r \middle| w_j; \eta\right) + \sum_{k=1}^{K} e_k t_\eta(s | x_k; \eta) \\
&\quad + \sum_{j=1}^{J} d_j \rho_{j,\eta}(\eta) + \sum_{k=1}^{K} e_k \lambda_{k,\eta}(\eta), \qquad \text{(A.6)}
\end{aligned}
$$

where

$$
\begin{aligned}
s_\eta\left(y \middle| w_j, x_k, r, s; \eta\right) &= \nabla_\eta \ln f\left(y \middle| w_j, x_k, r, s; \eta\right) \\
k_\eta\left(w_j, x_k, r, s; \eta\right) &= \nabla_\eta \ln p_{jk}(r, s; \eta) - \nabla_\eta \ln p_j(r; \eta) \\
&\quad - \nabla_\eta \ln p_k(s; \eta) \\
&\quad + \nabla_\eta \ln f(r, s; \eta) - \nabla_\eta \ln f(r; \eta) \\
&\quad - \nabla_\eta \ln f(s; \eta) \\
t_\eta\left(r \middle| w_j; \eta\right) &= \nabla_\eta \ln f\left(r \middle| w_j; \eta\right) \\
t_\eta(s | x_k; \eta) &= \nabla_\eta \ln f(s | x_k; \eta) \\
\rho_{j,\eta}(\eta) &= \frac{\partial \ln \rho_j(\eta)}{\partial \eta} \\
\lambda_{k,\eta}(\eta) &= \frac{\partial \ln \lambda_k(\eta)}{\partial \eta}.
\end{aligned}
$$

By the usual conditional mean zero property of the score function,

$$
\begin{aligned}
\mathbb{E}[s_\eta(Y | W, X, R, S) | W, X, R, S] &= 0 \\
\mathbb{E}[k_\eta(W, X, R, S)] &= 0 \\
\mathbb{E}[t_\eta(R | W) | W] &= 0 \\
\mathbb{E}[t_\eta(S | X) | X] &= 0, \qquad \text{(A.7)}
\end{aligned}
$$

where the suppression of η in a function means that it is evaluated at is population value (e.g., $t_\eta(S|X) = t_\eta(S|X; \eta_0)$).

From (A.6) and (A.7), the tangent set is therefore given by

$$
\begin{aligned}
\mathcal{T} = \Bigg\{ &\sum_{j=1}^{J}\sum_{k=1}^{K} d_j e_k s\left(y \middle| w_j, x_k, r, s\right) + \sum_{j=1}^{J}\sum_{k=1}^{K} d_j e_k k(w_j, x_k, r, s) \\
&+ \sum_{j=1}^{J} d_j t\left(r \middle| w_j\right) + \sum_{k=1}^{K} e_k t(s | x_k) + \sum_{j=1}^{J} d_j a_j + \sum_{k=1}^{K} e_k b_k \Bigg\}, \qquad \text{(A.8)}
\end{aligned}
$$

where a_j and b_k are finite constants for $j = 1, \ldots, J$ and $k = 1, \ldots, K$ and $s(y|w_j, x_k, r, s)$, $k(w_j, x_k, r, s)$, $t(r|w_j)$, and $t(s|x_k)$ satisfy

$$
\begin{aligned}
\mathbb{E}[s(Y | W, X, R, S) | W, X, R, S] &= 0 \\
\mathbb{E}[k(W, X, R, S)] &= 0 \\
\mathbb{E}[t(R | W) | W] &= 0 \\
\mathbb{E}[t(S | X) | X] &= 0.
\end{aligned}
$$

Step 2: Demonstration of pathwise differentiability

Under the parametric submodel $\beta(\eta)$ is identified by

$$
\begin{aligned}
&\beta(w, x; \eta) \\
&= \int\int\left[\int y f(y | w, x, r, s; \eta)\, dy\right] f(r | w; \eta) f(s | x; \eta)\, dr ds.
\end{aligned}
$$

Differentiating under the integral and evaluating at $\eta = \eta_0$ gives

$$
\begin{aligned}
&\frac{\partial \beta(w, x; \eta_0)}{\partial \eta'} \\
&= \int\int \mathbb{E}\left[Y s_\eta(Y | W, X, R, S) \middle| w, x, r, s\right] f(r | w; \eta_0) f(s | x; \eta_0)\, dr ds \\
&\quad + \int\int q(w, x, r, s) t_\eta(r | w)' f(r | w; \eta_0) f(s | x; \eta_0)\, dr ds \\
&\quad + \int\int q(w, x, r, s) t_\eta(s | x)' f(r | w; \eta_0) f(s | x; \eta_0)\, dr ds \\
&= \int\int \mathbb{E}\left[Y s_\eta(Y | W, X, R, S) \middle| w, x, r, s\right] f(r | w; \eta_0) f(s | x; \eta_0)\, dr ds \\
&\quad + \int e_S(w, x, r) t_\eta(r | w)' f(r | w; \eta_0)\, dr \\
&\quad + \int e_R(w, x, s) t_\eta(s | x)' f(s | x; \eta_0)\, ds \\
&= \int\int \mathbb{E}\left[Y s_\eta(Y | W, X, R, S) \middle| w, x, r, s\right] f(r | w; \eta_0) f(s | x; \eta_0)\, dr ds \\
&\quad + \mathbb{E}\left[e_S(w, x, R) t_\eta(R | w)' \middle| w\right] \\
&\quad + \mathbb{E}\left[e_R(w, x, S) t_\eta(S | x)' \middle| x\right], \qquad \text{(A.9)}
\end{aligned}
$$

where

$$e_S(w,x,r) = \int q(w,x,r,s)\, f(s|\, x;\eta_0)\, ds$$
$$e_R(w,x,s) = \int q(w,x,r,s)\, f(r|\, w;\eta_0)\, dr.$$

To demonstrate pathwise differentiability of $\beta_{jk} = \beta(w_j, x_k)$, we require $F(Y, w_j, x_k, R, S)$ such that

$$\frac{\partial\beta(w_j,x_k;\eta_0)}{\partial\eta'} = \mathbb{E}[F(Y,w_j,x_k,R,S)s_\eta(Y,w_j,x_k,R,S)']. \quad \text{(A.10)}$$

With some work it is possible to show that condition (A.10) holds for

$$\begin{aligned} F\left(Y,w_j,x_k,R,S\right) &= \frac{f\left(R|\, w_j\right) f\left(S|\, x_k\right)}{f(R,S)} \frac{D_j E_k}{p_{jk}(R,S)}\left(Y - q\left(w_j,x_k,R,S\right)\right) \\ &\quad + \frac{D_j}{\rho_j}\left(e_S\left(w_j,x_k,R\right) - \beta_{jk}\right) + \frac{E_k}{\lambda_k}\left(e_R\left(w_j,x_k,S\right) - \beta_{jk}\right). \quad \text{(A.11)} \end{aligned}$$

We evaluate the covariance of each of the three terms in (A.11) with $s_\eta(Y, w_j, x_k, R, S)$ in turn.

We begin with

$$\begin{aligned} &\mathbb{E}\Bigg[\frac{f\left(R|\, w_j\right) f\left(S|\, x_k\right)}{f(R,S)}\frac{D_j E_k}{p_{jk}(R,S)}\left(Y - q\left(w_j,x_k,R,S\right)\right) \\ &\quad \times D_j E_k s_\eta\left(Y,w_j,x_k,R,S\right)\Bigg] \\ &= \mathbb{E}\Bigg[\frac{f(R|w_j)f(S|x_k)}{f(R,S)}\frac{D_j E_k}{p_{jk}(R,S)} \\ &\quad \times \left(Y - q\left(w_j,x_k,R,S\right)\right) D_j E_k s_\eta\left(Y|\, w_j,x_k,R\right)\Bigg] \\ &\quad + \mathbb{E}\Bigg[\frac{f\left(R|\, w_j\right) f\left(S|\, x_k\right)}{f(R,S)}\frac{D_j E_k}{p_{jk}(R,S)} \\ &\quad \times \left(Y - q\left(w_j,x_k,R,S\right)\right) D_j E_k \frac{\partial \log f\left(w_j,x_k,R,S;\eta_0\right)}{\partial\eta'}\Bigg] \\ &= \mathbb{E}\left[\frac{f\left(R|\, w_j\right) f\left(S|\, x_k\right)}{f(R,S)}\frac{D_j E_k}{p_{jk}(R,S)} Y s_\eta\left(Y|\, w_j,x_k,R,S\right)\right] \\ &= \mathbb{E}\Bigg[\frac{f\left(R|\, w_j\right) f\left(S|\, x_k\right)}{f(R,S)}\frac{D_j E_k}{p_{jk}(R,S)} \\ &\quad \times \mathbb{E}\left[Y s_\eta\left(Y|\, w_j,x_k,R,S\right)\big|\, w_j,x_k,R,S\right]\Bigg] \\ &= \int\int \sum_{l=1}^{J}\sum_{m=1}^{K} \frac{f\left(r|\, w_j\right) f\left(s|\, x_k\right)}{f(r,s)}\frac{D_j(w_l)E_k(x_m)}{p_{jk}(r,s)} \\ &\quad \times \mathbb{E}\left[Y s_\eta\left(Y|\, w_j,x_k,R,S\right)\big|\, w_j,x_k,r,s\right] p_{lm}(r,s)\Big] f(r,s)\,\mathrm{d}r\mathrm{d}s \\ &= \int\int \mathbb{E}\left[Y s_\eta\left(Y|\, w_j,x_k,R,S\right)\big|\, w_j,x_k,r,s\right] \\ &\quad \times f\left(r|\, w_j\right) f\left(s|\, x_k\right) drds, \end{aligned}$$

which coincides with the first component of (A.9). The second equality above follows by iterated expectations and the conditional mean zero property of the score function. The third and fourth equalities follow from applications of iterated expectations.

To evaluate the covariance of the second two terms in (A.11) with $s_\eta(Y, w_j, x_k, R, S; \eta_0)$ the following alternative density factorizations will prove useful:

$$f(w,x,r,s;\eta) = f(r|\, w;\eta)\, f(x,s|\, w,r;\eta)\, f(w;\eta)$$
$$f(w,x,r,s;\eta) = f(s|\, x;\eta)\, f(w,r|\, x,s;\eta)\, f(x;\eta).$$

These give, in an abuse of notation, the score decompositions

$$\begin{aligned} s_\eta(Y,W,X,R,S;\eta) &= s_\eta(Y|\, W,X,R,S;\eta) + t_\eta(R|\, W;\eta) \\ &\quad + s_\eta(X,S|\, W,R;\eta) + s_\eta(W;\eta) \\ s_\eta(Y,W,X,R,S;\eta) &= s_\eta(Y|\, W,X,R,S;\eta) + t_\eta(S|\, X;\eta) \\ &\quad + s_\eta(W,R|\, X,S;\eta) + s_\eta(X;\eta). \end{aligned}$$

By the conditional mean zero property of the score function

$$\mathbb{E}\left[\frac{D_j}{\rho_j}\left(e_S\left(w_j,x_k,R\right) - \beta_{jk}\right) s_\eta(Y|\, W,X,R,S)\right] = 0.$$

Using iterated expectations further yields

$$\begin{aligned} &\mathbb{E}\left[\frac{D_j}{\rho_j}\left(e_S\left(w_j,x_k,R\right) - \beta_{jk}\right) s_\eta(W)\right] \\ &= \mathbb{E}\left[s_\eta(W)\,\mathbb{E}\left[\frac{D_j}{\rho_j}\left(e_S\left(w_j,x_k,R\right) - \beta_{jk}\right)\Big|\, W\right]\right] \\ &= s_\eta\left(w_j\right)\mathbb{E}\left[\left(e_S\left(w_j,x_k,R\right) - \beta_{jk}\right)\big|\, W = w_j\right] \\ &= s_\eta\left(w_j\right)\left(\beta_{jk} - \beta_{jk}\right) \\ &= 0. \end{aligned}$$

Similarly, using iterated expectations and the conditional mean zero property of the score function, yields

$$\begin{aligned} &\mathbb{E}\left[\frac{D_j}{\rho_j}\left(e_S\left(w_j,x_k,R\right) - \beta_{jk}\right) s_\eta(X,S|\, W,R)\right] \\ &= \mathbb{E}\left[\frac{D_j}{\rho_j}\left(e_S\left(w_j,x_k,R\right) - \beta_{jk}\right)\mathbb{E}\left[s_\eta(X,S|\, W,R)\big|\, W,R\right]\right] \\ &= \mathbb{E}\left[\frac{D_j}{\rho_j}\left(e_S\left(w_j,x_k,R\right) - \beta_{jk}\right)\cdot 0\right] \\ &= 0. \end{aligned}$$

Finally,

$$\begin{aligned} &\mathbb{E}\left[\frac{D_j}{\rho_j}\left(e_S\left(w_j,x_k,R\right) - \beta_{jk}\right) t_\eta(R|\, W)\right] \\ &= \mathbb{E}\left[\frac{D_j}{\rho_j}\left(e_S\left(w_j,x_k,R\right) - \beta_{jk}\right) t_\eta(R|\, W)'\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\frac{D_j}{\rho_j}\left(e_S\left(w_j,x_k,R\right) - \beta_{jk}\right) t_\eta(R|\, W)'\Big|\, W\right]\right] \\ &= \mathbb{E}\left[\left(e_S\left(w_j,x_k,R\right) - \beta_{jk}\right) t_\eta(R|\, W)'\big|\, W = w_j\right] \\ &= \mathbb{E}\left[e_S\left(w_j,x_k,R\right) t_\eta(R|\, W)'\big|\, W = w_j\right], \end{aligned}$$

again using the conditional mean zero property of the score function. Putting these results together gives

$$\mathbb{E}\left[\frac{D_j}{\rho_j}(e_S(w_j, x_k, R) - \beta_{jk})s_\eta(Y, W, X, R, S; \eta_0)\right]$$
$$= \mathbb{E}[e_S(w_j, x_k, R)t_\eta(R|W)'|W = w_j].$$

Analogous calculations yield the expression

$$\mathbb{E}\left[\frac{E_k}{\lambda_k}(e_R(w_j, x_k, S) - \beta_{jk})s_\eta(Y, W, X, R, S; \eta_0)\right]$$
$$= \mathbb{E}[e_R(w_j, x_k, S; \eta_0)t_\eta(S|X)'|X = x_k].$$

These expressions coincide with the second and third components of (A.9). Condition (A.10) then holds for $F(Y, w_j, x_k, R, S)$ as defined in (A.11).

Step 3: Calculation of projection

The semiparametric variance bound for β_{jk} is the expected square of the projection of $F(Y, w_j, x_k, R, S)$ onto $\mathcal{T}$. Since $F(Y, w_j, x_k, R, S) \in \mathcal{T}$ it coincides with the required projection and is therefore the efficient influence function as claimed. Here,

$$\frac{f\left(R|\, w_j\right) f\left(S|\, x_k\right)}{f\left(R, S\right)} \frac{D_j E_k}{p_{jk}\left(R, S\right)}(Y - q(w_j, x_k, R, S))$$

plays the role of $\sum_{j=1}^{J}\sum_{k=1}^{K} d_j e_k s(y|w_j, x_k, r, s)$ and

$$\frac{D_j}{\rho_j}(e_S(w_j, x_k, R) - \beta_{jk})$$

and

$$\frac{E_k}{\lambda_k}(e_R(w_j, x_k, S) - \beta_{jk})$$

the roles of, respectively, $d_j t(r|w_j)$ and $e_k t(s|x_k)$. Zeros play the role of the remaining terms.

ACKNOWLEDGMENTS

This article was presented at the 2012 European Summer Meetings of the Econometric Society, the December 2012 SFB 884 Research Conference on the Evaluation of Political reforms at the University of Mannheim, and at seminars hosted by the University of California - Berkeley, University of California - Davis, University of Wisconsin - Madison, Northwestern University, the University of Southern California, and Singapore Management University. The authors thank these seminar audiences for useful feedback. The authors are also grateful to Stephane Bonhomme, Richard Blundell, Konrad Menzel, and James Powell for useful discussions. The current draft reflects valuable suggests made by a co-editor, associate editor, and two referees. Financial support from the National Science Foundation (SES #0820361) for the first author's contribution is gratefully acknowledged. All the usual disclaimers apply.

FUNDING

Division of Social and Economic Sciences [0820361].

[Received February 2017. Revised May 2018.]

REFERENCES

Becker, G. S. (1973), "A Theory of Marriage: Part I," *Journal of Political Economy*, 81, 813–846. [307]

Bickel, P. J., Klaassen, C. A. J., Ritov, Y., and Wellner, J. A. (1993), *Efficient and Adaptive Estimation for Semiparametric Models*, New York: Springer-Verlag. [312]

Boyd, D., Lankford, H., Loeb, S., and Wyckoff, J. (2013), "Analyzing the Determinants of the Matching of Public School Teachers to Jobs: Disentangling the Preferences of Teachers And Employers," *Journal of Labor Economics*, 31, 83–117. [303]

Chamberlain, G. (1984), "Panel Data," in *Handbook of Econometrics* (Vol. 2), eds. Z. Griliches, and M. D. Intriligator, Amsterdam: North Holland, pp. 1247–1318. [304,309]

Chiappori, P.-A., and Salanié, B. (2016), "The Economics of Matching Models," *Journal of Economic Literature*, 54, 832–861. [311]

Chiappori, P.-A, Salanié, B., and Weiss, Y. (2015), "Partner Choice and the Marital College Premium," Mimeo, Columbia University, New York. [304]

Choo, E., and Siow, A. (2006a), "Who Marries Whom and Why?" *Journal of Political Economy*, 114, 175–201. [304,306,307,311]

——— (2006b), "Estimating a Marriage Matching Model with Spillover Effects," *Demography*, 43, 464–490. [304,306,307,311]

Dagsvik, J. K. (2000), "Aggregation in Matching Markets," *International Economic Review*, 41, 27–58. [304]

Dupuy, A., and Galichon, A. (2014), "Personality Traits and the Marriage Market," *Journal of Political Economy*, 122, 1271–1319. [304]

Freedman, D. (1997), "From Association to Causation via Regression," *Advances in Applied Mathematics*, 18, 59–110. [310]

Galichon, A., and Hsieh, Y. (2015), "Love and Chance: Equilibrium and Identification in a Large NTU Matching Market with Stochastic Choice," Mimeo, University of Southern California, Los Angeles, CA. [304]

Galichon, A., and Salanié, B. (2015), "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models," Mimeo, New York University, New York. [304,307,309]

Graham, B. S. (2008), "Identifying Social Interactions Through Conditional Variance Restrictions," *Econometrica*, 76, 643–660. [306,311]

——— (2011a), "Econometric Methods for the Analysis of Assignment Problems in the Presence of Complementarity and Social Spillovers," in *Handbook of Social Economics* (Vol. 1B), eds. J. Benhabib, A. Bisin, and M. Jackson, Amsterdam: North-Holland, pp. 965–1052. [303,305,306,311]

——— (2011b), "Efficiency Bounds for Missing Data Models with Semiparametric Restrictions," *Econometrica*, 79, 437–452. [310]

——— (2013), "Comparative Static and Computational Methods for an Empirical One-to-One Transferable Utility Matching Model," *Advances in Econometrics: Structural Econometric Models*, 31, 151–179. [304,308,309]

Graham, B. S., de Xavier Pinto, C. C., and Egel, D. (2012), "Inverse Probability Tilting for Moment Condition Models with Missing Data," *Review of Economic Studies*, 79, 1053–1079. [311]

——— (2016), "Efficient Estimation of Data Combination Models by the Method of Auxiliary-to-Study Tilting (AST)," *Journal of Business and Economic Statistics*, 34, 288–301. [311]

Graham, B. S., Imbens, G. W., and Ridder, G. (2007), "Redistributive Effects for Discretely-Valued Inputs," IEPR Working Paper No. 07.7., IEPR. [305,311]

——— (2010), "Measuring the Effects of Segregation in the Presence of Social Spillovers: A Nonparametric Approach," NBER Working Paper No. 16499, NBER. [310]

——— (2014), "Complementarity and Aggregate Implications of Assortative Matching: A Nonparametric Analysis," *Quantitative Economics*, 5, 29–66. [305,307,311]

Griliches, Z., and Mairesse, J. (1998), "Production Functions: The Search for Identification," in *Econometrics and Economic Theory in the 20th Century: The Ragner Frisch Memorial Symposium*, Cambridge University Press, pp. 169–203. [309]

Hahn, J. (1998), "On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects," *Econometrica*, 66, 315–331. [306]

Heckman, J. J., Smith, J., and Clements, N. (1997), "Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts," *Review of Economic Studies*, 64, 487–535. [304,309]

Hirano, K., Imbens, G. W., and Ridder, G. (2003), "Efficient Estimation of Average Treatment Effects using the Estimated Propensity Score," *Econometrica*, 71, 1161–1189. [311]

Holland, P. W. (1986), "Statistics and Causal Inference," *Journal of the American Statistical Association*, 81, 945–960. [304]

Imbens, G. W. (2004), "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review," *Review of Economics and Statistics*, 86, 4–29. [304,306,309]

Imbens, G. W., and Rubin, D. B. (2015), *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*, Cambridge: Cambridge University Press. [304]

Imbens, G. W., and Wooldridge, J. (2009), "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, 47, 5–86. [305]

Keiding, N., and Clayton, D. (2014), "Standardization and Control for Confounding in Observational Studies: A Historical Perspective," *Statistical Science*, 29, 529–558. [310]

Manski, C. F. (2007), *Identification for Prediction and Decision*, Cambridge, MA: Harvard University Press. [304]

Menzel, K. (2015), "Large Matching Markets as Two-Sided Demand Systems," *Econometrica*, 83, 897–941. [305]

Newey, W. K. (1990), "Semiparametric Efficiency Bounds," *Journal of Applied Econometrics*, 5, 99–135. [310,312]

——— (1994a), "Kernel Estimation of Partial Means and a General Variance Estimator," *Econometric Theory*, 10, 233–253. [311]

——— (1994b), "The Asymptotic Variance of Semiparametric Estimators," *Econometrica*, 62, 1349–1382. [310,312]

Olley, G. S., and Pakes, A. (1996), "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64, 1263–1297. [304,309]

Shapley, L. S., and Shubik, M. (1971) "The Assignment Game I: The core," *International Journal of Game Theory*, 1, 111–130. [307]

Yule, G. U. (1897), "An Investigation into the Causes of Changes in Pauperism in England, Chiefly during the Last Two Intercensal Decades (Part I)," *Journal of the Royal Statistical Society*, 62, 249–295. [305]