Comments for MFI Conference;

Chicago, May 20011;

Ariel Pakes, Harvard University.

The papers consider three “market structures”, with “matches” between two sides of a market.

Matches have externalities if the benefits from A’s match to B depend on whether C is matched to D (for $A \neq C$ and $B \neq D$), and peer effects if the benefits from A’s match to B depend on whether A is matched to C.

Market structures.

- A centralized matching market without "externalities" (and usually without peer effects). Students to schools. Residents to programmes.

- A decentralized matching market without "externalities". Marriage, and certain labor, housing, and used equiptment markets.

- A decentralized matching market with "externalities". I will focus on "Buyer-Seller" networks.

Centralized match. The mechansim is announced, at least one side submits preferences, and the mechanism yields possible outcomes. Easy to work with if submitted preferences are truthful, and the mechanism yields a unique outcome given the preferences. I start with that case.

Typical applied issues:

- evaluate policy changes that would change the underlying primitives of the problem (add a school...).

- Compare alternative mechanisms (market equilibrium where one side sets wages, ...). This will usually require additional assumptions; both informational and procedural (and may not yield a unique equilibrium).

For either of these we need to estimate the distribution of preferences (on each side of the market). If we went one step and wanted to analyze

- the gains from the match (in terms of human capital, or anything else society might think useful)

we would need to estimate a production function which conditions on the selection induced by the mechanism (unless the preferences are derived from the outcome of interest, in which case we may be able to unravel the production function from the match itself).

Typical Data. Includes characteristics of both sides of the market, actual match, and sometimes

- stated preferences by at least one side of the market.

If truth telling is a dominant strategy, stated preferences yields a discrete choice problem (independent of mechanism).
Data contains more information that usual (have an ordering rather than an indicator for maximal utility).
Generates an ability to estimate detailed preference intensities (e.g. MicroBLP).
If analyzed done jointly with data on match, should be able to determine preferences on the other side of the market.

Estimation from match data: outcome unique. Functional form assumptions + mechanism $\Rightarrow$ probability of an allocation.

Use simulation that mimics the preferences for different values of the parameter vector to obtain probabilities conditional on θ.

Simulated mle is problematic but lots of other estimators available.

Depending on complexity of mechanism and the number of agents on each side of the market, there may be a computational problem here, but there is no conceptual issue.

E.g. **Agarwal** (thesis in process). One side has vertical preferences (all agree the ranks of the other side). The order is determined by an unknown quality index

$$q = x\alpha + \nu, \ \ \nu \sim N(0, 1),$$

where α is a vector of parameters.

$s \in S$ is an ordering of the N individuals, with $s(j)$ the j^{th} individual in that order, and $M(\cdot)$ provides the observed match

Deferred acceptance algorithm plus vertical assumption gives likelihood

$$Pr(M|s) = \sum_{j} Pr\{M(s_j) = max_{C - \cup_{q=1}^{j-1} M(s_q)} U(s_j; \beta)\}$$

where $U(\cdot)$ is a discrete choice utility function, and

$$Pr(M) = \sum_{s \in S} Pr(M|s) P\{s|\alpha\}$$

Estimate: Simulate ns on $\nu^s \equiv [\nu_1^s, \ldots, \nu_N^s]$. For a fixed α run draws through mechanism to get order of choice, then use discrete choice model above to evaluate objective function. Search over α.

Consistency proof usually uses a uniform LLN which is not obvious here (at least if the limiting dimension is not the number of independent markets). There are other ways to get consistency,..., but I don't think we know much. Of course if we know the order of choice we can estimate the utility function parameters consistently.

Estimation from match data: if outcome is not unique. Only have pairwise stability. This + functional form assumptions lead to inequalities that only certain vectors of parameters will satisfy.

Typically produces a "set estimator" (i.e. we accept any value of the parameter which satisfies all the inequalities, or that could satisfy all the inequalites given the possibilities induced by sampling error).

Primitives need to be specified with disturbances (else zero probability events occur). Reasonable properties for disturbances vary with type of data, and determine how to estimate. I will come back to this in the context of buyer seller networks where uniqueness is highly unlikely.

Though estimation often feasible when there is no unique outcome, counterfactuals more problematic (work with sets, or some form of selection, or learning, rule). From an applied point of view, this is the major problem with non-uniqueness.

Estimation of output function. y = output measure and assume a production function which depends on q_i and the match $M(i) = m$

$$y_i = g(m, q_i; \gamma) + \epsilon.$$

where ϵ is conditional mean zero. Then

$$E[y|x_i, M(i) = m] =$$

$$\int_\nu g(m, x_i'\beta + \nu) dF(\nu|x_i, m = M(i)),$$

i.e. to compute the regression function we need to use the allocation mechanism. A person with a "low" x_i who got into a high quality m is expected to have a larger ν then a person who got in with a "high" x.

Note. Even if preferences do not have peer effects, you might want to allow for them here. Also room for regression discontinuity estimators here.

Decentralized match with transfers (but without externalities).

Rules of the game: implicit.

Sufficient for every agent to check with every agent on opposite side (could the two generate a match which improves both their utility), but in many applications this is hard to swallow.
Leads to question of what would happen with more detailed behavioral and/or informational assumptions.

Paper: Chiappori, McCann, Nesheim. Assumptions guarantees a unique pairwise stable match.
Formal comparison to matching literature.
Conditions when assumptions imply existence of (a unique) hedonic function which prices

characteristics of the sellers' products.
Some attention given to computation of match outcomes conditional on the joint distribution of buyers and seller types; as we will see important for applied work.

Note. If there is a unique hedonic function, *and* agents knew it plus the vectors of seller characteristics available, an assumption on maximizing utility rationalizes the match; but the reasoning is circular.

Typical empirical issues.

- If the hedonic function were known (or estimable) and *stable*, we could use it to provide a signal for investment decisions (schooling choices, construction of housing, ...).

- If the hedonic function were stable, or if we could *compute* what it would look like were we to perturb primitives, we could evaluate likely re-allocations after possible policy changes.

- If there is a production function for outputs delivered by the match, and we want to estimate it, we will need the matching mechansim to correct for selection (as above).

What has been learned in Industrial Organization which might be useful?

From consumer goods markets.

- Explicit game form: Nash in prices on producer side, consumers know prices and characteristics of goods marketed (sometimes extended to allow for search and/or variable choice sets). Does well within market and time period, but not so well over time.

- This, functional form assumptions, and standard market level data, enable estimation of distribution of consumer and producer attributes.

- Notice that we neither require uniqueness nor purity of strategies for identification or estimation (use necessary conditions).

Game form plus distribution of consumer and producer attributes can be used to generate a “hedonic” function (regress prices on characteristics). Result is used in price indices, and we have alot of empirical experience with them. Consistent with the theory, empirical work indicates that the hedonic function:

- Is not stable across markets or time periods.

Suggests: Focus applied work on

- Estimation of distribution of primitives and

- Computational algorithms for computing equilibria conditional on primitives.

and not on estimation of hedonic function.

Paper: Fox and Yang.

Focus: sufficient conditions for identifying the joint distribution of outcomes *from all possible* matches using only the data generated by the match and weak functional form restrictions.

No interest in transfer (here the hedonic) function $\Rightarrow$ particularly relevant for this setting.

Identification: across markets with different characteristics (assumes lots of markets with very different characteristics).

Alternative is to consider what can be learned from replications of "markets" in your data.

To use this with our data need:
(i) actual estimators (need relevant functional form and distributional assumptions), and
(ii) computational algorithm.

Industrial Organization on upstream (or vertical) markets.

Characterization of Market:" Producer" sells to "retailer" who re-markets the goods (or services) to consumers. Small number of agents and competition both between producing firms and between retailing firms generate "externalities".

- We are in a multi-lateral bargaining situation (not a large number of agents on one side). $\Rightarrow$ Nash in prices inappropriate.

- No agreement on game form except that it has to allow for externalities. $\Rightarrow$ multiplicity a bigger problem.

- Rules out computing probabilities of match outcomes (at least without assumptions on selection).

- "No profitable deviation at observed match" is not a necessary condition, but it is often assumed (at least in expectations). Meant to capture idea of a "rest point" (abstracts from staggered contracting and costs of renegotiation).

Feasible deviations define a less preferred counterfactual. Not always obvious what they are, and often not easy to compute. If we agree on them and can characterizes the counterfactual profits we can use the inequalities they generate to

- characterize the equilibria, and

- as a basis for esitmation.

Question: could you follow a similar strategy in transferable utility model (we know something about feasible choices...)?

Typical Data. Who matches with whom **plus**:
(i) quantities re-marketed by buyer to consumers and price re-marketed at (sometimes also characteristics of consumer who bought), and
(ii) either cost of production of seller or something about cost function's structure (e.g.: constant marginal cost).

Contracts between buyer and seller **typically proprietary**, so we do not know transfer function.

Use of match data in empirical work: Determinants of contract terms (often only part of the problem we do not have direct information on). Important because contracts split the profits from the relationship, and therefore determine

- producer investment incentives, and
- retailer's cost (a determinant of price remarketed to consumers at).

Makes analysis somewhat orthogonal to Fox and Yang (they assume different data and objects of interest).

E.g. Simple HMO-Hospital Networks.

HMO Profits. Premium revenues minus transfer to hospitals ($T_{\mathtt{m},h}(\cdot)$)

$$\pi^{\mathtt{m}}(H_{\mathtt{m}}, H_{-\mathtt{m}}, z) =$$
$$p_{\mathtt{m}}(H_{\mathtt{m}}, H_{-\mathtt{m}}, z) q_{\mathtt{m}}(H_{\mathtt{m}}, H_{-\mathtt{m}}, z) - \sum_{h \in H_{\mathtt{m}}} T_{\mathtt{m},h}(H_{\mathtt{m}}, H_{-\mathtt{m}}, z),$$

where H_m is the network of m's hospitals (vector of ones and zeroes), and $q_{\mathtt{m}}(\cdot)$, $p_{\mathtt{m}}(\cdot)$ and are computed from the solution to the revenue setting game and illness proclivities.

Hospital Profits. If c_h is the per patient costs of hospital h and M_h is network of HMO's

$$\pi^{h}(M_h, M_{-h}, z) =$$
$$\sum_{\mathtt{m} \in M_h} T_{\mathtt{m},h}(H_{\mathtt{m}}, H_{-\mathtt{m}}, z) - c_h \sum_{\mathtt{m} \in M_h} q_{\mathtt{m},h}(H_{\mathtt{m}}, H_{-\mathtt{m}}, z).$$

Simplest behavioral assumption. Negotiations accrue in separate rooms with no communication between them, and no re-negotiation.

If $\chi_{m,h}$ is the indicator function for contract

$$\chi_{m,h}\mathcal{E}_h[\pi^h(M_h, M_{-h}, z) - \pi^h(M_{h/m}, M_{-h}, z)] \geq 0$$

while if it is not made

$$[1-\chi_{m,h}]\mathcal{E}_h[\pi^h(M_h, M_{-h}, z)-\pi^h(M_{h\cup m}, M_{-h}, z)] \geq 0$$

and there are similar inequalities for the HMO.

Analysis. Assume we have a model which can compute HMO premiums and hospital costs from existing and counterfactual networks up to an error which is mean zero conditional on some information set available to agents when decisions were made; i.e. if $R(\cdot)$ is revenue from premiums

$$R^o_m(H_m, H_{-m}) = R_m(H_m, H_{-m}) + \nu_{1,H_m,H_{-m}},$$

and if $C^o_h(\cdot)$ is hospital costs then

$$C^o_h(M_h, M_{-h}) = C_h(M_h, M_{-h}) + \nu_{1,M_h,M_{-h}},$$

and that we have a reduced form model for transfers

$$T_{m,h}(\cdot) = X_{m,h}\beta$$

Then

$$\mathcal{E}_m \chi_{m,h}[R^o_m(H_m, H_{-m}) - R^o_m(H_{m/h}, H_{-m}) - \sum_h X_{m,h}\beta]$$

is positive and, so under standard regularity conditions sample averages, can be used to bound β in one direction, and the similar inequality for the hospital will bound it from the other.

If there is an error in your reduced form model, i.e. if

$$T_{m,h}(\cdot) = X_{m,h}\beta + \nu_{2,m,h}$$

where $\nu_{2,m,h}$ is known to the agents, but not to the econometrician, the positive inequality

$$\mathcal{E}_m \chi_{m,h}[R^o_m(H_m, H_{-m}) - R^o_m(H_{m/h}, H_{-m}) - \sum_h X_{m,h}\beta]$$

$$-\mathcal{E}_m \chi_{m,h} \sum_h \nu_{2,m,h}$$

and this last term will generate a selection bias. I.e. even if the $\nu_{2,m,h}$ are mean zero, those associated with positive profits for the HMO will be lower cost given $X_{m,h}$.

There are inequalities that will not have this selection bias in this model, but they are different.

Form unconditional averages; take difference between actual and counterfactual profits of the hospital when the contract is formed (gains transfer error) and difference between actual and counterfactual profits when HMO when no contract (saves transfer; see Pakes, 2010).

Final points.

- We are unlikely to make progress on these issues without bringing in direct data on demand and costs; it is just too much to ask from match data to give us that information as well as the determinants of contracts.

- Very careful attention needs to be paid to the source of error for these problems. Generally you will require both errors that are subject to selection **and** errors that are not.