# An Empirical, Repeated Matching Game Applied to Market Thickness and Switching

## PRELIMINARY AND INCOMPLETE

Jeremy T. Fox*

University of Chicago

March 2008

**Abstract**

I introduce a repeated two-sided matching game. Both workers and firms are forward looking, have perfect information about all potential partners, and can switch matches each period. Wages are determined each period in a transferable utility matching game that equates supply and demand for each position. Estimation takes about as much computer time as solving the game once. I use data on almost all elite Swedish engineers in the private sector from 1970–1990 to examine the relationship between market thickness and switching. A counterfactual experiment investigates how many workers a new entrant firm can hire through voluntary switching in the absence of incumbent firms shutting down and laying off workers.

*Thanks to seminar participants at Johns Hopkins and the Labour Market Dynamic Growth conference. This paper is part of the project Pay and Promotion in Swedish Private Industries 1970–1990, which is under the leadership of Eva Meyersson–Milgrom, with financial support from the Swedish Council for Research in the Humanities and Social Sciences. Thanks to Arie Hietasalo, Åke Kempe and Svenskt Näringsliv for generously providing the SAF data and guidance on their use. The Kapnick Foundation supported a much different version through a grant to the Stanford Institute for Economic Policy Research. My email address is fox@uchicago.edu.

# 1 Introduction

## 1.1 Empirical assignment games

Perfect information matching games are an important tool for analyzing data on relationships. Often we have data such as who is married to whom, which workers are employed by certain firms, which manufacturers supply particular retailers, which venture capitalists fund particular startups, which school districts consolidate, and so on. Matching games make the formation of these relationships the dependent variable: in equilibrium, certain assignments of agents to partners are stable. The stability of an assignment of matches is a function of the preferences of all agents, and the preferences of agents are functions of exogenous characteristics of agents that a researcher may have data on.

Structural estimation of matching games has the goal of estimating agent payoffs in order to understand the relative importance of each agent characteristic in forming the matches that we observe. Becker (1973) performed an empirical investigation into marriage markets where each potential spouse had a single exogenous characteristic. More recently, a growing literature uses various matching models and corresponding structural methods to estimate perfect information, two-sided matching games (Ahlin, 2006; Akkus and Hortacsu, 2006; Angelov, 2006; Boyd, Lankford, Loeb and Wyckoff, 2003; Choo and Siow, 2006; Dagsvik, 2000; Ferrall, Salavanes and Sørensen, 2004; Gordon and Knight, 2005; Park, 2007; Sørensen, 2007; Weiss, 2007; Yang, 2006).

My own work (Bajari and Fox, 2007; Fox, 2007) has focused on transferable utility matching games where agents exchange money as part of the terms of the match. For example, the utility of a worker $i$ matched to firm $j$ may be $\tilde{u}_{ij} - \gamma w_{ij}$, where $\tilde{u}_{ij}$ is an exogenous constant governing preferences and $w_{ij}$ is an endogenous wage. While Becker (1973) applied the model with endogenous prices to marriage, it is more obvious that many product and labor markets are best modeled using endogenous prices. The wide number of theoretical models that fit into this framework are often called assignment games (Tinbergen, 1947; Koopmans and Beckmann, 1957; Shapley and Shubik, 1972; Becker, 1973; Sattinger, 1979; Kelso and Crawford, 1982; Leonard, 1983; Demange and Gale, 1985; Sotomayor, 1992; Kovalenkov and Wooders, 2003; Ostrovsky, 2004; Garicano and Rossi-Hansberg, 2006). I keep the focus on assignment games here and do not consider other, related literatures: matching models without prices (Gale and Shapley, 1962) and imperfect information search models with heterogeneous agents (Shimer and Smith, 2000; Atakan, 2006).

## 1.2 The repeated matching game

The static empirical matching game literature uses data on the sorting pattern of agents at one point at time. However, often the researcher has data over time where agents are observed to end some relationships and start new ones. Married people divorce and remarry, employees quit and find new jobs, retailers change suppliers and independently operating firms decide to merge. Data on switching matches provides a different type of dependent variable: both the fraction of people who change matches each period and the old and new partners of the switchers. The closeness in terms of observables of an agent's old and new match partners can tell us a lot about the unobservables governing switching, such as constant or changing preferences and switching costs. Similarly, how the incidence of switching at all varies with the similarity of outside option partners can also inform us about these unobservable parameters.

The empirical application in this paper focuses on the role of labor market thickness in allowing voluntary switching to move workers to new establishments. Firm and jobs are differentiated products based on firm characteristics such as geographic location, industry, and occupation. The hypothesis is that firms closer to other firms in characteristic space operate in thicker labor markets, which will encourage workers to switch, other factors held constant. After estimating the model's parameters, I examine to what extent workers will move to open positions in a brand new firm. To do this, I compute the equilibrium to a labor market with the entrant firm, and examine how many of its jobs are likely to be filled. I return more to the market thickness application below.

Using data on agents switching matches typically requires a dynamic model. If an agent is forward looking, the agent will weight the utility from the destination partner by the endogenous distribution of lengths that the match will continue. The more switching predicted in the future, the lower the importance placed on the flow utility from a single match. Therefore, we might expect measurable changes in the characteristic of a match partner that could be several time periods away to play less of a role in determining matches when the (endogenous model outcome) switching probability is high. As examples, consider workers acquiring firm specific human capital or a supplier learning by doing about the needs of a retailer. In these examples, keeping a match for a long time may induce a higher flow utility itself. Properly evaluating the costs and benefits of a switch thus requires agents to be forward looking.

This paper introduces a dynamic, repeated matching game. Each time period, a matching market clears so that all agents are matched. Prices are chosen to that every period supply equals demand at all pairs of worker and firm states. Then all agents' state variables evolve and during the next period another matching game is held. All agents on both sides of the market are forward looking: they maximize the expected present discounted value of the utility from matches. This is not a search model: each period all agents observe all potential partners and form new matches. To my knowledge, my model of a repeated static matching game with forward-looking agents is new and not found in the theory literature.

## 1.3 Application to market thickness and employer switching

This paper examines the labor market for elite Swedish engineers from 1970–1990. I have a 21 year panel on most elite engineers in a medium sized national labor market. The Swedish data track the engineers when they switch firms. The data are collected for all the larger, and many of the smaller, private sector employers in Sweden. Therefore, I observe data on almost a complete national labor market. This is opposed to typical public US datasets, where individual workers are interviewed about their current jobs, but no data are collected about the other employers in the area. Without data on the other employers competing in a labor market, a researcher cannot examine whether the firm has low turnover because of say switching costs or just the lack of labor market competitors.

I model employers as existing in a product characteristic space. Establishments (employers) are distinguished by geographic location, industry, and corporate parent.[1] Within a firm, I use a detailed Swedish job title code (standardized across firms) to assign individual jobs to broad occupations. The standardized occupational codes are the critical advantage of the Swedish data over other national-level employer-employee matched datasets. In the model, the labor market clears so that the demand and supply of elite engineers at each plant each equals the number of engineers recorded in the data, at least in expectation.

I estimate the matching game using data on the yearly market assignment of elite engineers to jobs. I then use the estimated model to explore the role of market thickness in employer switching decisions. Market thickness for a given firm is a rough notion based on the characteristics of rival firms, including their workforce sizes, and the estimated worker preference parameters over firm characteristics. If Malmo has many manufacturing establishments, one such manufacturer in Malmo is in a thick labor market if geographic distance and industry are important characteristics in the utility gain or loss from switching firms.

To understand the economic interpretation of the estimated structural parameters, I compute counterfactual equilibria where the product characteristics of firms are altered along one dimension, such as a counterfactual where all firms are said to operate in the same industry. In this counterfactual, all firms are in a thick labor market, at least as far the product characteristic of industries are concerned. These counterfactual equilibria both provide an extreme example of varying market thickness and play roughly the same role as calculating marginal effects in a single-agent discrete choice model.

The overall goal of the application is to use the estimated model to answer the question of whether voluntary employer switching is enough to move workers from incumbent to entrant firms. The alternative is that workers will not move voluntarily, so only involuntary layoffs ("push" mobility) will be able to force workers to switch firms. As even a temporary state of unemployment can impose a psychological and financial cost on the workers involved, it would seem preferable to sort workers to growing firms with wages rather than layoffs.

I introduce an entrant firm with currently unfilled slots. After adjusting the model structure to allows jobs to be unfilled in expectation, I recompute the equilibrium wages and continuation values. I then explore how many of the entrant firm's positions are filled by voluntary worker movement one time period after entry. I vary the entrant experiment by placing the entrant in thinner or thicker markets. This experiment answers the question of how much pull mobility from higher wages can encourage workers to move to new sectors, and how much the effectiveness of

[1]Unlike many papers in the equilibrium search literature, firm size is fixed in the matching model because I believe unmodeled product market factors (how popular are Saab's cars) explain much more of the variation in employment across establishments than labor market factors. The matching equilibrium requires that all slots in a firm are filled, in expectation.

wages as a sorting mechanism varies with market thickness.

# 2 Overview of major model and empirical details

## 2.1 Computational curses of dimensionality

Both dynamic programming models and matching models suffer from computational curses of dimensionality. In single agent dynamic models, increasing the number of state variables quickly raises the number of unknown continuation values that must be solved for. If there are $d$ state variables with $x$ values each, there are $x^d$ total states and hence unknown continuation values. In matching games, increasing the number of agents in the model increases the number of potential matches and hence the number of potential market assignments than could be part of an equilibrium. In a one-to-one, two-sided matching game with $e$ agents on each side, there are $e^2$ individual matches and $e!$ possible equilibrium assignments. If $e = 100$, then $e^2 = 10,000$ and $e! \approx 9.33 \times 10^{157}$. As both single-agent dynamic programming and matching games suffer from computational curses of dimensionality, one can imagine that combining the two into a repeated matching game creates enormous additional computational difficulties.[2]

### 2.1.1 Dynamic programming and matching

First I address combining dynamic programming and matching games. The previous discussion emphasized dynamics in individual agents' paths or in individual matches. A perfect information matching model introduces another form of dynamics: how the overall matching market changes as other agents form matches. In the empirical industrial organization literature on dynamic Nash games, typically the players are a small number of firms in an industry and the each firm's state variable includes the observed states of all the rival firms (Ericson and Pakes, 1995). Making the state variable of a single player expand with the total number of players causes a computational curse of dimensionality, which some papers have proposed changes to the game being modeled in order to mitigate (Doraszelski and Judd, 2007; Weintraub et al., 2007). In a matching game, the equilibrium concept is not Nash equilibrium (it is a cooperative game theoretic model), but the overall distribution of rivals' matches does affect the probability of matching next period. Therefore, in a repeated matching game similar computational curse of dimensionality issues occur as in the dynamic Nash games literature. One approach suitable for small matching markets is to proceed as in the IO literature and add the states of all agents in the market to each agent's individual state. Mergers, R&D alliance, standards and other types of firm relationships can be modeled using the matching framework here, with no conceptual difficulties.

In a very large matching market, an alternative might be to appeal to theorems like those in Hopenhayn (1992)

[2]On top of the computational difficulties inherent in repeated matching games, I also want to estimate the model, so there must be econometric error terms in the model. If the error terms must be integrated out, with the integrand being a repeated matching game, the model will combine three curses of dimensionality: dynamic programming, matching games and now integrating out error terms for every period and possible match.

for industry dynamics to argue there exist sufficient conditions for the market to remain in an aggregate steady state, where an aggregate steady state involves idiosyncratic switching that does not change the aggregate distributions of agent states and matches. However, imposing such assumptions on payoff functions and the evolution of states would contradict my empirical goal of using the estimator to work with microdata on smaller matching markets. For example, my empirical application looks at thin labor markets, where there is often only a handful of firms employing workers in a given city or industry. Transforming this type of microdata into a continuum removes much of the empirical richness of the application. Also, the model assumptions that ensure an steady state are strong and do not allow one to look at data on interesting empirical phenomena that disturb the aggregate steady state, such as the entry of a new, large firm.

For my application to thin markets, I make a modeling compromise that fits in between the purist approaches in the small industry and atomistic agent models. I impose the behavioral assumption that agents believe the market will remain in steady state.[3] In particular, each agent believes the continuation value for being a particular individual state in the future will be the same as the continuation value for being at that state today. This is an assumption: in reality the market-clearing prices will change as other agents enter and exit and existing agents switch matches. This keeps the focus of the dynamic model on the trajectory of each individual agent rather than modeling each agent's uncertainty over the extremely complex economy-wide dynamics. This assumption fits the labor market application well, as individual workers could be reasonably be thought of as looking to the experiences of their elders for career advice, rather than recomputing equilibria to the entire labor market.

### 2.1.2 Empirical tractability, matching and dynamic programming

The worst curse of dimensionality for estimation would come if the solution to a repeated matching game was part of the integrand of the integral over a large number of matched-specific error terms.[4] The main issue is that the continuation value in a dynamic model involves not only the observed state variables, but also the expected value of the error term for taking a particular action, given that the action is the optimal one. The distribution of the error terms is required to calculate this conditional expectation and solve Bellman's equation.

The key solution to the computational curse of dimensionality from integrating out match specific shocks is to change the model's timing assumptions so that endogenous prices are computed before the realization of match-specific taste shocks. Taste shocks affect the equilibrium matches but not the computation of the equilibrium prices. If the discount factor is set to 0, my model reduces to a static logit matching game that has been applied to marriage by Choo and Siow (2006) and Weiss (2007).[5] Unlike the latter authors, I show how to compute the equilibrium to

[3] Krusell and Smith Jr (1998) study a heterogeneous-agent growth model with another behavioral assumption: agents need to keep track of only a finite list of the moments of the wealth distribution, rather than the entire distribution.

[4] The semiparametric matching game maximum score estimator in Fox (2007) that avoids recomputing equilibria as part of the estimation routine does not extend to a dynamic matching game. Maximum score is a partial identification approach, where the estimator is consistent in the presence of error terms but the distribution of error terms is not identifiable (under minimal assumptions) and not estimated.

[5] Dagsvik (2000) considers a static logit matching game where an equilibrium match contract can be more general than a monetary exchange. For example, a labor contract can specify hours of work in addition to the wage.

the model. My dynamic model can be expressed as a series of nonlinear equations: one Bellman's equation for each distinct agent state, and one supply equals demand equation for each pair of partner states. The endogenous variables in the model are the continuation values, one for each agent state, and the wages, one for each pair of partner states.

#### 2.1.3 Number of equilibrium equations and unknowns

A final computational curse of dimensionality exists: the number of equilibrium wages and corresponding supply equals demand conditions goes up with the number of states. I discuss this issue in more detail in Section 4.2.

### 2.2 Estimation

Choo and Siow and Weiss estimate their static models using aggregate data / market share inversion methods similar to those in Berry (1994) for single-agent demand estimation. As Berry, Linton and Pakes (2002) show in simulation studies, aggregate data share inversion methods have large finite sample biases if there are few agents for each combination of observable characteristics, so that market shares (choice probabilities) are estimated with statistical noise in a first stage.[6] Consequently, the market share inversion estimators in Choo and Siow and Weiss will have large finite sample biases if applied to many typical matching datasets with small numbers of agents at each observed state.

My application is to market thickness, where in thin markets by definition there are few observations for some agent states. More generally, my goal is to work with microdata on specific markets in industrial organization and labor economics. Therefore, I generalize the single agent, dynamic discrete choice estimator of Rust (1987) to to the case of a repeated matching game. For each guess of a vector of parameters and the empirical distribution of agents over states, the dynamic matching game in principle may be solved, so that the likelihood function of the model and data may be evaluated. As maximum likelihood is asymptotically efficient, no alternative estimator is preferable on statistical grounds, under the parametric assumptions. The estimator is compatible with microdata as no first-stage estimates of market shares are required, unlike in the share inversion methods of Choo and Siow and Weiss and the mathematically similar dynamic two-step estimators of Hotz and Miller (1993), Aguirregabiria and Mira (2002), Bajari, Benkard and Levin (2007), Pakes, Ostrovsky and Berry (2007), Aguirregabiria and Mira (2007), Pesendorfer and Schmidt-Dengler (2007), as well as Bajari and Hong (2005).

Computing a single equilibrium to my dynamic matching game (solving a system of nonlinear equations) can be time consuming, potentially taking hours or days for matching markets with large numbers of agents. Therefore, solving the model each time the likelihood function is evaluated, as in the nested fixed point approach of Rust (1987), may require years to numerically converge. I adopt a suggestion by Su and Judd (2007) to enforce the nonlinear equations that define the model as nonlinear constraints and to optimize the likelihood by picking both the unknown

[6] See the small $n$ cells in Table 1 of Berry, Linton and Pakes.

structural parameters and the unknown, endogenous objects in equilibrium: continuation values at all states and the unobserved wages at all pairs of potential partner states. I will show that maximizing a likelihood subject to the model's structure is almost the same computational procedure as solving the model once. Therefore, using the Su and Judd suggestion makes estimation about as costly as solving the model once, for known parameter values. One research style only estimates a structural model where the equilibrium is possible to compute, as otherwise the model cannot be used to construct counterfactual predictions. For my matching game, any model whose equilibrium can be computed once can be estimated using almost the same computer time. This is one of the first applications of the Su and Judd suggestion and shows the usefulness of the approach in a serious empirical application.

## 2.3 Data on prices vs. matches and matching model robustness

In some datasets, there is information on physical matches but not the transfers exchanged between partners. For example, Becker (1973) studies marriage using a transferable utility framework. Likewise, in many supplier-retailer applications, inspection of store shelves or catalogs reveals the suppliers matched to each retailer, but the wholesale prices may be private contractual details. Because of data availability, it is useful to allow estimation using only one of the two dependent variables in the model: matches but not prices.

In the empirical application, I have data on both matches and wages, with wages for chosen positions only. In principle, one could explore a selection likelihood for the wage data that incorporates the existing matches-only likelihood as the selection equation in Heckman (1979).[7]

The estimator based on the selection likelihood would be consistent under the assumption that the wages from the model are the wages in the data. However, I do not necessarily believe the wages I observe come from the model. In the matching game, there is, subject to some modeling discretion, a separate wage for each pair of a worker and a firm state. Firms that do not want to hire a particular worker can offer him or her a low wage. If some unqualified (in terms of observables) worker was employed in a prestigious job (perhaps because of a high worker match-specific taste shock), the model would predict that the unqualified worker would be paid less than his colleagues. In the repeated matching model, once worker-specific wages are set, there is no further screening of applicants.

However, my reading of the administrative rules of how the Swedish labor market works suggests that wages are tied very strongly to positions (job titles). One equilibrium, transferable utility matching model consistent with administrative wage setting is Garicano and Rossi-Hansberg (2006). Indeed, the Swedish data list a detailed four-digit code representing a nationally standardized categorization for job titles. This code strongly predicts wages. If the model is one of administrative wage setting, if an unqualified worker was assigned to a particular job (because of a taste shock by the firm, say), we would expect the unqualified worker to be paid roughly the same. In a world where

[7] Another approach would be to use the price data in a single agent discrete choice model, with instruments for prices (Bayer et al., 2002). It is not clear where the instruments would come from in this model.

wages are tied to jobs, screening (job interviews) is the main mechanism used to sort workers to firms.

The conjecture that wages are not generated from the matching game does not mean that the match data cannot be used to estimate the model. Both screening and individual-job specific wages can support the same equilibrium assignment of workers to jobs: the competitive assignment. In a two-sided, many-to-one matching game with additive separability across the match output from multiple workers at the same firm, Sotomayor (1992) proves that any pairwise stable match will maximize the total output of the economy. She generalizes results on one-to-one matching from Koopmans and Beckmann (1957), Shapley and Shubik (1972), Becker (1973) and others. This output maximizing equilibrium can be considered the first welfare theorem (any equilibrium is efficient) for models with transferable utility. The stage game of my repeated matching model is similar to the static matching model of Sotomayor (1992), up to a set of econometric assumptions that simplify estimation.

Because multiple mechanisms can sustain the same assignment of workers to jobs, using data on matches only provides an estimator that it robust to misspecifying the model of equilibrium wages. Wages can be attached to either worker-firm matches and no screening used, or wages can be attached to jobs and screening used to assign applicants to jobs. As long as both assignment mechanisms support the competitive assignment of workers to jobs, estimation using data on matches only is consistent while remaining agnostic about the market-clearing mechanism.

# 3 A Repeated Matching Game

I introduce a dynamic, repeated matching game. For concreteness, I use language from the worker-firm labor market application. However, following the exposition style of Rust (1987), I lay out the model using generic notation for utility flows, state vectors, state transition densities, and error term distributions. This illustrates the generality of the method to other fields and applications. In a few places, I use footnotes to document when an assumption might make more sense in the labor application than in another application, such as firm relationships in industrial organization.

## 3.1 Time and timing

A year $t$ describes calendar time: the matching market clears each period. $T$ is the last period of data. A worker $a$ has age $d_t$ in period $t$. A worker works until he is, say, 60 years old, which is the same age for all workers but a different calendar year $D_a$ for each worker. Firms are infinitely lived.[8]

The timing of the game within a period is as follows.

1. At the beginning of a period, all agents have states variables known to them and all other agents.

[8]The finite horizon nature of workers requires no special adjustment, other than a bookkeeping terminal state for year $D_a + 1$ with a continuation value of 0. As backwards recursion for a single agent is not a computationally effective method of solving a market equilibrium model, making both sides of the market infinitely lived would simplify, not complicate, matters.

2. Then equilibrium wages and (under one optimal assumption to come) the corresponding agent beliefs about future opportunities are simultaneously determined. Equilibrium wages ensure the expected (over the period $t$ taste shocks) number of workers at each job equals the expected demand for workers at that job.

3. After wages are computed, each worker and each job receives a taste shock to each potential job (for workers) or worker (for firms).

4. Given these taste shocks, workers, and workers only, unilaterally pick a firm to work at. If more or fewer workers then the expected number appear at a firm because of taste shocks, the firm adjusts its slots to accommodate them.

In the fourth step, the firm receives a taste shock but cannot screen (refuse to hire) workers if the taste shock is low, for example. The firm taste shock serves to create a firm labor demand function so that demand equals supply in the second step.

In estimation, the timing assumption that the taste shocks occur after wages are set allows the researcher not to have to embed a linear programming problem inside a numerical integral over all of the match-specific taste shocks of both workers and firms. Therefore, in this model wages are not correlated with the realized taste shocks. This timing assumption does not radically alter the equilibrium predictions in a static model: the equilibrium will still roughly maximize some notion of the sum of match outputs, as Koopmans and Beckmann (1957) and other papers show for games without taste shocks.

## 3.2 Matches and states

The notation $ai_t$ refers to the event that the match between worker $a$ and job $i$ occurred in the period $t$.

The state vector is the same for a worker and a job in a match. Say worker $a$ and job $i$ matched in period $t-1$, or the match $ai_{t-1}$ occurred. Then the period $t$ state is $s_{ait}$. The overall match state contains the vector of characteristics $x_{at}$ of worker $a$ and the vector of characteristics $x^f_{it}$ of job $i$. So $s_{ait} = \left\{x_{at}, x^f_{it}\right\}$. Occasionally I will use superscripts $f$ to refer to model elements for jobs. Also, to keep the definition of an equilibrium simple (in a few sections), I will constrain $x_{at}$ and $x^f_{it}$ to each take on a finite number of values.

Say worker $a$ makes a new match with job $j$ during period $t$, or the event $aj_t$ occurs. Previously, job $j$ was matched to worker $b$. Then the state transition distribution is

$$h_{\theta_1}\left(s_{aj,t+1} \mid s_{ait}, s_{bjt}, aj_t\right),$$

where, optionally, $\theta_1$ is an estimable vector of a finite number of parameters that govern the updating of states. Perhaps the simplest possibility is that $h$ inserts the time invariant characteristics of worker $a$ and job $j$ into a new $s_{aj,t+1} =$

$\left\{x_{at}, x^f_{jt}\right\}$. In this case, there are no transition parameters $\theta_1$ to estimate. For some applications, it is important that there the state transition of worker $a$ be a function of the worker's job. For example, the worker could be accumulating occupation, industry or firm specific human capital. In these cases, switching to another occupation, industry or firm might reset the respective stock of specific human capital to zero. If parameters on the human capital accumulation problem needed to be estimated, they would be in $\theta_1$.

A worker could have been unmatched since the end of period $t-1$: $s_{a\emptyset t} = \{x_{at}, \emptyset\}$. This can happen for several reasons: the worker has just entered or reentered employment or the worker's period $t-1$ employer eliminated his position. Likewise, a job $i$ could be unmatched: $s_{\emptyset it} = \left\{\emptyset, x^f_{it}\right\}$. This could happen if a firm is growing, adding new positions for period $t$, or if some workers retired or exited the labor market. If layoff risk is modeled explicitly, it enters $h_{\theta_1}$.

In the standard case, $x_{at}$ and $x^f_{jt}$ are recorded in the data. However, the theoretical model of a repeated matching game is still valid if $x_{at}$ and $x^f_{jt}$ are observed by the agents in the model but are not recorded in the data. Like in single-agent discrete choice methods, having unobserved, time-persistent states will typically require panel data approaches to be used in estimation. For an equilibrium model, this makes computation of a statistical objective function time consuming as there will be a different equilibrium for each realization of the unobserved states for all agents. In industrial organization, researchers typically deal with unobserved state variables by estimating these variables in a first stage. For example, one can use price and quantity data to estimate a structural demand system and back out a firm $j$'s unobserved product quality $\xi_{jt}$ as the econometric error term BLP. In the estimation of the matching game, $\xi_{jt}$ is then treated as an observable state variable.

## 3.3 Steady state beliefs

As I discussed in Section 2.1, in applications to firm relationships in small industries in industrial organization, each agent (a firm) keeps track of the states of all rival firms. In my labor application, only a worker $a$'s current (period $t-1$) employer's characteristics $x^f_{jt}$ enter $s_{ait}$. Embedding the characteristics of all other agents into the state will often be computationally infeasible because of a computational curse of dimensionality in the number of outcomes to the matching stage game. Consider a medium-sized matching game with 100 workers and 100 available jobs. There are $100! \approx 9.33 \times 10^{157}$, way more than the atoms in the universe, possible sets of matches that could occur next period. Having agents explore the combinations of moves of all workers and all jobs is infeasible computationally in medium-sized matching games.

The only computationally tractable approach for matching games of medium size is for agents to believe that the market is in an aggregate steady state, so that the overall distribution of potential partners and wages will remain the same as it is today. A rational expectations approach to justifying this agent belief is to tightly structure the model so that the economy actually is in an aggregate steady state. Imposing these strong theoretical assumptions is at odds with providing a general econometric framework that can be taken to many micro data sets in many fields. Instead, I assume that agents believe the continuation value for being at a particular state $s_{ait}$ will not change.

**Assumption 1.** *If a worker with state $s_{ait}$ in period $t$ has a continuation value $V(s_{ait})$, workers in period $t$ who might end up at state $s_{ai\tau} = s_{ait}$ in some period $\tau > t$ expect to have continuation value $V(s_{ai\tau}) = V(s_{ait})$ when they reach that state. An analogous property holds for jobs.*

The force of this assumption is that the experiences of workers of say 50 today are reflective of the outcomes workers of age 40 expect to receive in 10 years. Workers do not consider the aggregate dynamics of the economy but consider their own trajectory in that economy. For example, a worker moving to a large city may know the switch opens up the possibility of in the future switching to the many employers located in the large city. But the worker does not concern himself with the possibility that other workers and firms may move to the city, disturbing the market equilibrium. The assumption can be recast as one of bounded rationality: no individual agent can hope to work out all of the combinations of future combinations of matches, so they choose not to and consider only their own future trajectory at the current distribution of states and equilibrium wages.

The assumption that agents ignore economy-wide dynamics and consider only their own future matches seems relatively innocuous for some applications to the labor market, where agents are somewhat small (but not atomistic) relative to the market. However, if the assumption is dangerous, an aggregate state $A_t = \left\{ \left\{ s_{ajt} \right\}_{a \in N}, \left\{ s_{bit} \right\}_{i \in J} \right\}$ that describes the state of all agents can be added as an additional argument in all model functions.

## 3.4 Worker decisions

For a worker $a$ of age $d_t$ in year $t$ with the state vector $s_{ajt}$, his current period utility function if he choose employment at a job $i$ out of $J$ total jobs is

$$u_\beta (ai_t, s_{ajt}) + \varepsilon_{ait} = \tilde{u}_{\tilde{\beta}} (ai_t, s_{ajt}) + \beta_w w_{ait} + \varepsilon_{ait},$$

where the match $ai_t$ is his employment choice and $\varepsilon_{ait}$ is a worker's job specific stochastic disturbance in utility. Utility $u$ is parameterized by the parameter vector $\beta$, which is to be estimated. Notationally, worker age $d_t$ is part of $x_{at}$ in $s_{ajt} = \left\{ x_{at}, x_{jt}^f \right\}$.

The wage at job $i$, $w_{ait}$ is implied by $x_{at}$ in $s_{ajt}$ from the equilibrium structure of the model. As wages clear the labor market, I model workers as making a unilateral employment decision. The wage enters $u_\beta$ additively separably, with the parameter $\beta_w$ governing its importance in utility. I use tildes on objects to denote the non-wage portions of those terms.

The job match $ai_t$ is chosen to maximize the worker's expected, present discounted value of utility

$$E\left[\sum_{\tau=t}^{D_a}\delta^{\tau-t}\left(u_\beta\left(ai_\tau, s_{aj\tau}\right)+\varepsilon_{ai\tau}\right) \mid s_{ajt}, \varepsilon_{at}\right],$$

where $\delta$ is a discount factor between 0 and 1 and $\varepsilon_{at}$ is the vector of all employer-specific taste shocks at age $t$.

This is a finite-horizon and discrete-time dynamic programming problem. Given wages, I solve for the model's predicted probabilities of job choice $\Pr\left(i_{at} \mid s_{ait}\right)$ by backwards recursion. Recall $h_{\theta_1}\left(s_{ai,t+1} \mid s_{ajt}, s_{bit}, ai_t\right)$ is the transition density of the state variable $s_{ait}$ conditional on the worker's new match $ai_t$. Define a continuation value to be

$$\begin{aligned}
V\left(s_{ajt}\right) &= \int \max_{ai_t\in\{1,\ldots,J\}} E\left[\sum_{\tau=t}^{D_a}\delta^{\tau-t}\left(u_\beta\left(ai_\tau, s_{ajt}\right)+\varepsilon_{ai\tau}\right) \mid s_{ajt}, \varepsilon_{at}\right] g_{\theta_2}\left(\varepsilon_{at}\right) d\varepsilon_{at} \\
&= \int \max_{ai_t\in\{1,\ldots,J\}} \left[u_\beta\left(ai_t, s_{ajt}\right)+\varepsilon_{ait}+\delta\int_{s_{t+1}\in S} V\left(s_{ai,t+1}\right) h_{\theta_1}\left(s_{ai,t+1} \mid s_{ajt}, s_{bit}, ai_t\right) ds_{ai,t+1}\right] g_{\theta_2}\left(\varepsilon_{at}\right) d\varepsilon_{at} \\
&= \int \max_{ai_t\in\{1,\ldots,J\}} \left[v\left(ai_t, s_{ajt}\right)+\varepsilon_{ait}\right] g_{\theta_2}\left(\varepsilon_{at}\right) d\varepsilon_{at}, \qquad (1)
\end{aligned}$$

where $g_{\theta_2}$ is the joint density of the worker's vector of taste shocks $\varepsilon_{at}$, and where $v\left(ai_t, s_{ajt}\right)$ is the match $ai_t$-specific value function implicitly defined in the equation. $\theta_2$ is an optional vector of estimable parameters that may govern the distribution of taste shocks. To handle planned retirement, $V\left(s_{ajt}\right)\equiv 0\,\forall j$ if $t > D_a$.[9]

At the state $s_{ajt}$, and integrating out the taste shocks $\varepsilon_{at}$, the probability of worker $a$ picking the match $ai_t$ is

$$\Pr_t\left(ai_t \mid s_{ajt}\right) = \int 1\left[v\left(ai_t, s_{ajt}\right)+\varepsilon_{ait} > v\left(ak_t, s_{ajt}\right)+\varepsilon_{akt}\,\forall k\neq i\right] g_{\theta_2}\left(\varepsilon_{at}\right) d\varepsilon_{at}. \qquad (2)$$

The $t$ subscript emphasizes that this probability is for the market configuration at time $t$. If the match-specific shock $\varepsilon_{at}$ is iid with the logit distribution, McFadden (1973) shows the equilibrium probability of the match $ai_t$ for an age $t$ worker simplifies to

$$\Pr_t\left(ai_t \mid s_{ajt}\right) = \frac{\exp\left(v\left(s_{ajt}, ai_t\right)\right)}{\sum_{ak=1}^{J}\exp\left(v\left(s_{ajt}, ak\right)\right)}. \qquad (3)$$

The relevant choice's present discounted value of payoffs is in the numerator, normalized by the sum of the payoffs of all choices in the denominator. More generally, choosing a distribution for $\varepsilon_{at}$ in McFadden (1973)'s generalized extreme value (GEV) class ensures closed form integrals for both $\Pr_t\left(ai_t \mid s_{ajt}\right)$ and $V\left(s_{ajt}\right)$.[10]

A single-agent, dynamic programming, discrete choice model with many differentiated choices is found in Kennan and Walker (2002), who study the choice between the fifty American states. They do not model the preferences of

[9] Readers familiar with Rust (1987) will recognize his conditional independence assumption: taste shocks only affect state transitions by altering matches. Time persistent taste shocks enter $s_{ajt}$ and were discussed at the end of the previous subsection.

[10] The simplification $V\left(s_{ajt}\right)=\sum_{ai=1}^{J}\Pr_t\left(ai_t \mid s_{ajt}\right)\left\{v\left(s_{ajt}, ai_t\right)+E\left[\varepsilon_{ait} \mid ai_t^\star = ai_t\right]\right\}$ eases the calculation of $V\left(s_{at}\right)$. If the error terms have the extreme value (logit) distribution, the expected value of the error term conditional on a match is $E\left[\varepsilon_{ait} \mid ai_t^\star = ai_t\right] = 0.577 - \log\Pr_t\left(ai_t \mid s_{ajt}\right)$.

firms or market equilibrium.

## 3.5 Jobs

The timing assumptions show how workers unilaterally choose individual jobs in equilibrium. However, we derive each job's labor demand functions so we can define equilibrium.

Job $i$ at some firm hiring a worker $a$ in year $t$ produces current-period output, or profit, $\pi_\gamma(ai_t, s_{bit}) + \varepsilon^f_{ait}$, where $s_{bit}$ is the job's state vector from matching with worker $b$ in period $t-1$, $\varepsilon^f_{ait}$ is a match-$ai$ specific logit shock, and $\gamma$ is an estimable parameter vector. The wage $w_{ait}$ is implied by the equilibrium structure of the model and enters $\pi_\gamma$ additively separably,

$$\pi_\gamma(ai_t, s_{bit}) + \varepsilon^f_{ait} = \tilde{\pi}_\gamma(ai_t, s_{bit}) - \gamma_w w_{ait} + \varepsilon^f_{ait}.$$

For simplicity in dynamic programming, profits are additively separable across jobs (employment relationships at the firm) for a firm that hires multiple workers. Additive separability is also used in Sotomayor (1992).[11]

Jobs are infinitely lived and their maximization of the expected, present discounted value of profits for a particular job (employment slot)

$$E\left[\sum_{\tau=t}^{\infty} \delta^{\tau-t}\left(\pi_\gamma(ai_\tau, s_{bit}) + \varepsilon^f_{ai\tau}\right) \mid s_{bit}, \varepsilon^f_{it}\right],$$

is a stationary problem if $s_{bit}$ contains the states of all matches or the earlier steady state beliefs assumption is maintained. Here $\varepsilon^f_{it}$ is the vector of firm $i$'s taste shocks. The discount factor $\delta \in [0,1)$ is the same as for workers, for simplicity.

Let there be $N$ total workers. Jobs are forward-looking and have state-$s_{bit}$ value functions $V^f(s_{bit})$ that can be decomposed into choice-specific value functions $v^f(ai_t, s_{bit})$:

$$\begin{aligned}
V^f(s_{bit}) &= \int \max_{ai_t\in\{1,\dots,N\}} E\left[\sum_{\tau=t}^{\infty} \delta^{\tau-t}\left(\pi_\gamma(ai_\tau, s_{bit}) + \varepsilon^f_{ai\tau}\right) \mid r_{it}, \varepsilon^f_{it}\right] g^f_{\theta_3}\left(\varepsilon^f_{it}\right) d\varepsilon^f_{it} \\
&= \int \max_{ai_t\in\{1,\dots,N\}} \left[\pi_\gamma(ai_\tau, s_{bit}) + \varepsilon^f_{ait} + \delta \int_{s_{t+1}\in S} V(s_{ai,t+1})\, h_{\theta_1}(s_{ai,t+1} \mid s_{ajt}, s_{bit}, ai_t)\, ds_{ai,t+1}\right] g^f_{\theta_3}\left(\varepsilon^f_{it}\right) d\varepsilon^f_{it} \\
&= \int \max_{ai_t\in\{1,\dots,N\}} \left[v^f(ai_t, s_{bit}) + \varepsilon^f_{ait}\right] g^f_{\theta_3}\left(\varepsilon^f_{it}\right) d\varepsilon^f_{it}. \qquad (4)
\end{aligned}$$

$V^f(s_{bit})$ for all states $s_{bit}$ is a system of nonlinear equations, Bellman's equation. Conditional on the equilibrium wages, iterating the system creates a contraction mapping, with a unique solution for the $V^f(s_{bit})$'s.

Firms have a labor demand preference: how much they would prefer to employ a worker. This can be expressed

[11] For static matching games, Fox (2007) presents a feasible estimator for many-to-many two-sided matching games where profits are not additively separable across multiple matches involving the same agent.

as a quasi-demand probability, for worker $i$ and a particular slot in the firm,

$$\Pr_t^f(ai_t \mid s_{bit}) = \int 1\left[v^f(ai_t, s_{bit}) + \varepsilon_{ait}^f > v^f(ak_t, s_{bit}) + \varepsilon_{cit}^f \,\forall c \neq a\right] g_{\theta_3}^f\left(\varepsilon_{it}^f\right) d\varepsilon_{it}^f.$$

Again, this quasi-demand probability is the solution to a $N$-dimensional integral over the $\varepsilon_{ait}^f$'s.

### 3.6 Equilibrium in wages and match probabilities

Recall there are finite numbers of the worker and firm contributions to states, $x_{at}$ and $x_{it}^f$, so there is a finite number of states $s_{ait} = \left\{x_{at}, x_{it}^f\right\}$. There is also a finite number of workers $N(s_{ait})$ at state $s_{ait}$ and a finite number of jobs at state $s_{ait}$, $N^f(s_{ait})$. By feasibility, $N(s_{ait}) = N^f(s_{ait})$ unless one of the firm or worker's index is $\emptyset$, the partner representing being unmatched since the end of period $t-1$.

Wages are used to clear the market. All workers $a$ are offered wages $w_{ait}$ by all jobs $i$. By the timing assumptions, $\varepsilon_{ait}$ is realized after wages are set.[12] This is mainly for computational reasons: to avoid nesting an algorithm to compute the equilibrium inside the $J \cdot N$ integral of all match-specific taste shocks. Therefore, by payoff equivalence, it is without loss of generality to assume equilibrium wages are a function of only the non-taste shock states: $w_{ai}(s_{ajt}, s_{bit})$. The notational convention is the worker's state $s_{ajt}$ is listed first, followed by the job's state $s_{bit}$ in the second argument. The subscripts $ai$ emphasize that the wage is for the match $ai_t$, not any of the other three matches that could have occurred: $aj_t$, $bi_t$ and $bj_t$.

The following definition describes equilibrium.

**Definition 1.** *If Assumption 1 is maintained, an equilibrium in period $t$ is a function $w_{ai}(s_{ajt}, s_{bit})$ that describes the wages for all pairs of worker and job states and where*

- *Expected labor supply equals expected labor demand for all matches and pairs of states, meaning*

$$N(s_{ajt}) P(ai_t \mid s_{ajt}) = N^f(s_{bit}) P^f(ai_t \mid s_{bit}) \tag{5}$$

  *for all pairs $\left\{s_{ajt}, s_{bit}\right\}$ and all period $t$ matches $ai_t$ where $a \in N(s_{ajt})$ and $i \in N^f(s_{bit})$.*

- *All workers' and jobs' Bellman equations are satisfied.*

Let $S$ be the number of match states. Without aggregate states, the model is therefore described by a set of nonlinear equations:

1. $S \times S$ supply equals demand conditions, (5).

[12] Again, Fox (2007) presents an estimator for static matching games where equilibrium prices may be a function of taste shocks.

2. $S$ workers' Bellman equations (1).

3. $S$ jobs' Bellman equations (4).

Corresponding to these nonlinear equations, there are the following model unknowns

1. $S \times S$ wages $w_{ai}(s_{ajt}, s_{bit})$

2. $S$ workers' continuation values $V(s_{ajt})$

3. $S$ jobs' continuation values $V^f(s_{bit})$

As the number of unknowns equal the number of equations, it is possible that there exists a solution for any given parameterization.[13]

With aggregate states, there are the above equations and unknowns multiplied by the total number of aggregate states. Here we see the computational curse of dimensionality, and why the steady state beliefs assumption must be imposed on medium-sized matching games.

Once equilibrium wage functions $w_{ai}(s_{ajt}, s_{bit})$ and continuation values $V(s_{ajt})$ for workers have been calculated, it is easy to compute equilibrium match probabilities by substituting wages and continuation values back into labor supply probabilities, (2). These equilibrium match probabilities $\Pr_t^\star(ai_t \mid s_{ajt})$ describe the final pattern of sorting in period $t$: the equilibrium assignment of workers to jobs by their labor market histories and characteristics. Note these probabilities are a function of the entire set of matches, so they are subscripted by $t$.[14]

### 3.7 Equilibrium match probabilities for a special case

Let there be two workers and two jobs: $N = 2$ and $J = 2$. Let the workers be $a$ and $b$ and the jobs $i$ and $j$. Say the current states are $s_{ajt}$ and $s_{bit}$. Let the taste shocks for workers and firms both be logit with scale parameter 1. Let the Assumption 1 hold.

I work with the four supply equals demand equations and leave the Bellman's equations unsolved. Algebra can be used to show that, in equilibrium,

$$\Pr_t^\star(ai_t \mid s_{ajt}) = \Pr_t^\star(bj_t \mid s_{bit}) = \frac{A}{A+B} \tag{6}$$

[13] If instead all rivals' states enter an agent's state, so that the aggregate state is $A_t$, an equilibrium is a function $w_{ai}(s_{ajt}, s_{bit}, A_t)$ that describes the wages for all pairs of worker and job states given a certain aggregate state $A_t$, and where expected labor supply equals expected labor demand for all matches and pairs of states at all aggregate states and all workers' and jobs' Bellman equations are satisfied at all individual and aggregate states.

[14] The requirement that supply equals demand for every pair of a worker and firm state means the solution concept known as oblivious equilibrium cannot be applied to a matching game (Weintraub, Benkard and Roy, 2007). In an oblivious equilibrium, agents optimize while believing the aggregate market state is currently and will remain forever and some limiting, long-run state. In a matching game, if agents act according to some state that is not the true aggregate market state, then supply and demand will not equate at each period. For example, a consider a firm where all of its jobs are filled in the long-run state. Few workers will match there, even if in the real-world state none of the jobs are filled.

where

$$A = \exp\left(\frac{1}{2(1+\alpha_w)}\left(\alpha_w\left(\tilde{v}\left(ai_t, s_{ajt}\right) + \tilde{v}\left(bj_t, s_{bit}\right)\right) + \tilde{v}^f\left(ai_t, s_{bit}\right) + \tilde{v}^f\left(bj_t, s_{ajt}\right)\right)\right)$$

and

$$B = \exp\left(\frac{1}{2(1+\alpha_w)}\left(\alpha_w\left(\tilde{v}\left(aj_t, s_{ajt}\right) + \tilde{v}\left(bi_t, \tilde{s}_{bit}\right)\right) + \tilde{v}^f\left(bi_t, s_{bit}\right) + \tilde{v}^f\left(aj_t, s_{ajt}\right)\right)\right).$$

Here I have solved for the static equilibrium as a function of the choice-specific value functions that exclude the wage component of period $t$'s flow utility:

$$\tilde{v}\left(ai_t, s_{ajt}\right) = v\left(ai_t, s_{ajt}\right) - \beta_w w_{ai}\left(s_{ajt}, s_{bit}\right) \;\&\; \tilde{v}^f\left(ai_t, s_{bit}\right) = v^f\left(ai_t, s_{bit}\right) + \gamma_w w_{ai}\left(s_{ajt}, s_{bit}\right).$$

Wages are endogenous and can be substituted out of equilibrium match probabilities. Also, $\alpha_w = \gamma_w/\beta_w$ is the ratio of the wage parameters of the job and worker.

In this simple example, the endogenous wages support what is more or less the efficient outcome. The equilibrium match probabilities look like they are comparing the sums of the non-wage expected present discounted value of two assignments in period $t$: $ai_t$ and $bj_t$ as well as $aj_t$ and $bi_t$. Under the model's timing assumptions, workers make employment decisions after wages are set. Therefore, the taste shocks $\varepsilon_{ait}$ could be such that $a$ and $b$ could match with the job $i$ and the job $j$ could be unfilled.

However, the equilibrium match probabilities are written as if the outcome maximizes the sum of the payoffs of all four agents and as if each job can have only worker. It is similar to the result of Koopmans and Beckmann (1957) for a model without taste shocks: the decentralized stable matching equilibrium is equal to the solution of social planner's problem. Here, the only change might be the social planner has a unobserved payoff component for each of the two hypothetically exclusive assignments. The scaling $\alpha_w = \gamma_w/\beta_w$ uses money to equalize the units of the utility of workers, which are arbitrarily normalized to the variance of the logit taste shocks, $\pi^2/6$, with the units of the profits of jobs, which are also arbitrarily scaled.

The formula (6) is sensible for other reasons. First, by switching indices one can see the probabilities of the matches $ai_t$ and $aj_t$ sum to 1, as they must by market clearing. Second, consider a job $j$ and year $t$ only utility flow $\xi_{jt}$ that occurs to both workers $a$ and $b$. A firm fixed effect, or unobserved product characteristic, $\xi_{jt}$ is priced out in the equilibrium wages. Becker (1973) shows that only interactions between the characteristics of firms and workers affect matching probabilities. To see the invariance, add $\xi_{jt}$ to both $\tilde{v}\left(ai_t, s_{ajt}\right)$ and $\tilde{v}^f\left(ai_t, s_{bit}\right)$. There is now one $\xi_{jt}$ in the same place inside each of the three exp terms. Therefore, the $\xi_{jt}$ terms cancel by multiplying the numerator and denominator by $\exp\left(-\frac{\alpha_w}{2(1+\alpha_w)}\xi_{jt}\right)$.

The $\tilde{v}\left(ai_t, s_{ajt}\right)$'s and $\tilde{v}^f\left(ai_t, s_{bit}\right)$'s are not model primitives: they are choice-specific continuation values. The solution to the entire dynamic matching game is required to compute matching probabilities in terms of model primitives, such as $u_\beta$ and $\pi_\gamma$. I compute an equilibrium numerically from now on.[15]

# 4 Computing an equilibrium

## 4.1 Solving the system of equations

The system of nonlinear equations (1), (4) and (5) needs to be solved in order to compute an equilibrium. Many methods exist for solving systems of nonlinear equations. Methods applicable to single-agent dynamic programming problems, such as contraction mappings, are not guaranteed to converge in this equilibrium model. Therefore I use Newton's method, which converges quickly near a solution (SCHMEDDERS2007).

In practice, I use dedicated nonlinear programming solvers, which are used to minimize some function subject to nonlinear constraints. To compute an equilibrium, I minimize the constant function 0 subject to the set of equality constraints that the model's equilibrium equations are satisfied:

$$\max_{\left\{V(s),V^f(s),w_{s,s'}\right\}} 0 \text{ s.t. model equations.}$$

I have found the best success using the commercial solver KNITRO (Byrd, Hribar and Nocedal, 1999). Finding a solution may require restarts using randomly generated starting values if little information exists about the solution. Because of the many restarts (which KNITRO automates), computing an equilibrium once may take hours or days for large matching markets. Once a solution is found, computing new equilibria for model perturbations is much easier as good starting values are available. Typically only one starting value is needed.

## 4.2 Reducing the number of model equations and unknowns

Because all agents in a matching market are heterogeneous, studying matching markets suffers from combinatorics issues in computation. Likewise, dynamic programming problems suffer from scale issues. The combination of matching and dynamics, the repeated matching game, has the potential to be very computationally difficult. I have imposed the timing assumptions, like the taste shocks arising after wages are computed, as well as the assumption on the state space of agents, Assumption 1, to eliminate several forms of computational curses of dimensionality.

However, there is still at least one computational curse of dimensionality remaining. There are $S+S+S\times S=2S+S^2$ nonlinear equations (1), (4) and (5). The remaining computational curse comes from the $S\times S=S^2$ supply equals demand equations, (5), and corresponding equilibrium wages $w_{ai}(s_{ajt},s_{bit})$. Recall each state $s_{ait}$ is the unique combination of a vector of worker characteristics $x_{at}$ and a vector of firm characteristics, $x^f_{it}$. As the number of states $S$

[15] I have not been able to use (6) to guess and verify a closed form solution for match probabilities for $N>2$.

increases rapidly as more heterogeneity in either worker or firm characteristics is added, having the number of model equations increase at the rate $S^2$ creates a tremendous computational burden.

My solution is to make wages less flexible: to have wages be constant over some aggregations of workers and jobs. Correspondingly, supply must only equal demand at these aggregations. I group pairs of worker and jobs states $\{s_{ajt}, s_{bit}\}$ into equivalence classes $\mathscr{X}_t$. The notational convention is that the worker's state $s_{ajt}$ is listed before the job's state $s_{bit}$. Then there is an equilibrium wage function $w(\mathscr{X}_t)$ that replaces each of the wages $w_{ai}(s_{ajt}, s_{bit})$ for $\{s_{ajt}, s_{bit}\} \in \mathscr{X}_t$ and a corresponding supply equals demand equation

$$\sum_{\{s_{ajt} \mid \{s_{ajt}, s_{bit}\} \in \mathscr{X}_t\}} N(s_{ajt}) \Pr_t(ai_t \mid s_{ajt}) = \sum_{\{s_{bit} \mid \{s_{ajt}, s_{bit}\} \in \mathscr{X}_t\}} N^f(s_{bit}) \Pr_t^f(ai_t \mid s_{bit}).$$

Computationally, instead of $S^2$ supply equals demand equations, (5), we have only as many equations as equivalence classes. There are fewer supply equals demand equations, but each equation has many more terms. It is not obvious that this way of reducing the number of model equations reduces the computational burden of computing an equilibrium, but in my experiments it has.

In the empirical application, I use a simple scheme for the equivalence classes. An equivalence class is all the jobs at one firm and the set of workers either inside or outside the firm. In other words, each firm offers two wages, one to its current workers and one to potential new workers. I choose this simplification because in the model low wages are used to screen out workers; a firm has a fixed number of positions and so offers low wages to newcomers if its positions do not need to be filled. As Section 2.3 argues, the same assignment can be supported by both active screening programs and by match-specific wages. However, I do plan to consider alternatives. Note that the simplifying assumption says nothing about the Bellman equations; there is still one Bellman equation (1) for each worker state $s_{ait}$, for example.

## 4.3 Existence and uniqueness

In a static, one-to-many two-sided matching game with firms having additive payoffs with multiple workers, Sotomayor (1992) proves that any equilibrium assignment of workers to jobs will maximize the total output of the economy. The equilibrium is guaranteed to exist and the assignment of workers to firms is unique if no two assignments both maximize the total output of the economy. The equilibrium wages that support the competitive assignment are not unique; they lie in bounds because of the discrete nature of the inequalities that define an equilibrium.

I add taste shocks to the matching model's stage game for the econometric need for the model to be able to rationalize any data set. These taste shocks preserve many of the economic intuition about the pattern of sorting that may arise in equilibrium, as argued by Choo and Siow (2006) and Weiss (2007) for marriage, but do change the mathematical structure to make the matching game look more like a Bertrand Nash pricing game with differentiated products demand from industrial organization.

The literature on Bertrand pricing equilibria suggests that existence and uniqueness is largely a function of assumptions on the densities of taste shocks, $g_{\theta_2}(\varepsilon_{at})$ for workers and $g^f_{\theta_3}\left(\varepsilon^f_{it}\right)$ for firms (Anderson et al., 1992). These presumably differentiable densities smooth out the match probabilities as a function of the wages, $\Pr_t(ai_t \mid s_{ajt})$ for workers and $\Pr^f_t(ai_t \mid s_{bit})$ for firms. These probabilities are similar to reaction functions in a Bertrand Nash pricing game and their shapes determine whether an equilibrium exists at all and just how many equilibria there might be.

For computational reasons, I have limited myself to assuming that the $\varepsilon_{at}$ and $\varepsilon^f_{it}$ are both iid extreme value (yielding the logit). As discussed earlier, this assumption provides closed form solutions for all the integrals in the match probabilities (2) and the expectations over next period's states that appear in the model equations (1), (4) and (5). For these error densities and $N = J = 2$ workers and jobs, I algebraically derived the unique equilibrium match probabilities in Section 3.7. For more agents, I have found no examples where I have been able to numerically compute two sets of equilibrium matching probabilities. This includes the repeated matching game with forward looking agents. A recent literature uses numerical techniques known as homotopies to find more (but not always all) of the solutions to a game (Borkovsky, Doraszelski and Kryukov, 2007). Homotopies are much too slow to apply to a repeated matching game.

If there is no outside option of being unmatched that pays a wage $w_{a\emptyset}(s_{ajt}, \emptyset) = 0$, then the wages that support an equilibrium assignment are unique up to a common additive constant. This can be seen by adding $B$ to all continuation values in (3): match probabilities remain the same. Unlike Sotomayor (1992), the wages that support a given equilibrium set of match probabilities are otherwise more likely to be unique because of the smoothness from $g_{\theta_2}(\varepsilon_{at})$ and $g^f_{\theta_3}\left(\varepsilon^f_{it}\right)$.

# 5 Estimation

## 5.1 Dependent variables for choice model: matches

The data do not contain direct observations on the number of jobs a firm posts. Because of the worker taste shocks $\varepsilon_{ait}$, some jobs may remain vacant even though the expected number of holders of the job is 1. In each year $t$, an assignment is a vector $\{1i_t, 2i_t, \ldots Ni_t\}$ of the realized job matches of each worker $1, \ldots, N$. The assignment satisfies worker feasibility: each worker must have only one job. Under the model's timing assumptions, if more or fewer workers than the expected number accept a firm's job because of logit shocks, then the job is, for at least a year, filled by larger or smaller number than its targeted size of 1.

## 5.2 Data and the transition rule of the states

Assume that all non logit shock states are observable in the data. An observation $\{ai_t, s_{ajt}, s_{bit}, s_{ai,t+1}\}$ is a match $ai_t$, the worker $a$'s period $t$ state (based on his period $t-1$ match $aj_{t-1}$) $s_{ajt}$, the job $i$'s state $s_{bit}$ and the realized state after the match $ai_t$ occurs, $s_{ai,t+1}$. The period $t$ states are like regressors in a linear regression, and the period $t+1$ states and the match are the dependent variables. All states are discrete.

The likelihood contribution of this observation can be factored, using the laws of conditional probability, as

$$o\left(ai_t, s_{ai,t+1} \mid s_{ajt}, s_{bit}\right) = h_{\theta_1}\left(s_{ai,t+1} \mid s_{ajt}, s_{bit}, ai_t\right) \Pr\nolimits^{\star}_{t,\tilde{\beta},\tilde{\gamma},\theta_1,\theta_2,\theta_3}\left(ai_t \mid s_{ajt}\right).$$

Taking the log of $o$ will result in the log of the two components separately. By assumption the utility and profit structural parameters $\left\{\tilde{\beta}, \tilde{\gamma}, \theta_2, \theta_3\right\}$ do not enter the transition rule of the observable states. The parameters $\theta_1$ of the transition density of the worker states can be estimated by maximizing the partial log-likelihood

$$\sum_{t=1}^{T}\sum_{a=1}^{N_t} \log h_{\theta_1}\left(s_{ai,t+1} \mid s_{ajt}, s_{bit}, ai_t\right).$$

This estimator will be consistent. Substituting the estimate $\hat{\theta}_1$ into the likelihood contribution $\Pr^{\star}_{t,\tilde{\beta},\tilde{\gamma},\theta_1,\theta_2,\theta_3}\left(ai_t \mid s_{ajt}\right)$ will allow the other parameters $\left\{\tilde{\beta}, \tilde{\gamma}, \theta_2, \theta_3\right\}$ to be estimated in a second step. Given the computational cost of evaluating $\Pr^*$ and the the computational curse of dimensionality in nonlinear optimization, moving the transition density estimates to a first step reduces the computational cost of the second step. The argument for the consistency of the two-step procedure is standard in the literature (Rust, 1987). The partial likelihood procedure is consistent for all parameters, but not as efficient as the joint estimation of all parameters.[16]

## 5.3 Matching likelihood

Estimation maximizes a partial log-likelihood over all observed matches,

$$L\left(\tilde{\beta}, \tilde{\gamma}, \theta_2, \theta_3\right) = \sum_{t=1}^{T}\sum_{a=1}^{N_t} \log \Pr\nolimits^{\star}_{t,\tilde{\beta},\tilde{\gamma},\hat{\theta}_1,\theta_2,\theta_3}\left(ai_t \mid s_{ajt}\right). \tag{7}$$

The worker non-wage utility parameters $\tilde{\beta}$ , the firm non-wage profit parameters $\tilde{\gamma}$, and the error distribution parameters $\theta_2$ and $\theta_3$, if included, enter the model in the primitive forms specified earlier.[17]

[16] I have written the transition density of the non-wage state variables $h_{\theta_1}\left(s_{ai,t+1} \mid s_{ajt}, s_{bit}, ai_t\right)$ to exclude the endogenous wage data, to exploit the transferable utility structure of the model. Wages evolve in a worker's or job's single-agent dynamic programming problem, but they do so through the equilibrium structure of the matching game. Having wages affect the transition of non-wage states (through a direct channel, rather than indirectly through equilibrium matches) may be permissible, but any attempt to have wages enter $h_{\theta_1}$ will be application-specific.

[17] $\theta_2$ and $\theta_3$ parameterize the distributions of tastes for workers and jobs, respectively. Because of the computational need for closed form solutions to matching probabilities, the pure logit can easily only be generalized to the Generalized Extreme Value (GEV) class, which nests the logit and nested logit, among other distributions.

## 5.4 Computation in estimation

Evaluating the likelihood function (7) requires solving for an equilibrium to the matching game. Given the computational costs of solving for a matching equilibrium, which mainly involve trying many starting values, nesting the solution to a matching model inside the likelihood means that it could take hours or even days to evaluate the likelihood only once. The nested solutions method suggested by (Rust, 1987) for single agent dynamic programming models is computationally infeasible. Further, two-step estimators such as Hotz and Miller (1993) are statistically infeasible because they require first-stage nonparametric estimates of $\Pr_t^\star(ai_t \mid s_{ajt})$, which would require a thick matching market: a very large number observationally identical workers and jobs for each state. Asymptotics across matching markets or years will not provide the required number of observations, as the matching equilibrium will adjust to the new distribution of states. Unfortunately, my empirical application studies market thickness, where in thin markets almost by definition I cannot form a nonparametric estimate of $\Pr_t^\star(ai_t \mid s_{ajt})$.

Instead, I adopt a suggestion by Su and Judd (2007) and maximize the likelihood (7) subject to the nonlinear constraints of the model: the equations (5), (1), and (4). The control variables in the maximization routine are both the structural parameters $\left\{\tilde{\beta}, \tilde{\gamma}, \theta_2, \theta_3\right\}$ and the equilibrium model objects: for each market $t$ the $S$ workers' continuation values $V(s_{ajt})$, the $S$ jobs' continuation values $V^f(s_{bit})$, and the $S \times S$ wages $w_{ai}(s_{ajt}, s_{bit})$. Su and Judd (2007) show and it is simple to see that the maximum value of this constrained optimization problem is the same as the maximum value of (7). Also, the structural parameter estimates for $\left\{\tilde{\beta}, \tilde{\gamma}, \theta_2, \theta_3\right\}$ are numerically the same in both methods.[18]

For a matching market, the $T \cdot (S + S + S \times S)$ model unknowns typically far outnumber the number of structural parameters $\left\{\tilde{\beta}, \tilde{\gamma}, \theta_2, \theta_3\right\}$. Therefore, the computational time to estimate the model using maximum likelihood is about equal to the time it takes to solve the model once (or $T$ times), as described in Section 4.[19] Therefore, any model whose equilibrium can be computed can be estimated, and any model that can be estimated can be used to compute counterfactuals. Indeed, the computer code I use for equilibrium computation differs in only about five lines from the code I use for estimation. This is one of the first empirical applications using the suggestion of Su and Judd (2007), and I have found it to be very practical.[20]

## 5.5 Consistency

Sampling error in the matches likelihood arises only because of the workers' $\varepsilon_{ait}$ job-specific taste shocks. The taste shocks are independent across workers and occur after wages are set, so each match $ai_t$'s likelihood contribution is independent from the contribution of other matches. However, the states $s_{ajt}$ and $s_{bit}$ affect the equilibrium wages and hence the equilibrium match probabilities of all agents. Adding new workers and jobs to the market will probably

[18] Just to be sure, I coded the nested equilibrium computation estimator for small matching markets and verified the numerical equivalency of the estimation algorithms.

[19] One should first estimate the model for $T = 1$ to get good starting values for the full data set.

[20] Computing standard errors uses a method for constrained likelihood estimation (Aitchison and Silvey, 1958).

change the matches of agents. Asymptotics where the dependent variable change as new dependent variable data are collected are not addressed here.

The purest asymptotics are as $T \to \infty$: the number of distinct realizations of the repeated matching market increases. In Fox (2007), I argue that a matching estimator can be consistent if $N_t \to \infty$ if $N_t$ is only the number of workers with data. The fiction is that there is labor market with millions of workers, and we have only data on $N_t$ of them. This fiction is similar to asymptotics in the number of independent countries in cross-country regression: we do not model colonization of Antarctica or the Moon. Collecting more data on the same matching market does not change the matches of existing agents, so it is possible to prove consistency this way. However, this large-market fiction is slightly at odds with the equilibrium computation needed to evaluate the likelihood. Therefore, I focus on the simpler process of collecting more years of data here ($T \to \infty$), and refer readers to Fox (2007) for more details on the $N_t \to \infty$ argument.

The model treats different years of data as statistically independent for computing match probabilities and hence for consistency. This is standard in applications of the Rust (1987) single-agent dynamic discrete choice estimator, where his conditional independence assumptions means the taste shocks $\varepsilon_{ait}$ are independent across time. Here, Assumption 1 makes this easier: workers and jobs treat the aggregate matching equilibrium as fixed in their forecasts for the future.

### 5.6 Multiple equilibria

As discussed in Section 4.3, the existence and uniqueness of equilibria likely depends on the distribution of taste shocks. In a model where there are multiple equilibria (or multiple equilibrium match probabilities here as wage data are not used), Su and Judd (2007) point out that an equilibrium that does not generate the data is unlikely to be a local optimum to the likelihood in terms of both the structural parameters and the model's equilibrium objects. Thus, Su and Judd argue that their estimator is consistent if the underlying economic model has multiple equilibria. I have verified their conjecture computationally for private information Nash games with multiple equilibria.

## 6 Identification

### 6.1 Identification of flow utilities from continuation values

Choices in dynamic programming models are governed by choice-specific continuation values, such as $v(s_{at}, ai_t)$. In a single-agent model with some parametric error distribution such as the logit, revealed preference using data on choice probabilities identifies utility differences, such as $v(s_{at}, ai_t) - v(s_{at}, aj_t)$ for all pairs. It requires additional assumptions to identify flow utilities $u_\beta(s_{at}, ai_t)$ from these continuation value differences. There is a growing literature on this topic in both single agent and Nash games models: Rust (1994), Taber (2000), Magnac and Thesmar (2002), Aguirregabiria (2002), Bajari and Hong (2005), Pesendorfer and Schmidt-Dengler (2007), and Heckman and Navarro (2007). All of the lessons from this literature can be applied to repeated matching games.

### 6.2 Breaking a match surplus into worker and firm surpluses

In the match probabilities for the $N = J = 2$ case in (6), the continuation value of worker $a$ for match $ai$ always occurs in a matched pair with the continuation value of the firm $i$ for match $ai$. If this pairing occurs in models with more agents (and numerically I have verified that it appears to), then match probabilities identify the differences between the sum of worker and firm continuation values for a match, $\widetilde{m}(ai_t, s_{ajt}, s_{bit}) - \widetilde{m}(bj_t, s_{ajt}, s_{bit})$, where

$$\widetilde{m}(ai_t, s_{ajt}, s_{bit}) = \alpha_w \tilde{v}(ai_t, s_{ajt}) + \widetilde{v^f}(ai_t, s_{bit})$$

and $\alpha_w = \gamma_w / \beta_w$. Fox (2007) studies nonparametric identification in a static matching game with transferable utility. In a purely static matching game, $\delta = 0$, only the total surplus of a match, $f_{\tilde{\gamma}\tilde{\beta}}(ai_t, s_{ajt}, s_{bit}) = \alpha_w \tilde{u}_{\tilde{\beta}}(ai_t, s_{ajt}) + \tilde{\pi}_{\tilde{\gamma}}(ai_t, s_{bit})$, is identifiable from data on who matches with whom. To apply his results to a dynamic example, say one observes that not many switches occur between firms $i$ and $j$. In a switching costs model, one can identify that switching costs are high, but one cannot identify whether the switching costs are incurred by the worker, say as a psychic disruption cost from moving, or the firm, say as a retraining cost.[21]

In a static model, computing the equilibrium matching probabilities requires data on only the total surplus $f_{\tilde{\gamma}\tilde{\beta}}$. To compute the solutions to the separate worker and firm dynamic programming problems as part of computing an equilibrium to a repeated matching game, one needs to take a stand on breaking $f_{\tilde{\gamma}\tilde{\beta}}$ into $\tilde{u}_{\tilde{\beta}}$ and $\tilde{\pi}_{\tilde{\gamma}}$. Certain state variables may be payoff relevant for workers, and others for firms. These exclusion restrictions may be motivated by economic theory or other empirical evidence. Below I use results from a separate productivity study to argue that the output lost from having high turnover appears small, so I model switching costs as occurring to the worker.

### 6.3 Normalizations

$\alpha_w = \gamma_w / \beta_w$ is not separately identified from any parameters that enter linearly into $\tilde{u}_{\tilde{\beta}}(ai_t, s_{ajt})$. Using data on matches alone, it is not possible to tell whether workers have a higher value of money than firms or higher utility gains from observed states: only the product $\alpha_w \tilde{\beta}$ is identified if all elements of the vector $\tilde{\beta}$ enter flow utility linearly.

[21] Choo and Siow (2006) argue that data on non-matched or single people can identify a composite term $u_\beta(ai_t, s_{ait}) + \beta_w w_{ait}$, which shows that value of the match $ai_t$ over being single, which has a utility normalized to 0. This normalization is not innocuous and further does not identify a structural object, rather the sum of a structural object and an equilibrium object, the wage.

# 7 Estimating the Switching Costs of Engineers

The empirical application of the repeated matching game is to market thickness and switching in the labor market for elite Swedish engineers. I use data on most elite engineers in Sweden for 1970–1990. Key advantages include data on almost all employers, although the data lack firm level accounting data such as output. I do have a job title code that is standardized across firms, which is a rarity.

## 7.1 Current Period Utility Function

For a worker $i$ of age age (before the retirement age $D_a$) and with overall state vector $s_{at}$, his current period utility function if he chooses match $ai_t$ is

$$u_\beta\left(ai_t, s_{ait}\right) + \varepsilon_{ait} = \beta_w w_{ait} + \beta_{\text{dist,age}} \text{dist}\left(ai_{t-1}, ai_t\right) + \sum_{l=1}^{5} \beta_{\text{sc}_l,\text{age}} \text{sc}_l\left(ai_{t-1}, ai_t\right) + \varepsilon_{ait}. \tag{8}$$

The immediate goal is to estimate

$$\tilde{\beta} = \left\{ \left\{ \beta_{\text{dist,age}}, \left\{ \beta_{\text{sc}_5,\text{age}} \right\}_{l=1}^{5} \right\}_{\text{age}=25}^{60} \right\},$$

which describes the relative importance of the switching costs that enter the current period utility function. All the parameters vary by age, as switching behavior decreases with age. As before, $\beta_w$ is not estimated.

Switching costs depend on the relative locations, in employer characteristic space, of the origin and destination firms. The switching cost in geographic distance $\beta_{\text{dist,age}}\text{dist}\left(ai_{t-1}, ai_t\right)$, is an estimated parameter times the log distance in kilometers between the capitals of the Swedish counties in which the old, $ai_{t-1}$, and potential new, $ai_t$, firms are located.[22] The terms $\left\{\text{sc}_l\left(ai_{t-1}, ai_t\right)\right\}_{l=1}^{3}$are indicator variables that refer to employer changes that cause non-geographic switching costs. One term is a base switching cost, which a worker incurs even if he moves to a new employer across the street. An additional switching cost is incurred if the worker switches between firms in different industries. Also, the data show that 21% of observed switches for Swedish engineers in the estimation sample are between establishments owned by the same legal corporation.[23] Another switching cost is incurred if a worker transfers outside of the boundaries of his corporation. Finally, an attractive element of the Swedish data is that data on the occupation of each job are recorded. The final switching cost deals with workers switching occupations.

Larger employers have lower turnover. This can rise endogenously in the model, in part because more jobs are located inside the firm and workers receive a separate shock for each job. Also recall the job-specific working conditions that are valued the same by all workers do not effect equilibrium match probabilities.

[22] $\text{dist}\left(ai_{t-1}, ai_t\right) = 0$ for firms in the same county.

[23] It is important to distinguish between establishments owned by the same legal corporation in order to preserve the geographic nature of an employment choice. The model assumes that a worker makes all within-corporation transfers through the matching game.

The final term in the current period utility function is a worker, age and employer-specific taste shock $\varepsilon_{ait}$. Empirically, the shock allows the model to match the data, where observationally equivalent workers make different choices. Economically, the relative importance of the taste shock compared to other parameters represents the contribution of idiosyncratic life developments to employer switching. If the patterns of turnover in the data are not driven by differences in wages and other characteristics, it is likely that employer switching is caused by the idiosyncratic developments captured by $\varepsilon_{ait}$.[24]

In the model, the main reason for a worker to switch matches is because of taste shocks. There is no scope for experimentation or working at different firms to accumulate different types of general human capital, as postulated by Lazear (2003). Also, the occupational stratification of workers into entry-level employees and managers across firms of different size in some static matching models like Rosen (1982) and Garicano and Rossi-Hansberg (2006) suggest, for example, that as workers gain human capital with experience they may move from being entry-level workers at large firms to managers at small firms. There is some evidence in favor of these deterministic career paths in the data; for example workers do tend to disproportionately move from large to small firms with age. With my 21 year panel these career paths are observable for many engineers and could be included as observable states $s_{ait}$ without conceptual difficulty, but with some computational difficulty. However, I keep the focus on market thickness and switching rather than modeling these deterministic career paths explicitly to keep the empirical work focused on the effectiveness of thin and thick matching markets.

## 7.2 Job profit functions

The empirical specification of the flow profits of a job is trivial:

$$\pi_\gamma(ai_t, s_{bit}) + \varepsilon^f_{ait} = -\gamma_w w_{ait} + \varepsilon^f_{ait}.$$

As $\gamma_w$ cannot be estimated using data only on matches, in practice almost all of the structural surplus is assumed to accrue to firms. Even given this simple firm profit function, the model still has firms being forward looking as part of equilibrium.

## 7.3 Evidence in favor of switching costs independent of stocks of tenure

Fox and Smeets (2007) have firm-level output data and regress firm output on disaggregated labor inputs for firms in large groups of industries. Broadly speaking, manufacturing is the industrial sector employing the most engineers, and so the results from that sector are the most relevant. Fox and Smeets (Table 3) find that workers with one or more

[24]If $\varepsilon_{ijt}$ is autocorrelated across a career for the same employer (an assumption that complicates computation), the model is a statistical matching model where the match quality reflects the worker's idiosyncratic evaluation of the employer's workplace environment. This contrasts to other models of statistical matching, where career-employer shocks primarily affect worker productivities and therefore wages (Jovanovic, 1979a).

years of tenure at a firm produce more output than newcomers, who are workers with zero years of tenure at the firm. However, there is no statistical evidence in favor of an increasing wage premium with more years of tenure than just one. The output data are consistent with a one-time training or adjustment cost. The primitive firm output data seems to support the switching cost rather than slightly more complex stock of tenure specification.

Fox and Smeets use firm output data and it will be more typical to place the switching costs in the profit function of the firm than the utility function of the worker. However, workers can have additional psychic disruption switching costs, so an exclusion restriction for one type of switching cost is not natural. Section (6.2) argued that match data alone cannot distinguish whether the payoff to a match occurs to a worker or to a firm in transferable utility matching games. For this identification reason, I make the firm's profit function not include any data or estimable parameters.

Fox and Smeets also consider industry (but not occupational) tenure. The coefficients on industry tenure are lower than the coefficients on firm tenure. Again, there is no statistical evidence in favor of a sharp returns to industry specific tenure. A switching cost that does not depend on the stock of tenure in an industry seems sufficient. However, as stocks of tenure are often observed in panel data, including worker tenure presents no conceptual difficulty, but does not increase the computational burden of estimation by adding to the number of worker characteristics and hence total model states.

## 7.4 Other models of persistence in employment

The focus in the empirical application is on market thickness and switching. Reliance on switching costs in the model to explain the persistence (not-switching) of employment relationships is just one way to fit the data on market thickness and switching in an equilibrium matching model with forward looking agents. Alternative theories that could explain the evidence include search costs / lack of information about the wages at other firms (Burdett and Mortensen, 1998) as well as time persistence in unobserved, match-specific taste or productivity shocks (Jovanovic, 1979b).

The focus is on switching and market thickness. My goal is not to tell apart these alternative models that can explain switching and market thickness. There are datasets where this might be possible under additional assumptions. For example, Dubé, Hitsch and Rossi (2007) use high frequency price variation (sales) to shift purchase behavior and then to see whether the new products continue to be purchased after the sales end. If so, consumers are said to have switching costs rather than persistent preferences (or search costs / new information). However, there is no such high-frequency price variation in labor markets for elite workers that could identify the differences between switching costs and time persistence of preferences.

Another explanation for more switching in a thicker labor market is correlated preferences across similar choices. This is somewhat easier to add to a matching model. For example, one can use a nested logit rather than the pure logit for $g_{\theta_2}(\varepsilon_{at})$. Thompson (1989) studies the nonparametric identification of the distribution of choice-specific error terms in a single-agent multinomial choice model. With good variation in the states of workers, one might be able to identify these correlated error terms separately from switching costs. I do not pursue this.

### 7.5 Comparisons to some single-agent and some search models

If I modeled only the worker's problem and not the matching equilibrium, one would have to use a statistical selection correction procedure to infer the offered wage distribution from the accepted wage distribution. With 300 or so employers potentially offering different wages, the selection probability would be a function of 300 indices. Without 300 independently moving instruments, one for each firm, any result of a selection procedure would be sensitive to functional form, independence or auxiliary assumptions (Dahl, 2002; Bayer, Khan and Timmins, 2007).

Another issue with some single-agent methods and some equilibrium search models is that the dependent variable is often a worker selecting a certain firm, so in effect the dependent variable is the size of the firm, just like the dependent variable in consumer demand estimation is the market share of a product. The first order explanation for plant and firm size variation in a cross section probably has to do with product market issues such as the demand curves of consumers for different products, and not labor market issues. Single-agent methods have to explain why workers do not pick high-wage firms, while the matching estimator only needs to explain why one worker is a better match than another for a fixed position. Equilibrium search models explain failure to take high wage jobs as an informational friction. This seems like a poor explanation for Swedish graduate engineers, who work in a medium-sized nation for many internationally recognized firms and probably are well aware of the various employers. In any case, their labor union would collect data on wage rates and provide it to their members if informational frictions would otherwise be a major welfare issue.

## 8 Data

In order to estimate my employer choice model, I need data that cover most employers that hire engineers in a national labor market, that track workers as they switch employers, and that have multiple workers per employer. In this section I detail data with these properties.

### 8.1 The Swedish Labor Market for Engineers

This paper uses data on the wages and employer matches of engineers who work as white collar workers in the Swedish private sector. The data are collected by the Swedish Employers' Federation (SAF in Swedish). I use engineers who have completed a five-year undergraduate degree from one of several university equivalents. The official translation of this degree is "graduate engineer." The term "graduate" distinguishes these engineers from those attended who attended less prestigious trade schools. Especially during the sample period 1970–1990, these engineers were among the most highly educated workers in the Swedish private sector. The number of Swedish engineers is constrained by enrollment limitations at the fixed number of government-run technical institutions allowed to offer five-year engineering degrees. Therefore, these workers were likely highly valuable employees that were important to many firms' successes. The use of engineers contrasts with studies of worker behavior that focus on all workers, less-skilled workers, or on unusual workers such as CEOs and professional athletes.

An outcome of both negotiations between firms and unions and the 1974 Law on Employment Security is layoffs in Sweden happen in the inverse order of seniority.[25] This lack of discretion implies that firms may be reluctant to lay off workers. Firing for cause is also very difficult in Sweden, because firings must be justifiable to a union. Therefore, Sweden is an excellent place to study voluntary employer switching.

My focus is on switching between relatively stable firms. I do not model layoffs due to a dramatic downsizing in a firm. These layoffs largely happen for product market and not labor market reasons The product market is not part of the matching model. I clean the data by eliminating establishments that shut down during a particular year. Workers frequently transfer between establishments owned by the same legal company, so the model includes a switching cost term in the current period utility function to fit this aspect of the data.

During the sample period 1970–1990, firms and unions collectively bargained over wages for engineers at the national, industry, and firm levels. National bargaining for white collar workers ended in 1991, after the sample period. Engineers belong to the CF[26] union, which is part of the SACO university graduate union federation, which itself bargained with the Swedish Employers' Federation (SAF) at the national level as part of the PTK white collar workers' cartel. A large fraction of nominal wage increases is due to wage drift, or firm level wage increases. Union bargaining is not explicitly part of the transferable utility matching game.

This paper estimates the relative importance of switching costs in employer switching decisions. In the model, workers have perfect informed about expected outcomes at all employers. I therefore view switching costs as most likely representing psychic costs of disrupting work and living routines, rather than informational frictions. While I believe most switching costs are psychic disruption costs, institutional features can also affect switching costs. In Sweden, the government provides or heavily regulates health care and pension plans, and switching firms does not reduce these welfare state benefits. On the other hand, capital gains taxes on housing may interfere with geographic mobility (Lundborg and Skedinger, 1998). Likewise, some Swedes live in rent-controlled apartments, which the renters cannot use if they switch to employers in different regions.

[25] The legal relationship between employer seniority and layoff risk means that switching costs may be higher in Sweden. However, unemployment for engineers is close to zero over the sample period, so the cost of a layoff is more likely to be a psychic disruption cost, rather than a large salary decrease.

[26] I do not translate Swedish acronyms for labor unions.

## 8.2 Data overview

The data come from the Swedish Employers' Federation (SAF in Swedish), an organization that represents firms in negotiations with labor unions. The SAF is composed of industry-level employers' federations, and the data list the smaller federation an establishment belongs to. I use these sub-affiliations as my measure of an establishment's, and hence a worker's, industry. The five largest employers of graduate engineers are the basic manufacturing sector ("engineering" sector in Swedish), a broad chemicals and related group called ALMEGA, steel and metal, building and construction and finally forestry.

The SAF data aim to cover all workers in member companies of the SAF from 1970–1990. Every year, each establishment owned by a member company is asked to report information on the worker ID code, age, monthly salary, education and sex of all white collar employees.[27] As a result, the data track workers as they move from employer to employer in the private sector, but cannot track workers if they leave the labor force or switch to an uncovered sector.[28] In practice, collection problems mean that not all establishments report in every year. The data contain information on roughly 60% of the Swedish private sector work force.[29] The coverage rate for engineers in the private sector should be higher than 60% because they are more likely to be employed by large manufacturing firms, which choose to join the SAF. 79% of engineers work in the private sector according to the current website of the labor union for engineers,

The most serious problem affecting data coverage is that not all establishments report the educational background of their workers. About 50% of workers in the data have reported schooling. The establishments reporting on schooling tend to focus on the need to employ highly educated workers. Therefore, studying a group of highly educated workers such as engineers minimizes the coverage problem.[30]

I model workers as making employer decisions at each successive career stage. While I do not directly observe total labor market experience, age is a good proxy for experience when workers have roughly the same background and do not take time off from the labor market. Sweden had extremely low reported rates of unemployment during the period (although some workers were in government training programs), and the rate for the most highly educated workers is presumably lower. I consider only male engineers, who are less likely to take time off from work for unmodeled family reasons. A typical male engineer needs to complete his five years of university education and around a year of military service before entering the labor market. For this reason I only include engineers aged 26 and higher. Many Swedish workers choose early retirement plans, and I lose track of employees whose firms promote them to executive positions. For this reason, I do not model workers older than 61, even though the mandatory retirement age is 65.

[27] The data do not contain information on the personal characteristics of workers, such as their family status and nonwage income, or on the financial performance of firms.

[28] There is no particular age pattern for workers who leave the data, except for the expected spike at retirement.

[29] I calculate this based on numbers in Calmfors and Forslund (1990). The main exceptions are the banking industry, which is represented by a different employers' federation, firms that are cooperatively owned by their workers, and firms that do not belong to any employers' federation. The data coverage is much better for workers in manufacturing.

[30] My attempts to accurately predict schooling from other covariates failed.

Table 1: Facts About Establishments: Mean Over Years

| Variable | All Estabs. | Estimation Sample |
|---|---|---|
| # of Establishments | 1816 | 464 |
| Minimum Size for Inclusion | N/A | 4 |
| Mean # Engineers per Establishment | 8.42 | 26.5 |
| Mean # Total White-Collar Workers | 112 | 283 |
| # of Estabs. in Stockholm County | 497 | 151 |
| # of Estabs. in Small Counties | 879 | 191 |
| # of Estabs. in Manufacturing | 513 | 167 |
| # of Estabs. in Industry and Chemicals | 603 | 151 |

Large counties are the 1990 administrative boundaries that include the cities of Goteborg, Stockholm and Malmo. Small counties excluded those three cities.

Finally, I only include engineers who work full-time, which is defined by the SAF as working more than 35 hours in a week.[31]

Swedish firms pay engineers a straight monthly salary. While the data report contractual hours (usually 40) for white collar workers, contractual hours may be a poor proxy for actual hours of work in an office environment.

The SAF data do not follow workers before they enter the labor market, so this paper does not investigate the initial decision of where to work. In a dynamic setting, all workers must pay a switching cost for this initial employer choice. For brand new workers, the elasticity of labor supply with respect to the stochastic process for age–wage profiles at a particular firm may be especially high, and firms may strongly compete to attract new workers into potentially long careers inside a single firm's internal labor market.

## 8.3 Firm Estimation Sample

I use the term *firm* as a generic term to refer to any employer. Throughout most of this paper, the basic unit of analysis is a physical establishment, which I sometimes still refer to as a firm for simplicity. I focus on establishments because geography is a key component in switching costs. I use the term *company* when I need to explicitly deal with issues surrounding a legal organization that owns more than one establishment.

Table 1 lists sample statistics for establishments in my estimation sample. There are an average of 1816 establishments per year. Handling this many choices in a nonlinear discrete choice model is computationally intensive. In order to increase the speed of the estimation procedure, I narrow the sample by eliminating firms with less than four engineers in either the start or end year of a two-year period. Establishments that shut down are excluded as a result. Table 1 labels this truncated sample the estimation sample. From now on, all results use the estimation sample. Table 1 shows that many of the firms are in Stockholm County, and presents the number of establishments in the two largest industries: "manufacturing" and the ALMEGA "industry and chemicals."

[31] Sweden was not a participant in World War II, so the war disrupted the careers of the relevant age cohorts less than the careers of workers in some other countries. Sweden had no baby boom, although over time the government established new technical universities, increasing the number of graduate engineering entering the labor market.

Table 2: Descriptive Statistics about Worker Sample

| Variable | Mean | Std. Dev. |
|---|---|---|
| Yearly Salary in 1990 Crowns | 266,300 | 81,100 |
| Yearly Salary in Approximate US Dollars | 38,000 | 11,600 |
| Normalized Salary Percentage Increase (%) | 3.93 | 5.86 |
| Switching Rate (%) | 4.08 | 19.8 |
| Age | 38.1 | 8.71 |
| Large County (%) | 59.7 | 49.0 |
| Establishment Size | 1463 | 1857 |
| Fellow Engineers at the Establishment | 208 | 284 |
| Total Worker Years | 195,113 | |
| Total Unique Workers | 28,251 | |
| Mean Years per Worker | 6.91 | |

This table uses the estimation sample. Large counties are the 1990 administrative boundaries that include the cities of Goteborg, Stockholm and Malmo.

## 8.4 Worker Estimation Sample

The estimation sample covers 28,251 unique Swedish engineers. Each engineer is observed for an average of 6.9 years.

Table 2 lists descriptive statistics for the engineers in the estimation sample. Monthly salaries are in Swedish crowns normalized to their value in 1990. The yearly percentage growth in salaries reflects the average growth in the normalized wage levels; in this case it is 3.9%. The yearly switching rate is the percentage of workers who switch to another firm, with both the new and old firms remaining in business over the period. In the estimation sample, engineers have a 4.1% probability of switching employers over a yearly period.[32] The variable large county is an indicator variable for working in one of the three largest metropolitan areas in Sweden. A majority of Swedish engineers work in large counties. Establishment size is the number of other white collar workers at a worker's place of employment, while the number of engineers is the number of employees with a five-year engineering degree at that establishment.[33] Clearly, many of the engineers are employed at large establishments. Note that throughout the remainder of the paper, worker age will refer to a worker's age at the end a yearly observation.

[32]In the complete sample of full-time, male engineers, the yearly switching rate is 11.7%. When eliminating establishments that enter or leave the data, the number drops to 4.08%. The matching game does not model product market changes.

[33]The estimation sample does not include all engineers as listed in Table 2, as I delete workers if they work part time, are female, or if they were not in the data in the previous period.

### 8.5 Computational burden of the empirical application

## 9 Descriptive statistics about switching

### 9.1 Plant closings vs. idiosyncratic switching

### 9.2 Industries

Lots in construction

### 9.3 Counties

Lots in Stockholm and its suburb Uppsala, not so much in Goteborg and Malmo. More concentration in Goteborg and Malmo.

### 9.4 Worker career stages

Older workers very unlikely to switch

### 9.5 Firm sizes

Large firms have lower turnover

## 10 Switching cost estimates

### 10.1 Point Estimates

- Relative to standard deviation ($\approx$ 1.28) of logit error term

| Switching cost | Point estimate | Standard error |
|---|---|---|
| Plant | -5.19 | 0.638 |
| Corporation | -1.89 | 0.600 |
| Industry | -0.988 | 0.799 |
| County | -3.72 | 0.376 |
| Occupation | -5.06 | 0.123 |

## 10.2 Statistical fit

Counties, 1978–1983, switch job at all (occupation, plant, etc.)

| County | Data | Model |
|---|---|---|
| Stockholm | 0.27 | 0.34 |
| Jankoping | 0.38 | 0.37 |
| Malmohus | 0.09 | 0.19 |
| Goteborg | 0.04 | 0.04 |
| Kopparberg | 0.27 | 0.25 |
| Gavleborg | 0 | 0.03 |
| Vasternorrland | 0.28 | 0.28 |

Industries

| Industry | Data | Model |
|---|---|---|
| Chemicals | 0.34 | 0.36 |
| Construction | 0.09 | 0.19 |
| Forestry | 0 | 0.03 |
| Manufacturing | 0.30 | 0.30 |
| Petroleum | 0.13 | 0.23 |
| Textiles | 0.04 | 0.04 |

# 11 Counterfactuals

## 11.1 Alternative market structures

Predicted switching rate in Sweden in 1978–1983 if all jobs (plants & occupations) in

| Counterfactual | Change | Switching rate |
|---|---|---|
| County | Same | 0.35 |
| | Different | 0.16 |
| Industry | Same | 0.30 |
| | Different | 0.24 |
| Occup. | Same | 0.51 |
| | Different | 0.17 |
| Plant | Same | 0.67 |
| | Different | 0.28 |
| Corporation | Same | 0.33 |
| | Different | 0.21 |

### 11.2 Entrant firm with unfilled jobs

## 12 Conclusions

## References

**Aguirregabiria, Victor**, "Another Look at the Identification of Dynamic Discrete Decision Processes," February 2002. working paper.

_ **and Pedro Mira**, "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models," *Econometrica*, July 2002, *70* (4), 1519–1543.

_ **and** _ , "Sequential Simulated-Based Estimation of Dynamic Discrete Games," *Econometrica*, January 2007, *75* (1), 1–53. working paper.

**Ahlin, Christian**, "Risk-Matching and Diversification in Joint Liability Borrowing Groups," 2006. working paper.

**Aitchison, J. and SD Silvey**, "Maximum-Likelihood Estimation of Parameters Subject to Restraints," *The Annals of Mathematical Statistics*, 1958, *29* (3), 813–828.

**Akkus, Oktay and Ali Hortacsu**, "The Determinants of Bank Mergers: A Revealed Preference Analysis," 2006. working paper.

**Anderson, S.P., A. de Palma, and J.F. Thisse**, "Discrete Choice Theory of Product Differentiation," 1992.

**Angelov, Nikolay**, "MODELLING FIRM MERGERS AS A ROOMMATE PROBLEM," 2006.

**Atakan, Alp E.**, "Assortative matching with explicit search costs," *Econometrica*, 2006. forthcoming.

**Bajari, Patrick and Han Hong**, "Semiparametric Estimation of a Dynamic Game of Incomplete Information," 2005. working paper.

_ **and Jeremy T. Fox**, "Measuring the Efficiency of an FCC Spectrum Auction," 2007. working paper.

_ **, C. Lanier Benkard, and Jonathan Levin**, "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 2007.

**Bayer, Patrick, Robert McMillan, and Kim Reuben**, "The Causes and Consequences of Residential Segregation, an Equilibrium Analysis of Neighborhood Sorting," January 2002. working paper.

_ **, Shakeeb Khan, and Christoper Timmins**, "Nonparametric Identification and Estimation in a Generalin a Generalized Roy Model," 2007. working paper.

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

**Berry, Steven T.**, "Estimating Discrete-Choice Models of Product Differentiation," *The RAND Journal of Economics*, Summer 1994, *25* (2), 242–262.

_ **, Oliver Linton, and Ariel Pakes**, "Limit Theorems for Estimating Parameters of Differentiated Products Demand Systems," May 2002. working paper.

**Borkovsky, Ron N., Ulrich Doraszelski, and Yaroslav Kryukov**, "A User's Guide to Solving Dynamic Stochastic Games Using the Homotopy Continuation Method," 2007. working paper.

**Boyd, D., H. Lankford, S. Loeb, and J. Wyckoff**, "Analyzing the Determinants of the Matching Public School Teachers to Jobs: Estimating Compensating Differentials in Imperfect Labor Markets," 2003.

**Burdett, Kenneth and Dale T. Mortensen**, "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, May 1998, *2* (39), 257–273.

**Byrd, Richard H., Mary E. Hribar, and Jorge Nocedal**, "An Interior Point Algorithm for Large-Scale Nonlinear Programming," *SIAM Journal on Optimization*, 1999, *9* (4), 877–900.

**Calmfors, Lars and Anders Forslund**, "Wage Formation in Sweden," in Lars Calmfors, ed., *Wage Formation and Macroeconomic Policy in the Nordic Countries*, SNS Forlag, 1990, chapter 2.

**Choo, Eugene and Aloysius Siow**, "Who Marries Whom and Why," *The Journal of Political Economy*, November 2006.

**Dagsvik, John K.**, "Aggregation in Matching Markets," *International Economic Review*, February 2000, *41* (1), 27–57.

**Dahl, Gordon B.**, "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets," *Econometrica*, November 2002, *70* (6), 2367–2420.

**Demange, Gabrielle and David Gale**, "The Strategy structure of two-sided matching markets," *Econometrica*, 1985, *53*, 873–888.

**Doraszelski, Ulrich and Kenneth L. Judd**, "Avoiding the Curse of Dimensionality in Dynamic Stochastic Games," 2007. working paper.

**Dubé, Jean-Pierre, Günter J. Hitsch, and Peter E. Rossi**, "Do Switching Costs Make Markets Less Competitive?," 2007. working paper.

**Ericson, Richard and Ariel Pakes**, "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *The Review of Economic Studies*, January 1995, *62* (1), 53–82.

**Ferrall, Christopher, Kjell G. Salavanes, and Erik Ø. Sørensen**, "Joint Production at the Workplace: Evidence from Norwegian Establishment Data," 2004. working paper.

**Fox, Jeremy T.**, "Estimating Matching Games with Transfers," 2007. working paper.

_ **and Valérie Smeets**, "Do Input Quality and Structural Productivity Estimates Drive Measured Differences in Firm Productivity?," 2007. working paper.

**Gale, David and Lloyd S. Shapley**, "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 1962, *69*, 9–15.

**Garicano, Luis and Esteban Rossi-Hansberg**, "Organization and Inequality in a Knowledge Economy," *The Quarterly Journal of Economics*, 2006. forthcoming.

**Gordon, Nora and Brian Knight**, "The Political Economy of School District Mergers," October 2005. working paper.

**Heckman, James J.**, "Sample Selection Bias as a Specification Error," *Econometrica*, 1979, *47* (1), 153–162.

_ **and Salvador Navarro**, "Dynamic Discrete Choice and Dynamic Treatment Effects," *Journal of Econometrics*, 2007.

**Hopenhayn, Hugo A.**, "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, September 1992, *60* (5), 1127–1150.

**Hotz, V. Joseph and Robert A. Miller**, "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, 1993, *60*, 497–529.

**Jovanovic, Boyan**, "Firm-specific Capital and Turnover," *Journal of Political Economy*, December 1979, *87* (6), 1246–1260.

__ , "Job Matching and the Theory of Turnover," *Journal of Political Economy*, October 1979, *87* (5), 972–990.

**Kelso, Alexander S. and Vincent P. Crawford**, "Job Matching, Coalition Formation, and Gross Substitutes," *Econometrica*, November 1982, *50* (6), 1483–1504.

**Kennan, John and James R. Walker**, "The Effect of Expected Income on Individual Migration Decisions," December 2002. working paper.

**Koopmans, Tjalling C. and Martin Beckmann**, "Assignment Problems and the Location of Economic Activities," *Econometrica*, January 1957, *25* (1), 53–76.

**Kovalenkov, Alexander and Myrna Holtz Wooders**, "Advances in the Theory of Large Cooperative Games and Application to Club Theory: the Side Payments Case," in Carlo Carraro, ed., *The Endogenous Formation of Economic Coalitions*, Edward Elgar: Chelenham, 2003.

**Krusell, P. and A.A. Smith Jr**, "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 1998, *106* (5), 867.

**Lazear, E.P.**, "Firm-Specific Human Capital: A Skill-Weights Approach," 2003.

**Leonard, Herman B.**, "Elicitation of Honest Preferences for the Assignment of Individuals to Positions," *The Journal of Political Economy*, June 1983, *91* (3), 461–479.

**Lundborg, Per and Per Skedinger**, "Capital gains taxation and residential mobility in Sweden," *Journal of Public Economics*, 1998, *67*, 399–419.

**Magnac, Thierry and David Thesmar**, "Identifying Dynamic Discrete Decision Processes," *Econometrica*, 2002, *70*, 801–816.

**McFadden, D.**, *Conditional Logit Analysis of Qualitative Choice Behavior*, Institute of Urban and Regional Development, University of California, 1973.

**Ostrovsky, Michael**, "Stability in Supply Chain Networks," December 2004.

**Pakes, Ariel, Michael Ostrovsky, and Steve Berry**, "Simple Estimators for the Parameters of Dynamic Discrete Games (with Entry/Exit Examples)," *RAND Journal of Economics*, 2007. working paper.

**Park, Minjung**, “M and A Incentives and Outcomes: Evidence from the Mutual Fund Industry,” 2007. working paper.

**Pesendorfer, Martin and Philipp Schmidt-Dengler**, “Asymptotic Least Squares Estimators for Dynamic Games,” April 2007. working paper.

**Rosen, Sherwin**, “Authority, control, and the distribution of earnings,” *The Bell Journal of Economics*, Autumn 1982, *13* (2), 311–323.

**Rust, John**, “Optimal Replacement of the GMC Bus Engines: An Empirical Model of Harold Zurcher,” *Econometrica*, 1987.

_ , “Structural Estimation of Markov Decision Processes,” in R.F. Engle and D.L. McFadden, eds., *Handbook of Econometrics*, Vol. IV, Elsevier Science, 1994, chapter 51, pp. 3081–3143.

**Sattinger, Michael**, “Differential Rents and the Distribution of Earnings,” *Oxford Economic Papers*, March 1979, *31* (1), 60–71.

**Shapley, Lloyd S. and Martin Shubik**, “The assignment game I: the core,” *International Journal of Game Theory*, 1972, *1*, 111–130.

**Shimer, Robert and Lones Smith**, “Assortative Matching and Search,” *Econometrica*, 2000, *68* (2), 343–369.

**Sørensen, Morten**, “How Smart is Smart Money? A Two-Sided Matching Model of Venture Capital,” *Journal of Finance*, 2007.

**Sotomayor, Marilda**, “The multiple partners game,” in Mukul Majumdar, ed., *Equilibrium and dynamics: essays in honour of David Gale*, Macmillan, 1992, chapter 17, pp. 322–336.

**Su, Che-Lin and Kenneth L. Judd**, “Constrained Optimization Approaches to Estimation of Structural Models,” 2007. working paper.

**Taber, Christopher R.**, “Semiparametric identification and heterogeneity in discrete choice dynamic programming models,” *Journal of Econometrics*, 2000, *96* (2), 201–229.

**Thompson, T. Scott**, “Identification of Semiparametric Discrete Choice Models,” September 1989. never published working paper.

**Tinbergen, Jan**, “Some Remarks on the Distribution of Labour Incomes,” in “Proceedings of the Royal Netherlands Academy of Sciences,” International Economic Papers (1951, No 1), 1947.

**Weintraub, Gabriel Y., C. Lanier Benkard, and Benjamin Van Roy**, “Markov Perfect Industry Dynamics with Many Firms,” 2007. working paper.

**Weiss, Yoram**, "Sharing Rules," 2007. unpublished book chapter.

**Yang, Yupin**, "Empirically Identifying the Value of a Brand Alliance: An Application to Professional Team Sports," 2006. working paper.