Developing a Matching Algorithm for Agricultural Labor Markets: A Technical Note

Authors: Umut Dur, Siddharth Hari and Efsan Nas Ozen[1]

1. Background and Context

Employment in Turkish agriculture continues to remain highly informal owing to both to the nature of work in the sector as well as structural factors. Estimates suggest that in 2020 while 31 percent of all employment in the country was informal, the number was 84 percent in the agriculture sector. The nature of work in the sector which is predominantly seasonal, short-term, and low-skilled, the use of unpaid family labor on smaller farms, and the costs of writing formal labor contracts and registering workers with the Social Security Institution (SGK) contributes to the low rates of formalization. Further, long-standing structural, institutional, and policy constraints such as low productivity, limited domestic market integration, and distinct dualistic pattern in farm structures act as a binding constraint on growth and limit the sector's ability to create good quality jobs amenable to formalization.

Informal employment in agriculture is associated with a reliance on labor market intermediaries resulting in inefficiencies and are constraint to finding quality employment. Brokers fulfill the job matching function in agriculture in Turkey. Their activities fill an important gap as public employment services provided by the Turkish Employment Agency do not extend to the rural areas. In agricultural labor markets, farmers are likely to use brokers to recruit particularly short-term workers, for example, farmers in Bursa report to have hired about 40% of their workers through a broker[2]. A recent assessment suggests that about 80% of surveyed farmers are satisfied by the performance of labor brokers.[3] Brokers act as the bridge between the farmer and the worker and arrange all the logistical requirements for workers to complete their tasks in the farm during the agreed time period. In exchange, when a broker is used for job matching, farmers provide the total payment to the broker, who extracts a share of the total payment (about 5 to 10%)[4] for his or her services, further reducing the final amount received by the worker. Broker activities in agriculture fills a gap in the agricultural labor market, particularly for seasonal workers as they lack the knowledge to find where agricultural jobs are located. However, they create an additional cost item for workers who already are cash-constrained. There is also evidence that brokers delay payments to workers despite the fact that they receive the relevant payments on time from the farmers.[5] Furthermore, labor brokers stand as an important obstacle in accessing longer term employment with better conditions by breaking any direct relationship between the farmer and the worker, and making it difficult for the worker to negotiate his or her rights with the farmer, or for the farmer to get to know the worker better to be able to offer longer-term (or formal) employment. The inability to form a direct

[1] The authors would like to thank Laurent Bossavie and Mauro Testaverde for helpful discussions

[2] Farmer survey conducted in the Bursa Province between November 2021 and February 2022.

[3] ACC Project Training Needs Assessment

[4] Source: Yılmaz, M. , Sayın, C., and Bozoğlu, M. 2021. "Mevsimlik gezici tarım işçiliğinde tarım aracılığının rolü: Giresun ili fındık örneği". *Mediterranean Agricultural Sciences*, 34 (2), 189-194.

[5] ACC Project Training Needs Assessment

relationship with the worker causes inefficiencies for the farmer as well, because a worker trained to complete tasks in their farm may not show up for work the next day if the labor broker assigns him or her to a different farm on that day.

The *Agricultural Employment Support for Refugees and Turkish Citizens Through Enhanced Market Linkages* Project aims to improve working conditions in the agriculture sector and promote formalization through a combination of demand and supply side interventions. Funded by the European Union through the Facility for Refugees in Turkey (FRIT) with a budget of 46.8 million Euros, the Project is implemented by the Turkish Agricultural Credit Cooperatives (ACC) under the administrative management of the World Bank. On the labor demand side, the Project includes technical assistance for participating farmers as well as wage subsidies to hire workers formally.[6] Labor supply side interventions include screening of potential workers to assess their interest in and suitability for work in agriculture as well as technical, soft-skills and Turkish language skills training. The Project which began implementation in June 2021 with a pilot in the province of Mersin has now been scaled up to five other provinces – Adana, Adiyaman, Bursa, Gaziantep and Izmir.

The Project will also develop and implement an algorithm, based on insights from matching theory and using information on worker preferences and farmer needs, improve the quality of matches in the labor market. As part of this Project, detailed information on worker preferences (for example over work location, types of tasks, crops as well as times of the year they are available to work) as well as skills and experience is being collected to create a database of workers. Similarly, a farmer profiling exercise is being carried out to collect information on hiring needs for their farm (e.g. tasks they are looking to hire for, skill requirements, preferences over worker attributes etc.) which will be used to create a database of vacancies. Using this information, a matching algorithm will be developed which will support the allocation of workers to the most suitable jobs taking into account both the preferences of workers as well as the needs of the vacancy.

This note provides an overview of the proposed algorithm and discusses its potential to improve efficiency in the agricultural labor market as well its limitations. The rest of this note is structured as follows: Section 2 provides an introduction to matching theory and its applications. Section 3 describes the information required to implement a matching algorithm in the context of agricultural labor markets. Section 4 provides a conceptual framework guiding the algorithm while section 5 describes the algorithm in detail. Section 6 discussed the potential of such an algorithm to improve efficiency in the labor market, its limitations as well as the some implementational details.

2. An Introduction to Matching

Matching theory has been one of the most widely used applications of microeconomics and game theory. It examines questions related to two-sided markets such as students applying to schools or colleges, residency candidates applying for positions at hospitals, etc. and proposes algorithms to improve efficiency in these markets. In the last two decades, economists have worked on the design of matching markets and introduced matching algorithms which take into account policymakers' objectives and yield desired outcomes.

[6] This includes payment of (at least) the minimum wage and registration with Turkey's social security agency (SGK).

Most matching markets which have been studied function in a centralized manner. There is a central authority (e.g., school district, National Residency Matching Program) who runs the *"clearing house"*. The market is divided into two sets of participants (e.g., students and schools or doctors and hospitals). In order to match participants in both sides of the market, the central authority collects preference rankings of each participant (e.g. doctors) over the participants on the other side of the market (e.g., hospitals). Then an algorithm is developed which incorporates these preferences rankings to assign matches.

In this study, the tools of matching theory are applied to the agricultural labor markets in Turkey, which as described above is characterized by high levels of informal markets and where labor brokers are central in addressing informational frictions. This market differs significantly from the other settings where matching theory has been successfully applied. First, there is no central authority to whom farmers and workers can provide their preference rankings. Second, participants do not know each other *a priori*, therefore, preparing a preference ranking would require extensive interviewing process for both sides of the market which is impractical. Due to these critical differences, direct application of matching procedures used in other settings will not be sufficient, and important modifications are required.

3. Information Needs for Matching

To implement the matching algorithm described below, information regarding the characteristics and preferences of workers and farmers needs to be collected. This will be done through a *worker screening tool* and a *farmer profiling* tool which are briefly described in this section.

Worker Screening Tool: The objective of this tool, which will be administered to workers who express interest in participating in the Project, is to collect information on education, skills, past work experience and job preferences. The tool includes the following modules:

- Demographics: This module will collect information on age, gender, citizenship, address, access social assistance etc.

- Education, Skills and Labor Market Constraints: Information on highest completed level of education, training in agriculture, Turkish language skills as well physical disabilities which may limit a worker's ability to perform certain agricultural tasks will be collected through this module.

- Work History: This module will be used to collect detailed information on workers' employment histories. For refugee workers this section will include their work experience both in Turkey and in their country of origin. For workers with experience in agriculture, this module includes detailed questions about specific tasks performed, crops worked on etc.

- Work Preferences: This module includes questions on workers' availability to work during different months of the year, work locations as well as the kinds of tasks they would like to work on.

Farmer Profiling Tool: Mirroring the worker screening tool, the farmer profiling tool is used to collect information about vacancies identified by farmers. This tool includes the following modules:

- Demographics and Farm Characteristics: Basic information about the farmer (such as age, gender and location) as well as about the farm (size of farm, registration with the SGK, crops produced etc.) will be collected through this module.
- Information on Vacancies: This module will collect information on the number of workers to be hired, the crops and tasks they will work on, as well as importance of different characteristics (e.g. ability to understand Turkish, having previous experience working on the specific crop or specific task, formal training etc.).

4. Conceptual Framework

Consider a matching market composed of farmers and workers and denote the sets of workers and farmers W and F, respectively. Let $w \in W$ and $f \in F$ be a representative worker and farmer, respectively. Let $A = W \cup F$ denote the set of all participants. Let $a \in A$ be representative agent. Let q_f denote the maximum number of workers farmer f would like to hire.

Each participant a has her own characteristics such as work experience, skills, education, crop produced etc. Let c_a be the vector including information about the characteristics of participant a. Each participant a has taste over the characteristics of the participants from the other side of the market. Let t_a be the vector including information about the taste of participant a. Let $t = (t_a)_{a \in A}$ and $c = (c_a)_{a \in A}$ be the taste and characteristics profiles of the participants, respectively. Notice that, all information included t and c are collected in the farmer profiling and worker screening tools.

Given t and c, the objective is to construct a *proxy* preference ranking for each participant over the participants from the other side of the market. To this end, a score function denoted s will be used which takes taste and characteristics vectors as input and produce a score for each pair of participants from different sides of the markets. In particular, the participant $b's$ score for participant a is denoted by $s(t_b, c_a, c_b)$. If $s(t_b, c_a, c_b) > s(t_b, c_{a'}, c_b)$, then participant b ranks a over a' under her proxy preference ranking. Participants may consider specific characteristics of a participant from other side of the market as a "must have" characteristic – for example, a farmer would not consider hiring a workers who cannot speak Turkish. If participant b does not satisfy some must have characteristics based on t_a, then we say b is unacceptable for a and we set $s(t_a, c_b, c_a)$=-∞. We denote the proxy preference ranking of participant a over the participants on the other side of the market with P_a and:

$$s(t_a, c_b, c_a) > s(t_a, c_{b'}, c_a) \Leftrightarrow bP_a b'.$$

It should be emphasized that P_a represents the *proxy* preference ranking of participant a. That is, it might be different from the actual preference ranking. However, as explained above, it is not feasible to collect data on the *actual* preference ordering of farmers over workers and worker over farmers given information asymmetries and difficulties in conducting interviews with all participants. Hence, the rest of this note focuses on defining matching procedures using proxy preference rankings.

Definitions

A ***matching*** $\mu: W \cup F \to W \cup F$ is a function where $\mu(f) \subseteq W \cup \{f\}$, $\mu(w) \subseteq F \cup \{w\}$, $|\mu(f)| \leq q_f$ and $|\mu(w)| \leq 1$ for all $f \in F$ and $w \in W$. If $\mu(a) = a$, then participant a is unmatched. A ***mechanism*** is a systematic way which selects a matching for any instance.

A matching μ is ***stable*** if no participant is assigned to unacceptable participant and whenever there exists a farmer-worker pair (f, w) such that w prefers f to her match $\mu(w)$, then all vacancies at f are filled, $|\mu(f)| = q_f$, with workers who are ranked over w under P_f . The importance of achieving stability in this market should be highlighted. A matching fails to be stable if either (i) a participant is matched with an unacceptable partner or (ii) a farmer has not filled their capacity and a worker would like to match with that farmer or (iii) a farmer would like to replace one of her current matches with another worker whom would like to match with that farmer. Hence, when a matching fails to be stable, then the outcome of market will not be sustainable and there is an avoidable welfare loss in the market.

5. Description of the Matching Algorithm to be Implemented

In the matching markets studied in the literature, a matching mechanism takes submitted preference rankings of all participants as well as capacities into account and selects an outcome by following the objectives of the policy maker. In the setting being studied here complete preference rankings of all participants are unavailable and instead proxy preference rankings are constructed using the information collected through the farmer profiling and worker screening tools. This constitutes the first step of the matching procedure. Then, by using the constructed preferences the matching algorithm will be run.

Step 1: Constructing Proxy Preference Rankings

Recall that in order to construct proxy preference rankings, the scores of each participant for the participants on the other side of the market are calculated by using their responses in the profiling and screening tools. The score of a participant for another participant in the other side of the market increases as her responses satisfy more criteria listed. Moreover, we use different weights based on the importance level of the criteria considered. The exact weights used can be determined by the policy maker and can be easily incorporated into the matching procedure. The following example illustrates how proxy preference rankings are constructed (the example below shows the construction of proxy preference ranking for farmers over workers. The proxy preferences of workers over farmers can be constructed analogously).

Example: There are four workers and an apple farmer. The responses to the profiling and screening tools are given as follows:

Table 1: An example of farmer preferences and worker characteristics

	Daily working hours required	Ability to speak Turkish	Education level	Experience in apple production
Farmer	Min 8hrs	Important	Not important	Extremely Important
Worker 1	Max 10hrs	Yes	No	Yes
Worker 2	Max 6hrs	Yes	Yes	No
Worker 3	Max 8hrs	No	Yes	Yes

Worker 4	Max 8hrs	Yes	Yes	No

Notice that the responses to the question on daily working hours determine the acceptability of a worker for the farmer. In particular, independent of her responses to the other questions, if a worker cannot work at least 8 hours a day, then she is unacceptable for the farmer. The other questions do not determine acceptability, rather the responses to those questions determine the score a worker gets for the farmer. If a worker satisfies the criteria considered, then she gets a score depending on the importance of that criteria for the farmer. For simplicity, we score the responses to those questions as follows: if the corresponding criteria is "extremely important" and "important" then the worker gains X points and Y points with X>Y whenever she satisfies that factor, respectively. Table 2 shows the scores each worker would be assigned for this farmer.

Table 2: Each worker's score for the farmer

	Daily working hours	Speak Turkish	Education	Experience in apple	TOTAL
Worker 1	0	Y	0	X	X+Y
Worker 2	-∞	Y	0	0	-∞
Worker 3	0	0	0	X	X
Worker 4	0	Y	0	0	Y

Notice that, Worker 2 is unacceptable for the farmer, and therefore is assigned a score of -∞. The farmer's proxy preference ranking over the acceptable workers is:

Worker 1 P_f Worker 3 P_f Worker 4

Notice that in addition to the proxy preference ranking obtained through the profiling and screening tools a farmer may want to employ her current workers. In that case, these current workers can be moved to the top of the farmer's proxy preference ranking.

Finally, it should be emphasized that the proxy preference rankings of both sides are based on their responses to the profiling and screening tools. Since it is possible that multiple workers may provide the same responses to the questions, they all obtain the same score for a farmer (or vice versa). In these cases a tie-breaking rule will be used to construct the rankings.

Step 2: Matching Workers and Farmers Based on Proxy Preference Rankings

Three main algorithms that can be used to match workers and farmers in an employment relationship are defined. In the first algorithm, the match is determined randomly. In the second algorithm - *random serial dictatorship mechanism* – only the preferences of the workers are taken into account. In the third algorithm - *deferred acceptance algorithm* - preferences of both the farmers and workers are considered. The third algorithm is the preferred approach which will be implemented. The definitions of the other alternatives are provided for the sake of completeness and as a point of comparison.

The first algorithm ignores the preference orderings of both sides and only constraint is "acceptability" – i.e any worker can be matched with any farmer as long as they are both acceptable to each other.

The preferences used in the second and third algorithm are constructed based on the responses given by the workers and the farmers in the profiling tools. These algorithms are defined as below:

Option 1: Random Matching Algorithm:

Step 0: Select a random order over the workers.

Step 1: The first worker in this random order is randomly matched with one of the farmers who considers him/her acceptable.

In general,

Step k: The kth worker in this random order is randomly matched with one of his/her acceptable farmers with remaining capacity considering him/her acceptable.

This algorithm ignores all the information collected through the profiling and screening tool (except those which determine acceptability). This procedure can be thought as pure lottery with the only constraint being acceptability.

Alternatively, instead of choosing a random order over the workers, random order over the farmers can be chosen in step 0, and the above steps repeated. The equilibria (in expectation) of both these approaches are identical.

In a way this procedure imitates the matching procedure used by ISKUR with one difference. Employers submit requests to ISKUR which specify the characteristics of the workers they need. ISKUR matches a pool of workers to that job and then randomly draws the names of the workers for the job considered. If the arrival of employers to ISKUR is random, then the algorithms would be similar to the random matching algorithm.

Option 2: Serial Dictatorship Algorithm:

Step 0: We select a random order over the workers.

Step 1: The first worker in this random order selects his/her best farmer among those who consider him/her acceptable.

In general,

Step k: The kth worker in this random order selects his/her best farmer remaining in the pool among those considering him/her acceptable.

In this case, the information on the proxy preferences of the workers is considered, and the information collected from farmers is not utilized (except to determine acceptability). Variants of this procedure is used to match students to dorms and students to courses at colleges. The algorithm can be implemented analogously starting with farmers in Step 0.

Deferred Acceptance Algorithm:

Step 1: Every worker applies to the best farmer according to his/her proxy preference ranking. Every farmer considers his/her applicants in this step and tentatively accepts the best ones (up to the capacity) according to his/her proxy preference ranking. The workers who are not tentatively accepted are rejected.

In general,

Step k: Every worker applies to the best farmer according to his/her proxy preference ranking who has not rejected him/her yet. Every farmer considers his/her applicants in this step and tentatively accepts the best ones up to capacity according to his/her proxy preference ranking. The workers who are not tentatively accepted are rejected.

This procedure imitates the decentralized labor market dynamics where workers apply to the jobs one by one and the employers accept or decline applications. The main objective of this mechanism is maintaining a stability in the assignment while considering preferences of both sides of the market. In other words, the information on the preferences of both the farmers and the workers (obtained from the profiling and screening tools) are used and, hence, the matching process is more informationally efficient. Examples of the implementation of deferred acceptance algorithm include, but not limited to, National Residency Matching Program in the US and college admissions in Turkey, China and Hungary. Here, stability can be interpreted as follows: If a worker prefers another farmer to his current match, then all the workers hired by that farmer are better than him. Different from the decentralized labor markets, we do not have frictions due to application, search and replacing workers with better ones.

6. Discussion

Given the high levels of informality and frictions in the agricultural labor market, an algorithm which matches workers and farmers taking into account their preferences and needs offers several advantages including better quality matches, longer job tenures and reduced worker exploitation. As described above, labor brokers play a central role in bringing workers to farmers. The matching algorithm, which makes systematic use of the preferences of workers and the needs of farmers is likely to result in higher quality matches and improved worker productivity. Agricultural labor markets are characterized by very

short-term work and high turnover. By incorporating the preferences of workers and farmers the matching algorithm is likely to increase satisfaction among both workers and farmers, thereby increasing job tenures. This, in turn, may make farmers more willing to invest in their workers and could result in improved work conditions as well as increased formalization. Finally, the use of the algorithm will also reduce the exploitation of workers by labor brokers and allow them to form relationships directly with their employers.

However, this approach is not without its challenges. First, given the nature of agricultural work, collecting information on the needs of each vacancy and assessing the skills and preferences of workers might be difficult. The worker screening and farmer profiling tools are critical inputs for the matching algorithm, but farmers and workers may find it challenging to respond to these questions accurately. Efforts are being made to maximize the quality of data by incorporating insights from implementers on the ground who are familiar with the local context. These tools will be updated based on learnings from initial rounds of data collection. Second, given the rural context, manpower required to collect this information from farmers and workers who are spread out in different rural areas incurs a nonnegligible cost. Finally, the algorithm is likely to be most useful for jobs which require specific skills since farmers can clearly specify the tasks to be performed and the characteristics of the workers they would want to hire. However, several jobs in the sector are short-term (for example farmers hire several workers during the harvest season) for which farmers may not be able to specify the skills or worker characteristics they are looking for.

Developing a matching algorithm for the agriculture sector is an innovative approach to address some of the challenges associated with informal labor markets. Significant efforts will be made to learn from the implementation of the algorithm to understand which aspects work well and which may need to be modified. If successful, this approach has the potential to be adopted and expanded to other contexts faced with similar labor market frictions.