Skill Mismatch Unemployment

Marianna Kudlyak  Erin Wolcott
Federal Reserve Bank of San Francisco  Middlebury College

PRELIMINARY AND INCOMPLETE. NOT FOR DISTRIBUTION

October 2018

Abstract:

What is the contribution of mismatch between the skills demanded by vacancies and the skills supplied by the unemployed to the unemployment rate? We measure skill mismatch unemployment using a search and matching framework in which unemployed workers and vacancies meet randomly, but the probability of the meeting resulting in a match depends on how closely the unemployed worker’s skillset aligns with the job requirements. Using the matched CPS, O*NET, and HWOL data from 2005 to 2018, we find that skill matching efficiency between vacancies and the unemployed decreased during the Great Recession, but has since increased back to the pre-recession levels.

JEL: E24, E32, J30, J41, J63, J64
Keywords: Unemployment. Mismatch. Skills

The views expressed here are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of San Francisco, the Federal Reserve System, or any other institution with which the authors are affiliated.
1. Introduction

What determines whether an unemployed worker matches with a job opening? Many factors are at play, including: how many job openings are available, how many other workers are unemployed, but most importantly whether the skill possessed by the unemployed meet the requirements of the opening. We measure skill mismatch unemployment—unemployment due to mismatch between the demand and supply of skills—using an extended search and matching framework.

The most intuitive way of thinking why a worker is not hired for a particular vacancy is because the worker’s skills does not match the requirements. Recent studies document significant secular shifts in the skill requirements due to automation and job polarization (Autor and Dorn, 2013), demand for interpersonal skills (Deming, forthcoming), or software substitution of skilled jobs (Aum, 2017). The skills of laid-off workers are often tied to the disappearing occupations. Skills of new labor market entrants reflect current schooling curriculum that might not be up to speed with the requirements of the new jobs.

In this paper, we measure mismatch unemployment via changes in aggregate matching efficiency resulting from a mismatch between the skills of job seekers and skill requirements of the vacancies. The lower the matching efficiency, the higher the mismatch. We employ a search and matching function that allows for nontrivial probability of a meeting between job seeker and vacancy with different skill characteristics resulting in a match.

In a standard textbook random search and matching framework (Pissarides, 2000) all vacancies and unemployed are homogeneous. Any unemployed worker can match with any vacancy once the job seeker and vacancy meet. This is counterfactual. Consequently, the standard random search framework is not suited for studying the role that skill mismatch plays in driving aggregate matching efficiency. The directed search approach such as that employed by Sahin et al. (2014) can potentially account for skills by segregating workers and job openings onto islands (occupations) and only workers with a particular skillset (i.e., in a given occupation) can search for jobs on an associated island. However, this approach assumes that job seekers can only search in one island, i.e., effectively setting to zero the probability of a match of a vacancy and a worker from different occupation.
We study skill mismatch unemployment in a framework where unemployed workers and vacancies meet randomly but the probability of the match depends on the skill match, i.e., distance, between the unemployed and the vacancy. In other words, we account for skill mismatch by assuming all job seekers and all job openings can meet with each other and potentially create a match, but the pairs that are further apart in terms of skill content have a lower probability of creating successful match. We then use counts of vacancies and the unemployed with different skillsets and aggregate over match probabilities of all potential pairs to create a new measure of aggregate matching efficiency that directly accounts for skill differences. This measure is close to one when unemployed workers have the skills firms are looking for. This measure is close to zero when unemployed workers have skills that are very different from what firms are looking for.

To obtain data on the skillsets of job seekers, we merge the CPS with skill data from O*NET. To obtain data on the skill requirements of job openings, we merge The Conference Board’s help-wanted online (HWOL) data with skill data from O*NET. Since HWOL is available starting with 2005, our sample period covers 2005-2018.

We find that the skill matching efficiency declined during the Great Recession, but has since recovered to pre-recession levels. Through counterfactual exercises, we find much of this recovery is from low-skill unemployed and marginally attached workers. In other words, following the recession the supply of skills such as “equipment maintenance,” “repairing,” and “operation and control” decreased relative to the demand. Since 2015, matching efficiency has tapered off a bit. This is largely driven by high-skill vacancy creation that is not being met by high skill labor supply, meaning firms’ demand for skills such as “programming,” “science,” and “technology design” has recently increased relative to supply.

1.1. Related Literature

TBA

The rest of the paper is organized as follows. Section 2 describes a theoretical framework for a matching function with skill heterogeneity. Section 3 describes the data. Section 4 provides descriptive statistics about the supply and demand of skills. Section 5 describes the skill mismatch results. Section 6 concludes.
2. Matching Function with Skill Heterogeneity

This Section describes random search with probabilistic matching.

2.1. Theoretical Framework

Consider an economy with heterogeneous pool of job seekers and vacancies. Let $k$ index a vacancy characterized by an $S$-dimensional vector of requirements $v^{1}, v^{2}, ..., v^{S}$. Let $i$ indexes a non-employed job seeker characterized by an $S$-dimensional vector of characteristics $u^{1}, u^{2}, ..., u^{S}$ that correspond to the vacancy requirements. These vacancy requirements can encompass a vector of skills or an amount of experience.

Search is random. The number of meetings is described by the standard Cobb-Douglas matching function

$$m_{i} = \mu v_{i}^{\alpha} u_{i}^{1-\alpha},$$

where $\alpha$ is the matching function elasticity with respect to the number of vacancies and $\mu$ is the matching efficiency.

The probability that a job seeker $i$ meets a vacancy is

$$\frac{m_{i}}{u_{i}} = \mu \left( \frac{v_{i}}{u_{i}} \right)^{\alpha},$$

The probability that job seeker $i$ meets vacancy $j$ and the meeting results in a match is

$$\lambda_{ij} = \mu \left( \frac{v_{i}}{u_{i}} \right)^{\alpha} \frac{P_{ij}}{v_{i}}, \quad (1)$$

where $P_{ij}$ is the probability that job seeker of type $i$ is hired by vacancy of type $j$. Note that the probability is time-invariant. We relax this assumption later.

The job finding rate of job seeker $i$ is then
\[ \lambda_v = \sum_{j=1}^{J} \lambda_{v_j} = \mu \left( \frac{v}{u} \right) \sum_{j=1}^{J} \frac{v_j}{v_i} P_j. \]

The average job finding rate in the economy is

\[ \lambda_v^* = \sum_{j=1}^{J} \frac{u_j}{u_i} \lambda_{v_j} = \mu \left( \frac{v}{u} \right)^v \sum_{j=1}^{J} \frac{u_j}{u_i} \frac{v_j}{v_i} P_j, \quad (2) \]

Or we can re-write

\[ \lambda_v = \mu \left( \frac{v}{u} \right)^v \]

where

\[ \mu^* = \mu \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{u_i}{u_j} \frac{v_j}{v_i} P_j \quad (3) \]

and \( \mu \) is the aggregate matching efficiency in the standard models, without job seeker or vacancy heterogeneity.

The term \( \sum_{j=1}^{J} \sum_{j=1}^{J} \frac{u_j}{u_i} \frac{v_j}{v_i} P_j \) is the skill mismatch term that depends on the probability of a successful match between job seeker \( i \) and vacancy \( j \) and on the relative shares of different types of job seekers and vacancies in the aggregate search pool.

If all meetings between job seekers \( i \) and vacancies \( j \) are equally likely to turn into hiring, i.e., \( P_v = 1 \forall i, j \), then \( \mu^* = \mu \), i.e.,

\[ \mu^* = \mu \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{u_i}{u_j} \frac{v_j}{v_i} = \mu \sum_{i=1}^{I} \frac{u_i}{u_j} \sum_{j=1}^{J} \frac{v_j}{v_i} = \mu \sum_{i=1}^{I} \frac{u_i}{u_j} = \mu. \]

However, generally, \( \mu^* \neq \mu \). For example, if only meetings between job seekers and vacancies with equal skills vectors turn into hiring, i.e., \( P_v = 1 \forall i = j \) and \( P_v = 0 \forall i \neq j \), then, i.e.,

\[ \mu^* = \mu \sum_{i=1}^{I} \frac{u_i}{u_j} \frac{v_i}{v_j}. \quad (4) \]
Clearly, the skill mismatch term in Eq. 3 is time-varying. Even if the probability of a match between job seeker $i$ and vacancy $j$ is constant, the shares of different types of job seekers and vacancies in the entire searching pool vary. This variation contributes to the variation in $\mu_t$.

In what follows, we analyze how much of the change in the aggregate matching efficiency in the standard matching function can be attributed to the change in the skill mismatch. In other words, we normalize $\mu_t = 1$ and we compute $\mu_t'$ from the sample period our data allows: May 2005 to August 2018.

3.2. Measuring a Probability of a Match

To compute the probability a worker in occupation $i$ matches with an opening in occupation $j$, we turn to data on skills and calculate how different the occupations are in skill space. The literature provides different ways to calculate distance between skills.

For example, Guvenen, Kuruscu, Tanaka, and Wiczer (2015) estimate mismatch between a worker and the job that worker is employed. Their measure of mismatch is the weighted sum of the absolute difference between possessed and required skills,

$$ mis'_v = \sum_{i} \omega_i |\mu'_i - v'_i|, \quad (6) $$

where the weights are estimated empirically from wages. To construct this distance, we need to actually observe workers and vacancies in match. However, is the probability of a match is low, such matches are rare in the data or never observed.

Most closely related to ours is the measure used in Macaluso (2017). Macaluso (2017) constructs a measure of skill remoteness that measures the distance between worker’s current occupation and the rest occupations in the city in which others are employed. Macaluso’s distance is a simple average of absolute differences between skills.

$$ mis'_s = \frac{1}{S} \sum_{i} |\mu'_i - v'_i| \quad (7) $$
We analyze skill mismatch under this alternative specification where instead of comparing a worker’s current occupation with the rest of the occupations in the city, we compare a worker’s previous occupation with the potential job opening’s occupation.

3. Data

3.1. Datasets

To create a measure of skill mismatch we merge data on skills from the U.S. Department of Labor's O*NET 4.0 database with (1) the Current Population Survey (CPS) and (2) The Conference Board’s help-wanted online dataset (HWOL). For each occupation, O*NET assigns ratings to a set of skills. Specifically, for each occupation, O*NET rates the importance of 35 skills on a relative scale. Examples of these skills include reading comprehension, writing, critical thinking, coordination, and others. We merge these datasets through occupations. We use 2-digit HWOL data from HAVER and CPS data. We convert 8-digit SOC codes to 2-digit occupation codes, taking the simple average, in order to merge this data with the other two datasets.

3.2. Measurement of Job Seeker Characteristics

We assign skills to job seekers based on their most recent occupation. Two challenges arise in doing this: (1) we do not have information on previous occupation for new labor market entrants (students etc.), and (2) we do not have recent employment information for individuals who left the labor force a while ago.

We consider two definitions of job seekers: (1) the unemployed which is the narrow definition since respondents must report looking for a job in the last four weeks to be counted as unemployed and (1) the non-employed which is a broader definition because it includes all non-working respondents (i.e. the unemployed and those out of the labor force).

We calculate shares of job-seekers (unemployed and non-employed) in each occupation using the CPS. We restrict attention to prime-age respondents (25-54). Of the unemployed, 98 percent of our sample has a valid occupation entry. However, only 2.4 percent of observations labeled out of the labor force have valid occupation entry. To address these missing occupations, we use the
matched structure of the CPS. We look across the 8 months individuals are surveyed. First, we carry occupations forward: if a respondent has a valid occupation entry in month t-1, but not in t, we use the occupation in month t-1 for t. Second, we carry occupations backwards: if a respondent has a valid occupation in t+1, but invalid in t, we use t+1 for t. This procedure results in 89 percent of the non-employed having valid occupations for 2005-2017 compared to 80 percent before the adjustment. The vast majority of respondents who do not have a valid occupation entry (and we therefore drop), are NNNNNNNNN. A handful are UUUUUUUU or some combination of the two. We do not worry much about these dropped observations because these individuals are very marginally attached the labor force.

3.3. Skills

There are two skills rating measures in O*NET. One measure rates the “importance” of each skill to a specific occupation on a scale of 0 to 5 and another measure rates the "level" used of each skill on a scale of 1 to 7. We use the “importance” measure and rescale it so it ranges from 0 to 1. We calculate the Manhattan distance (L1 norm) between all occupation pairs based on their 35-element skill vector. We construct probabilities $P_{ij}$ by subtracting theses distances from 1 so that low distances translate into high probabilities and high distances translate into low probabilities.

4. Supply and Demand of Skills

Figure 1 plots the average vacancy skill profile (black) and average non-employed skill profile (blue) for the entire sample period. We do this by averaging across all occupations using vacancy shares and unemployed shares as weights. Skills are ordered according to decreasing tightness (i.e. vacancy-nonemployment ratios). Unsurprisingly the three tightest skills are programming, science, and technology design. The three slackest skills are operation and control, repairing, and equipment maintenance. In other words, there is more demand relative to supply for higher type skills and more supply relative to demand for lower type skills.

Figure 2 then plots tightness ratios for the three tightest and three slackest skills over time. The horizontal line represents where the tightness ratio is one, meaning average supply and demand of
these skills is equivalent. The higher type skills are all above one, indicating demand for these skills exceed supply. Conversely, the lower type skills are all below one, indicating the supply of these skills exceed demand.

Figure 2 shows movements in these six selected skills over the business cycle. Leading up to and during the recession, demand of the high skills increasingly exceed supply, while supply of the low skills increasingly exceeded demand. This flipped following the recession: the high skills became less tight and the low skills became tighter.

5. Skill Mismatch

5.1. Aggregate Matching Efficiency

Figure 3 plots the component of matching efficiency that accounts for skill heterogeneity ($\mu'$). In other words, the inverse of skill mismatch. We first do this defining job seeker as non-employed in Figure 3 and then unemployed in Figure 4. The fact that both lines are less than one means that the average job seeker is not perfectly suited for the average job opening—the average job seeker has a set of skills that overlaps with the average job opening, but is not a perfect match. Using our measure of the distance, we find that the skill matching efficiency is 86% of the matching efficiency under homogeneous or perfect matching (i.e., when all vacancies and job seekers match with probability 1).

Both figures (for the unemployed and for all the non-employed) display a similar cyclical pattern. Leading up to and during the Great Recession, job seekers and vacancies were less well suited. During the recovery, job seekers were a better fit for the set of job openings. More recently, this component of matching efficiency has tapered off.

The matching efficiency for the unemployed declined 0.63 percent from May 2006 to trough, while the matching efficiency for the non-employed only declined by 0.57 percent. This indicates that in the recession the pool of the unemployed (who by definition are actively searching for work) turned particularly ill-suited for the available vacancies. That is, the individuals who are most interested in employment in recession were those with the worse prospects. This contradicts prevailing existing intuition that in recessions the pool of unemployed is of better “quality “on
average. However, it is consistent with a notion that these workers are unemployed because they are not well-suited for the jobs that firms create.

5.2. Demand or Supply – What Drives Changes in the Skill Matching Efficiency

Do changes in vacancies requirements or in the skill sets of the job seekers drive the cyclical pattern that we observe? To test this we construct counterfactual measures of the skill matching efficiency in which we hold either the vacancy or the unemployed skill sets constant.

We first construct the “average” vacancy over the sample period. $P_{ij}$ is a 22-element vector representing the probability an unemployed worker in occupation $i$ is a good match with the average vacancy in the economy $j$. We then plug this set of matching probabilities into our matching efficiency measure and examine to what extent changes in the supply of skills (the number of the unemployed with different skill sets) drive the change in the matching efficiency.

$$P_{ij} = 1 - \frac{1}{35} \sum_{s=1}^{S} p_{i}^{s} - \frac{1}{T} \sum_{t=1}^{T} \frac{v_{jt}}{v_{j}} p_{i}^{s}$$

$$\mu_{i}^{U} = \sum_{j=1}^{J} \frac{u_{it}}{u_{i}} p_{ij}$$

We do the analogous exercise to determine if vacancy shares is what is driving the cyclical pattern of the aggregate skill matching efficiency. Here we average over occupations and skills among the unemployed.

$$P_{ij} = 1 - \frac{1}{35} \sum_{s=1}^{S} p_{j}^{s} - \frac{1}{T} \sum_{t=1}^{T} \frac{u_{it}}{u_{i}} p_{j}^{s}$$

$$\mu_{i}^{V} = \sum_{j=1}^{J} \frac{v_{jt}}{v_{j}} p_{ij}$$

Figures 5 and 6 plot the contribution of unemployment shares $\mu_{i}^{U}$ and vacancy shares $\mu_{i}^{V}$ to overall matching efficiency, for the unemployed and for all the non-employed. We find that both
vacancies and non-employed/unemployed workers are responsible for the rise in matching efficiency immediately following the Great Recession. Figure 7 investigates this further by plotting the contribution of the three tightness and three slackest skills. To understand this figure, imagine averaging over all panels in Figure 7 (and the other skills not shown). Doing so would result in Figure 5. This new Figure 7 shows that changes in high skill nonemployment (top row) cannot explain much of the fluctuations in aggregated matching efficiency however high skill vacancies contributed to its rise post-2009. Turning now to the bottom row, vacancies and nonemployment seem to work in opposite directions. For the most part, changes in the composition of low skill nonemployed workers is what contributed to the post-recession recovery.

Since 2015, matching efficiency has tapered off a bit. This is largely driven by high-skill vacancy creation, meaning firms’ demand for O*NET skills such as “programming,” “science,” and “technology design” has recently increased relative to supply.

5.3. Skill Mismatch Unemployment

TBA

6. Conclusion

TBD
References


Figure 1: Average vacancy-nonemployed ratios by skill: 2005m5-2018m8
Figure 2: Vacancy-Nonemployed Ratios by Skill
Figure 3: Matching Efficiency with Skill Heterogeneity for the Nonemployed

Figure 4: Matching Efficiency with Skill Heterogeneity for the Unemployed *

*Data is missing because for a select number of years at least one occupation in the CPS does not have any unemployed workers
Figure 5: Counterfactual Skill Matching Efficiency, Non-employed

Figure 6: Counterfactual Skill Matching Efficiency, Unemployed
Figure 7: Counterfactuals by Skill