Learning Through Transient Matching*

Andrew Ferdowsian†

April 18, 2022

Abstract

I study a model of dynamic matching with overlapping generations, in which workers are born with incomplete preference information. To learn their preferences, workers must temporarily match with firms. Workers freely choose a firm to apply to each period, and firms hire their top applicants, up to a capacity constraint. I develop an algorithm extending techniques from the bandit literature to characterize the unique matching equilibrium. In general, equilibrium outcomes fail to satisfy standard notations of stability; furthermore, equilibrium search patterns differ from results in the directed search literature.

Extended Abstract

Motivation

Standard models of matching often assume that matches are permanent, and furthermore, that agents are aware of their preferences before matching. In practice, this isn't always the case. In many markets, such as job markets or even marriage markets, participants aren't endowed with knowledge of their preference rankings over all possible matches. Consider, for instance, a software engineer freshly graduated from college. While she might have broad preferences, say for location; she is unlikely to be certain that Meta's work culture better suits her than Google's. Instead, agents learn their preferences through interacting with potential partners. Critically, this necessarily entails spending time with a possible match partner, and therefore accruing flow utility from that match. "On average, it takes 4 months for a firm and a worker

*I thank Dorisz Albrecht, Robert Bell, Sylvain Chassang, Navin Kartik, Kwok Hao Lee, Alessandro Lizzeri, Lukas Mann, Daniel McGee, Xiaosheng Mu, Pietro Ortoleva, Evgenii Safonov, João Thereze, Tim Wang, and audiences at the Princeton for their helpful advice. Special thanks go to Can Urgun and Leeat Yariv for their continued support. I gratefully acknowledge financial support from the Dietrich Economic Theory Center.

†Department of Economics, Princeton University.

to learn the quality of their match" Menzio et al. (2012). This generates a classical exploitation versus exploration tradeoff, wherein participants may choose to propose to match partners with lower immediate expected value.

In this environment, predictions of standard matching models such as stability become tenuous. Several recent papers explore the outcomes of matching in environments with incomplete information, under the assumption that the market has already achieved stability, (Liu et al., 2014; Liu, 2020). However, the process by which stability might be reached hasn't been explored yet. Furthermore, these works assume that the matching market is a one-shot affair, wherein no new participants enter the market after the game begins. In many inherently dynamical environments, this fails to be the case. Returning to the example of an engineer, perhaps after a year spent working at various firms, she has learned that she prefers working at Google to Meta. During this timeframe, a new class of software engineers will have graduated and enter the market, unaware of their exact preferences. Since all agents compete in the same market, they exert strategic influence over each other's actions. This motivates the overlapping generations setup. Every period, workers stochastically retire while other workers enter the market, and all workers compete for the same pool of jobs.

Motivating Examples

As touched on, one key motivating example is that of software engineers. I focus on this case because of the large level of turnover, suggestive of learning. Over 50% of programmers leave their firms within two years, and over 75% leave within four years (Entreprogrammer, 2021).

The model also describes the myriad of online matching markets that have sprung up: Rover for pet sitting, Upwork for contracting, and Custom-Made for design, among others. In each of these, agents repeatedly match in short term contracts. Since reviews can only tell part of the story and necessarily won't match a prospective customer's exact preference, incomplete information plays a role. For instance, Rover.com displays ratings and number of repeat customers for each of its pet sitters. The variation in the number of repeat customers by sitter indicates that customers make strategic decisions regarding whether to hire their previous sitter again.

For a more concrete example, consider the market for French public school teachers. Every year, public school teachers in France are allowed to apply for placement in any region throughout France. Wages are fixed across regions, and so the "value" of a match depends solely on the teacher's fit or quality. Again, these matches involve teachers learning about their preferences for various regions, and using those preferences as a basis for applications.

Design and Results

I develop a model of transient matching with learning. To fix ideas, I describe a job market. A continuum of workers search for desirable jobs at a finite number of firms. A worker begins the game with incomplete information regarding his match values. To learn his match value at a firm, a worker must apply to that firm and be hired. Each period, every worker chooses a single firm to apply to. Each firm observes her set of applicants and hires the most qualified workers, subject to a capacity constraint. Hired workers learn their exact match values, while rejected workers do not. All workers have a small chance of retiring, after which a new worker of the same type is born. I assume that a form of Nash bargaining occurs, workers and firms equally split the surplus from matching.[1]

This setup features several technical challenges. Firstly, agents face non-trivial tradeoffs between exploiting their current knowledge about their maximal expected match value and exploring their match values at new firms. This generates a multiarmed bandit problem for each worker with collisions, wherein one worker applying to a firm may preclude another worker from matching with that firm. This prevents standard techniques such as the Gittins index from being directly applicable, since a worker's actions affect other workers actions and therefore change the rewards from each firm. I show that an extension of the Gittins index suffices to determine the optimal worker strategy in equilibrium.

The main result of this paper is the development of an algorithm which characterizes the unique equilibrium of the model. The algorithm uses insights from the bandit literature while utilizing a fixed point characterization of firms' hiring thresholds in order to construct an equilibrium. Utilizing this result, I show that increasing safety nets, the benefit from being rejected, fails to always incentivize learning. Relatedly, I characterize the conditions under which firms benefit from incomplete information. Last, I extend the model to heterogeneous discount factors, showing that the model is robust to extensions such as firm specific burn rates.

Related Literature

This paper touches on three distinct literatures: Bandits with collisions, dynamic matching with incomplete information, and directed search. Several papers on multi-armed bandits with collisions have recently emerged in the computer science literature, (see Liu et al. 2020, Liu et al. 2021). In these papers, the authors assume that agents know their priorities. In many practical situations, such as a job market, workers face uncertainty regarding their hiring prospects not only because they are unaware of that firm's set of applicants, but also because they do not know

[1] In the appendix, I consider an extension where wage transfers are strategically chosen by firms in an equilibrium of a supergame.

how they will be rated. Another key difference between this paper and the current literature, is that the solutions thus far fail to be incentive compatible when firms disagree on the ranking of workers. This paper focuses on the Markovian equilibrium of the game when all agents are strategic, lending itself to an economic interpretation of the situation.

As touched on, there is a burgeoning literature in matching with incomplete information. For instance, Anderson and Smith (2010) examine matching where agents form reputations regarding their quality over time and show that positive assortative matching emerges over time. To contrast, I instead investigate large markets where firms already have established reputations and workers instead aim to learn their own preferences. Ferdowsian, Niederle and Yariv (2022) explores matching under incomplete information in decentralized settings and shows that when preferences are aligned and agents are patient, stability is an equilibrium outcome. In this paper, I focus on cases where agents feature impatience, in particular, this implies that they will not wait arbitrarily long to determine their optimal match value, and instead they must decide when to stop learning and take their best offer. Kotowski et al. (2015) considers stability in long run matchings when agents have preferences over all possible sequences of matchings, and are aware of those preferences from the outset. Immorlica et al. (2020) is the closest to this paper in the economics literature. They consider students with similar incomplete information over preferences. The two key differences between my paper and theirs is that 1) they consider a centralized one-shot school matching market wherein once the match is finalized the market terminates; 2) the cost of acquiring information in their model is exogenous, while in this model it is the opportunity cost of leaving one's current position.

Last, there is an extensive existing literature in directed search. The motivation for this paper draws heavily on Menzio et al. (2012). The authors estimate that the time required to learn the quality of a match is 4 months on average. Becker (2011) shows that as much as 40% of benefits from employment are non-wage based, implying great value to proper matching. Chade et al. (2017) provides a useful survey that covers the directed search literature. This paper connects two distinct subliteratures, the search with marriage matching subliterature along with the literature on search with frictions. Miller (1984); Dagsvik et al. (1985); Jovanovic (1979) consider a one-agent version of this model, using multiarmed bandit techniques to describe the optimal search pattern for a single agent. I characterize the results found which generalize to a general equilibrium setting, and provide examples contradicting those that are specific to the individual worker setting. For instance, I find sharper conditions regarding when exploration is encouraged by safety nets. Of course, this section wouldn't be complete without noting that several of the techniques utilized in this paper are based on the work of Weitzman (1979).

References

Anderson, Axel and Lones Smith, "Dynamic matching and evolving reputations," *The Review of Economic Studies*, 2010, *77* (1), 3–29.

Becker, Dan, "Non-wage job characteristics and the case of the missing margin," *Available at SSRN 1805761*, 2011.

Chade, Hector, Jan Eeckhout, and Lones Smith, "Sorting through search and matching models in economics," *Journal of Economic Literature*, 2017, *55* (2), 493–544.

Dagsvik, John, Boyan Jovanovic, and Andrea Shepard, "A foundation for three popular assumptions in job-matching models," *Journal of Labor Economics*, 1985, *3* (4), 403–420.

Entreprogrammer, "Why programmers shouldn't stay in one company for a long time," July 2021.

Ferdowsian, Andrew, Muriel Niederle, and Leeat Yariv, "Decentralized Matching with Aligned Preferences," Technical Report, mimeo 2022.

Immorlica, Nicole, Jacob Leshno, Irene Lo, and Brendan Lucier, "Information acquisition in matching markets: The role of price discovery," *Available at SSRN 3705049*, 2020.

Jovanovic, Boyan, "Job matching and the theory of turnover," *Journal of political economy*, 1979, *87* (5, Part 1), 972–990.

Kotowski, Maciej H et al., *A note on stability in one-to-one, multi-period matching markets*, Harvard Univ., John F. Kennedy School of Government, 2015.

Liu, Lydia T, Feng Ruan, Horia Mania, and Michael I Jordan, "Bandit learning in decentralized matching markets," *Journal of Machine Learning Research*, 2021, *22* (211), 1–34.

_ , Horia Mania, and Michael Jordan, "Competing bandits in matching markets," in "International Conference on Artificial Intelligence and Statistics" PMLR 2020, pp. 1618–1628.

Liu, Qingmin, "Stability and bayesian consistency in two-sided markets," *American Economic Review*, 2020, *110* (8), 2625–66.

_ , George J Mailath, Andrew Postlewaite, and Larry Samuelson, "Stable matching with incomplete information," *Econometrica*, 2014, *82* (2), 541–587.

Menzio, G, IA Telyukova, and L Visschers, "Directed Search over the Life Cycle," Technical Report, mimeo 2012.

Miller, Robert A, "Job matching and occupational choice," *Journal of Political economy*, 1984, *92* (6), 1086–1120.

Weitzman, Martin L, "Optimal search for the best alternative," *Econometrica: Journal of the Econometric Society*, 1979, pp. 641–654.