cemmap

centre for microdata methods and practice

Repeated matching games an empirical framework

Pauline Corblet
Jeremy T. Fox
Alfred Galichon

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP18/25

UKRI Economic and Social Research Council

Repeated Matching Games
An Empirical Framework

Pauline Corblet
NYU Abu Dhabi

Jeremy T. Fox
Rice University & NBER

Alfred Galichon
NYU and Sciences Po*

October 1, 2025

Abstract

We introduce a model of dynamic matching with transferable utility, extending the static model of Shapley and Shubik (1971). Forward-looking agents have individual states that evolve with current matches. Each period, a matching market with market-clearing prices takes place. We prove the existence of an equilibrium with time-varying distributions of agent types and show it is the solution to a social planner's problem. We also prove that a stationary equilibrium exists. We introduce econometric shocks to account for unobserved heterogeneity in match formation. We propose two algorithms to compute a stationary equilibrium. We adapt both algorithms for estimation. We estimate a model of accumulation of job-specific human capital using data on Swedish engineers.

Keywords: Matching, repeated games, stationary equilibrium, empirical estimation

*We thank seminar participants at Aarhus, Bristol, Caltech, Chicago, Columbia, Exeter, Fudan, Irvine, Jinan, Johns Hopkins, MIT, North Carolina, Princeton, Rice, Texas and Tilburg. We thank Stefan Hubner for working on an early application of our method. Fox thanks the National Science Foundation for funding.

1 Introduction

This paper introduces a tractable model of one-to-one, two-sided dynamic matching. Relationship formation is pervasive in economics, and appears in a wide range of settings, such as marriage, labor or health care. Matching models are a key class of analytical tools that predict the formation of relationships. Consequently, they have become important empirical tools alongside the availability of datasets on formed relationships. In these models, agents on both sides on the market are paired based on their observable characteristics, or types. The value generated from a match, usually referred to as the match surplus, depends critically on the interaction of these types. This interdependence makes the two sides of the market economic complements. For instance, on labor markets, where workers match with firms, the match surplus depends on the worker's level of human capital and the firm's productivity. The surplus created through matching is not equally divided but is instead competitively allocated across agents based on their desirability (how much surplus they can potentially produce) and their scarcity (the relative number of agents of a given type present on the market). In the labor market, the share of the surplus obtained by workers takes the form of wages. The scarcer the workers with high human capital, the higher their wages.

Importantly, many matching markets are dynamic in that they involve repeated interactions over time. Couples divorce and remarry, workers change jobs, and patients switch doctors – highlighting the dynamic nature of matching processes in real-world settings. However, much of the traditional literature, such as the celebrated Becker model of marriage, models matching at a point in time, treating each period as an independent market and abstracting away from potential intertemporal linkages. This overlooks an important feature of dynamic environments: while agents' current types drive present matches, those matches, in turn, influence the future evolution of types. For example, a worker's human capital evolves as a function of their employment history, implying that past matches shape future opportunities. Workers know this and account for this future change in their match decision.

In this paper, we develop a model that explicitly incorporates this dynamic feedback into matching games with transferable utility. Agents are forward-looking and have complete information about their potential partners. They internalize how their current matching decisions affect the future trajectories of their types. The model results in a novel framework that we call repeated matching games. The solution concept in this model is a dynamic competitive equilibrium, which can be viewed as an extension of a Walrasian equilibrium to a dynamic setting with complete information. Our approach combines the stable matching problem in the tradition of Becker (1973) and Shapley and Shubik (1971) with Markov decision processes akin to Rust (1987). Static, transferable utility matching games have productively formed the basis for many papers that structurally estimate models of relationship formation (e.g., Dagsvik, 2000; Choo and Siow, 2006; Chiappori et al., 2017; Dupuy and Galichon, 2014; Fox et al., 2018; Galichon and Salanié, 2022). Our repeated matching game can similarly be used in structural work. After assuming that observed matches are the solution to a matching equilibrium and adding appropriate error terms, the model enables us to structurally estimate underlying preference parameters using observed data on matches, generalizing the seminal framework by Choo and Siow (2006).

Our repeated matching game operates in discrete time. Each period, there is a set of active agents. Each agent has a state variable, which is also the type of an agent in the language of static matching games. Making a match or remaining unmatched can affect the evolution of this agent state variable or agent type. Each period, there is a matching market with prices or transfers for different matches. These prices clear the market. Given these prices, each agent selects the best partner in a forward-looking manner. Agents have complete information about the state variables' distributions over time. In other words, each agent picks a partner today taking into account both current structural payoffs and transfers as well as how the relationship choice affects the agent's own state variable and hence the profitability of possibly all matches in future periods. Next period the matching market reopens, new prices are stated and new matches form. Each period should be thought of as long enough for all agents to consider exiting a current match and choosing a new partner. Frictions such as switching costs can be included if desired, for example as one explanation for sticky matches that last multiple periods.

A repeated matching game has both individual and aggregate dynamics. At the individual level, each agent is solving a single-agent dynamic programming problem, where at each period the agent's action is to choose a partner to match with. At the aggregate level, the aggregate state variable of the matching market is the active agents' current set of types or state variables. This aggregate state variable evolves with the decisions of the individual agents. We prove that a decentralized competitive equilibrium exists, meaning that there exist prices in each period so that forward-looking workers and firms make profit-maximizing yet feasible matches. We also prove that the assignment portion of the competitive equilibrium to the decentralized economy satisfies a social planner's problem, as in static, one-to-one matching games with transferable utility (Shapley and Shubik, 1971), thereby generalizing an important characterization of the equilibrium matching in static matching games to our dynamic setting. Solving the social planner's dynamic optimization problem obtains the aggregate matching distribution.

Another important theoretical result is that a stationary equilibrium exists: there is a distribution of individual states such that, after optimal matches are chosen by forward-looking agents in a decentralized competitive equilibrium, the same distribution of states occurs next period. The existence of a stationary equilibrium holds for any admissible parameter vector satisfying the usual finiteness and discounting assumptions and lets the researcher optionally ignore aggregate dynamics by imposing that the matching game is at a stationary equilibrium.

A repeated matching game can be a useful empirical framework for structural estimation of the production function or match surplus function that is the sum of payoffs of the workers ans firms for a given match. We introduce a version of the repeated matching game with econometric errors representing unobserved heterogeneity in the preferences of agents for partner types. The repeated matching game with econometric errors can best be explained as the combination of two touchstone papers in the literature. Choo and Siow (2006) proposes an estimator for static matching games with logit errors. Rust (1987) proposes an estimator for single agent, dynamic discrete choice models, often using logit errors. In our repeated matching game, an agent's discrete choice each period includes whom to match with and faces possibly logit errors for each type of partner. The agent's type in the matching game is also its state variable, as in dynamic discrete choice models. After computing the prices in a competitive equilibrium, our model of an individual agent's behavior coincides with the dynamic discrete choice model in Rust (1987). If we set the discount factors to zero, a repeated version of Choo and Siow (2006) arises.

For the model with econometric errors, we prove that a decentralized competitive equilibrium with time-varying aggregate states exists and the matching in such an equilibrium can be computed by a social planner's dynamic problem. We also prove that a stationary equilibrium exists, which is important for many empirical applications in empirical micro that do not focus on aggregate dynamics.

We introduce and benchmark computational methods to compute both an equilibrium for the model with time-varying aggregate states as well as to directly compute a stationary equilibrium. For both the models with and without econometric errors, we compute the equilibrium matching for the time-varying aggregate state by solving the social planner's Bellman equation for the planner's value function. We approximate the value function using function approximation techniques, such as deep learning, inside value function iteration. Our two algorithms for computing a stationary equilibrium are more novel. One method, MPEC, solves a system of nonlinear equations using a nonlinear programming solver. The second method uses a primal-dual algorithm by Chambolle and Pock (2011). Our benchmarks show that both these methods can scale to problems with many agent types, and the primal-dual algorithm scales better.

Our Supplementary Material discusses structural estimation using data on matches and agent states from a stationary equilibrium. Both the MPEC and the primal-dual algorithms can be extended from equilibrium computation to structural estimation by adding appropriate terms to the mathematical programs to be solved. We also benchmark our two estimators and show that a similar conclusion holds: the primal-dual algorithm scales better with the number of structural parameters.

In an empirical illustration, we use panel data on Swedish engineers who work at private-sector employers to estimate match production as a function of worker and firm types. The engineers' time-varying states are overall experience as well as recent experience in, separately, technical and managerial jobs. The two job-specific measures of human capital accumulate when the worker matches to a job of the relevant type. We estimate the match production function as a function of these worker experience variables and the type of the job, technical or managerial.

To our knowledge, there is not a useful off-the-shelf model from the theory literature that generalizes static matching games such as the ones developed in Gale (1989); Koopmans and Beckmann (1957); Becker (1973); Shapley and Shubik (1971) to a dynamic setting. Yet such a generalization is useful to study a wide array of markets. For instance, entrepreneurs might be generalists who require experience in several roles before launching their own firms (Lazear, 2009). In supplier/assembler matching, lower-quality car part suppliers participating in Toyota's Supplier Development Program might raise the quality of future parts (Fox, 2018).

We use econometric assumptions from the literature on estimating static matching games with a continuum of agents (Choo and Siow, 2006; Chiappori et al., 2017; Fox, 2018; Galichon and Salanié, 2022). Our individual agent problems are dynamic discrete choice models (Miller, 1984; Wolpin, 1984; Pakes, 1986; Rust, 1987). More recently Rosaia (2021) links undiscounted Markov decision processes to static discrete choice models. In terms of dynamic matching, Choo (2015) derives closed-form formulas for a model where matched agents are exogenously separated from the pool of agents who can match. By contrast in our models' equilibrium, agents endogenously separate based in part on the availability of attractive partners. Erlinger et al. (2015) and McCann et al. (2015) use two-period models, where in the first period an agent goes to school and in the second period the agent participates in the labor market. Peski (2021) also focuses on the evolution of individual agent state variables, in his case with a dynamic search model where each period each unmatched agent meets another and accepts or rejects the match. Separations are exogenous and hence unrelated to attractive potential partners, unlike our model. Our model with econometric errors is perhaps mathematically most closely linked to the model of trade in used cars by Gillingham et al. (2022). Used cars in their model are not forward looking. By contrast, both sides of the market are forward looking in our approach. Anderson and Smith (2010) propose a model of dynamic matching where types are fixed, but reputations evolve according to Bayesian updating. We adopt a different focus, in that our model captures agents' anticipation in a change of their own type.

Our model is a strong departure from the large and influential literature on search models (e.g., Burdett and Mortensen, 1998), in which frictions arise from imperfect meeting technologies: agents encounter one another randomly rather than being instantaneously matched at no cost. In particular, Shimer and Smith (2000) and Atakan (2006) combine matching *à la* Becker with search frictions. As already mentioned, a special case of our model includes switching costs by defining the agent states to include previous matches.

Section 2 presents the baseline model of repeated matching games, theoretically showing the existence of both an equilibrium with time-varying aggregate states and a stationary equilibrium. Section 3 describes the model with econometric shocks. Section 4 presents our methods for equilibrium computation. Section 5 is our empirical application to Swedish engineers switching employers. Section 6 concludes.

2 The Baseline Model

2.1 Set Up

Agents match in a one-to-one, two-sided market.[1] We refer to one side of the market as workers and to the other side as firms. Let $x \in \mathcal{X}$ be the state of the worker, with the set of worker states $\mathcal{X}$ being finite. We also call x the type of the worker, recognizing the type can change over time. Let $y \in \mathcal{Y}$ be the firm state, with $\mathcal{Y}$ also finite. A worker with state x can match with any y firm, but the worker also has the option to remain unmatched, which we denote by 0. The choice set of workers is therefore $\mathcal{Y}_0 = \mathcal{Y} \cup \{0\}$ and the choice set of firms is $\mathcal{X}_0 = \mathcal{X} \cup \{0\}$.

Our model is one of large numbers of both workers and firms, which we conjecture is required for results that rely on real numbers, such as the coming result on the existence of a stationary equilibrium. Therefore, we assume that there is a continuum of workers and a continuum of firms.

We consider an infinite-horizon model in which periods are discrete and the matching market takes place every period. Workers and firms discount the future at rate $\beta < 1$. Note that the horizon is the horizon for the entire economy, rather than the horizon for an individual worker or firm, which can be finite by placing worker or firm age in the state variables. The worker and firm states evolve according to known transition rules that are functions of the current match (x, y). The conditional probability mass function for the worker's next state x' if at current state x and matched to state y is

$$P_{x'|xy}$$

and the transition rule for firm state y is

$$Q_{y'|xy}$$

Not restricting these transition rules further is a key aspect of generality relative to some prior work.

In the aggregate economy, we keep track of the masses of workers and firms of each

[1] We conjecture that the results in this paper could be extended to the fairly general case of trading networks, where an agent can in generality make multiple trades/matches as both a buyer and a seller simultaneously, as in Hatfield et al. (2013) and in Section 6 of Azevedo and Hatfield (2018).

type. Let m_x^t be the mass of workers of type x in period t, with $m^t = (m_x^t)_{x\in\mathcal{X}}$ being the vector of masses for all worker states. Likewise, let n_y^t be the mass of firms of type y, with $n^t = (n_y^t)_{y\in\mathcal{Y}}$. The aggregate state of the economy in period t is (m^t, n^t), which contains the masses of all worker and firm types. Additional macro states, like demand shifters for the industry being studied, can be added to the aggregate state with little conceptual difficulty, although we do not pursue that extension. The total masses of workers and firms M and N remain constant over time, i.e. it must always be the case that the aggregate state lies in a bounded set L, as in

$$(m^t, n^t) \in L \equiv \left\{ (m,n) \geq 0 \,\middle|\, \sum_{x\in\mathcal{X}} m_x = M, \sum_{y\in\mathcal{Y}} n_y = N \right\} \quad \forall t.$$

In a proposed outcome to the model in period t, let μ_{xy}^t be the the mass of matches between workers of state x and firms of state y. Likewise μ_{x0}^t is the mass of workers of type x who are unmatched and μ_{y0}^t is the mass of vacant firms. Let $\mu^t = \left(\mu_{xy}^t\right)_{xy\in\mathcal{X}_0\,\mathcal{Y}_0}$ be the matrix of masses of matches, where $\mathcal{X}_0\,\mathcal{Y}_0 = \{(x,y)\,|\,x\in\mathcal{X}_0,\ y\in\mathcal{Y}_0,\ (x,y)\neq(0,0)\}$. In our discussion of estimation in the Supplemental Material, we will have data randomly sampled from μ^t.

Matched agents exchange monetary transfers in equilibrium. Let w_{xy}^t be the monetary transfer paid by y to x when the two are matched. Agents who remain unmatched do not receive or pay any transfers. Let $w^t = (w_{xy}^t)_{xy\in\mathcal{X}\,\mathcal{Y}}$ be the vector of (endogenously determined) wages in period t, which we also refer to as a wage menu. In estimation, we will not use data on monetary transfers, as using data on matches and types only has been the most common data scheme for transferable utility matching games since the early work of Becker (1973) and early structural empirical work by Choo and Siow (2006).

An outcome to the model has matches $\mu(m,n)$ and transfers $w(m,n)$ for all possible aggregate states (m,n). The aggregate state transitions using the matches and the individual state transition rules. We use the shorthand notation $(P\mu, Q\mu)$ for the next period's aggregate state:

$$(P\mu)_x = \sum_{x'\in\mathcal{X},y'\in\mathcal{Y}_0} P_{x|x'y'}\mu_{x'y'} \quad \text{and} \quad (Q\mu)_y = \sum_{x'\in\mathcal{X}_0,y'\in\mathcal{Y}} Q_{y|x'y'}\mu_{x'y'}.$$

Aggregate transitions are deterministic, although adding stochasticity at the aggregate level is conceptually straightforward in our framework. At the individual level, transitions are stochastic according to the rules P and Q. Individual workers may gain or lose human capital in various occupations. At the aggregate level, the total masses M and N and transition probabilities P and Q are exogenously given, while wages w and matches μ are endogenously determined. At the individual level, the wage schedule w is taken as exogenous and determines the matching choice. We describe these mechanisms in the next subsection.

In our empirical application to Swedish engineers, we augment the repeated matching model to include the arrival and departure of workers and jobs in each period in order to match the data.

2.2 Dynamic Competitive Equilibrium

In this section, we start by describing the matching problem solved by individual agents. We then define our solution concept for the model, which we call a dynamic competitive equilibrium.

If a worker of state x matches to a firm of state y in period t, the worker receives flow profit

$$\alpha_{xy} + w_{xy}^t,$$

where α_{xy} is a structural parameter measuring the worker's non-monetary utility and w_{xy}^t is the equilibrium wage paid by firm y to worker x. α_{xy} is the same in every period and it captures amenities perceived by the worker, while w_{xy}^t can change over time. If the worker is unmatched, he or she does not receive a transfer and we also assume zero amenities, $\alpha_{x0} = 0$.

The wages in period t are $w^t = (w_{xy}^t)_{xy}$, the tuple of wages for all xy pairs. Let $w = (w^t)_t$ be the wage schedule, the infinite series of wages. The worker is forward looking and chooses a partner y^t in every period t to maximize his or her expected present discounted value of lifetime profit, or

$$\mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t \left(\alpha_{x^t y^t} + w_{x^t y^t}^t\right) | x^0 = x\right],$$

where x^t is the worker's state variable in period $t \geq 1$ and y^t is the firm partner type picked that period. Wages $(w^t)_t$ are taken as given by the individual. Because the individual state transitions are stochastic, future states are random variables. The expectation is taken over the future sequence of individual states x. In the next section, we detail how wages are set in every period depending on the aggregate state (m, n).

The problem can be analyzed recursively using the workers' Bellman equation:

$$U_x^t\left(w^{(t)}\right) = \max_{y \in \mathcal{Y}_0} \left\{ \alpha_{xy} + w_{xy}^t + \beta \sum_{x' \in \mathcal{X}} P_{x'|xy} U_{x'}^{t+1}\left(w^{(t+1)}\right) \right\}, \tag{1}$$

where $w^{(s)} = (w^t)_{t \geq s}$. The function $U_x^t\left(w^{(t)}\right)$ is the continuation value for a worker with state variable x choosing from a menu of wages $w^{(t)}$. The sum $\sum_{x' \in \mathcal{X}} P_{x'|xy} U_{x'}\left(w^{(t+1)}\right)$ is the expected continuation value in the next period.

Symmetrically, a firm of type y has flow profit

$$\gamma_{xy} - w_{xy}^t,$$

where γ_{xy} is the non-transfer portion of profit accruing directly to the firm, its output. If the firm is unmatched, it pays no wages and has no output, $\gamma_{0y} = 0$. The firm's Bellman equation is

$$V_y^t\left(w^{(t)}\right) = \max_{x \in \mathcal{X}_0} \left\{ \gamma_{xy} - w_{xy}^t + \beta \sum_{y' \in \mathcal{Y}} Q_{y'|xy} V_{y'}^{t+1}\left(w^{(t+1)}\right) \right\}, \tag{2}$$

where $V_y^t\left(w^{(t)}\right)$ is the continuation value for a firm with state variable y choosing from a menu of wages $w^{(t)}$.

Given the series of wages $w^{(t)}$, the worker's and firm's problems are akin to one-sided problems. Each worker and each firm is solving a dynamic discrete choice problem, where the discrete choice is a partner type. In the next section, we specify how wages adjust to clear the market at the aggregate level and are taken as given by individual agents. Other discrete choices, like the decision to undertake an explicit investment to improve a state variable, can be added to the model without changing its basic mathematical structure.

Example. To illustrate the model, we use the following example later in this section. Consider two types of workers and two types of firms. The total masses of workers and firms are 1 each. Worker and firm types are either high, h, or low l. We use the vector of amenities:

$$\alpha = [\alpha_{ll}\,\alpha_{lh}\,\alpha_{hl}\,\alpha_{hh}\,\alpha_{l0}\,\alpha_{h0}] = [1\,2\,2\,4\,0\,0]\,.$$

Workers and firms are subject to transition rules P and Q. Matching with a given type in period t gives agents a high probability to transition to this type themselves in period $t+1$. Let

$$P = \begin{bmatrix} P_{l|ll} & P_{l|lh} & P_{l|hl} & P_{l|hh} & P_{l|l0} & P_{l|h0} \\ P_{h|ll} & P_{h|lh} & P_{h|hl} & P_{h|hh} & P_{h|l0} & P_{h|h0} \end{bmatrix} = \begin{bmatrix} .8 & .3 & .6 & .2 & .9 & .1 \\ .2 & .7 & .4 & .8 & .1 & .9 \end{bmatrix}$$

For simplicity, we set firms' outputs γ and transitions Q to be the same as workers' amenities and transitions:

$$\gamma = \alpha \text{ and } Q = P.$$

□

As in the static matching game literature such as Shapley and Shubik (1971), the solution concept for our model is competitive equilibrium, which we refer to as dynamic competitive equilibrium.

Matching masses are $\mu = (\mu^t)_t$, where μ^t is the tuple $\big((\mu^t_{xy})_{xy}, (\mu^t_{x0})_x, (\mu^t_{0y})_y\big)$ in a time period t. We say that matching μ is feasible for an aggregate state (m, n) if it satisfies

$$\sum_{y\in\mathcal{Y}_0} \mu^0_{xy} = m_x \quad \text{and} \quad \sum_{x\in\mathcal{X}_0} \mu^0_{xy} = n_y \tag{3}$$
$$\sum_{x'\in\mathcal{X},y'\in\mathcal{Y}_0} P_{x|x'y'}\mu^t_{x'y'} = \sum_{y\in\mathcal{Y}_0} \mu^{t+1}_{xy} \quad \text{and} \quad \sum_{x'\in\mathcal{X}_0,y'\in\mathcal{Y}} Q_{y|x'y'}\mu^t_{x'y'} = \sum_{x\in\mathcal{X}_0} \mu^{t+1}_{xy}.$$

The first two equations ensure that μ^0 sums to the initial aggregate masses m and n. The last two impose that μ^{t+1} sums to the aggregate masses in $t+1$ since

$$\sum_{x'\in\mathcal{X},y'\in\mathcal{Y}_0} P_{x|x'y'}\mu^t_{x'y'} = m^{t+1}_x \text{ and } \sum_{x'\in\mathcal{X}_0,y'\in\mathcal{Y}} Q_{y|x'y'}\mu^t_{x'y'} = n^{t+1}_y.$$

In the following Definition 1, (μ, w) is a tuple of matches and wages.

Definition 1. (μ, w) is a dynamic competitive equilibrium (DCE) if it is feasible for all $(m, n) \in L$ and for all $\bar{x} \in \mathcal{X}$, $\bar{y} \in \mathcal{Y}$, a positive matching mass $\mu^t_{\bar{x}\bar{y}} > 0$ in period t

implies the match between worker $\bar{x}$ and firm $\bar{y}$ maximizes both agents' profits as in

$$\mu^t_{\bar{x}\bar{y}} > 0 \Rightarrow \begin{cases} \bar{y} \in \arg\max_{y \in \mathcal{Y}_0} & \alpha_{\bar{x}y} + w^t_{\bar{x}y} + \beta \sum_{x' \in \mathcal{X}} P_{x'|\bar{x}y} U^{t+1}_{x'} \left(w^{(t+1)}\right) \\ \bar{x} \in \arg\max_{x \in \mathcal{X}_0} & \gamma_{x\bar{y}} - w^t_{x\bar{y}} + \beta \sum_{y' \in \mathcal{Y}} Q_{y'|x\bar{y}} V^{t+1}_{y'} \left(w^{(t+1)}\right) \end{cases}, \tag{4}$$

where U^{t+1} and V^{t+1} are the agents' continuation values given w, as defined in (1) and (2).

In a DCE, each agent is maximizing its expected, present-discounted sum of profits.

2.3 The Social Planner Problem

Solving for the decentralized dynamic competitive equilibrium using Definition 1 presents significant challenges, as it requires computing agents' continuation values on the entire space of possible wage menus. Instead of directly computing the decentralized equilibrium, we extend a key result from static matching games to our repeated matching game. We consider a social planner's problem, show the existence of a solution, and show that the solution is also the matching portion of a decentralized dynamic competitive equilibrium. Therefore, we prove the equivalent of the social planner's property in the static model of Shapley and Shubik (1971) for our dynamic model. If we set the discount factor β to be zero for both workers and firms, the static social planner result of Shapley and Shubik would apply to each period separately.

Static matching games with transferable utility often highlight the importance of the total flow surplus or production of a match:

$$\Phi_{xy} = \alpha_{xy} + \gamma_{xy}.$$

We will also focus on the match production.

Given an initial aggregate state (m, n), the social planner's problem is to maximize the present discounted value of economywide profits $W(m, n)$:

$$W(m,n) = \max_{\mu^t \geq 0} \left\{ \sum_{t=0}^{\infty} \beta^t \sum_{xy \in \mathcal{X}_0 \times \mathcal{Y}_0} \mu^t_{xy} \Phi_{xy} \right\}$$

$$\text{s.t} \sum_{y \in \mathcal{Y}_0} \mu^0_{xy} = m_x \quad \text{and} \quad \sum_{x \in \mathcal{X}_0} \mu^0_{xy} = n_y \tag{5}$$

$$\sum_{x' \in \mathcal{X}, y' \in \mathcal{Y}_0} P_{x|x'y'} \mu^t_{x'y'} = \sum_{y \in \mathcal{Y}_0} \mu^{t+1}_{xy} \quad \text{and} \quad \sum_{x' \in \mathcal{X}_0, y' \in \mathcal{Y}} Q_{y|x'y'} \mu^t_{x'y'} = \sum_{x \in \mathcal{X}_0} \mu^{t+1}_{xy}.$$

The constraints are the same as the feasibility constraints (3).

The primal problem can be analyzed recursively using the social planner's Bellman equation

$$W(m,n) = \max_{\mu \in \mathcal{M}(m,n)} \left\{ \sum_{xy \in \mathcal{X}_0 \mathcal{Y}_0} \mu_{xy} \Phi_{xy} + \beta W\left(P\mu, Q\mu\right) \right\}, \tag{6}$$

where $\mathcal{X}_0 \mathcal{Y}_0 = \{x \in \mathcal{X}_0, y \in \mathcal{Y}_0 \mid (x,y) \neq (0,0)\}$ and $\mathcal{M}(m,n)$ is the set of matchings that satisfy the feasibility constraints

$$\mathcal{M}(m,n) = \left\{ (\mu_{xy})_{xy \in \mathcal{X}_0 \mathcal{Y}_0} \geq 0 \,\middle|\, \sum_{y \in \mathcal{Y}_0} \mu_{xy} = m_x, \sum_{x \in \mathcal{X}_0} \mu_{xy} = n_y \right\}.$$

The present discounted value of economy-wide profits $W : L \to \mathbb{R}$ is a function from the space of aggregate states $L = \left\{ (m,n) \geq 0 \,\middle|\, \sum_{x \in \mathcal{X}} m_x = M, \sum_{y \in \mathcal{Y}} n_y = N \right\}$ to $\mathbb{R}$. The notation $W\left(P\mu, Q\mu\right)$ is shorthand for evaluating the next period's continuation value by applying the worker and firm transition rules, P and Q, to the match masses in the tuple μ.

2.3.1 Solving the Social Planner Problem

We will show that the recursive formulation lets us prove that a unique present discounted value for economywide profit for each aggregate state (m,n) exists across all equilibria.

The social planner problem is a dynamic program with continuous states (m,n) and continuous controls μ. Therefore, the social planner problem fits into classic reference works on such single-agent dynamic programs, such as Stokey et al. (1989).

Proposition 1. There is a unique function $W : L \to \mathbb{R}$ that is solution to equation (6) and it is continuous, bounded, concave, and defined on the entire set L.

Sketch of proof. The proof is in the Supplemental Material, Section S.1.2. It builds on similar results in Stokey et al. (1989). □

A useful corollary of Proposition 1 is that an optimal matching policy $\mu(m,n)$ exists for any aggregate state in L but is not necessarily unique.

Corollary 1. Given the aggregate state (m,n), an optimal matching policy $\mu(m,n)$ exists.

Proof. Existence of an optimal policy derives from the theorems cited in the proposition's proof. $\square$

2.3.2 Using the Social Planner Problem to Solve for the DCE

Solving for the social planner's optimal policy yields a optimal matching policy. We now show that this optimal matching policy is also compatible with a decentralized competitive equilibrium, in two steps. First we derive the dual of the social planner's problem, which allows the calculation of optimal monetary transfers. Second, we show that the optimal matching policy and the optimal monetary transfers obtained in the social planner's primal and dual problems are together a decentralized competitive equilibrium (μ, w).

We define the social planner's cost minimization problem at aggregate state (m,n)

$$
\begin{aligned}
&\min_{U^t,V^t}\left\{\sum_{x\in\mathcal{X}} m_x U_x^0 + \sum_{y\in\mathcal{Y}} n_y V_y^0\right\}\\
&\text{subject to}\\
&U_x^t + V_y^t \geq \Phi_{xy} + \beta\sum_{x'\in\mathcal{X}} P_{x'|xy} U_{x'}^{t+1} + \beta \sum_{y'\in\mathcal{Y}} Q_{y'|xy} V_{y'}^{t+1} \quad \forall t\geq 0,\, x\in\mathcal{X},\, y\in\mathcal{Y}\\
&U_x^t \geq \beta\sum_{x'\in\mathcal{X}} P_{x'|x0} U_{x'}^{t+1} \quad \forall t,\, x\in\mathcal{X}\\
&V_y^t \geq \beta\sum_{y'\in\mathcal{Y}} Q_{y'|0y} V_{y'}^{t+1} \quad \forall t,\, y\in\mathcal{Y}\,.
\end{aligned}
$$

Proposition 2. The social planner's cost minimization problem is the dual of the primal problem and strong duality holds, so that the value of the dual objective at a solution is the same as the value of the primal problem's objective at a solution to that problem, $W(m,n)$.

Sketch of proof. The primal problem is a linear program with a countable number of controls in the objective function and a countable number of constraints. The paper Romeijn and Smith (1998) provides a formulation of the dual and sufficient conditions for strong duality in such countable linear programs. The complete proof is in Appendix A.1.1. □

Strong duality holding is critical for standard equilibrium properties such as the coming existence of a competitive equilibrium. Strong duality holds in our model in part because of time discounting. Strong duality in linear and nonlinear programs with countably infinite controls and constraints is a non-trivial extension over results for finite programs and is still an active area of research in mathematics, as many problems with countably infinite controls that we do not study actually do not satisfy strong duality. While the references in mathematics that we cite in proofs do not explicitly state what further properties hold once strong duality is established, the news is good. For example and just like in the finite controls case, it is simple to show that strong duality holding implies that the Lagrange multipliers of the primal problem constraints are the solutions to the dual problem.

Given an aggregate state (m, n), the social planner's problem admits at least an optimal policy $\mu^*(m, n)$ and its associated Lagrange multipliers $(U^*(m, n), V^*(m, n))$. The next period's aggregate state is $(m', n') = (P\mu^*(m, n), Q\mu^*(m, n))$, for which there is once again an optimal policy $\mu^*(m', n')$ and Lagrange multipliers $(U^*(m', n'), V^*(m', n'))$. Applying the optimal policy successively, starting from $(m^0, n^0) = (m, n)$ and such that $(m^{t+1}, n^{t+1}) = (P\mu^*(m^t, n^t), Q\mu^*(m^t, n^t))$, yields an infinite series of optimal matchings $(\mu^t)_t$ and Lagrange multipliers $(U^t, V^t)_t$. These are the solutions to the infinite horizon formulations of the social planners' primal and dual. In Theorem 1, we show how to use these series to obtain a dynamic competitive equilibrium (DCE).

Theorem 1. *Let $(\mu^t)_t$ and $(U^t, V^t)_t$ be the series of optimal matchings and Lagrange multipliers that solve the social planner problem for the series of aggregate states $(m^t, n^t)_t$. Define transfers $w = (w^t)_t$ that satisfy*

$$\begin{aligned} -V^t_y + \gamma_{xy} + \beta \sum_{y' \in \mathcal{Y}} Q_{y'|xy} V^{t+1}_{y'} &\leq w^t_{xy} \\ &\leq U^t_x - \alpha_{xy} - \beta \sum_{x' \in \mathcal{X}} P_{x'|xy} U^{t+1}_{x'} \quad \forall x \in \mathcal{X}, y \in \mathcal{Y} \end{aligned} \tag{7}$$

Then the tuple (μ, w) is a dynamic competitive equilibrium. Conversely, let (μ, w) be a DCE and let $(U^t, V^t)_t$ be the associated continuation values as given in definition 1.

Then $(\mu^t)_t$ *and* $(U^t, V^t)_t$ *solve the social planner problem.*

The proof of Theorem 1 is in appendix A.1.2. Note that if $\mu^t_{xy} > 0$, then the upper bound of w^t_{xy} coincides with the lower bound, so in this case, the value of w is unambiguously defined.

The social planner's optimal matching policy μ given aggregate state (m, n) is therefore part of a dynamic competitive equilibrium at time t when the aggregate state is $(m^t, n^t) = (m, n)$. Note that the social planner problem gives us a practical way of finding a dynamic competitive equilibrium. Once we have solved for W by value function iteration, we need only solve the primal for a given aggregate state to obtain a matching policy that is part of a dynamic competitive equilibrium. We refer to such a policy as $\mu(m, n)$.

The dynamic competitive equilibrium on the space of aggregate states L depends only on the model parameters: M, N, α, γ, P, Q and β. Typically, the aggregate state (m, n) varies from period to period. The time series $(m^t, n^t)_t$ is deterministic given a starting value (m^0, n^0) for the aggregate state, with the transition rule determined by the optimal matching policy: $(m^{t+1}, n^{t+1}) = (P\mu(m^t, n^t), Q\mu(m^t, n^t))$. In the next section, we show that there exists a constant aggregate state such that $(m^{t+1}, n^{t+1}) = (m^t, n^t)$.

Example. Let us return to our earlier example. Choose $\beta = 0.95$. Given the values of β, α, γ, P and Q that we described previously, we can solve for the social planner's function W (see Section 4 for more details on how we solve for W numerically). Once we have computed W, we can compute the dynamic competitive equilibrium for all aggregate states $(m, n) \in L$.

To illustrate, let us choose three different aggregate states at time $t = 0$:

$$(m^1, n^1) = (.05, .95, .05, .95) \quad (m^2, n^2) = (.95, .05, .95, .05) \quad (m^3, n^3) = (.5, .5, .5, .5).$$

For each of (m^1, n^1), (m^2, n^2) and (m^3, n^3) we can solve for the optimal matching policy and wage: (μ^1, w^1), (μ^2, w^2) and (μ^3, w^3). Given these, we obtain three different next-period aggregate states, for which we can again solve for the optimal policy, and so forth for future periods. The left pane of Figure 1 plots the evolution of the aggregate state of low-type workers l over 15 periods in the model, starting from each of the three aggregate states m^1_l, m^2_l, and m^3_l. All three time series converge to the same constant aggregate state (.46, .54). We discuss constant aggregate states in the next subsection.

At the aggregate level, the aggregate state is a deterministic time series, as shown in the left pane of Figure 1. However, at the individual level, the state variable that the agent reaches next period is stochastic, because of the transition rules P and Q. Consider an initially low-type worker in a world where (m^3, n^3) is the starting aggregate state. The right pane of Figure 1 illustrates three paths the worker can take over time. Many more paths are possible, because the next period's state for each worker is random, as it depends on the current match and the transition rules.

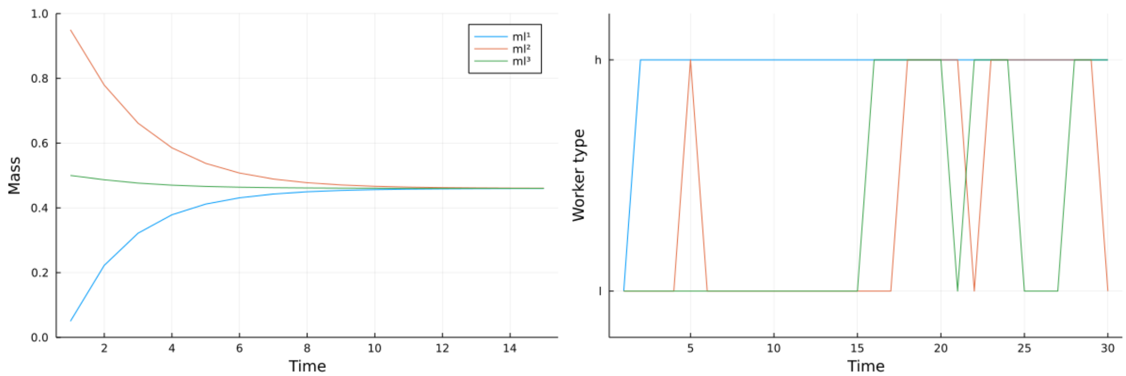


Figure 1: Workers' Aggregate State Evolution (left) and Individual Worker's Possible State Variable Paths (right)

□

2.4 Constant Aggregate State

An aggregate state is constant if it remains the same next period. The matching policy associated to this aggregate state is then stationary.

Definition 2. A constant aggregate state is an aggregate state (m, n) such that there exists a matching μ solution to (6) given (m, n) that satisfies the stationarity conditions

$$m_x = \sum_{y \in \mathcal{Y}_0} \mu_{xy} = \sum_{x' \in \mathcal{X}, y' \in \mathcal{Y}_0} P_{x|x'y'} \mu_{x'y'} \quad \forall x \in \mathcal{X}$$
$$n_y = \sum_{x \in \mathcal{X}_0} \mu_{xy} = \sum_{x' \in \mathcal{X}_0, y' \in \mathcal{Y}} Q_{y|x'y'} \mu_{x'y'} \quad \forall y \in \mathcal{Y} .$$

The constant aggregate state in Definition 2 is such that there is a matching that solves the social planner problem that results in the same aggregate state next period. We refer to the matching policy and wages schedule (μ, w) at a constant aggregate state (m, n) as a stationary equilibrium. Note that the dynamic competitive equilibrium is defined at every aggregate state $(m, n) \in L$ while the stationary equilibrium is only defined at a constant aggregate state.

We show the existence of a constant aggregate state in Theorem 2.[2]

Theorem 2. *A constant aggregate state exists.*

Sketch of proof. We rely on Proposition 1 to show that the set-valued function that yields the next-period aggregate states (m', n') given (m, n) satisfies the conditions of Kakutani's theorem and as such admits a fixed point. The complete proof is in Appendix A.1.3. □

Note that it is straightforward to have individual workers and firms with finite horizons, even when the economy has a finite horizon: one can adapt the transition rules to include absorbing states that effectively end the game for workers or firms. However, introducing a finite horizon to the entire economy changes how one solves the primal problem: It can be solved by backward induction. It is unlikely that there is a constant aggregate state in the finite-horizon problem.

3 Model with Econometric Errors

In the previous section, we defined a dynamic competitive equilibrium in our model and showed that such a equilibrium could be computed by solving a social planner problem. Finally we demonstrated that a constant aggregate state existed. In this section, we introduce econometric shocks to our model and show that the same results hold.

3.1 Specification

The previous model often predicts that some matches never occur, meaning $\mu_{xy} = 0$ for some types x and y. This contradicts available datasets where, with enough

[2]In some cases a constant aggregate state will put all mass at the boundary of the state space. For example, if workers are infinitely lived and accumulate experience deterministically and monotonically, the constant aggregate state will involve all workers having the upper bound on experience.

observations, it is the case that μ_{xy} is rarely or never zero. This contradiction is solved by including random utility terms in the flow profits of both workers and firms. These random utility terms matter for match formation but are unobserved to the econometrician. They can also be called econometric errors, preference shocks, preference heterogeneity, or unobserved state variables.

Let the per period flow profit for worker i of type x matched with a type y firm be

$$u_{xy} + \epsilon_{iy} = \alpha_{xy} + w_{xy} + \epsilon_{iy},$$

where ϵ_{iy} is worker i's preference shock for type y partners. Worker i is indifferent between all partners of the same observed type y. The flow profit from being unmatched is $\alpha_{x0} + \epsilon_{i0}$. Let the flow profit for a firm j of type y matched with a worker of type x be

$$v_{xy} + \eta_{xj} = \gamma_{xy} - w_{xy} + \eta_{xj},$$

where η_{xj} is firm j's preference for workers of type x. The flow profit to a firm for being unmatched is $\gamma_{xy} + \eta_{j0}$.

When we turn to estimation in a later section, the econometrician only knows the distribution of the econometric errors, but not their realizations. We make similar assumptions to Choo and Siow (2006) for static matching games and Rust (1987) for single-agent dynamic discrete choice models.

Assumption 1. The econometric errors satisfy the following assumptions:

1. The distribution of the random vector $(\epsilon_{iy})_{y\in\mathcal{Y}_0}$, where i is randomly drawn within workers of type x, is $\mathcal{L}_{\epsilon|x}$ in every period t. The distribution of the random vector $(\eta_{jx})_{x\in\mathcal{X}_0}$ given individual firm j drawn within firms of type y is $\mathcal{L}_{\eta|y}$ in every period t. The vectors $(\epsilon_{iy})_{y\in\mathcal{Y}_0}$ and $(\eta_{jx})_{x\in\mathcal{X}_0}$ have finite first moments for all y and x.

2. For a single worker i in the two time periods t and $t+1$ with measured states x_i^t and x_i^{t+1}, the distribution of $\left(\epsilon_{iy}^{t+1}\right)_{y\in\mathcal{Y}_0}$ satisfies the following conditional independence property: $\mathcal{L}\left(\left(\epsilon_{iy}^{t+1}\right)_{y\in\mathcal{Y}_0} \middle| x_i^t, x_i^{t+1}, \left(\epsilon_{iy}^{t}\right)_{y\in\mathcal{Y}_0}\right) = \mathcal{L}\left(\left(\epsilon_{iy}^{t+1}\right)_{y\in\mathcal{Y}_0} \middle| x_i^{t+1}\right)$. A similar conditional independence assumption holds for firms.

Under our model with a large number of agents, agents have no market power and therefore it is irrelevant whether these econometric errors are public or private information to the market participants. Also, Part 2 of the assumption states that preferences are drawn anew each time period conditional on measured states x or y, rather than being possibly correlated over time. Considering identification and estimation with unmeasured states that are persistent over time is left to future work.

To ensure the existence of economy-wide profits in the setting with econometric shocks, we also make the following assumption.

Assumption 2. $\forall x \in \mathcal{X}, y \in \mathcal{Y}$, the distributions $\mathcal{L}_{\epsilon|x}$ and $\mathcal{L}_{\eta|y}$ have full support and are absolutely continuous with respect to the Lebesgue measure.

Unmeasured preferences in the literature on estimating static matching games with a small number of matching markets, each with a continuum of agents, are typically preferences over measured partner types x or y rather than unmeasured preferences attributes (Choo and Siow, 2006; Dupuy and Galichon, 2014; Chiappori et al., 2017; Fox, 2018; Galichon and Salanié, 2022). This contrasts with a data scheme of many smaller markets, where agents could have preferences over unmeasured (in data) attributes of partners (Fox et al., 2018).

To ensure that no masses in the constant aggregate state are 0, we also require that all state variables are visited from one period to the next. This assumption ensures the social planner problem with econometric shocks is well defined (see Lemma 3 below).

Assumption 3. For all $x \in \mathcal{X}$ there exists (x, y) such that $P_{x'|xy} > 0$. For all $y \in \mathcal{Y}$ there exists (x, y) such that $Q_{y'|xy} > 0$.

3.2 The Dynamic Competitive Equilibrium with Preference Shocks

3.2.1 General Econometric Shocks

In the model with econometric preference shocks, the worker and firm Bellman equations are changed to make the preference shock realization part of the current period's state variable for each agent:

$$
\begin{aligned}
U_x^t\left(w^{(t)},\epsilon^t\right) &= \max_{y\in\mathcal{Y}_0}\left\{\alpha_{xy}+w_{xy}^t+\epsilon_y^t+\beta\sum_{x'\in\mathcal{X}}P_{x'|x_iy}\mathbb{E}\left[U_{x'}^{t+1}\left(w^{(t+1)},\epsilon^{t+1}\right)\right]\right\}\\
V_y^t\left(w^{(t)},\eta^t\right) &= \max_{x\in\mathcal{X}_0}\left\{\gamma_{xy}-w_{xy}^t+\eta_x^t+\beta\sum_{y'\in\mathcal{Y}}Q_{y'|xy}\mathbb{E}\left[V_{y'}^{t+1}\left(w^{(t+1)},\eta^{t+1}\right)\right]\right\}
\end{aligned}, \tag{8}
$$

where ϵ_y and η_x are the realized econometric shocks, and the expected values in U_x^{t+1} and V_y^{t+1} are taken with respect to the distributions of next period's econometric shocks ϵ^{t+1} and η^{t+1}.

With the worker and firm value functions, we can adapt the notion of a dynamic competitive equilibrium (DCE) from Section A.1 to the model with unobserved heterogeneity. It is computationally attractive to work with aggregates over the realizations of the unobserved heterogeneity terms ϵ_x and η_y. A dynamic competitive equilibrium (DCE) is defined as follows.

Definition 3. In the framework with econometric errors, the tuple (μ,w) is a dynamic competitive equilibrium (DCE) if μ^t corresponds to the probability that each $\tilde{x}$ is optimal for each $\tilde{y}$ and conversely, given the wage schedule w. That is

$$
\begin{aligned}
\tfrac{\mu_{\tilde{x}\tilde{y}}^t}{\sum_{y\in\mathcal{Y}_0}\mu_{\tilde{x}y}^t} &= \Pr\big(\tilde{y}\in\arg\max\nolimits_{y\in\mathcal{Y}_0}\alpha_{\tilde{x}y}+w_{\tilde{x}y}^t+\epsilon_y^t\\
&\qquad +\beta\textstyle\sum_{x'\in\mathcal{X}}P_{x'|\tilde{x}y}\mathbb{E}\left[U_{x'}^{t+1}\left(w^{(t+1)},\epsilon^{t+1}\right)\right]\big)\\
\tfrac{\mu_{\tilde{x}\tilde{y}}^t}{\sum_{x\in\mathcal{X}_0}\mu_{x\tilde{y}}^t} &= \Pr\big(\tilde{x}\in\arg\max\nolimits_{x\in\mathcal{X}_0}\gamma_{x\tilde{y}}-w_{x\tilde{y}}^t+\eta_x^t\\
&\qquad +\beta\textstyle\sum_{y'\in\mathcal{Y}}Q_{y_j'|x\tilde{y}}\mathbb{E}\left[V_{y'}^{t+1}\left(w^{(t+1)},\eta^{t+1}\right)\right]\big),
\end{aligned} \tag{9}
$$

where the expected values are taken over future draws of preference shocks ϵ^{t+1} and η^{t+1}.

As in the case with no econometric shocks, we will use the social planner's problem to compute a DCE. In the case with econometric shocks the social planner problem is said to be regularized, meaning that the objective function contains an additional term that accounts for the econometric shocks. A key role is played by the so-called general entropy function that quantifies the effect of the econometric shocks. The generalized entropy is defined in the following steps. First, introduce the expected indirect payoff functions G_x and H_y given the distributions for the econometric shocks:

$$G_x(u) = \mathbb{E}\left[\max_{y \in \mathcal{Y}_0} \{u_{xy} + \epsilon_y\}\right] \quad \text{and} \quad H_y(v) = \mathbb{E}\left[\max_{x \in \mathcal{X}_0} \{v_{xy} + \eta_x\}\right],$$

where Assumption 2 ensure the max is well defined. Their population counterparts G and H are

$$G(u, m) = \sum_{x \in \mathcal{X}} m_x G_x(u) \text{ and } H(v, n) = \sum_{y \in \mathcal{Y}} n_y H_y(v).$$

The total expected indirect payoffs let us express the generalized entropy as:

$$\mathcal{E}(\mu) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}_0} \mu_{xy} G_x^* \left(\frac{\mu_{x.}}{\sum_{y \in \mathcal{Y}_0} \mu_{xy}} \right) + \sum_{y \in \mathcal{Y}} \sum_{x \in \mathcal{X}_0} \mu_{xy} H_y^* \left(\frac{\mu_{.y}}{\sum_{x \in \mathcal{X}_0} \mu_{xy}} \right),$$

where $\mu_{x.} = (\mu_{xy})_{y \in \mathcal{Y}_0}$, $\mu_{.y} = (\mu_{xy})_{x \in \mathcal{X}_0}$, and G^* and H^* are the Fenchel-Legendre transforms of G and H (Galichon and Salanié, 2022). Our initial requirement that all state variables are visited from one period to the next ensures that $\sum_{y \in \mathcal{Y}_0} \mu_{xy}^t > 0$ and $\sum_{x \in \mathcal{X}_0} \mu_{xy}^t > 0$, so that $\mathcal{E}$ is defined at every time period t. Introduce $\mathcal{M}$ as the set of $\mu \geq 0$ such that μ such that $\sum_{xy \in \mathcal{X}_0 \mathcal{Y}} \mu_{xy} = M$ and $\sum_{xy \in \mathcal{X} \mathcal{Y}_0} \mu_{xy} = N$. The set $\mathcal{M}$ is a closed set, whose interior is the set of vectors μ such that $\mu_{xy} > 0$ for all $xy \in (\mathcal{X}_0 \times \mathcal{Y}) \cup (\mathcal{X} \times \mathcal{Y}_0)$. The transforms G^* and H^* are not defined outside of the interior of $\mathcal{M}$. When some of the $\mu_{xy} = 0$, the following lemma shows that the generalized entropy is bounded on the interior of $\mathcal{M}$.

Lemma 3. *The function $\mathcal{E}$ is defined, continuous, and bounded on the interior of $\mathcal{M}$.*

The proof is in the Supplemental Material, Section S.2.1, and rests on Assumptions 1, 2 and 3. As a result, the function $\mathcal{E}$ can be extended by continuity to the entire set $\mathcal{M}$, and in the sequel we will denote by the same notation $\mathcal{E}$ that extension.

Consider a social planner starting at the aggregate state (m, n). The generalized entropy function $\mathcal{E}$ as defined us allows to write down the social planner's primal problem with econometric errors as the maximization of the expected, present-discounted sum of economywide production under the chosen matching policy, minus the generalized entropy penalty function $\mathcal{E}$:

$$\max_{\mu_{xy}^t \geq 0} \left\{ \sum_{t=0}^{\infty} \beta^t \left(\sum_{x,y \in \mathcal{X}_0 \mathcal{Y}_0} \mu_{xy}^t \Phi_{xy} - \mathcal{E}\left(\mu^t\right) \right) \right\}, \tag{10}$$

subject to the same feasibility constraints and transition rules as in the model without preference shocks,

$$\sum_{y \in \mathcal{Y}_0} \mu_{xy}^0 = m_x \quad \text{and} \quad \sum_{x \in \mathcal{X}_0} \mu_{xy}^0 = n_y, \tag{11}$$

$$\sum_{x' \in \mathcal{X}, y' \in \mathcal{Y}_0} P_{x|x'y'} \mu_{x'y'}^t = \sum_{y \in \mathcal{Y}_0} \mu_{xy}^{t+1} \quad \text{and} \quad \sum_{x' \in \mathcal{X}_0, y' \in \mathcal{Y}} Q_{y|x'y'} \mu_{x'y'}^t = \sum_{x \in \mathcal{X}_0} \mu_{xy}^{t+1}. \tag{12}$$

The regularized social planner's Bellman equation is

$$W(m,n) = \max_{\mu \in \mathcal{M}(m,n)} \left\{ \sum_{xy \in \mathcal{X}_0 \, \mathcal{Y}_0} \mu_{xy} \Phi_{xy} - \mathcal{E}(\mu) + \beta W\left(P\mu, Q\mu\right) \right\}. \tag{13}$$

The solution $W(m,n)$ is unique.

Proposition 3. There exists a unique function $W : L \to \mathbb{R}$ that satisfies (13). It is defined on the entire set L, continuous, bounded and strictly concave.

The proof, which is included in the Supplemental Material, Section S.2.2 for completeness, uses the same reasoning as for Proposition 1.

Corollary 2. The optimal matching policy μ that solves the social planner problem (13) exists and is unique.

Proof. The function $\mu \to \sum_{xy \in \mathcal{X}_0 \, \mathcal{Y}_0} \mu_{xy} \Phi_{xy} - \mathcal{E}(\mu) + \beta W\left(P\mu, Q\mu\right)$ is continuous and strictly concave. We are maximizing on the compact set $\mathcal{M}(m,n)$. Therefore, there exists a unique maximum to the regularized social planner problem. $\square$

It is precisely the nonlinear entropy term $\mathcal{E}(\mu)$ that makes the social planner's problem strictly concave, ensuring that the social planner's problem has a unique solution.

The social planner's problem is a nonlinear program with countably infinite controls and constraints. Mathematical knowledge about this class of problems has recently been growing. We are able to use the recent literature to establish that strong duality holds for the relationship between the social planner's primal problem and an appropriate dual. The dual that we will derive in the proof of the following proposition is

$$
\begin{aligned}
&\inf_{u^t, v^t} \sum_{x\in\mathcal{X}} m_x^0 G_x(u^0) + \sum_{y\in\mathcal{Y}} n_y^0 H_y(v^0) \\
&\text{s.t } u_{xy}^t + v_{xy}^t = \Phi_{xy} + \beta \left(P^\top G(u^{t+1}) + Q^\top H(v^{t+1})\right)_{xy} \\
&\quad u_{x0}^t = \Phi_{x0} + \beta \left(P^\top G(u^{t+1}\right)_{x0} \\
&\quad v_{0y}^t = \Phi_{0y} + \beta \left(Q^\top H(v^{t+1})\right)_{0y}
\end{aligned}
\tag{14}
$$

where $G(u^{t+1})$ and $H(v^{t+1})$ are the stacked vectors of $(G_x(u^{t+1}))_x$ and $(H_y(v^{t+1}))_y$.

Proposition 4. The social planner's primal problem (10) is dual to problem (S.2). Both problems have optimal solutions and strong duality holds, i.e. their value coincide.

The proof is in the Supplemental Material, Section S.2.3, and is the lengthiest in this paper. The key paper Luc and Volle (2021) that the proof uses actually shows a weak notion of strong duality and some effort in the proof is spent building on that work to show a stronger notion of strong duality. Establishing strong duality means that many of the key results from nonlinear programs with a finite number of controls and constraints immediately extend to our nonlinear programs with countably infinite numbers of controls and constraints.

As in the model without econometric errors, solving the social planner problem with entropy is the same as solving for the matching that is part of a decentralized dynamic competitive equilibrium.

Theorem 4. *Let $(\mu^t)_t$ and $(u^t, v^t)_t$ be the series of optimal match masses and Lagrange multipliers that solve the social planner problem for the series of aggregate states $(m^t, n^t)_t$. Define transfers $w = (w^t)_t$ that satisfy*

$$
\begin{aligned}
w_{xy}^t &= u_{xy}^t - \alpha_{xy} - \beta \left(P^\top G_x(u^{t+1})\right)_{xy} \\
&= -v_{xy}^t + \gamma_{xy} + \beta \left(Q^\top H_y(v^{t+1})\right)_{xy}.
\end{aligned}
\tag{15}
$$

Then the tuple (μ, w) is a dynamic competitive equilibrium, DCE. Conversely, let (μ, w) be a DCE and let $(U^t, V^t)_t$ be the associated continuation values. Then $(\mu^t)_t$

and $(U^t, V^t)_t$ solve the social planner problems.

The proof is in the Supplemental Material, Section S.2.4. The proof rests on the use of the strong duality shown in Proposition 4.

3.2.2 The Logit Case

In order to express the competitive matching policy in closed form, we now assume a particular, well-known distribution for the econometric errors, following the literature, and in particular Choo and Siow (2006) and Rust (1987).

Assumption 4. The econometric errors ϵ and η have the type one extreme value (also called the Gumbel) distribution.

Under Assumption 4, the indirect payoffs have the logit form (see Galichon and Salanié (2022) for derivations):

$$G_x(u) = \log \sum_{y \in \mathcal{Y}_0} \exp(u_{xy}) \quad \text{and} \quad H_y(v) = \log \sum_{x \in \mathcal{X}_0} \exp(v_{xy}).$$

Also, the entropy $\mathcal{E}$ penalty term under logit errors is

$$\mathcal{E}(\mu) = \sum_{xy \in \mathcal{X}\,\mathcal{Y}_0} \mu_{xy} \log \frac{\mu_{xy}}{\sum_{y \in Y_0} \mu_{xy}} + \sum_{xy \in \mathcal{Y}\,\mathcal{X}_0} \mu_{xy} \log \frac{\mu_{xy}}{\sum_{x \in \mathcal{X}_0} \mu_{xy}}.$$

Given the logit set up, we can compute certain equations that hold in equilibrium. These are not solutions to the social planner's Bellman equation (13), but more a reformulation of Bellman's equation given the logit errors, as the terms depend on the expected, present-discounted profits U_x and V_y, which are themselves equilibrium objects.

Proposition 5. Under Assumption 4, the dynamic competitive equilibrium matching μ^t in any given period t where the aggregate state is (m^t, n^t) satisfies for all $x \in \mathcal{X}$

and $y \in \mathcal{Y}$

$$\mu^t_{xy} = \sqrt{m^t_x n^t_y} \exp\left(\frac{\Phi_{xy} + \beta \sum_{x' \in \mathcal{X}} U^{t+1}_{x'} P_{x'|xy} + \beta \sum_{y' \in \mathcal{Y}} V^{t+1}_{y'} Q_{y'|xy} - U^t_x - V^t_y}{2}\right)$$

$$\mu^t_{x0} = m^t_x \exp\left(\beta \sum_{x' \in \mathcal{X}} U^{t+1}_{x'} P_{x'|x0} - U^t_x\right)$$

$$\mu^t_{0y} = n^t_y \exp\left(\beta \sum_{y' \in \mathcal{Y}} V^{t+1}_{y'} Q_{y'|0y} - V^t_y\right)$$

where U^t, V^t are the Lagrange multipliers on constraints (11), (12).

Proof. The equilibrium matches arise from calculating the first order conditions of the social planner's primal problem (10). The calculations are omitted for space reasons. □

We use Proposition 5 in our empirical application in Section 5 where we assume a logit set up when econometric shocks are present.

3.3 The Constant Aggregate State with Econometric Errors

We define a constant aggregate state and a stationary equilibrium as in the setting without econometric shocks (Definition 2). We can also show that a constant aggregate state exists in the setting with shocks. The proof uses a different fixed point theorem than the corresponding proof for the model without econometric errors.

Theorem 5. *A constant aggregate state exists in the model with general econometric errors.*

Sketch of proof. We rely on Brouwer's fixed point theorem. The complete proof is in Appendix A.2.1. □

4 Methods for Equilibrium Computation

This section develops methods for computing dynamic competitive equilibria, meaning equilibria with aggregate dynamics, and stationary equilibria, meaning equilibria with a constant aggregate state. We develop algorithms for both the models without and with econometric shocks.

In the non-stationary environment, we are solving a single-agent dynamic programming problem for the social planner and rely on value function iteration, making use of the social planner Bellman equations (6) and (13), as detailed in Section 4.1.

Computing a stationary equilibrium is less computationally intensive since we only need to compute an optimal matching policy for the constant aggregate state, which we solve for. In the Supplemental Material, Section S.3.1, we show that the constant aggregate state without econometric errors can be computed using quadratic optimization. Section 4.1 focuses on the aggregate dynamics with econometric errors, and Section 4.2 leverages two strategies to solve for the constant aggregate state with econometric errors. The first strategy uses the Mathematical Programming with Equilibrium Constraint (MPEC) formulation of our problem (Su and Judd, 2012). The second strategy reformulates the stationary equilibrium equations as a min-max problem and solves it using techniques from convex optimization (Chambolle and Pock, 2011).

Both algorithms for computing a stationary equilibrium for the model with econometric errors are easy to adapt to estimating the model's structural parameters using data on matches from a stationary equilibrium. We discuss structural estimation in the Supplemental Material, Section S.4.

4.1 Aggregate Dynamics

The social planner's Bellman equations with and without econometric errors (6) and (13) are Bellman equations from a single-agent dynamic programming problem. Such problems are most classically solved using value function iteration, exploiting the property that the right side of the Bellman equation is a contraction. In what follows, we explore value function iteration in the model with econometric errors, but most details around using value function iteration also apply to the model without econometric errors.

The state for the social planner's problem is (m, n), the vector of the masses of each worker type and each firm type. Dynamic programming methods of all sorts suffer from a curse of dimensionality in the number of continuous state variables, which for the non-stationary case is equal to the number of worker plus the number of firm types.

Value function iteration operates on a grid $((m_g, n_g))_{g\in\{1,\dots,G\}}$ of nodes, where each node is an aggregate state and G is the chosen number of points in the grid. The $k+1$ iteration of value function iteration is

$$W^{k+1}(m_g, n_g) = TW^k(m_g, n_g)$$

where $TW(m,n) = \max_{\mu \in \mathcal{M}(m,n)} \left\{ \sum_{xy \in \mathcal{X}_0 \mathcal{Y}_0} \Phi_{xy}\mu_{xy} + \beta W(P\mu, Q\mu) - \mathcal{E}(\mu) \right\}$ in the model with econometric errors.

Because the map T is a contraction, $\left(W^k\right)_k$ eventually converges to the fixed point of the social planner's Bellman equation (13). Once the social planner's W is known, the matches $\mu\left(m_g, n_g\right)$ can be computed as the optimal policy of the social planner given W at every node g. Note that because $(P\mu, Q\mu)$ does not necessarily land on a point of the grid $\left((m_g, n_g)\right)_{g \in \{1,\dots,G\}}$, some interpolation technique is needed to compute the value of $W\left(P\mu, Q\mu\right)$ at this point. Both polynomials and deep nets can be used as approximation schemes.

Each iteration of value function iteration has a computational cost that is proportional to the size G of the grid $\left((m_g, n_g)\right)_{g \in \{1,\dots,G\}}$. We can parallelize each iteration across nodes. Additionally, for each node (m_g, n_g) and at each iteration, the continuous optimization problem over match masses μ_{xy} must be calculated. We use the NLOPT solver, which can be called from a wide array of programming languages, and find that the maximization on each point of the grid is solved quickly.[3]

The numerical analysis literature provides a wide array of methods to accelerate fixed-point iterations (Fang and Saad, 2009; Walker and Ni, 2011). Our implementation uses the Anderson acceleration method. Its main idea is to use not only W^k to update to W^{k+1}, but also the values from the previous iterations W^{k-1}, W^{k-2}, ... up to some threshold decided by the analyst. With an aggregate state of dimension 2×2, we run a value function iteration on the $[.01, 1]^2 \times [.01, 1]^2$ grid in 32 minutes.[4]

[3] The algorithm we call through NLOPT is the sequential least-square quadratic programming algorithm (Johnson (2007), Kraft (1994)).

[4] Ran in Julia on 4 threads, on a M2 chip Macbook Pro.

4.2 Constant Aggregate State

We assume a logit set up, so that the optimal policy μ has a closed-form solution, as described in Proposition 5. Let

$$\mu(U,V,U'V',m,n) = (\mu_{xy}(U,V,U'V',m,n), \\ \mu_{x0}(U,V,U'V',m,n), \mu_{0y}(U,V,U'V',m,n))_{x,y\in\mathcal{X}\,\mathcal{Y}}$$

be the function that returns the closed form expressions in Proposition 5, where the first two arguments are the current period payoffs, and the next two are the next period payoffs.[5] We present two numerical methods to compute a constant aggregate state and its associated stationary equilibrium in the model with econometric errors.

4.2.1 Mathematical Programming with Equilibrium Constraints

Mathematical Programming with Equilibrium Constraints, or MPEC, has been used by Su and Judd (2012) to estimate the single-agent dynamic discrete choice model of Rust (1987) and by Dubé et al. (2012) to estimate the aggregate demand model of Berry et al. (1995). MPEC formulates the model as a set of constraints and solves the set using nonlinear programming.

Instead of solving for optimal matching policy μ at each iteration on W, as is done in value function iteration, MPEC focuses on the control variables (m,n,U,V) in the search for a constant aggregate state. The primitives (m,n,U,V) have to satisfy a number of constraints: the feasibility conditions and stationary transition rules outlined previously.[6] The constraints are:

$$\begin{aligned} &\text{Feasibility} \left\{ \begin{array}{l} \sum_{y\in\mathcal{Y}_0} \mu_{xy}(U,V,U,V,m,n) = m_x \,\forall x\in\mathcal{X} \\ \sum_{x\in\mathcal{X}_0} \mu_{xy}(U,V,U,V,m,n) = n_y \,\forall y\in\mathcal{Y} \end{array} \right. \\ &\text{Stationary Transitions} \left\{ \begin{array}{l} \sum_{xy\in\mathcal{X}_0\,\mathcal{Y}} P_{x'|xy}\mu_{xy}(U,V,U,V,m,n) = m_{x'} \,\forall x'\in\mathcal{X} \\ \sum_{xy\in\mathcal{X}\,\mathcal{Y}_0} Q_{y'|xy}\mu_{xy}(U,V,U,V,m,n) = n_{y'} \,\forall y'\in\mathcal{Y}\,. \end{array} \right. \end{aligned} \tag{16}$$

To solve the system of equations (16), we use a nonlinear solver to maximize a

[5] These correspond to (U^t,V^t) and (U^{t+1},V^{t+1}) in Proposition 5, respectively.

[6] The optimal matching policy should also sum to the total masses on each side of the market: $\sum_{xy\in\mathcal{X}\,\mathcal{Y}_0} \mu_{xy}(U,V,m,n) = M$ and $\sum_{xy\in\mathcal{X}_0\,\mathcal{Y}} \mu_{xy}(U,V,m,n) = N$. Because the logit formulas are homogenous of degree 1/2 in (m,n), these two conditions are straightforward to satisfy by adding a constant to U and V.

constant function (say 0) subject to the constraints (16). The problem definition language JuMP in Julia along with the solvers IPOPT and KNITRO are well-suited to the task.[7]

4.2.2 Primal-Dual Method

Our second method relies on the same observation as the first: that the constant aggregate state relies on a set (U, V, m, n) that satisfies the feasibility conditions and the stationary transition rules.[8] The primal-dual approach can be explained in two steps. First we show that when $\beta = 1$, the three sets of conditions mentioned before are the first order conditions to a min-max optimization problem. Such an optimization problem can be solved using a primal-dual algorithm (Chambolle and Pock, 2011). Second, we modify the algorithm to accommodate that agents do discount the future, or that $\beta < 1$. In practise, the algorithm converges to a solution to (16). To vary β between our two steps, we augment the set of parameters of μ by the discount factor β, and consider the following function Z:

$$Z\left(U, V, U', V', m, n, \beta\right) = \sum_{xy \in \mathcal{X}_0 \mathcal{Y}_0} w_{xy} \mu_{xy}(U, V, U', V', m, n, \beta) - \sum_{x \in \mathcal{X}} m_x - \sum_{y \in \mathcal{Y}} n_y,$$

where $w_{xy} = 2$ for $x \in \mathcal{X}, y \in \mathcal{Y}$, $w_{x0} = 1$ for $x \in \mathcal{X}$ and $w_{0y} = 1$ for $y \in \mathcal{Y}$.

The min-max problem we solve for $\beta = 1$ is the following:

$$\min_{U,V} \max_{m,n} Z\left(U, V, U, V, m, n, 1\right). \tag{17}$$

Z is concave in (m, n) and convex in (U, V). Taking first order conditions for the max in (17) yields the feasibility conditions, and doing the same to the min obtains the stationary transition rules. Problem (17) can be solved numerically using the primal-dual algorithm. For the max-min problem (17), the algorithm takes starting values (U^0, V^0, m^0, n^0) and $(m^1, n^1) = (m^0, n^0)$. Given a small increment τ and a

[7]Both solvers use interior point methods.

[8]Total mass normalization can be enforced with the primal-dual method too.

threshold δ, it iterates on $k \geq 1$ according to the following:

$$\begin{aligned}
&\text{Intermediary } (\tilde{m}, \tilde{n}) \begin{cases} \tilde{m}_x^k = 2m_x^k - m_x^{k-1} & \forall x \in \mathcal{X} \\ \tilde{n}_y^k = 2n_y^k - n_y^{k-1} & \forall y \in \mathcal{Y} \end{cases} \\
&(U, V) \text{ update } \begin{cases} U_x^{k+1} = U_x^k - \tau\left(\partial_{U_x} Z^k + \beta^{-1} \partial_{U'_x} Z^k\right) & \forall x \in \mathcal{X} \\ V_y^{k+1} = V_y^k - \tau\left(\partial_{V_y} Z^k + \beta^{-1} \partial_{V'_y} Z^k\right) & \forall y \in \mathcal{Y} \end{cases} \\
&(m, n) \text{ update } \begin{cases} m_x^{k+1} = m_x^k + \tau \partial_{m_x} Z^{k+1} & \forall x \in \mathcal{X} \\ n_y^{k+1} = n_y^k + \tau \partial_{n_y} Z^{k+1} & \forall y \in \mathcal{Y} \end{cases},
\end{aligned} \tag{18}$$

where

$$\begin{cases} Z^k = Z(U^k, V^k, U^k, V^k, \tilde{m}^k, \tilde{n}^k, \beta) \\ Z^{k+1} = Z(U^{k+1}, V^{k+1}, U^{k+1}, V^{k+1}, m^k, n^k, \beta). \end{cases}$$

The stopping criteria is

$$\max\left(\left|U^{k+1} - U^k\right|, \left|V^{k+1} - V^k\right|, \left|m^{k+1} - m^k\right|, \left|n^{k+1} - n^k\right|\right) < \delta.$$

The main feature of the primal-dual algorithm is that it uses $(\tilde{m}^k, \tilde{n}^k)$, an average of (m^k, n^k) and (m^{k-1}, n^{k-1}), to compute the next (U^{k+1}, V^{k+1}). This ensures stability in the algorithm. Chambolle and Pock (2011) show that the algorithm converges when $\beta = 1$, meaning when the feasibility and stability conditions are the first order conditions to optimization problem (17). In practise, we have found that the algorithm converges to (U, V, m, n) that solve the three sets of conditions even when $\beta < 1$.

4.2.3 Methods Comparison

Table 1 describes equilibrium computation performance measures for both MPEC and the primal-dual algorithm, depending on the size of $\mathcal{X}$ and $\mathcal{Y}$, meaning depending on the number of types on both sides of the market.

The MPEC method is faster than the primal-dual method on an equilibrium computation for a small number of types ($\mathcal{X} \times \mathcal{Y} = 2 \times 2$ and $\mathcal{X} \times \mathcal{Y} = 10 \times 10$), but is slower with a large number of types ($\mathcal{X} \times \mathcal{Y} = 30 \times 30$ and $\mathcal{X} \times \mathcal{Y} = 100 \times 100$). An iteration in the primal-dual method is a step in algorithm (18), while an iteration in MPEC is a step in the gradient descent algorithm used by the solver. An iteration in MPEC involves the evaluation of the constraints in (16) as well as the computation of their gradient. The primal-dual algorithm requires many more iterations than MPEC, but each iteration takes up less time.

Table 1: MPEC and Primal-Dual Performance - Equilibrium Computation

$\#\mathcal{X} \times \#\mathcal{Y}$	2×2	10×10	30×30	100×100
		MPEC		
Min iterations	3	5	6	8
Max iterations	4	6	7	8
Mean time elapsed	.0025	.0210	1.342	87.29
		Primal-Dual		
Min iterations	5979	2271	2309	5701
Max iterations	9266	2420	2710	6940
Mean time elapsed	.0242	.0510	.7685	60.93

Notes: Code run in Julia on a Macbook Pro with an M2 chip, 16GB of RAM, and 8 cores. The nonlinear solver for MPEC is KNITRO. Statistics computed on 10 replications. The convergence tolerances are set to 10e-6.

A similar comparison is performed for both MPEC and the primal-dual method adapted to structural estimation in the Supplemental Material, Section S.4.

5 Empirical Application

To illustrate the usefulness of our model applied to labor data, we estimate returns to occupation-specific experience for elite Swedish engineers, those with five-year engineering degrees, in the 1970s and 1980s. Depending on which occupation they are employed in, workers can accumulate different types of human capital. Engineers in particular can hold both technical and managerial human capital. A engineer's productivity in a given occupation depends on the type of human capital acquired and also on his or her total experience in the labor market, meaning across all types of jobs. We estimate our model on Swedish administrative data to measure the respective contributions of occupation-specific human capital and total labor market experience to employer-employee match formation.

We use data on observed matchings between Swedish engineers and firms from 1970 to 1990 from the Swedish Employer's Federation (SAF). The dataset is a yearly panel that follows engineers through time and allows us to reconstruct their past experiences from 1970 onward. We refer the reader to Fox (2009, 2010a) for more background on the data, and see Section S.5 in the Supplemental Matirial for details on the data cleaning. We parameterize the match surplus as a function of the engineer's years of experience in each occupation and his or her age (to proxy for total labor market experience). We structurally estimate the surplus function parameters with the MPEC method and maximum likelihood.

There are some necessary differences in this application from the setup in the previous sections. Our administrative data on employed engineers do not contain vacant jobs or unemployed workers, so our model in this section does not allow for these possibilities. Also, we augment the model to allow workers and jobs to enter and leave the labor market, in order to match the data.

5.1 Model Parameterization

To parameterize the model, we first define workers' and jobs' state variables, or types. A job here is characterized solely by its occupation, general G or technical T:

$$y = 1 \text{ if the job is general, 0 otherwise.}$$

A worker's state variable is three-dimensional: potential experience x_a measured as the difference between the worker's age and 26, technical experience x_t measured as the number of years employed in a technical occupation in the past 5 years, and general experience x_g measured as the number of years employed in a general occupation in the past 5 years:

$$x = (x_e, x_t, x_g) \text{ where } x_e \in \{0, \ldots, 38\},\ x_g \in \{0, \ldots, 5\},\ x_t \in \{0, \ldots, 5\}.$$

We restrict occupation-specific experience to five years because it allows us to use match data from 1975 on, as we cannot measure occupation-specific experience before the start of the panel in 1970. Note that if a worker has been employed in both technical and general jobs in the past five years, he or she holds both technical and general experience: $x_t > 0$ and $x_g > 0$.

Given workers' and jobs' state variables, we parameterize match surplus or production as:

$$\Phi_{xy}(a, b) = a(\tilde{x} - y)^2 + b x_e y, \tag{19}$$

where $\tilde{x} = \frac{x_t}{x_t + x_g}$ is the share of years employed in a technical occupation in the past five years. The share $\tilde{x}$ is a measure of specialization into the technical occupation. We rescale x_e to be between 0 and 1, instead of 0 and 38, so that both $\tilde{x}$ and x_e are between 0 and 1. Also, recall that $y \in \{0, 1\}$.

If the coefficients a and b are both positive, match production is higher when a worker with a general or managerial job has lots of managerial experience and total labor market experience. Fox (2010b) discusses the nonparametric identification of static matching games; parameters like the ratio $2a/b$ are identified in a static matching game without data on unemployed workers and vacant jobs and without relying on the parametric assumption of type 1 extreme value (logit) errors. The ratio $2a/b$ is related to the importance of occupation-specific human capital versus the importance of total labor market experience.[9] Leaving the ratio $2a/b$ aside, the values a and b convert the production from matches between types x and y to standard logit units, as in Choo and Siow (2006) and much followup work, such as Chiappori et al. (2017).

5.2 Estimation

To estimate $\lambda = (a, b)$, we assume logit errors, a discount factor of $\beta = 0.95$ for our annual data, and build on the MPEC strategy exposed in Section S.4 of the Supplemental Material, with two alterations. The first alteration adapts the strategy to the absence of unemployed workers and vacancies in the our administrative dataset.[10] The second alteration accounts for an incoming or outgoing flow of workers to and from the labor market every year, which are added to the stationary transition conditions in equations (16). The flows are introduced to account for workers entering

[9] To gain intuition, consider a static labor market without logit shocks with one managerial job $y = 1$, one technical job $y = 0$, and two workers with types $\left(\tilde{x}^1, x_e^1\right)$ and $\left(\tilde{x}^2, x_e^2\right)$ respectively. Using the social planner result in Shapley and Shubik (1971), algebra shows that worker 1 will match to the managerial job when

$$\frac{x_e^1 - x_e^2}{\tilde{x}^1 - \tilde{x}^2} > \frac{2a}{b}.$$

[10] Sweden had an incredibly low, by modern standards, unemployment rate during 1970–1990 and elite Swedish engineers were even less likely to be unemployed than typical workers. One may presume that there were unfilled vacancies, although our empirical model does not allow for unfilled vacancies due to a lack of data on them.

and exiting the labor market. They are imposed during estimation, such that

$$\sum_{x',y'} P_{x|x'y'}\mu_{x'y'}(U,V,n,m) + i_x = \sum_{y} \mu_{xy}(U,V,n,m)$$
$$\sum_{x',y'} Q_{y|x'y'}\mu_{x'y'}(U,V,n,m) + i_y = \sum_{x} \mu_{xy}(U,V,n,m),$$

where i_x and i_y are net changes in the number of jobs and workers of a given type from year to year.

The observed stationary matching $\hat{\mu}$ is the ratio of the number of observed matches (x,y) over the total number of matches between 1975 and 1990:

$$\hat{\mu}_{xy} = \frac{\sum_{t=1975}^{1990} N_{xy}^t}{\sum_{t=1975}^{1990}\sum_{x,y\in\mathcal{X}\,\mathcal{Y}} N_{xy}^t},$$

where N_{xy}^t is the number of observed (x,y) matches in year t. We choose the $\{1975,\ldots,1990\}$ interval because we use the five years between 1970 and 1974 to measure workers' past occupational experience. By construction, $\hat{\mu}$ sums to 1:

$$\sum_{x,y\in\mathcal{X}\,\mathcal{Y}} \hat{\mu}_{xy} = 1.$$

Therefore we also impose that $\mu(\lambda,U,V,m,n)$ also sums to 1.

The workers' transition matrix P is deterministic: x_t (resp. x_g) is increased by 1 if the worker was employed in a technical occupation in the previous year, with a cap at 5. Workers all gain one year of potential experience every year. Given our characterization of jobs, they do not change types: transition matrix Q has entries $Q_{y'|xy} = 1$ if $y' = y$, and 0 otherwise.

5.3 Results

Table 2 describes observed matching between job types (occupations) and worker types (potential experience and specialization). Workers employed in technical occupations tend to have less total labor market experienced (11.6 years versus 13.5 years) and are a bit more specialized in terms of recent job-specific experience (84.3% of recent experience in technical jobs versus for those currently in technical jobs versus $100\% - 16.6\% = 83.4\%$ in general job experience for those currently in general jobs).

Table 2: Matching Statistics

	In General Job	In Technical Job
Years potential labor market experience	13.5	11.6
% technical experience in last 5 years	16.6	84.3

Authors' calculations from SAF data. See Supplemental Material Section S.5 for details on the data construction.

Table 3: Point Estimates and Standard Errors

	a	b	$\frac{2a}{b}$
Point Estimate	1.83	0.75	4.87
Standard Error	(.001)	(.004)	(.342)

Authors' estimation from SAF data. Bootstrap standard errors computed with 50 bootstrap replications. Potential experience is normalized between 0 and 1.

In our dynamic model, the relative strength of observed matching between workers x and firms y compared to another match can be driven by two factors: a relatively higher surplus Φ_{xy} and next period's workers' expectation $P_{x'|xy}$ to transition to a high return state variable x'. Here, specialized workers could be employed in a technical occupation either because this match has high flow surplus or because workers expect high returns from specializing further in the future. Given the estimated transition matrices, our estimation is able to disentangle the two and estimate the surplus parameters free from the bias of anticipation.

The estimation results for equation (19) are reported in Table 3. The ratio $2a/b$ is equal to 4.87, indicating workers' occupation-specific experience in the past five years matters roughly four times more (subtracting 1 from 4.87 which is approximately 5) for matching into a general or managerial job over a technical job compared to overall potential experience in the labor market.

6 Conclusion

This paper introduces a new repeated matching games that generalizes the static, transferable-utility matching games of Shapley and Shubik (1971) and related work to a repeated matching game, where each period prices and matches form, flow profits are realized by the forward-looking agents, and agent state variables evolve stochastically as a function of current matches. We prove existence of both a time-varying, dynamic competitive equilibrium as well as a stationary equilibrium. We prove the key result that the dynamic competitive equilibrium solves a social planner's problem. Our results are shown both for the baseline model without econometric errors, as in the static Shapley and Shubik (1971) and related work, and a model with econometric errors, as in the static Choo and Siow (2006) and related work.

We provide computational tools for determining both a dynamic competitive equilibrium and a stationary equilibrium, for both the models with and without econometric errors. We show how to modify the models for stationary equilibrium for structural estimation of the parameters in the production to a match, using data on changing relationships over time. Our empirical illustration for elite Swedish engineers finds that recent work experience is roughly four times more important than total labor market experience in sorting into managerial instead of technical jobs.

A Proofs

A.1 Proofs of Results in Section 2

A.1.1 Proposition 2

Proof. We rely on Romeijn and Smith (1998), a paper on linear programming with a countable number of terms in the objective function and a countable number of constraints. This paper reverses the term primal and dual from our usage. In other words, the maximization problem, our primal, is called the dual in Romeijn and Smith. To avoid confusion, we will use the terms from our paper in this proof.

The formal dual from Romeijn and Smith uses control variables $\tilde{U}_x^t$ and $\tilde{V}_y^t$ that are the present discounted values of future utility from the viewpoint of the initial period, so that $\tilde{U}_x^t = \beta^t U_x^t$. Our dual program then becomes

$$\begin{aligned}
&\min_{\tilde{U}^t, \tilde{V}^t} \left\{ \sum_{x \in \mathcal{X}} m_x \tilde{U}_x^0 + \sum_{y \in \mathcal{Y}} n_y \tilde{V}_y^0 \right\} \\
&\text{subject to} \\
&\tilde{U}_x^t + \tilde{V}_y^t \geq \beta^t \Phi_{xy} + \sum_{x' \in \mathcal{X}} P_{x'|xy} \tilde{U}_{x'}^{t+1} + \sum_{y' \in \mathcal{Y}} Q_{y'|xy} \tilde{V}_{y'}^{t+1} \quad \forall t \geq 0,\, x \in \mathcal{X},\, y \in \mathcal{Y}
\end{aligned}$$

$$\tilde{U}_x^t \geq \sum_{x' \in \mathcal{X}} P_{x'|x0} \tilde{U}_{x'}^{t+1} \quad \forall t,\, x \in \mathcal{X}$$

$$\tilde{V}_y^t \geq \sum_{y' \in \mathcal{Y}} Q_{y'|0y} \tilde{V}_{y'}^{t+1} \quad \forall t,\, y \in \mathcal{Y}\,.$$

To verify that this is indeed the dual under the definition used in Romeijn and Smith, we need to show a crosswalk between our paper's notation and the notation used in Romeijn and Smith. We use bars to refer to the symbols in Romeijn and Smith within this proof, as some of their symbols are the same as symbols we use for other purposes.

The crosswalk with Romeijn and Smith is $\bar{i} \to t+1$, $\bar{c}_1 \to (m^\top, n^\top)^\top$, $\bar{c}_{\bar{i}} \to 0 \forall \bar{i} \geq 2$, $\bar{x}_{\bar{i}} \to ((\tilde{U}^{\bar{i}-1})^\top, (\tilde{V}^{\bar{i}-1})^\top)^\top$, $\bar{A}_{\bar{i},\bar{i}-1}$ is a matrix (defined independently of i) of size $(|\mathcal{X}||\mathcal{Y}| + |\mathcal{X}| + |\mathcal{Y}|) \times (|\mathcal{X}| + |\mathcal{Y}|)$ that is such that $\bar{A}_{\bar{i},\bar{i}-1}\binom{\tilde{U}}{\tilde{V}}$ is the vector obtained by stacking $\tilde{U}_x + \tilde{V}_y$, $\tilde{U}_x$ and $\tilde{V}_y$ in the row-major order.[11] Next, $\bar{A}_{\bar{i},\bar{i}} \to -(P^\top, Q^\top)$, $b_{\bar{i}} \to \beta^{\bar{i}-1}\Phi$ where Φ is the vectorized version of the match production matrix including values of 0 for being unmatched, and $\bar{y}_{\bar{i}} \to \mu^{\bar{i}-1}$ where μ is the vectorized version of the matching matrix. As we include outside options with production levels of zero, the nonnegativity constraints on the control variables in Romeijn and Smith will be satisfied in our dual.

Now, that we have derived the dual, we wish to prove strong duality, which means that the optimized objective function values of the primal and dual are equal. We refer to Corollary 3.9 in Romeijn and Smith to prove strong duality. This corollary requires upper bounds on the control variables each period. We let $\bar{u}_{\bar{i}} = (\beta^{\bar{i}-1} \max \Phi)\mathbf{1}_{|\mathcal{X}|+|\mathcal{Y}|}$ be an upper bound on the discounted payoffs of agents of any type. Similarly, we let $\bar{v}_{\bar{i}} = (\max(M, N))\mathbf{1}_{|\mathcal{X}||\mathcal{Y}|+|\mathcal{X}|+|\mathcal{Y}|}$ be an upper bound on the mass of matches of any type.

The condition in Corollary 3.9 is

$$\lim_{\bar{i}\to\infty} \bar{v}_{\bar{i}+1}^\top |\bar{A}_{\bar{i},\bar{i}-1}| \bar{u}_{\bar{i}} = 0$$

where $|\bar{A}_{\bar{i},\bar{i}-1}|$ is the matrix obtained by taking the absolute values of all entries in $\bar{A}_{\bar{i},\bar{i}-1}$ term by term. By substituting in the definitions of the terms, one can see that the this expression is of the order of β^i, and by Corollary 3.9, strong duality is proved.

[11] For example, if $\left|\tilde{X}\right| = |\mathcal{Y}| = 2$, we have $\bar{A}_{\bar{i},\bar{i}-1}\binom{\tilde{U}}{\tilde{V}} = (\tilde{U}_1 + \tilde{V}_1, \tilde{U}_1 + \tilde{V}_2, \tilde{U}_2 + \tilde{V}_1, \tilde{U}_2 + \tilde{V}_2, \tilde{U}_1, \tilde{U}_2, \tilde{V}_1, \tilde{V}_2)^\top$.

$\square$

A.1.2 Theorem 1

Proposition 2 states that strong duality holds for the primal and dual countably infinite linear programs and the text after the theorem states that many properties familiar from the analysis of finite linear programs immediately apply to our problem, given that strong duality holds.

Let's first show the forward direction, where we start with primal and dual solutions and show that a DCE can be found. At time t, let μ^t be the optimal policy of the social planner for the aggregate state (m^t, n^t) and let (U^t, V^t) be the associated Lagrange multipliers on the primal constraints. Then (U^t, V^t) satisfies the dual's inequality conditions, which imply

$$-V_y^t + \gamma_{xy} + \beta \sum_{y' \in \mathcal{Y}} Q_{y'|xy} V_{y'}^{t+1} \leq U_x^t - \alpha_{xy} - \beta \sum_{x' \in \mathcal{X}} P_{x'|xy} U_{x'}^{t+1} \quad \forall x \in \mathcal{X}, y \in \mathcal{Y} \tag{20}$$

with equality if $\mu_{xy}^t > 0$. Thus w^t defined in (7) is indeed well defined, and implies that

$$U_x^t \geq \alpha_{xy} + w_{xy}^t + \beta \sum_{x' \in \mathcal{X}} P_{x'|xy} U_{x'}^{t+1}.$$

holds for all x and t, with an equality if $\mu_{xy}^t > 0$, which implies that

$$U_x^t = \max_{y \in \mathcal{Y}_0} \{\alpha_{xy} + w_{xy}^t + \beta \sum_{x' \in \mathcal{X}} P_{x'|xy} U_{x'}^{t+1}\}.$$

A similar statement can be made for V_y^t, and thus equation (4) in definition 1 is satisfied, which shows that the tuple (μ, w) is a DCE.

For the other direction of the argument, we need to show that a DCE satisfies the optimality conditions of a linear program. Let (μ, w) be a DCE and let $(U^t, V^t)_t$ be the associated continuation values, as in Definition 1. Then every μ^t is feasible by definition, and hence satisfies the social planner's primal problem's feasibility conditions. To check the dual problem's feasibility conditions, note that by definition of the DCE:

$$U_x^t \left(w^{(t)}\right) \geq \alpha_{xy} + w_{xy}^t + \beta \sum_{x' \in \mathcal{X}} P_{x'|xy} U_{x'}^{t+1} \left(w^{(t+1)}\right) \quad \forall x, y \in \mathcal{X}\, \mathcal{Y}_0$$

and

$$V_y^t\left(w^{(t)}\right) \geq \gamma_{xy} - w_{xy}^t + \beta \sum_{y' \in \mathcal{Y}} Q_{y'|xy} V_{y'}^{t+1}\left(w^{(t+1)}\right) \quad \forall x, y \in \mathcal{X}_0 \mathcal{Y}.$$

Adding these two inequalities for every $x, y \in \mathcal{X}\,\mathcal{Y}$ gives

$$\begin{aligned} U_x^t\left(w^{(t)}\right) + & V_y^t\left(w^{(t)}\right) \geq \Phi_{xy} \\ & + \beta\left(\sum_{x' \in \mathcal{X}} P_{x'|xy} U_{x'}^{t+1}\left(w^{(t+1)}\right) + \sum_{y' \in \mathcal{Y}} Q_{y'|xy} V_{y'}^{t+1}\left(w^{(t+1)}\right)\right), \end{aligned} \tag{21}$$

which is the first inequality in the social planner's dual problem as shown in the main text. The next two inequalities in the dual problem are also satisfied by similar reasoning.

Finally, there remain to check the complementary slackness conditions:

$$\begin{aligned} &\mu_{xy}\Bigg(\Phi_{xy} - U_x^t\left(w^{(t)}\right) - V_y^t\left(w^{(t)}\right) \\ &\quad + \beta\left(\sum_{x' \in \mathcal{X}} P_{x'|xy} U_{x'}^{t+1}\left(w^{(t+1)}\right) + \sum_{y' \in \mathcal{Y}} Q_{y'|xy} V_{y'}^{t+1}\left(w^{(t+1)}\right)\right)\Bigg) = 0 \quad \forall x, y \in \mathcal{X}\,\mathcal{Y} \\ &\mu_{x0}\left(-U_x^t\left(w^{(t)}\right) + \beta \sum_{x' \in \mathcal{X}} P_{x'|x0} U_{x'}^{t+1}\left(w^{(t+1)}\right)\right) = 0 \quad \forall x \in \mathcal{X} \\ &\mu_{0y}\left(-V_y^t\left(w^{(t)}\right) + \beta \sum_{y' \in \mathcal{Y}} Q_{y'|0y} V_{y'}^{t+1}\left(w^{(t+1)}\right)\right) = 0 \quad \forall y \in \mathcal{Y}. \end{aligned}$$

These are obtained by the definition of a DCE: if $\mu_{xy}^t > 0$, then option x is optimal for y and conversely, and thus (21) holds as an equality. $(\mu^t)_t$ and $(U^t, V^t)_t$ therefore satisfy the primal equalities, the dual inequalities, and the complementary slackness conditions. Therefore, the components of the DCE are optimal for the social planner problem.

A.1.3 Theorem 2

Proof. Define the set-valued function $\varphi : L \to L$ by

$$\varphi : (m, n) \to \{(P\mu, Q\mu) \mid \mu \,\text{solution to (6) given}\, (m, n)\},$$

where we recall that $L = \Big\{(m,n) | \sum_x m_x = M,\, \sum_y n_y = N\Big\}$. This associates the next period's population counts to the present period ones. There can be multiple solutions μ to problem (6), so φ is a set-valued function.

We show that φ admits a fixed point using Kakutani's theorem. To apply the theorem we need the following:

(1) L is non-empty, compact and convex.

(2) φ has closed graph, where the graph of φ is

$$\mathrm{Gr}_\varphi = \{(m, n, m', n') \in L \times L \,|\, (m', n') \in \varphi(m, n)\}\,.$$

(3) The set $\varphi(m, n)$ is non-empty and convex.

Consider point (1). Clearly L is non-empty. Compactness arises because L is closed and bounded. For convexity, consider two aggregate states (m, n), $(m', n') \in L$. Then, on the worker side, $\sum_x \theta m_x + (1-\theta)m'_x = \theta M + (1-\theta)M = M$ and the same applies on the firm side. Hence the linear combination $(\theta m+(1-\theta)m', \theta n+(1-\theta)n')$ also belongs to L.

To show point (2), we use the closed graph theorem (recalled as Theorem 17.11 in Aliprantis and Border (2006)) for set-valued functions, which states that if $\varphi : L \to L$ is upper hemicontinuous and $\varphi(m, n)$ is a closed set for all $(m, n) \in L$ then Gr_φ is closed.

We use Berge's maximum theorem (Aliprantis and Border (2006), Theorem 17.31) which states that if

(a) the correspondence $\mathcal{C}(m, n) \rightrightarrows \{\mu | \mu \in \mathcal{M}(m, n)\}$ is compact-valued and continuous and

(b) the objective function $\mu \to \sum_{xy \in \mathcal{X}_0\, \mathcal{Y}_0} \Phi_{xy}\mu_{xy} + \beta W(P\mu, Q\mu)$ is continuous,

then the set of solutions μ is upper hemicontinuous in the argument (m, n), with non-empty and compact values. Because the set of solutions is compact and lies in a metric space, the set is also closed, the other condition of the closed graph theorem for set-valued functions.

We now show points (a) and (b). Lemma 6 in the Supplementary Material shows point (a). Point (b) is straightforward because μ enters the per-period payoffs linearly, we know W is uniquely defined across across all solutions and continuous from

Proposition 1, sums like $P\mu$ are themselves linear (continuous) functions of μ, and compositions of continuous functions like $W(P\mu, Q\mu)$ are continuous.

Point (3) of Kakutani's theorem is that $\varphi(m,n)$ is non-empty and convex. We just used the maximum theorem to show that $\varphi(m,n)$ is non-empty and compact. Convexity follows from the fact that $\varphi(m,n)$ is the set of maximizers of the function $\mu \to \sum_{xy\in\mathcal{X}_0\mathcal{Y}_0} \Phi_{xy}\mu_{xy} + \beta W(P\mu, Q\mu)$, which is concave, and therefore, is a convex set.

□

A.2 Proofs of Results in Section 3

A.2.1 Theorem 5

Proof. As in the proof of Theorem 2, define function $\varphi : L \to L$ by

$$\varphi : (m,n) \to (P\mu(m,n), Q\mu(m,n))$$

where $\mu(m,n)$ is the social planner's optimal policy given aggregate state (m,n), i.e. $\mu(m,n)$ solves (13). There is a unique social planner solution to the regularized problem.

We show that φ admits a fixed point using Brouwer's theorem. To apply the theorem we need the following:

(1) L is non-empty, compact and convex.

(2) φ is continuous in (m,n).

Point (1) was shown in the proof of Theorem 2. To show point (2), we need the function $(m,n) \to \mu(m,n)$ to be continuous in (m,n), as $P\mu$ and $Q\mu$ are linear functions of μ. We show continuity of $\mu(m,n)$ using Berge's maximum theorem. We have shown in Corollary 2 that $\mu(m,n)$ is the unique maximizer. To apply the maximum theorem we need the following:

(a) $\mathcal{C} : (m,n) \rightrightarrows \{\mu \,|\, \mu \in \mathcal{M}(m,n)\}$ is a compact-valued and continuous correspondence. This was shown in the proof of Theorem 2.

(b) W is continuous. This is shown in Proposition 3.

Since we have shown (1) and (2), we obtain with Brouwer that φ admits a fixed point. □

References

Aliprantis, C. D. and K. C. Border (2006): *Infinite Dimensional Analysis*, Berlin/Heidelberg: Springer-Verlag. 41

Anderson, A. and L. Smith (2010): "Dynamic Matching and Evolving Reputations," *The Review of Economic Studies*, 77, 3–29. 6

Atakan, A. E. (2006): "Assortative Matching with Explicit Search Costs," *Econometrica*, 74, 667–680. 6

Azevedo, E. M. and J. W. Hatfield (2018): "Existence of Equilibrium in Large Matching Markets With Complementarities," *University of Pennsylvania working paper*. 7

Becker, G. S. (1973): "A Theory of Marriage: Part I," *Journal of Political Economy*, 81, 813–846. 3, 5, 8, 64

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841–890. 29

Burdett, K. and D. T. Mortensen (1998): "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 39, 257–273. 6

Chambolle, A. and T. Pock (2011): "A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging," *Journal of Mathematical Imaging and Vision*, 40, 120–145. 5, 27, 30, 31

Chiappori, P.-A., B. Salanié, and Y. Weiss (2017): "Partner Choice, Investment in Children, and the Marital College Premium," *American Economic Review*, 107, 2109–2167. 3, 6, 20, 34

Choo, E. (2015): "Dynamic Marriage Matching: An Empirical Framework," *Econometrica*, 83, 1373–1423. 6

Choo, E. and A. Siow (2006): "Who Marries Whom and Why," *Journal of Political Economy*, 114, 175–201. 3, 4, 6, 8, 19, 20, 25, 34, 37, 64

Dagsvik, J. K. (2000): "Aggregation in Matching Markets," *International Economic Review*, 41, 27–57. 3

Dubé, J.-P., J. T. Fox, and C.-L. Su (2012): "Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation," *Econometrica*, 80, 2231–2267. 29

Dupuy, A. and A. Galichon (2014): "Personality Traits and the Marriage Market," *Journal of Political Economy*, 122, 1271–1319. 3, 20, 61

Erlinger, A., R. J. McCann, X. Shi, A. Siow, and R. Wolthoff (2015): "Academic Wages and Pyramid Schemes: A Mathematical Model," *Journal of Functional Analysis*, 269, 2709–2746. 6

Fang, H.-r. and Y. Saad (2009): "Two Classes of Multisecant Methods for Nonlinear Acceleration," *Numerical Linear Algebra with Applications*, 16, 197–221. 28

Fox, J., C. Yang, and D. H. Hsu (2018): "Unobserved Heterogeneity in Matching Games," *Journal of Political Economy*, Vol 126, No 4. 3, 20

Fox, J. T. (2009): "Firm-Size Wage Gaps, Job Responsibility, and Hierarchical Matching," *Journal of Labor Economics*, 27, 83–126. 33

——— (2010a): "Estimating the Employer Switching Costs and Wage Responses of Forward-Looking Engineers," *Journal of Labor Economics*, 28, 357–412. 33

——— (2010b): "Identification in Matching Games," *Quantitative Economics*, 1, 203–254. 34

——— (2018): "Estimating Matching Games with Transfers," *Quantitative Economics*, 9, 1–38. 5, 6, 20

Gale, D. (1989): *The Theory of Linear Economic Models*, Chicago, IL: University of Chicago Press. 5

Galichon, A. and B. Salanié (2022): "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models," *The Review of Economic Studies*, 89, 2600–2629. 3, 6, 20, 22, 25, 51, 56, 61

Gillingham, K., F. Iskhakov, A. Munk-Nielsen, J. Rust, and B. Schjerning (2022): "Equilibrium Trade in Automobiles," *Journal of Political Economy*, 130, 2534–2593. 6

Hatfield, J. W., S. D. Kominers, A. Nichifor, M. Ostrovsky, and A. Westkamp (2013): "Stability and Competitive Equilibrium in Trading Networks," *Journal of Political Economy*, 121, 966–1005. 7

Johnson, S. G. (2007): "The NLopt Nonlinear-Optimization Package," https://github.com/stevengj/nlopt. 28

Koopmans, T. C. and M. Beckmann (1957): "Assignment Problems and the Location of Economic Activities," *Econometrica*, 25, 53–76. 5

Kraft, D. (1994): "Algorithm 733: TOMP–Fortran Modules for Optimal Control Calculations," *ACM Transactions on Mathematical Software*, 20, 262–281. 28

Lazear, E. P. (2009): "Firm-Specific Human Capital: A Skill-Weights Approach," *Journal of Political Economy*, 117, 914–940. 5

Luc, D. T. and M. Volle (2021): "Duality for Extended Infinite Monotropic Optimization Problems," *Mathematical Programming*, 189, 409–432. 24, 51, 52, 54, 55

McCann, R. J., X. Shi, A. Siow, and R. Wolthoff (2015): "Becker Meets Ricardo: Multisector Matching with Communication and Cognitive Skills," *Journal of Law, Economics, and Organization*, 31, 690–720. 6

Miller, R. A. (1984): "Job Matching and Occupational Choice," *Journal of Political Economy*, 92, 1086–1120. 6

Pakes, A. (1986): "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," *Econometrica*, 54, 755–784. 6

Peski, M. (2021): "Tractable Model of Dynamic Many-to-Many Matching," *American Economic Journal: Microeconomics*. 6

Romeijn, H. E. and R. L. Smith (1998): "Shadow Prices in Infinite-Dimensional Linear Programming," *Mathematics of Operations Research*. 15, 37, 38

Romeijn, H. E., R. L. Smith, and J. C. Bean (1992): "Duality in Infinite Dimensional Linear Programming," *Mathematical Programming*, 53, 79–97. 57

Rosaia, N. (2021): "Duality and Estimation of Undiscounted Markov Decision Processes," *Working Paper*. 6

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, 999–1033. 3, 4, 6, 19, 25, 29, 64

Shapley, L. S. and M. Shubik (1971): "The Assignment Game I: The Core," *International Journal of Game Theory*, 1, 111–130. 1, 3, 4, 5, 11, 12, 34, 36, 37

Shimer, R. and L. Smith (2000): "Assortative Matching and Search," *Econometrica*, 68, 343–369. 6

Stokey, N. L., R. E. Lucas, and E. C. Prescott (1989): *Recursive Methods in Economic Dynamics*, Harvard University Press. 13, 49, 50, 51

Su, C.-L. and K. L. Judd (2012): "Constrained Optimization Approaches to Estimation of Structural Models," *Econometrica*, 80, 2213–2230. 27, 29

Walker, H. and P. Ni (2011): "Anderson Acceleration for Fixed-Point Iterations," *SIAM J. Numerical Analysis*, 49, 1715–1735. 28

Wolpin, K. I. (1984): "An Estimable Dynamic Stochastic Model of Fertility and Child Mortality," *Journal of Political Economy*, 92, 852–874. 6

Supplement to "Repeated Matching Games: An Empirical Framework"

S.1 Proof of Results in Section 2

S.1.1 Properties of the action set

We need to prove some properties of the social planner's action set $\mathcal{M}(m,n)$ and the related correspondence $\mathcal{C}(m,n) \rightrightarrows \{\mu|\mu \in \mathcal{M}(m,n)\}$. We start with the following lemma.

Lemma 6. *1. $\mathcal{C}$ is compact-valued.*

2. The correspondence $\mathcal{C} : (m,n) \rightrightarrows \{\mu \,|\, \mu \in \mathcal{M}(m,n)\}$ is both lower and upper hemicontinuous. Therefore it is continuous.

Proof. The conclusion about being compact valued in the lemma is easy to see as $\mathcal{M}(m,n)$ is closed and bounded for all $(m,n) \in L$. Hence $\mathcal{C}$ is compact-valued.

The rest of the proof focuses on showing that the correspondence $\mathcal{C}$ is continuous. First, to show that $\mathcal{C}$ is upper hemicontinuous, take two sequences $\{(m^j,n^j)\} \in L$ and $\{\mu^j\} \in \mathcal{M}(m^j,n^j)$ that converge to (m,n) and μ, respectively. We have to show that $\mu \in \mathcal{M}(m,n)$. This is straightforward since

$$m_x^j = \sum_{y \in \mathcal{Y}_0} \mu_{xy}^j \to \sum_{y \in \mathcal{Y}_0} \mu_{xy} = m_x \text{ and } n_y^j = \sum_{x \in \mathcal{X}_0} \mu_{xy}^j \to \sum_{x \in \mathcal{X}_0} \mu_{xy} = n_y.$$

The definition of $\mathcal{M}(m,n)$ requires all workers to be matched or single and all firms to be matched or single and finite sums of elements in the sequence converge, so $\mu \in \mathcal{M}(m,n)$, as desired.

Showing that $\mathcal{C}$ is lower hemicontinuous is lengthier. Fix $(m,n) \in L$ and $\mu \in \mathcal{M}(m,n)$. Let $\{(m^j,n^j)\}$ be a sequence that converges to (m,n). We will find a sequence $\{\mu^j\}$ that converges to μ and such that $\mu^j \in \mathcal{M}(m^j,n^j)$ for all j.

First, note that $\mathcal{M}(m,n)$ is a set defined by a finite number of linear inequalities and linear equalities. As such, it is a convex polyhedron and by Carathéodory's theorem every $\mu \in \mathcal{M}(m,n)$ can be written as

$$\mu = \sum_k \alpha_k \mu^k,$$

where the μ^k are the extreme points of $\mathcal{M}(m,n)$ and the coefficients α_k are all non-negative and sum to 1. Each extremal point μ^k is the unique (by being an extremal point) solution to

$$\min_{\mu \geq 0} \mu^\top \Psi^k \text{ s.t. } \sum_{y \in \mathcal{Y}_0} \mu_{xy} = m_x, \sum_{x \in \mathcal{X}_0} \mu_{xy} = n_y \tag{S.1}$$

for some appropriate choice of a direction vector Ψ^k. Note that the coefficients α_k are specific to the limit point (m,n) for the social planner's state variable and the desired limit point μ for the social planner's choice variable fixed above.

Define our candidate sequence $\{\mu^j\}$ as follows:

$$\mu^j = \sum_k \alpha_k \mu^{j,k} \text{ where each } \mu^{j,k} \text{ solves } \min_{\mu \geq 0} \mu^\top \Psi^k \text{ s.t. } \sum_{y \in \mathcal{Y}_0} \mu_{xy} = m_x^j, \sum_{x \in \mathcal{X}_0} \mu_{xy} = n_y^j.$$

The coefficients α_k are specific to the points (m,n) and the μ fixed above and so this is not another application of Carathéodory's theorem. Also note that the direction vectors Ψ^k are from the direction vectors corresponding to the extreme points of the limit set $\mathcal{M}(m,n)$, as implicit in (S.1). Because the coefficients α_k are nonnegative and sum to 1, μ^j as defined is a convex combination of points in $\mathcal{M}(m^j, n^j)$ and is in that set.

If we can show that for each k, $\mu^{j,k}$ converges to the solution of (S.1) as $j \to \infty$, then we will have shown that $\mu^j \to \mu$. Note that if we could apply the theorem of the maximum to (S.1), we would be done, as the unique solution to (S.1) would by that theorem be continuous in (m,n). However, applying the theorem of the maximum requires the output of Lemma 6, which we are trying to prove, so this is not a profitable direction.

Because no element of the tuple μ can be more than the number of workers or more than the number of firms, there exists a compact superset that contains all the sequence sets $\mathcal{M}(m^j, n^j)$ as well as the sequence limit $\mathcal{M}(m,n)$. Because μ^j lives in this compact superset, it converges to some $\tilde{\mu}$, up to extraction, within this superset.[12] This compactness argument does not show that $\tilde{\mu} \in \mathcal{M}(m,n)$.

We show that $\tilde{\mu}$ is a solution to problem (S.1) by using the compacity of (S.1)'s

[12] The phrase "by extraction" means that a convergent subsequence can be found. The phrase avoids needing to introduce separate notation for this convergent subsequence.

feasible set and its dual's feasible set's compacity. Problem (S.1)'s dual is

$$\max_{u,v} \sum_{x \in \mathcal{X}} m_x u_x + \sum_{y \in \mathcal{Y}} n_y v_y \text{ s.t } u_x + v_y \leq \Psi^k_{xy},\ u_x \leq \Psi^k_{x0}, v_y \leq \Psi^k_{0y}\ \forall x, y,$$

whose feasible set can without loss be made closed and bounded and hence compact.[13]

The Lagrange multipliers u^j and v^j for the primal problem for index j are the solutions to the dual for index j. The dual for the jth index is where each m_x is updated to m^j_x and each n_y is updated to n^j_y.

As by a previous argument the feasible set for the dual is compact, we can construct a product space of compact supersets where the tuples (u^j, v^j, μ^j) converge by extraction within this product space of compact supersets to the tuple $(\tilde{u}, \tilde{v}, \tilde{\mu})$. This is reminiscent of econometric theory, where parameter spaces are often compact in order to ensure convergence of an optimization-based estimator.

We wish to show that this tuple satisfies the optimality conditions for the primal and dual problems

$$\sum_{y \in \mathcal{Y}_0} \mu^j_{xy} = m^j_x \text{ and } \sum_{x \in \mathcal{X}_0} \mu^j_{xy} = n^j_y\ \forall x \in \mathcal{X}, y \in \mathcal{Y}$$
$$\mu^j_{xy} \geq 0\ \forall x \in \mathcal{X}_0, y \in \mathcal{Y}_0$$
$$u^j_x + v^j_y \leq \Psi^k_{xy},\ u^j_x \leq \Psi^k_{x0} \text{ and } v^j_y \leq \Psi^k_{0y}\ \forall x \in \mathcal{X}, y \in \mathcal{Y}$$
$$\sum_{x,y} \mu^j_{xy} \left(u^j_x + v^j_y - \Psi^k_{xy}\right) + \sum_x \mu^j_{x0} \left(u^j_x - \Psi^k_{x0}\right) + \sum_y \mu^j_{0y} \left(v^j_y - \Psi^k_{0y}\right) = 0\ \forall x \in \mathcal{X}, y \in \mathcal{Y}\,.$$

By inspection, these optimality conditions are continuous in $(m^j, n^j, u^j, v^j, \mu^j)$ and so converge to the optimality conditions for problem (S.1). Because the solution to problem (S.1) is unique, $\mu = \tilde{\mu}$.

$\square$

S.1.2 Proposition 1

Proof. The properties of being continuous and bounded arise from Theorem 4.6 of Stokey et al. One conditions of Theorem 4.6 is that the per-period objective function is

[13] There are many references on how to construct the dual of a finite-dimensional linear program. These primal and duals look deceptively like static, two-sided matching problems but are not matching problems.

bounded and continuous; see the argument just given. Another condition of Theorem 4.6 is that the social planner's action space $\mathcal{C} : (m,n) \rightrightarrows \{\mu \,|\, \mu \in \mathcal{M}(m,n)\}$, as seen as a correspondence with argument (m,n), is non-empty, compact valued, and continuous in (m,n). Non-emptiness is easy to verify by inspection. Our Lemma 6 shows that $\mathcal{C}$ is compact valued and continuous.

Let $B(L)$ be the space of bounded continuous functions $V : L \to \mathbb{R}$ with the sup norm, denoted $\|.\|$. To show that $W(m,n)$ satisfying (6) exists, we follow the same proof technique as in Stokey et al. (1989), Theorem 4.6 and show that operator $T : B(L) \to B(L)$ defined by:

$$(TV)(m,n) = \max_{\mu \in \mathcal{M}(m,n)} \left\{ \sum_{xy \in \mathcal{X}_0 \mathcal{Y}_0} \Phi_{xy}\mu_{xy} + \beta V(P\mu, Q\mu) \right\},$$

is a contraction. Because $V \in B(L)$, TV is also in $B(L)$. Next, we argue that T is a contraction in $B(L)$ for the sup norm using Blackwell's theorem, which, as we recall, states that if T is order-preserving and satisfies $T(V + c) = TV + \beta c$, then T is a contraction of modulus β for the sup norm. These two conditions are easily satisfied, and thus T is a contraction for the sup norm, which shows that equation (6) has a unique solution in $B(L)$.

Next, we need to show that the solution W to equation (6) is concave. To do this, we follow again the argument in Stokey et al. (1989), Theorem 4.8 and we introduce the space $CVB(L)$ of functions that are concave and bounded on L. This space is a subset of $B(L)$, and we show that it is stable by T. Indeed, the Lagrangian formulation for $TV(m,n)$ yields the following expression for $TV(m,n)$:

$$\begin{aligned}
&= \max_{\mu \in \mathcal{M}(m,n)} \left\{ \sum_{xy \in \mathcal{X}_0\mathcal{Y}_0} \mu_{xy}\Phi_{xy} + \beta V(P\mu, Q\mu) \right\} \\
&= \max_{\mu_{xy} \geq 0} \min_{u_x, v_y} \sum_{x \in \mathcal{X}} m_x u_x + \sum_{y \in \mathcal{Y}} n_y v_y + \sum_{xy \in \mathcal{X}_0\mathcal{Y}_0} \mu_{xy}(\Phi_{xy} - u_x - v_y) + \beta V(P\mu, Q\mu) \\
&= \min_{u,v} \sum_{x \in \mathcal{X}} m_x u_x + \sum_{y \in \mathcal{Y}} n_y v_y + \max_{\mu_{xy} \geq 0} \left\{ \sum_{xy \in \mathcal{X}_0\mathcal{Y}_0} \mu_{xy}(\Phi_{xy} - u_x - v_y) + \beta V(P\mu, Q\mu) \right\},
\end{aligned}$$

where strong duality applies between lines two and three. This shows that $TV(n,m)$ is concave in (n,m). As a result, T has a unique fixed point in $CVB(L)$, which coincides with W. This shows that W is an element of $CVB(L)$, and therefore, that it is a concave function.

$\square$

S.2 Proof of Results in Section 3

S.2.1 Lemma 3

Proof. From equation (3.5) in Galichon and Salanié (2022), we have that $G_x^*(\frac{\mu_x}{\sum_{y \in \mathcal{Y}_0} \mu_{xy}}) = -\mathbb{E}[\epsilon_Y]$, where Y is the alternative $y \in \mathcal{Y}$ chosen by an agent of type x with unobservable heterogeneity vector $(\epsilon_{iy})_{y \in \mathcal{Y}_0}$. Assumption 3 ensures the ratio $\frac{\mu_x}{\sum_{y \in \mathcal{Y}_0} \mu_{xy}}$ is well defined as $\sum_{y \in \mathcal{Y}_0} \mu_{xy} > 0$. Hence G^* and H^* are well defined by Assumption 2. By taking absolute values and applying the triangle inequality, one has $|G_x^*(\frac{\mu}{\sum_{y \in \mathcal{Y}_0} \mu_{xy}})| \leq \sum_{y \in \mathcal{Y}} \mathbb{E}[|\epsilon_y|]$, where this upper bound depends only on the distribution $\mathcal{L}_x$. It is finite by Assumption 1. A similar bound can be computed for H_y^*. This proves the result, as is a weighted sum of the functions G_x^* and H_y^*. G^* and H^* are continuously differentiable by Assumption 2, hence $\mu \to \mathcal{E}(\mu)$ is continuous $\square$

S.2.2 Proposition 3

Proof. We follow the same steps as in the proof of Proposition 1. Showing the conditions of monotonicity and discountability are straightforward. Lemma 3 ensures the objective function is continuous and bounded. A key difference with Proposition 1 is that the entropy function $\mu \to \mathcal{E}(\mu)$ is a strictly convex function, and since the difference between a concave and a strictly convex function is strictly concave, we obtain the strict concavity of W by directly citing Theorem 4.8 of Stokey et al. (1989). $\square$

S.2.3 Proposition 4

Proof. The main challenge here is that there is a nonlinear program with a countably infinite number of controls and a countably infinite number of constraints. The main paper we refer to is Luc and Volle (2021). In Luc and Volle, the primal is

$$\inf_{(\lambda^t)_{t \in K}} \sum_{t=0}^{\infty} f_t(\lambda^t) \qquad \text{(Plv)}$$

where K is a closed convex set.

Let the per-period objective function contribution be

$$f_t(\lambda^t) = \left(\mathcal{E}(\lambda^t) - \lambda^{t\top}\Phi\right),$$

where the entropy term is

$$\mathcal{E}(\lambda^t) = \sum_{x\in\mathcal{X}} \left(\sum_{y\in\mathcal{Y}_0} \lambda^t_{xy}\right) G^*_x \left(\frac{\lambda^t_{x.}}{\sum_{y\in\mathcal{Y}_0} \lambda^t_{xy}}\right) + \sum_{y\in\mathcal{Y}} \left(\sum_{x\in\mathcal{X}_0} \lambda^t_{xy}\right) H^*_y \left(\frac{\lambda^t_{.y}}{\sum_{x\in\mathcal{X}_0} \lambda^t_{xy}}\right).$$

Note the change of variable from the main text: $\lambda^t = \beta^t \mu^t$. Let also

$$K = \left\{\lambda \geq 0 : \begin{array}{l} \sum_{y\in\mathcal{Y}_0} \lambda^0_{xy} = m^0_x \text{ and } \sum_{x\in\mathcal{X}_0} \lambda^0_{xy} = n^0_y \\ \beta\left(P\lambda^{t-1}\right)_x = \sum_{y\in\mathcal{Y}_0} \lambda^t_{xy} \text{ and } \beta\left(Q\lambda^{t-1}\right)_y = \sum_{x\in\mathcal{X}_0} \lambda^t_{xy}, t \geq 1 \end{array}\right\}.$$

We omit a formal argument that K is closed and convex, but the latter property arises from the linearity of the terms in the definition of K. Our social planner problem (10) under (11) and (12) is the inverse of Luc and Volle (2021)'s primal:

$$-\sup_{(\lambda^t)_t\in K} \sum_{t=0}^{\infty} -f_t(\lambda^t).$$

Let f^*_t and δ^*_K be the convex conjugates of f_t and δ_K, where $\delta_K(\lambda) = \mathbb{1}\{\lambda \in K\}$. Luc and Volle (2021)'s dual is, for an arbitrary control variable d^t,

$$\sup_{(d^t)_t} - \left(\sum_{t=0}^{\infty} \left\{f^*_t(d^t)\right\} + \delta^*_K(-(d^t)_t)\right), \tag{Dlv}$$

which with our change of sign becomes

$$-\inf_{(d^t)_t} \left(\sum_{t=0}^{\infty} \left\{f^*_t(-d^t)\right\} + \delta^*_K((d^t)_t)\right),$$

with this equivalence shown using standard properties found in convex analysis textbooks. Theorem 4 in Luc and Volle (2021) states that if

1. The functions f_t are proper, convex, lower semicontinuous and have compact domains $\forall t$.

2. $(f^t)_t$ satisfies $\sum_t \|f_t\| < \infty$ for the sup norm.

3. f_t is non negative $\forall t$.

4. The primal has a feasible solution.

then $\min$ (Plv) $=$ $\sup$ (Dlv), meaning that the duality gap is zero and the optimum is attained for the primal.

Note that a proper function in convex analysis is a function that is never $-\infty$ and is not $+\infty$ at at least one argument.

We now show the four conditions above hold. Starting with the first condition, for all t, f^t is defined on the following compact domain:

$$\mathrm{dom}(f_t) = \left\{ \lambda \geq 0 \,\middle|\, \sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}_0} \lambda_{xy} = \beta^t M, \sum_{y\in\mathcal{Y}} \sum_{x\in\mathcal{X}_0} \lambda_{xy} = \beta^t N \right\}.$$

The only nonlinear part of f_t is the the entropy term $\mathcal{E}(\lambda^t)$. This entropy term $\mathcal{E}(\lambda^t)$ is comprised of a weighted sum, with nonnegative weights, of the convex conjugates of social surplus functions. If $G(u)$ is a social surplus function, then its convex conjugate is

$$G^*(p) = \sup_{u\in J_{\mathcal{Y}_0}} \{p \cdot u - G(u)\},$$

where $J_{\mathcal{Y}_0}$ is the number of y-side types in $\mathcal{Y}_0$. A convex conjugate of a social surplus function has a domain equal to the probability simplex of dimension $J_{\mathcal{Y}_0}$.

Note that the convex conjugates of particular social surplus functions are themselves sometimes called negative generalized entropy terms. The full-support Assumption 2 needs to be invoked to ensure that the social surplus functions $G(u)$ have good properties and that, as we will show, the convex conjugates $G^*(p)$ of social surplus functions. To reduce the length of the proof, we assume knowledge of the properties of social surplus functions $G(u)$ under Assumption 2 and focus on deriving properties of the convex conjugates of social surplus functions. These properties are then inherited by $\mathcal{E}(\lambda^t)$, the weighted sum, with nonnegative weights, of the convex conjugates of social surplus functions. If $\mathcal{E}(\lambda^t)$ has a certain property, then likely f_t will inherit that property as $\mathcal{E}(\lambda^t)$ enters f_t linearly with a positive multiplicative coefficient and the other term in $\mathcal{E}(\lambda^t)$ is itself linear in the argument λ^t and λ^t is restricted to the set K.

We now show that $G^*(p)$ is proper, as defined above, taking as known from prior literature that $G(u)$ is proper under Assumption 2. By the latter property, there exists a value u such that $G(u) < \infty$. By inspecting the definition of $G^*(p)$, we can see that $G^*(p)$ cannot be $-\infty$ for any p as the goal is to take a supremum over u of a term where $G(u)$ enters with a minus sign and we know that there is at least one u where $G(u) < \infty$. Also, choosing $p = 0$ clearly shows that $G^*(0) < \infty$. So $G^*(p)$ is proper and hence $\mathcal{E}(\lambda^t)$ is proper. Inspecting the definitions of f_t and the constraint set K, we can see that f_t cannot be made to be $-\infty$ everywhere on a value in K, so f_t is proper.

Since the function $G^*(p)$ is convex, it is continuous in the interior of its domain. Using this result and then inspecting the relevant formulas, the weighted sum $\mathcal{E}(\lambda^t)$ is continuous and the linear function f_t of $\mathcal{E}(\lambda^t)$ is continuous. A continuous function is automatically lower semicontinuous.

The convex conjugate $G^*(p)$ is almost by definition convex, as it is a convex conjugate. Each $G^*(p)$ enters positively into the otherwise linear $\mathcal{E}(\lambda^t)$ and f_t, so both $\mathcal{E}(\lambda^t)$ and f_t are convex.

For the second condition in Theorem 4 of Luc and Volle (2021), we need to show that $\sum_t \|f_t\| < \infty$ for the supremum norm. The argument in the proof of Lemma 3 implies that $G^*(p) < \infty$ for all p. From the definition of the domain of f_t, we know that any $\lambda_t \in \mathrm{dom}(f_t)$ is such that $\sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}_0} \lambda^t_{xy}$ and $\sum_{y\in\mathcal{Y}} \sum_{x\in\mathcal{X}_0} \lambda^t_{xy}$ tend to 0 as t tends to infinity. Since $\lambda^t_{xy} \geq 0$ for all t, this implies $\lim_{t\to\infty} \lambda^t_{xy} = 0$ for all x and y. As G^* is bounded, this implies that $\lim_{t\to\infty} f_t(\lambda_t) = 0$ for all $\lambda_t \in \mathrm{dom}(f_t)$, and therefore $\sum_t \|f_t\| < \infty$.

For the third condition in Theorem 4 of Luc and Volle (2021), f^t as such is not nonnegative for all t: it could be that for some λ in $\mathrm{dom}(f)$, $\mathcal{E}(\lambda) < \lambda^\top \Phi$, in which case $f^t(\lambda) < 0$. However, we can modify f^t such that the solutions to the primal and dual and the properties of f^t shown above remain unchanged. Add the positive constant $\beta^t \max\{M, N\} \max\{\Phi\}$ to every f^t. The newly redefined function f^t is non-negative for every $\lambda \in \mathrm{dom}(f_t)$ and adding a positive, finite constant does not alter the previously established properties.

For the fourth condition in Theorem 4 of Luc and Volle (2021), we need to show that the primal has a feasible solution. A simple feasible solution for (Plv) is λ such that:

$$\lambda_{xy}^0 = \delta\,, \quad \lambda_{x0}^0 = m_x^0 - \sum_{y\in\mathcal{Y}} \lambda_{xy}^0\,, \quad \lambda_{0y}^0 = n_y^0 - \sum_{x\in\mathcal{X}} \lambda_{xy}^0 \quad \forall x \in \mathcal{X}, y \in \mathcal{Y}$$

$$\lambda_{xy}^t = \delta\,, \quad \lambda_{x0}^t = \beta(P\lambda^{t-1})_x - \sum_{y\in\mathcal{Y}} \lambda_{xy}^t\,, \quad \lambda_{0y}^t = \beta(Q\lambda^{t-1})_y - \sum_{x\in\mathcal{X}} \lambda_{xy}^t \quad \forall x \in \mathcal{X}, y \in \mathcal{Y}\,, \forall t > 0,$$

where δ is an arbitrarily small positive number.

The conditions for Luc and Volle (2021)'s Theorem 4 are all satisfied, so the objective functions of the primal and dual are equal at solutions:

$$\max_{(\mu^t)_t\in K} \sum_{t=0}^{\infty} -f_t(\mu^t) = \inf_{(d^t)_t} \left(\sum_{t=0}^{\infty} f_t^*(-d^t) + \delta_K^*((d^t)_t) \right)$$

at the optima. This result is the main notion of strong duality shown in Luc and Volle (2021).

However, there are limitations of this strong duality result compared to the strong duality result in the statement of the proposition being proved. Therefore, we need two more parts to this proof.

(1) We need to transform the dual(Dlv) into a more intuitive, computationally attractive form, as shown in the statement of the proposition.

(2) Luc and Volle (2021)'s theorem does not specify if the infinum is attained in the dual, which we need to later show the equivalence of a dynamic competitive equilibrium to the dual solution and to show that computational methods have a well-defined target to compute.

We start with goal (1), deriving a more intuitive dual. Recall that the existing dual (Dlv) contains two terms for each t. Let us tackle them one at a time.

The per-period primal objective contribution f_t's convex conjugate is

$$\begin{aligned} f_t^*(-d) &= \max_{\lambda}\{-\mathcal{E}(\lambda) + \lambda^{t\top}(\Phi - d)\} \\ &= \mathcal{E}^*(\Phi - d). \end{aligned}$$

While $\mathcal{E}(\lambda)$ itself is a comprised of a weighted sum of convex conjugates of social surplus functions, here the new $\mathcal{E}^*(\lambda)$ is the convex conjugate of $\mathcal{E}(\lambda)$.

An important result is that $\mathcal{E}^*$ is either 0 or ∞, as shown in the following lemma.

Lemma 7. *$\mathcal{E}^*(\zeta) = 0$ if there exists $\tilde{u} = (\tilde{u}_{xy})_{x\in\mathcal{X},y\in\mathcal{Y}}$ such that $G_x(\tilde{u}) \leq 0$ and $H_y(\zeta - \tilde{u}) \leq 0$ for all $x \in \mathcal{X}, y \in \mathcal{Y}$, and $\mathcal{E}^*(\zeta) = +\infty$ otherwise.*

Proof. The definition of a convex conjugate $\mathcal{E}^*(\zeta)$ involves maximizing over some value μ. Based on the definition of $\mathcal{E}(\mu)$, it is instructive to thinking of this as an inner problem of maximizing over the set of feasible matching probabilities in some matching problem, $M(m, n) = \left\{\mu \geq 0 | \sum_{y\in\mathcal{Y}} \mu_{xy} = m_x \text{ and } \sum_{x\in\mathcal{X}} \mu_{xy} = n_y\right\}$ and an outer problem where the masses of types (m, n) are maximized over. The pair of the outer and inner maximization problems searchers over all μ, as in the original definition of $\mathcal{E}^*(\zeta)$. Exploiting the fact that μ enters linearly into the definition of $\mathcal{E}(\mu)$, the inner and outer maximization problems let us rewrite $\mathcal{E}^*(\zeta)$ as

$$\mathcal{E}^*(\zeta) = \max_{(m,n)} \max_{\mu\in M(m,n)} \mu^\top \left\{\zeta - \sum_{x\in\mathcal{X}} m_x G_x^* \left(\frac{\mu_{x.}}{m_x}\right) - \sum_{y\in\mathcal{Y}} n_y H_y^* \left(\frac{\mu_{.y}}{n_y}\right)\right\}.$$

Consider the inner optimization problem. By Theorem 3 in Galichon and Salanié (2022):

$$\begin{aligned}\max_{\mu\in M(m,n)} \{\mu^\top\zeta &- \sum_{x\in\mathcal{X}} m_x G_x^* \left(\frac{\mu_{x.}}{m_x}\right) - \sum_{y\in\mathcal{Y}} n_y H_y^* \left(\frac{\mu_{.y}}{n_y}\right)\Bigg\} \\ &= \min_{\tilde{u}} \left\{\sum_{x\in\mathcal{X}} m_x G_x(\tilde{u}) + \sum_{y\in\mathcal{Y}} n_y H_y(\zeta - \tilde{u})\right\},\end{aligned}$$

giving a reformulated inner optimization problem. Return to the outer optimization problem over (m, n). If the minimum in $\tilde{u}$ for the inner optimization problem is such that $G_x(\tilde{u}) > 0$, or $H_y(\zeta - \tilde{u}) > 0$ for all x, y, then in the outer problem one can choose m_x or n_y arbitrarily large so that $\mathcal{E}^*(\zeta) = \infty$. Hence $\mathcal{E}^*(\zeta) = 0$ if there exists a $\tilde{u}$ such that both $G_x(\tilde{u})$ and $H_y(\zeta - \tilde{u})$ are below 0.

□

Now we turn to the other term in the existing dual (Dlv), δ_K^*. This is the convex conjugate of an indicator function. The convex conjugate of an indicator function is sometimes called a support function for K. The first equality in the below algebra introduces the support function. Note that as $\lambda = (\lambda^t)_t \in l^1$, its dual variable d is in l^∞ (its dual space) so the support function is well-defined.

The second equality below introduces the Lagrangian with respect to the constraints K. The requirement from K that $\lambda \geq 0$ is kept in the supremum and not incorporated as a separate Lagrangian term. This means that the remaining terms in K are all equality constraints. With only equality constraints in the Lagrangian, the Lagrange multipliers U, V can be both negative and positive. As usual, the Lagrange multipliers do not enter the original problem's objective, here the support function. One can see that an infinum over the Lagrange multipliers has been added. Using generic notation, the problem of $\max_{z,\nu} g(x) + \nu h(x)$ is equivalent to $\max_z g(z) + \min_\nu \nu h(z)$, as either way the solution involves setting the Lagrange multiplier λ to 0 when the equality constraints are satisfied and picking λ to drive the entire objective to $-\infty$ when the constraints are violated.

The third inequality involves swapping the inner and outer problems, with the infinum now being the outer problem and the supremum being the inner problem. This is arising from the strong duality proved in Romeijn et al. (1992), as the objective function is linear and the constraints are affine. The nonnegativity constraints are satsified, and because we have already shown that $\lambda^t \to 0$ entrywise as t tends to infinity, so is the stationarity constraint.

The fourth equality holds because the outer optimization problem will make the parenthetical terms inside the supremum negative, so the inner problem is solved by setting $\mu = 0$ and the problem can be rewritten as a single-layer constrained optimization problem.

$$
\begin{aligned}
\delta_K^*(d) &= \sup_{\lambda \in K} \left\{ \sum_{t \geq 0} \lambda^{t\top} d^t \right\} \\
&= \sup_{\lambda \geq 0} \Bigg\{ \sum_{t \geq 0} \lambda^{t\top} d^t \\
&\qquad + \inf_{U,V} \Bigg\{ U^{0\top} m^0 + V^{0\top} n^0 - \sum_{x \in \mathcal{X}} U_x^0 \left(\sum_{y \in \mathcal{Y}_0} \lambda_{xy} \right) - \sum_{x \in \mathcal{X}} V_y^0 \left(\sum_{x \in \mathcal{X}_0} \lambda_{xy} \right) \\
&\qquad + \sum_{t \geq 1} \sum_{x \in \mathcal{X}} U_x^t \left(\beta (P\lambda^{t-1})_x - \sum_{y \in \mathcal{Y}_0} \lambda_{xy}^t \right) \\
&\qquad + \sum_{t \geq 1} \sum_{y \in \mathcal{Y}} V_y^t \left(\beta (Q\lambda^{t-1})_y - \sum_{x \in \mathcal{X}_0} \lambda_{xy}^t \right) \Bigg\} \Bigg\}
\end{aligned}
$$

$$
\begin{aligned}
&= \inf_{U,V} \left\{ U^{0\top} m^0 + V^{0\top} n^0 \right. \\
&\quad + \sup_{\lambda \geq 0} \left\{ \sum_{xy \in \mathcal{X}\mathcal{Y}} \lambda_{xy}^t \left(\beta (P^\top U^{t+1} + Q^\top V^{t+1})_{xy} - U_x^t - V_y^t + d_{xy}^t \right) \right. \\
&\qquad + \sum_{x \in \mathcal{X}} \lambda_{x0}^t \left(\beta (P^\top U^{t+1})_{x0} - U_x^t + d_{x0}^t \right) \\
&\qquad \left. \left. + \sum_{y \in \mathcal{Y}} \lambda_{0y}^t \left(\beta (Q^\top V^{t+1})_{0y} - V_y^t + d_{0y}^t \right) \right\} \right\} \\
&= \inf_{U,V} \left\{ U^{0\top} m^0 + V^{0\top} n^0 \right\} : U_x^t + V_y^t \geq d_{xy}^t + \beta \left(P^\top U^{t+1} + Q^\top V^{t+1} \right)_{xy}, \; t \geq 0 \\
&\qquad\qquad\qquad\qquad\qquad\qquad U_x^t \geq d_{x0}^t + \beta \left(P^\top U^{t+1} \right)_{x0}, \; t \geq 0 \\
&\qquad\qquad\qquad\qquad\qquad\qquad V_y^t \geq d_{0y}^t + \beta \left(Q^\top V^{t+1} \right)_{0y}, \; t \geq 0.
\end{aligned}
$$

Next, we put f_t^* and δ_K^* together by ensuring the constraints in Lemma 7 are met, so that f_t^* is 0. Then the dual is

$$
\begin{aligned}
&\inf_{d^t, U^t, V^t} \inf_{\tilde{u}^t} \sum_{x \in \mathcal{X}} m_x^0 U_x^0 + \sum_{y \in \mathcal{Y}} n_y^0 V_y^0 \\
&\quad \text{s.t } U_x^t + V_y^t \geq d_{xy}^t + \beta \left(P^\top U^{t+1} + Q^\top V^{t+1} \right)_{xy}, \; t \geq 0 \\
&\quad U_x^t \geq d_{x0}^t + \beta \left(P^\top U^{t+1} \right)_{x0}, \; t \geq 0 \\
&\quad V_y^t \geq d_{0y}^t + \beta \left(Q^\top V^{t+1} \right)_{0y}, \; t \geq 0 \\
&\quad G_x \left(\tilde{u}^t \right) \leq 0, \; t \geq 0 \\
&\quad H_y \left(\Phi - d^t - \tilde{u}^t \right) \leq 0, \; t \geq 0.
\end{aligned}
$$

The first three constraints in the problem above are saturated. We will use a proof by induction. The objective is minimized if U_0 and V_0 are as small as possible. Hence the constraints are saturated at $t = 0$. As part of the proof by induction, assume the constraints are saturated at $t - 1$ and show the constraints must be also saturated at t. This is because if the constraints are not saturated, then U^t and/or V^t are larger than they could be, which translates to U^{t-1} and/or V^{t-1} also being larger than they could be through the constraint, and down the line until U^0 and V^0 are larger than they code be, contradicting the previously established property for U^0 and V^0.

Together all constraints imply that

$$
G_x \left(\tilde{u}^t \right) \leq 0 \text{ and } H_y \left(\Phi - \left(U_x^t + V_y^t - \beta P^\top U^{t+1} - \beta Q^\top V^{t+1} \right) - \tilde{u}^t \right) \leq 0.
$$

where we have solved for d^t using the first three saturated constraints above and plugged it into the last constraints. Let $u^t_{xy} = U^t_x + \tilde{u}^t_{xy}$ and $v^t = -u^t_{xy} + \Phi + \beta\left(P^\top U^{t+1} + Q^\top V^{t+1}\right)$ then

$$G_x\left(u^t - U^t_x\right) \leq 0 \text{ and } H_y\left(v^t - V^t_y\right) \leq 0,$$

which implies

$$G_x(u^t) \leq U^t_x \text{ and } H_y\left(v^t\right) \leq V^t_y.$$

We will change the control variables to be something more interpretable. The dual becomes

$$\begin{aligned}
\inf_{U^t,V^t} \inf_{u^t,v^t} & \sum_{x\in\mathcal{X}} m^0_x U^0_x + \sum_{y\in\mathcal{Y}} n^0_y V^0_y \\
\text{s.t } & u^t_{xy} + v^t_{xy} = \Phi_{xy} + \beta\left(P^\top U^{t+1} + Q^\top V^{t+1}\right)_{xy} \\
& u^t_{x0} = \Phi_{x0} + \beta\left(P^\top U^{t+1}\right)_{x0} \\
& v^t_{0y} = \Phi_{0y} + \beta\left(Q^\top V^{t+1}\right)_{0y} \\
& U^t_x \geq G_x(u^t) \\
& V^t_y \geq H_y(v^t),\ t \geq 0.
\end{aligned}$$

The last two inequality constraints are also saturated, for the same reason as above. So the dual rewrites as

$$\begin{aligned}
\inf_{u^t,v^t} & \sum_{x\in\mathcal{X}} m^0_x G_x(u^0) + \sum_{y\in\mathcal{Y}} n^0_y H_y(v^0) \\
\text{s.t } & u^t_{xy} + v^t_{xy} = \Phi_{xy} + \beta\left(P^\top G(u^{t+1}) + Q^\top H(v^{t+1})\right)_{xy} \\
& u^t_{x0} = \Phi_{x0} + \beta\left(P^\top G(u^{t+1}\right)_{x0} \\
& v^t_{0y} = \Phi_{0y} + \beta\left(Q^\top H(v^{t+1})\right)_{0y},
\end{aligned} \tag{S.2}$$

where $G(u^{t+1})$ and $H(v^{t+1})$ are the stacked vectors of $(G_x(u^{t+1}))_x$ and $(H_y(v^{t+1}))_y$. That concludes point (1), to derive the more computational attractive form of the dual stated in the main text.

Now on point (2): write dual (S.2)'s Lagrangian

$$
\begin{aligned}
\inf_{u^t,v^t} & \sum_{x\in\mathcal{X}} m_x^0 G_x(u^0) + \sum_{y\in\mathcal{Y}} n_y^0 H_y(v^0) \\
& + \sum_t \sum_{xy} \lambda_{xy}^t \left(\Phi_{xy} + \beta\left(P^\top G(u^{t+1}) + Q^\top H(v^{t+1})\right)_{xy} - u_{xy}^t - v_{xy}^t\right) \\
& + \sum_t \lambda_{x0}^t \left(\Phi_{x0} + \beta\left(P^\top G(u^{t+1})\right)_{x0} - u_{x0}^t\right) \\
& + \sum_t \lambda_{0y}^t \left(\Phi_{0y} + \beta\left(Q^\top H(v^{t+1})\right)_{0y} - v_{0y}^t\right),
\end{aligned} \tag{S.3}
$$

By implication of strong duality, we know that the Lagrange multipliers are the solutions to the primal problem, so we can write that at the solution $\lambda_{xy}^t = \beta^t \mu_{xy}^t$, where μ are the optimal match masses at t. Next, We can see that the first order conditions (FOCs) to the dual's Lagrangian are, for the initial period 0,

$$
m_x^0 \frac{\partial G_x(u_{xy}^0)}{\partial u_{xy}^0} = \mu_{xy}^0 \quad \text{and} \quad n_y^0 \frac{\partial H_y(v_{xy}^0)}{\partial v_{xy}^0} = \mu_{xy}^0. \tag{S.4}
$$

For periods $t \geq 1$, the FOCs become

$$
\sum_{\bar{x}\bar{y}} \left(\beta^{t-1}\mu_{\bar{x}\bar{y}}^{t-1}\left(\beta P_{x|\bar{x},\bar{y}} \frac{\partial G_x(u_{xy}^t)}{\partial u_{xy}^t}\right)\right) = \beta^t \mu_{xy}^t \text{ and } \sum_{\bar{x}\bar{y}} \left(\beta^{t-1}\mu_{\bar{x}\bar{y}}^{t-1}\left(\beta Q_{y|\bar{x},\bar{y}} \frac{\partial H_y(v_{xy}^t)}{\partial v_{xy}^t}\right)\right) = \beta^t \mu_{xy}^t.
$$

Consider the x-side FOC, on the left. The partial derivative $\frac{\partial G_x(u_{xy}^t)}{\partial u_{xy}^t}$ is a multiplicative constant that factors out of the left side of the FOC. The discount factor β cancels on both sides, giving

$$
\frac{\partial G_x(u_{xy}^t)}{\partial u_{xy}^t} \sum_{\bar{x}\bar{y}} \left(\mu_{\bar{x}\bar{y}}^{t-1}\left(P_{x|\bar{x},\bar{y}}\right)\right) = \mu_{xy}^t \text{ and } \frac{\partial H_y(v_{xy}^t)}{\partial v_{xy}^t} \sum_{\bar{x}\bar{y}} \left(\mu_{\bar{x}\bar{y}}^{t-1}\left(Q_{y|\bar{x},\bar{y}}\right)\right) = \mu_{xy}^t.
$$

For the x-side, the term $\sum_{\bar{x}\bar{y}}\left(\mu_{\bar{x}\bar{y}}^{t-1}\left(P_{x|\bar{x},\bar{y}}\right)\right)$ can be seen to equal m_x^t, the mass of workers of type x in period t. So the FOCs become

$$
\frac{\partial G_x(u_{xy}^t)}{\partial u_{xy}^t} m_x^t = \mu_{xy}^t \quad \text{and} \quad \frac{\partial H_y(v_{xy}^t)}{\partial v_{xy}^t} n_y^t = \mu_{xy}^t. \tag{S.5}
$$

By standard arguments in convex optimization, we can define u^t and v^t using G's and H's Legendre transforms:

$$u^t_{xy} = \frac{\partial G^*_x\left(\frac{\mu}{m}\right)}{\partial \frac{\mu_{xy}}{m_x}} \quad \text{and} \quad v^t_{xy} = \frac{\partial H^*_y\left(\frac{\mu}{n}\right)}{\partial \frac{\mu_{xy}}{n_y}}$$

u and v are then solutions to the dual (S.2).

$\square$

S.2.4 Theorem 4

Proof. This proof rests on the use of duality: we show that the dynamic competitive equilibrium is equivalent to the social planner problem through its dual, in the same fashion as Dupuy and Galichon (2014) and Galichon and Salanié (2022). Unlike these papers however, our social planner problem has countably infinite controls and constraints. Proposition 4 shows strong duality for our social planner primal and dual problems. Also, we should point out again that our social planner problem is not a linear program as in the model without econometric errors.

First we show that the direction that says solutions to the primal and dual problems yield a DCE. Given a solution $(u^t, v^t)_t$ to the dual (S.2), let the wages be given as (15) in the statement of the theorem to be proved. Rearranging the wages in (15) to solve for u^t_{xy} and v^t_{xy} will give the same utilities as on the right sides of (9), up to the error terms.

Let us consider the dual problem's Lagrangian (S.3), as in the previous proof. The dual Lagrangian's FOCs (S.4) and (S.5) remain as in the previous proof. By the Williams-Daly-Zachary Theorem, the partial derivatives in the FOCs equal the choice probabilities

$$\Pr\left(\tilde{y} \in \arg\max_{y \in \mathcal{Y}_0} u^t_{xy} + \epsilon^t_y\right).$$

Using the proposal for equilibrium wages w^t_{xy} in the statement of the theorem, we can solve for u^t_{xy} and substitute in for that term in the choice probability, giving

$$\Pr\left(\tilde{y} \in \arg\max_{y \in \mathcal{Y}_0}\left(\alpha_{xy} + w^t_{xy} + \epsilon^t_y + \beta\left(P^\top G_x(u^{t+1})\right)_{xy}\right)\right).$$

The term $\left(P^\top G_x(u^{t+1})\right)_{xy}$ involves an integral over ϵ's and it can be expanded to match the notation in the formula in the definition of a DCE,

$$\Pr\left(\tilde{y} \in \arg\max_{y \in \mathcal{Y}_0} \alpha_{\tilde{x}y} + w^t_{\tilde{x}y} + \epsilon^t_y + \beta \sum_{x' \in \mathcal{X}} P_{x'|\tilde{x}y} \mathbb{E}\left[U^{t+1}_{x'}\left(w^{(t+1)}, \epsilon^{t+1}\right)\right]\right).$$

Returning to the x-side in equation (S.4), we divide μ^0_{xy} in the FOC by $m^0_x = \sum_{y \in \mathcal{Y}} \mu^0_{xy}$ and, after reversing the direction of the equation as written, we end up with

$$\frac{\mu^0_{\tilde{x}\tilde{y}}}{\sum_{y \in \mathcal{Y}_0} \mu^0_{\tilde{x}y}} = \Pr\left(\tilde{y} \in \arg\max_{y \in \mathcal{Y}_0} \alpha_{\tilde{x}y} + w^0_{\tilde{x}y} + \epsilon^0_y + \beta \sum_{x' \in \mathcal{X}} P_{x'|\tilde{x}y} \mathbb{E}\left[U^1_{x'}\left(w^{(1)}, \epsilon^1\right)\right]\right),$$

as defined in the definition of a DCE, Definition 3. Therefore, we satisfy the requirements for Definition 3 for matched agents in period 0. The argument for the unmatched x-side agents and the y-side agents are quite parallel.

For periods $t \geq 1$, applying the Williams-Daly-Zachary theorem indicates that the partial derivatives are choice probabilities as in the definition of a DCE. For the x-side in equation (S.5), diving by m^t_x gives the choice probability equality in the definition of a DCE. The arguments for unmatched x-side workers and all y-side firms are parallel. So we have verified Definition 3 and we have a DCE.

The other direction of the theorem starts with a DCE (μ, w) and ends with μ solving the primal problem and the u^t_{xy} and v^t_{xy} calculated using the expressions for the expected present discounted lifetime utilities solving the dual problem. All we need to show is that u^t_{xy} and v^t_{xy} solve the dual problem with the Lagrange multipliers μ^t_{xy}, as by the strong duality Proposition 4 the primal problem will also be solved.

The argument that the dual FOCs are solved is almost immediate from prior arguments, as we can simply multiply the probabilities in Definition 3 by the masses of each agent type, m^t_x or n^t_y, and get the dual FOCs derived above. We do not need to check second-order conditions by convexity. Therefore, the DCE yields a solution to the dual and hence the primal.

□

S.3 Methods for Equilibrium Computation

S.3.1 Constant Aggregate State Without Econometric Errors

Computing the constant aggregate state and associated stationary equilibrium for the model without econometric errors can be done by solving the following quadratic problem:

$$
\begin{aligned}
\min_{\mu\geq 0,U,V} & \sum_{xy\in\mathcal{X}_0\mathcal{Y}_0} \mu_{xy}\left(U_x+V_y-\Phi_{xy}-\beta\left(\sum_{x'\in\mathcal{X}} U_{x'}P_{x'|xy}+\sum_{y'\in\mathcal{Y}} V_{y'}Q_{y'|xy}\right)\right)\\
\text{s.t} & \sum_{y\in\mathcal{Y}_0}\mu_{xy}=\sum_{x',y'\in\mathcal{X}\mathcal{Y}_0} P_{x|x'y'}\mu_{x'y'}\ \forall x\in\mathcal{X}\\
& \sum_{x\in\mathcal{X}_0}\mu_{xy}=\sum_{x',y'\in\mathcal{X}_0\mathcal{Y}} Q_{y|x'y'}\mu_{x'y'}\ \forall y\in\mathcal{Y}\\
& U_x+V_y\geq\Phi_{xy}+\beta\left(\sum_{x'\in\mathcal{X}} U_{x'}P_{x'|xy}+\sum_{y'\in\mathcal{Y}} V_{y'}Q_{y'|xy}\right)\ \forall x\in\mathcal{X}, y\in\mathcal{Y}\\
& U_x\geq\beta\sum_{x'\in\mathcal{X}} U_{x'}P_{x'|x0}\ \forall x\in\mathcal{X}\\
& V_y\geq\beta\sum_{y'\in\mathcal{Y}} V_{y'}Q_{y'|0y}\ \forall y\in\mathcal{Y}\,.
\end{aligned}
\tag{S.6}
$$

The formulation as a quadratic problem ensures that the complementary slackness condition in Definition 1 is enforced. The constraints yield a constant aggregate state. Solving this quadratic problem is relatively straightforward, using a package such as Gurobi. We show that the optimal μ obtained in (S.6) is part of a stationary equilibrium in the following proposition.

Proposition S.6. The solution μ to problem (S.6) is part of a stationary equilibrium. The resulting aggregate state $(m,n)=\left(\sum_{y\in\mathcal{Y}_0}\mu_{xy},\,\sum_{x\in\mathcal{X}_0}\mu_{xy}\right)$ is a constant aggregate state.

Proof. Let (μ,U,V) solve problem (S.6). Then the tuple (μ,U,V) satisfies the primal feasibility, dual feasibility, and complementary slack conditions for the social planner problem at $(m,n) = \left(\sum_{y\in\mathcal{Y}_0} \mu_{xy}, \sum_{x\in\mathcal{X}_0} \mu_{xy}\right)$. It is straightforward that

$$(m,n) = \left(\sum_{x',y' \in \mathcal{X}\,\mathcal{Y}_0} P_{x|x'y'} \mu_{x'y'}, \sum_{x',y' \in \mathcal{X}_0\,\mathcal{Y}} Q_{y|x'y'} \mu_{x'y'} \right)$$

and therefore (m,n) is a constant aggregate state. □

S.4 Structural Estimation

It is easy to transition from computing a constant aggregate state and its associated stationary equilibrium to estimating the model parameters, if one assumes that the data come from a constant aggregate state. The minimal data set comes from one market in stationary equilibrium and has cross-sectional data on x, y, x', y' for a random sample of matches (x, y) from that market, where x', y' are the states of the two matched agents at the beginning of the next period. Data sets with longer panels can also be used, as we illustrate in our empirical application to the careers of Swedish engineers in Section 5.

We use the data on matches (x, y) to calculate the observed matching probabilities $\hat{\mu}_{xy}$, as well as $\hat{\mu}_{x0}$ and $\hat{\mu}_{0y}$ for unemployed workers and vacant jobs, respectively. Throughout this section, we assume that $\hat{\mu}$ is a stationary equilibrium.

The structural parameters are α, γ, P, Q and β. In estimation, we assume that the transition rules for the individual worker and firm states P and Q are estimated in a first stage, as is often done in dynamic discrete choice models (e.g., Rust, 1987).

We focus on a second stage in which structural parameters are estimated. Based on identification results for static matching games, starting with the theoretical characterizations of Becker (1973) and the logit model of Choo and Siow (2006), we estimate parameters in $\Phi = \alpha + \gamma$, the match production function, defined to be the sum of the worker amenities and the firm output for a given match. We parameterize match production Φ by

$$\Phi_{xy}(\lambda) = \sum_{l=1}^{L} \lambda^l \phi^l_{xy},$$

where each ϕ^l is a basis function of types x and y, and $\lambda = (\lambda^l)_{1\le l\le L}$ are the parameters we estimate. We fix the discount factor β, as is common in single-agent dynamic discrete choice models such as Rust (1987).

S.4.1 MPEC for Estimation

To estimate the model on cross-sectional data x, y, x', y', assuming that the observed equilibrium is stationary, one simply needs to augment the primitives (m, n, U, V) with parameters λ, and maximize a log-likelihood function on matching μ under the feasibility and stationarity constraints from (16). The log likelihood is

$$l(\lambda, U, V, m, n) = \sum_{xy \in \mathcal{X}\mathcal{Y}} \hat{\mu}_{xy} \log \mu_{xy}(\lambda, U, V, U, V, m, n) - \widehat{\mathcal{N}} \log \mathcal{N}(\lambda, U, V, U, V, m, n).$$

$\mathcal{N}(\lambda, U, V, m, n)$ counts the total mass of matches and unmatched agents on the market:

$$\begin{aligned}\mathcal{N}(\lambda, U, V, m, n) = & \sum_{xy \in \mathcal{X}\mathcal{Y}} \mu_{xy}(\lambda, U, V, U, V, m, n) \\ & + \sum_{x \in \mathcal{X}} \mu_{x0}(\lambda, U, V, U, V, m, n) + \sum_{y \in \mathcal{Y}} \mu_{0y}(\lambda, U, V, U, V, m, n)\end{aligned}$$

and $\widehat{\mathcal{N}}$ is its observed equivalent. The product $\widehat{\mathcal{N}} \log \mathcal{N}$ converts counts from $\hat{\mu}$ and μ to probabilities required in the likelihood function. Maximum likelihood is statistically efficient except for the first-stage estimation of the transition rule P and Q. Full efficiency can be gained by simultaneously estimating λ, P and Q using maximum likelihood.

S.4.2 Primal-Dual Algorithm for Estimation

To adapt the primal-dual algorithm for estimation, we augment the set of primitives with the structural parameters λ and replace function Z by Z_{est}:

$$\begin{aligned}Z_{\text{est}}(\lambda, U, V, U', V', m, n, \beta) = & Z(\lambda, U, V, U', V', m, n, \beta) \\ & - \sum_{xy \in \mathcal{X}\mathcal{Y}} \hat{\mu}_{xy} \Phi_{xy}(\lambda).\end{aligned} \tag{S.7}$$

When $\beta = 1$ we solve an augmented version of the min-max problem (17):

$$\min_{U,V,\lambda} \max_{m,n} Z_{\text{est}}(\lambda, U, V, U, V, m, n, 1).$$

The difference between Z and Z_{est} resides in the additional term $-\sum_{xy\in\mathcal{X}\mathcal{Y}}\hat{\mu}_{xy}\Phi_{xy}(\lambda)$. This term produces the following moment conditions when computing the first order conditions with respect to λ for problem (S.7):

$$\sum_{xy\in\mathcal{X}\mathcal{Y}}\hat{\mu}_{xy}\phi^l_{xy}(\lambda)=\sum_{xy\in\mathcal{X}\mathcal{Y}}\mu_{xy}(\lambda,U,V,U,V,m,n,\beta)\phi^l_{xy}(\lambda)\quad\forall l=1,\ldots,L.$$

When $\beta<1$, we use the Chambolle-Pock algorithm applied to Z_{est} to converge to the feasibility, stationarity and moment conditions. Choose an increment τ, a threshold δ, and initial values $(\lambda^0,U^0,V^0,m^0,n^0)$ and (m^1,n^1). Iterate on k according to:

$$\begin{aligned}
&\text{Intermediary }(\tilde{m},\tilde{n}) &&\begin{cases}\tilde{m}^k_x=2m^k_x-m^{k-1}_x\quad\forall x\in\mathcal{X}\\ \tilde{n}^k_y=2n^k_y-n^{k-1}_y\quad\forall y\in\mathcal{Y}\end{cases}\\
&(\lambda,U,V)\text{ update} &&\begin{cases}\lambda^{l,k+1}=\lambda^{l,k}-\tau\left(\partial_{\lambda^l}Z^k_{\text{est}}\right)\quad\forall l=1,\ldots,L\\ U^{k+1}_x=U^k_x-\tau\left(\partial_{U_x}Z^k_{\text{est}}+\beta^{-1}\partial_{U'_x}Z^k\right)\quad\forall x\in\mathcal{X}\\ V^{k+1}_y=V^k_y-\tau\left(\partial_{V_y}Z^k_{\text{est}}+\beta^{-1}\partial_{V'_y}Z^k\right)\quad\forall y\in\mathcal{Y}\end{cases}\\
&(m,n)\text{ update} &&\begin{cases}m^{k+1}_x=m^k_x+\tau\partial_{m_x}Z^{k+1}_{\text{est}}\quad\forall x\in\mathcal{X}\\ n^{k+1}_y=n^k_y+\tau\partial_{n_y}Z^{k+1}_{\text{est}}\quad\forall y\in\mathcal{Y},\end{cases}
\end{aligned}\tag{S.8}$$

where

$$\begin{cases}Z^k_{\text{est}}=Z(\lambda^k,U^k,V^k,U^k,V^k,\tilde{m}^k,\tilde{n}^k,\beta)\\ Z^{k+1}_{\text{est}}=Z(\lambda^{k+1},U^{k+1},V^{k+1},U^{k+1},V^{k+1},m^k,n^k,\beta).\end{cases}$$

The stopping criteria is

$$\max\left(\left|\lambda^{k+1}-\lambda^k\right|,\left|U^{k+1}-U^k\right|,\left|V^{k+1}-V^k\right|,\left|m^{k+1}-m^k\right|,\left|n^{k+1}-n^k\right|\right)<\delta.$$

S.4.3 Estimation Method Comparison

As we do for equilibrium computation, we compare the MPEC and primal-dual methods in structural estimation in Table S.1. We vary the number of types on each side of the market, as well as the number of parameters to estimate. MPEC proves once again faster than the primal-dual method in markets with low numbers of types and parameters to estimate, L, and is slower when markets have a large number of types and the model has a large number of parameters to estimate.

Table S.1: MPEC and Primal-Dual Performance - Structural Estimation

$\#\mathcal{X} \times \#\mathcal{Y} \times L$	$2 \times 2 \times 2$	$10 \times 10 \times 10$	$30 \times 30 \times 30$	$100 \times 100 \times 100$
		MPEC		
Min iter. nb.	4	6	8	9
Max iter. nb.	46	7	12	12
Mean time elapsed	.0028	.0342	2.494	161.8
		Primal-Dual		
Min iter. nb.	6033	3329	3528	7980
Max iter. nb.	19443	5286	4072	8315
Mean time elapsed	.0347	.0791	1.391	94.09

Notes: Program ran in Julia on a Macbook Pro with a M2 chip, 16GB of RAM, and 8 cores. The nonlinear solver for MPEC is KNITRO. Statistics computed on 10 replications. The convergence tolerance is 10e-6.

S.5 Data for Empirical Application

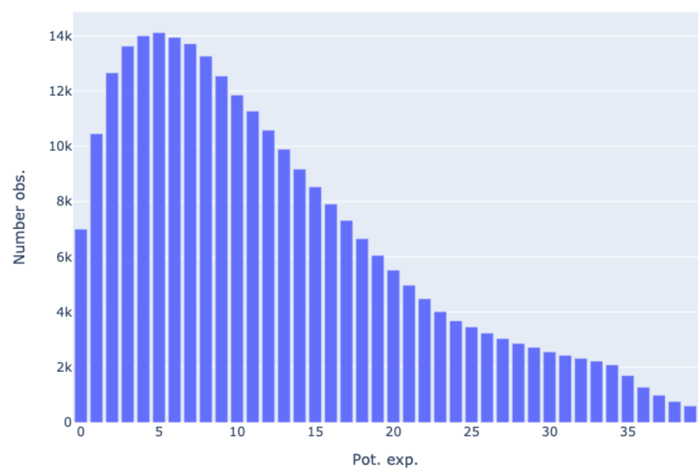
Our data is collected by the Swedish Employer's Federation (SAF in Swedish). The sample period is 1970–1990. Observations in the panel are at an individual times year level. Workers and firms have a unique identifier. The data contains a number of characteristics on both workers and firms, among which worker's age and job occupation. We do not consider observations that report less than 35 hours worked per week, nor workers below 25 years old. This is because becoming an engineer in Sweden requires at least five years of studying, and as all Swedish males must complete their military service, observations below 25 years old are scarce in the data. The worker population is mostly male (only 9.1% of engineers are women). We use the entire panel to estimate transitions and the stationary matching, as described in the main text.

We use information on occupation at the 1-digit level, and classify these occupations into two types: Technical and General. The Technical occupation are: Research and Development, Construction and Design, Technical Methodology, Planning, Control, Service and Industrial Preventive Health Care). The General occupations are: Administrative Work, Production Management, Communication, Library and Archival Work, Personnel Work, General Services, Business and Trade, Financial

Work and Office Services. Between 1975 and 1990, 64% of observations are technical occupation matches.

Figure S.1 shows the potential experience distribution across workers from 1975 to 1990. Most workers in the dataset have between 1 and 5 years of potential experience.

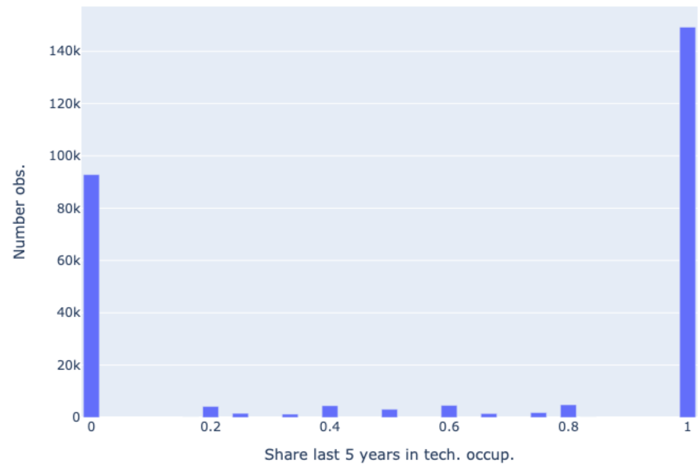
Figure S.1: x_e's distribution, 1975-1990



Authors' calculations from SAF data.

The share of a worker's time spent in a technical occupation in the past 5 years $\tilde{x}$ is measured starting in 1975 from the years 1970 to 1974, and over a sliding window for the years after 1975. Since the data is a panel, one simply counts the number of observations employed in a technical occupation by individual worker, and takes the ratio over the number of all observations in the past five years. Figure S.2 shows the distribution of this share across all workers between 1975 and 1990. Overall, 34.9% of observations are workers with no technical experience, 55% are workers with only technical experience, and 10.1% are workers with some level of technical experience.

Figure S.2: $\tilde{x}$'s distribution, 1975-1990



Authors' calculations from SAF data.