

# Employability

Defining Employability Coefficients for Careers and Professional Attributes

Atlas Research Team

2023-06-12

# Contents

<b>1</b>	<b>Mathematical Definition of Employability</b>	<b>3</b>
1.1	Initial Insights . . . . .	3
1.1.1	Probability of Finding a Job . . . . .	3
1.1.2	Adjusting for Professional Similarity . . . . .	3
1.2	Additional Metrics . . . . .	6
1.2.1	Revised Interchangeability Coefficients . . . . .	6
1.2.2	Utility-Consistent Employability Coefficients . . . . .	7
<b>2</b>	<b>Estimating Employability Coefficients</b>	<b>8</b>
2.1	For Occupations . . . . .	8
2.2	For Professional Attributes . . . . .	8
2.3	For Any Professional Profile . . . . .	8
<b>3</b>	<b>Professional Competitiveness</b>	<b>8</b>
3.1	Pressupposed Metrics . . . . .	8
3.2	Job Competitiveness in a Simulated Economy . . . . .	10

# 1 Mathematical Definition of Employability

## 1.1 Initial Insights

### 1.1.1 Probability of Finding a Job

Employability, as the word itself indicates, is *the ability to find a job*. This is very obvious, indeed, but it bears mentioning in order to quantify the concept properly. From this tautological notion, then, we could define, in mathematical terms, that employability is not exactly the *ability* to find a job, but rather the *probability* of being employed, given a particular set of professional attributes. That is, if we measure professional attributes on a scale, say, of 0 to 100, calling each attribute  $a_i^u$  the  $i$ -th attribute of person  $u$ , such that  $a_i^u \in [0, 100] \forall i, u$ , then it follows that employability is:

$$\Lambda(\mathbf{a}_u) = \Pr(l = 1 \mid \mathbf{a}_u),$$

where  $\mathbf{a}_u = (a_1^u, \dots, a_m^u)$  is the vector of professional attributes  $a_i^u$  that characterizes person  $u$ , and  $l \in \{0, 1\}$  is a dummy variable denoting whether the individual is employed or not.

Now, suppose every individual employed in an occupation  $q$  are characterized by the very same  $\mathbf{a}_q = (a_1^q, \dots, a_m^q)$  vector of professional attributes. Additionally, suppose we know the  $n_q \geq 0$  employment level of this occupation, and the total labor force  $N = \sum_j n_j$ . With this, it is then straightforward to estimate the probability  $\Lambda(\mathbf{a}_q)$  defined above as:

$$\begin{aligned} \Lambda(\mathbf{a}_q, n_q, N) &= \Pr(l = 1 \mid \mathbf{a}_q) \\ &= \Pr(l = 1 \mid n_q, N) \\ &= \frac{n_q}{N} \end{aligned}$$

In other words, the employability of occupation  $q$  is the probability of any person in the total work force  $N$  being one of the  $n_q$  people employed in that occupation. Pretty simple.

However, this formulation fails to be an adequate employability metric for two main reasons. The first is that it yields values too small to be easily interpretable. Since any given occupation is but a really tiny fraction of the total labor market,  $\Lambda$  values calculated this way are proportionally small. The second problem is a more interesting one, and will be treated in the following section.

### 1.1.2 Adjusting for Professional Similarity

The second problem with this rather naive employability coefficient is that people can find jobs in many different fields, provided they are sufficiently qualified. If an individual, for instance, is employed in occupation  $q$  and occupation  $q$  is very similar to another occupation  $\tilde{q}$ , this individual could also find employment as a  $\tilde{q}$ . Thus, their employability coefficient would need to account for the probability of them being employed either as  $q$  or  $\tilde{q}$  — again, supposing they are sufficiently qualified.

Mathematically, if occupations  $q$  and  $\tilde{q}$  are said to be similar, then the probability of a person with a professional profile identical with  $q$  finding a job either as  $q$  or  $\tilde{q}$  could be intuitively written as the probability of two mutually exclusive events, adjusted by the probability of two other independent events, namely: the probability of finding a job as a  $q$  or a  $\tilde{q}$ , adjusted by the probability of being *recognized* by the labor market as either one of these occupations to begin with.

The first part of this solution has to do with the amount of job posts available for each occupation, and it was already explained above. The second part has to do with professional similarity, as two occupations are recognized as identical by the labor market to the degree they are similar in the Euclidean sense. Hence, we may define a similarity metric like:

$$s(\mathbf{a}_q, \mathbf{a}_{\tilde{q}}) = 1 - \frac{d(\mathbf{a}_q, \mathbf{a}_{\tilde{q}})}{\max d(\mathbf{a}_j, \mathbf{a}_k)} = 1 - \frac{\sqrt{\sum_{i=1}^m (a_i^q - a_i^{\tilde{q}})^2}}{\max \sqrt{\sum_{i=1}^m (a_i^j - a_i^k)^2}},$$

where  $d(\mathbf{a}_q, \mathbf{a}_{\tilde{q}})$  is the Euclidean distance between the professional attribute vectors of occupations  $q$  and  $\tilde{q}$ , and  $\max d(\mathbf{a}_j, \mathbf{a}_k)$ , the maximum Euclidean distance between any two occupations  $j$  and  $k$ . From this, we understand that the overlap between vectors  $\mathbf{a}_q$  and  $\mathbf{a}_{\tilde{q}}$  is proportional to the interchangeability of occupations  $q$  and  $\tilde{q}$ . And so, as  $s(\mathbf{a}_j, \mathbf{a}_k) \in [0, 1] \forall j, k$ , this metric could be understood as an initial formulation for the probability of sufficient professional overlap:

$$I(\mathbf{a}_j, \mathbf{a}_k) = \Pr(j \approx k \mid \mathbf{a}_j, \mathbf{a}_k) = s(\mathbf{a}_j, \mathbf{a}_k)$$

The coefficient above refers to the interchangeability of occupations, and can be defined as the similarity between them, or any increasing function of it. Its interpretation is as follows: if employers are interested in a particular subset of professional attributes, which can be attained by any person, regardless of their career path, then the probability of a person having the desired competencies is the overlap between their own competencies and those of a specific occupation, or professional profile, displaying these competencies to the desired degree; or, alternatively, the probability of two professional profiles being interchangeable is the percentage of overlap between them.

The reasoning behind this apparently hasty conclusion is simple: all the job activities of an occupation require of individuals a particular set of skills, which is quantified by their professional profile. If a person has all these competencies, then the probability of being unable to perform a given task on the job is, *ceteris paribus*, precisely zero. But, conversely, if they lack any of the traits required to perform these tasks, then the probability they are unable to perform them properly increases by the exact measure of their under-qualification. Thus, the interchangeability of professional profiles is proportional to their similarity.

Now, of course, this initial formulation is a deliberate simplification, and we need to adjust for the possibility of *over*-qualification, as well as the importance of each professional attribute to each career path. So, we should write a revised interchangeability metric like

$$I(\mathbf{a}_q, \mathbf{a}_{\tilde{q}}) = \Pr(q \approx \tilde{q} \mid \mathbf{a}_q, \mathbf{a}_{\tilde{q}}) = S(\mathbf{a}_q, \mathbf{a}_{\tilde{q}}) = 1 - \sum_{i=1}^m \alpha_i^{\tilde{q}} \delta_i^q(a_i^q, a_i^{\tilde{q}}),$$

$$\alpha_i^k = \frac{a_i^k}{\sum_{i=1}^m a_i^k} \implies \sum_{i=1}^m \alpha_i^k = 1,$$

$$\delta_i^j(a_i^j, a_i^k) = \max\left(\frac{a_i^k - a_i^j}{100}, 0\right), \forall j, k,$$

where  $\delta_i^j(a_i^j, a_i^k)$  is the over-qualification-adjusted competency gap function, which returns the normalized distance between the  $i$ -th attributes of occupations  $j$  and  $k$ , if this gap is positive (i.e. if occupation  $j$  is under-qualified in this regard), and returns zero otherwise. This means the algorithm above does not “punish” over-qualification. Furthermore, these  $\delta_i^j(a_i^j, a_i^k)$  competency gaps are weighted by the  $\alpha_i^k$  relevance of each attribute to the comparison occupation. This way, competency gaps in the *core* competencies are given more importance.

It is evident this coefficient is mathematically consistent, as

$$\delta_i^q(a_i^q, a_i^{\tilde{q}}) = 0 \forall i \implies \sum_{i=1}^m \alpha_i^{\tilde{q}} \delta_i^q(a_i^q, a_i^{\tilde{q}}) = \sum_{i=1}^m \alpha_i^{\tilde{q}} \times 0 = 0 \implies S(\mathbf{a}_q, \mathbf{a}_{\tilde{q}}) = 1,$$

and, also, that

$$\delta_i^q(a_i^q, a_i^{\tilde{q}}) = 1 \forall i \implies \sum_{i=1}^m \alpha_i^{\tilde{q}} \delta_i^q(a_i^q, a_i^{\tilde{q}}) = \sum_{i=1}^m \alpha_i^{\tilde{q}} \times 1 = 1 \implies S(\mathbf{a}_q, \mathbf{a}_{\tilde{q}}) = 0.$$

[insert analogy here]

Another way of thinking about this concept would be to pose that some — if not most — of the jobs generally associated with a given career path do not actually require its entire set of skills, but only a fraction of it; so that any “sufficiently qualified” job applicant could be hired to perform such jobs. Yet another way of thinking about the interchangeability coefficient is to conjecture employers may be willing to hire under-qualified personnel, probably by less than their more qualified counterparts, and allow their skill set to develop on the job, and/or provide training themselves. In this sense, the interchangeability coefficient (i.e. the probability of being recognized by recruiters as “sufficiently qualified”)

A principal ideia do coeficiente de intercambialidade é que nem todos os empregos de uma profissão requerem 100% das competências, ou alternadamente, nem todos os empregadores requerem que você já esteja 100% pronto, permitindo portanto que desenvolva-se no trabalho, ou ainda que uma parcela das funções de determinado trabalho pode ser terceirizada a pessoas que não têm 100% das competências profissões requeridas, enquanto a parte mais importante do trabalho fica para as pessoas que têm essas competências. Basicamente, estamos dividindo o tempo entre pessoas.

Ou ainda que os empregos em si podem ser uma mistura desses perfis profissionais canônicos, de modo que as profissões mesmo não existam, apenas uma série de  $N$  empregos que requerem diferentes parcelas de cada profissão.

Having understood this, at last we can estimate the probability of finding a job as either  $q$  or  $\tilde{q}$ , adjusted by the similarity between these occupations:

$$\begin{aligned}\Lambda(\mathbf{a}_q, \mathbf{a}_{\tilde{q}}, n_q, n_{\tilde{q}}, N) &= \Pr(l = 1 \mid \mathbf{a}_q) \\ &= \Pr(q \approx q \mid \mathbf{a}_q, \mathbf{a}_q) \Pr(l = 1 \mid n_q, N) + \Pr(q \approx \tilde{q} \mid \mathbf{a}_q, \mathbf{a}_{\tilde{q}}) \Pr(l = 1 \mid n_{\tilde{q}}, N) \\ &= I(\mathbf{a}_q, \mathbf{a}_q) \Pr(l = 1 \mid n_q, N) + I(\mathbf{a}_q, \mathbf{a}_{\tilde{q}}) \Pr(l = 1 \mid n_{\tilde{q}}, N) \\ &= S(\mathbf{a}_q, \mathbf{a}_q) \Pr(l = 1 \mid n_q, N) + S(\mathbf{a}_q, \mathbf{a}_{\tilde{q}}) \Pr(l = 1 \mid n_{\tilde{q}}, N) \\ &= 1 \times \left(\frac{n_q}{N}\right) + S(\mathbf{a}_q, \mathbf{a}_{\tilde{q}}) \times \left(\frac{n_{\tilde{q}}}{N}\right) \\ &= \frac{n_q + S(\mathbf{a}_q, \mathbf{a}_{\tilde{q}})n_{\tilde{q}}}{N}\end{aligned}$$

Put simply, to the exact measure that occupations  $q$  and  $\tilde{q}$  are interchangeable, employment opportunities for either occupation should include the other: as their professional attributes overlap, so too their employment opportunities also overlap. In fact, we could say people with professional profiles  $\mathbf{a}_q$  and  $\mathbf{a}_{\tilde{q}}$  compete for the same  $n_q + n_{\tilde{q}}$  job posts.

In this sense, we notice, for example, that occupations which are completely dissimilar have no employment overlap, as

$$S(\mathbf{a}_q, \mathbf{a}_{\tilde{q}}) = 0 \implies \Lambda(\mathbf{a}_q, \mathbf{a}_{\tilde{q}}, n_q, n_{\tilde{q}}, N) = \frac{n_q + 0 \times n_{\tilde{q}}}{N} = \frac{n_q}{N},$$

and, on the other hand, that completely similar occupations have a complete overlap of employment opportunities as well:

$$S(\mathbf{a}_q, \mathbf{a}_{\tilde{q}}) = 1 \implies \Lambda(\mathbf{a}_q, \mathbf{a}_{\tilde{q}}, n_q, n_{\tilde{q}}, N) = \frac{n_q + 1 \times n_{\tilde{q}}}{N} = \frac{n_q + n_{\tilde{q}}}{N}$$

This said, it is important to note that this so-called “overlap” of employment opportunities in no way implies the total work force actually increases beyond  $N$ , since  $N$  is calculated by summing up the  $n_j$  *workers* employed in each occupation, not the total amount of *people* in the work force. That is, if a person chooses to work another job, they are counted as a separate worker in that particular occupation. Thus,  $N$  already accounts for people with multiple careers. This is important because, otherwise, we would need to adjust the denominator in this mathematical formulation, and in doing so we could fall into the first problem mentioned above (viz. coefficients of employability being too small).

The employability coefficient, therefore, only indicates which percentage of the  $N$  available jobs would be *suitable* to a given person, adjusting for their similarity with each occupation. It does not mean that said person in fact ends up working at the associated jobs, nor that the total amount of available job posts increases simply because workers could be employed elsewhere. Put another way, while employment opportunities do overlap, actual employment does not.

For example, Economists, Accountants, Executives, and Actuaries have a reasonably high professional similarity. And in practice, we do find these occupations competing for the same job posts. Engineers, and other STEM-related careers, likewise also have significant overlap in their employment opportunities. And the pattern repeats itself for every group of similar occupations. However, in any of these cases, no matter how many people are suitable for a given job post, only one of them is actually hired. So, despite the sum of overlapping employment opportunities being indeed much higher than available jobs, when each individual is finally employed, the total work force equals  $N$ .

Thus, at least so far, this metric seems intuitive and consistent. Therefore, we can write a provisional generalization for any number of occupations:

$$\begin{aligned}\Lambda(\mathbf{a}_u, \mathbf{A}, \mathbf{n}, N) &= \Pr(l = 1 \mid \mathbf{a}_u) \\ &= \sum_j \Pr(u \approx j \mid \mathbf{a}_u, \mathbf{a}_j) \Pr(l = 1 \mid n_j, N) \\ &= \frac{\sum_j I(\mathbf{a}_u, \mathbf{a}_j) n_j}{N} \\ &= \frac{\sum_j S(\mathbf{a}_u, \mathbf{a}_j) n_j}{N},\end{aligned}$$

where  $\mathbf{A}$  is the matrix of all  $\mathbf{a}_j$  professional profiles, and  $\mathbf{n} = (n_1, \dots, n_j)$ , the vector of job posts for the corresponding occupations, such that  $N = \sum_j n_j$ .

## 1.2 Additional Metrics

### 1.2.1 Revised Interchangeability Coefficients

Now, even though this adjusted metric is much more adequate than the previous one, it still has some minor flaws. For instance, it isn't reasonable to assume linear behavior in the professional interchangeability coefficient. For assuming such a relationship would imply that occupations with small similarity to other occupations, say in the 0 to 30% range, could be seen by the labor market as "still somewhat similar", and thus compete for a fraction of available job posts equal to their similarity coefficients; while, in practice, job applicants with such small similarities would be immediately dismissed by HR recruiters.

Thus, a scaling factor should be applied to the interchangeability metric, in order to reduce the probability of being recognized as "similar" to a given occupation, specially in the smaller ranges of professional similarity. Again, as  $S(\mathbf{a}_j, \mathbf{a}_k) \in [0, 1] \forall j, k$ , a simple way of implementing this adjustment would be:

$$\Pr(j \approx k \mid \mathbf{a}_j, \mathbf{a}_k) = I(\mathbf{a}_j, \mathbf{a}_k) = S(\mathbf{a}_j, \mathbf{a}_k)^\alpha, \quad \alpha > 1$$

as  $x^\alpha \leq x \forall \alpha > 1, x \in [0, 1]$ . However, this metric is also inadequate, as it can be "overly severe" in the higher ranges of similarity. A linear interpolation then is applied so as to smooth the transition between the adjusted equation above and the linear one already described:

$$\Pr(j \approx k \mid \mathbf{a}_j, \mathbf{a}_k) = I(\mathbf{a}_j, \mathbf{a}_k) = S(\mathbf{a}_j, \mathbf{a}_k)S(\mathbf{a}_j, \mathbf{a}_k) + [1 - S(\mathbf{a}_j, \mathbf{a}_k)]S(\mathbf{a}_j, \mathbf{a}_k)^\alpha, \quad \alpha > 1$$

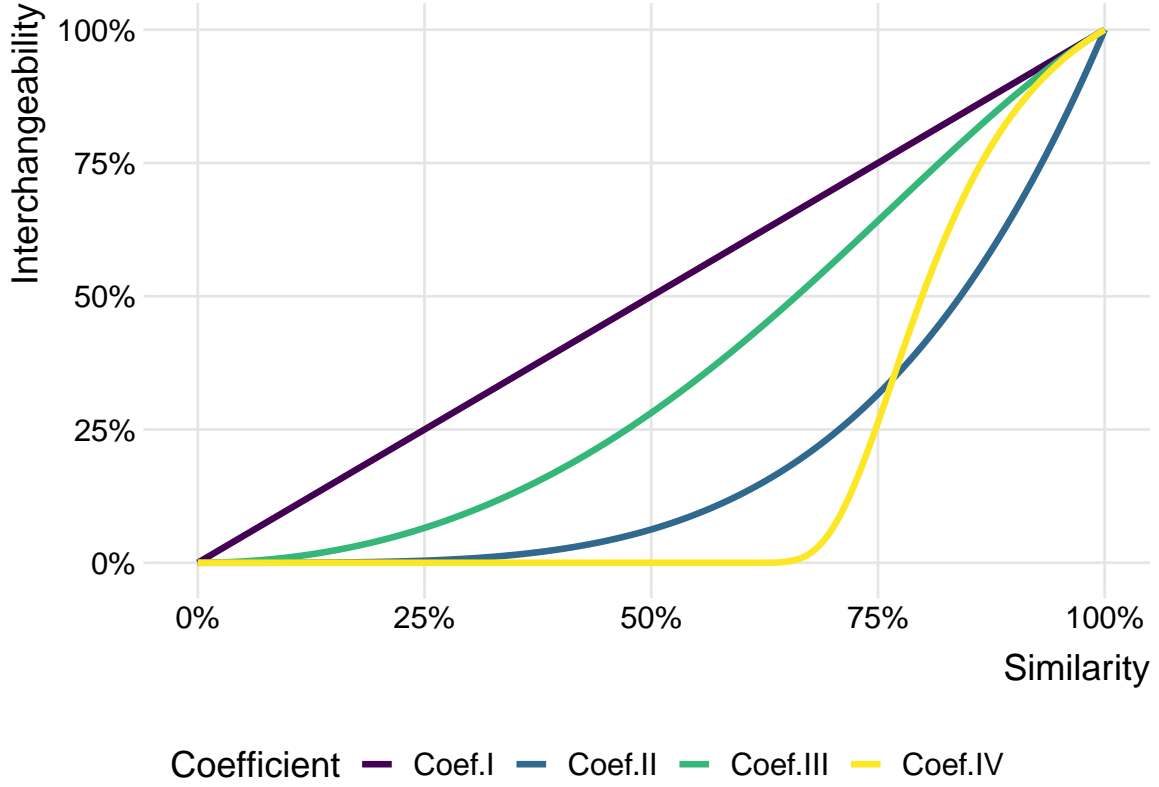
And an additional adjustment could provide an even more realistic behavior to the interchangeability function:

$$\Pr(j \approx k \mid \mathbf{a}_j, \mathbf{a}_k) = I(\mathbf{a}_j, \mathbf{a}_k) = S(\mathbf{a}_j, \mathbf{a}_k)S(\mathbf{a}_j, \mathbf{a}_k) + [1 - S(\mathbf{a}_j, \mathbf{a}_k)]S(\mathbf{a}_j, \mathbf{a}_k)^{\left[\frac{1}{S(\mathbf{a}_j, \mathbf{a}_k)}\right]^{\left[\frac{1}{S(\mathbf{a}_j, \mathbf{a}_k)}\right]}}$$

Here, the "severity" of the adjustment is entirely proportional to the similarity between occupations, and does not depend on arbitrarily set parameters. In either case, the more similarity approaches 1, the more interchangeability becomes linear.

Finally, these interchangeability metrics are illustrated in the diagram below:

Figure 1: Interchangeability Coefficients I–IV



From this figure, we immediately conclude that coefficients of interchangeability III and IV exhibit a very natural behavior, allowing for some overlap between occupations, without being “too severe” or “too lenient” in their compensation. Therefore, these mathematical formulations appear to adequately express what we referred to in the beginning of Section 1.2. as “sufficient qualification”. However, it is evident that coefficient IV is indeed the most realistic of them all.

### 1.2.2 Utility-Consistent Employability Coefficients

We can make the employability coefficient consistent with utility theory by applying the following adjustment:

$$\begin{aligned}
 \Lambda(\mathbf{a}_u, \mathbf{A}, \mathbf{n}, N, \mathbf{w}) &= \Pr(l = 1, U = 1 \mid \mathbf{a}_u) \\
 &= \sum_{k=1}^j U(\mathbf{a}_u, \mathbf{a}_k, w_u, w_k) \Pr(u \approx k \mid \mathbf{a}_u, \mathbf{a}_k) \Pr(l = 1 \mid n_k, N) \\
 &= \frac{\sum_{k=1}^j U(\mathbf{a}_u, \mathbf{a}_k, w_u, w_k) I(\mathbf{a}_u, \mathbf{a}_k) n_k}{N}, \\
 U(\mathbf{a}_u, \mathbf{a}_k, w_u, w_k) &= \begin{cases} 1, & \text{if } I(\mathbf{a}_u, \mathbf{a}_k) w_k > w_u. \\ 0, & \text{otherwise.} \end{cases},
 \end{aligned}$$

where the binary utility function  $U(\mathbf{a}_u, \mathbf{a}_k, w_u, w_k)$  indicates whether an economic agent would benefit from switching jobs, based on their current wage  $w_u$ , and the wages other occupations. We also assume here that their expect payment, if they choose to change jobs, is a  $I(\mathbf{a}_u, \mathbf{a}_k)$  fraction of market wages, so they should be paid in accordance with their interchangeability levels. This way, one can only attain the full salary of a given occupation, if considered 100% interchangeable with that occupation. We could interpret this mechanism as employer’s willingness to outsource a portion of a job’s activities that could be performed by lower-paid workers.

## 2 Estimating Employability Coefficients

We now estimate the employability metrics developed above in three instances: for occupations; for professional attributes; and for any professional profile. In order to estimate the employability of occupations we need only apply the  $\Lambda(\mathbf{a}_j, \mathbf{A}, \mathbf{n}, N)$  function for every  $\mathbf{a}_j$  vector of already registered career competencies, utilizing one of the four  $I(\mathbf{a}_j, \mathbf{a}_k)$  interchangeability coefficients illustrated in Figure 1. The estimation of employability coefficients for individual professional attributes involves carrying out a linear regression of some sort (either OLS or NNLS). And, provided this model is reliable, we can proceed to calculate employability coefficients for any arbitrary  $\mathbf{a}_u$  vector of professional attributes (i.e. a professional profile) by multiplying the model's parameters with individuals's attribute scores.

### 2.1 For Occupations

dsds

### 2.2 For Professional Attributes

P.S.: use  $\lambda_i$  for the employability of the  $i$ -th attribute, since  $\Lambda_u$  is the employability of professional profiles.

### 2.3 For Any Professional Profile

dsds

## 3 Professional Competitiveness

### 3.1 Pressupposed Metrics

From the employability metrics previously developed, arises the question of career competitiveness, that is: "What is the level of competition for job posts in a given occupation?" Now, this question is two-fold, for as already stated, people may compete for jobs with people from other fields (e.g. Accountants with Economists, Executives, and Actuaries, etc), again, provided the labor market recognizes these applicants are "sufficiently qualified" (i.e. being interchangeable, as described above).

If we think about this problem thoroughly, it is easy to see that while some job posts for an occupation are indeed too specialized to allow for *external* competition (i.e. applicants from other professional backgrounds), the other portion of jobs are suitable for any person with high enough professional similarity, as their overlapping competencies make them interchangeable to some degree, and therefore eligible to compete for a fraction of available job posts outside their fields of expertise.

Mathematically, we can say that a portion of job posts are subject only to *internal* competition, with applicants completely interchangeable with the "canonical" or even "platonic" representation of a particular occupation's professional profile  $\mathbf{a}_q$ , and the remaining jobs are subject to *external* competition with any  $k$  occupation to the  $I(S(\mathbf{a}_q, \mathbf{a}_k))$  degree of professional interchangeability. More precisely, the portion of these jobs (viz. those subject to external competition), is the maximum value of interchangeability with an occupation, since every job applicant said to be interchangeable with this occupation is eligible to compete for job posts below this maximum interchangeability level, above which only "canonical" (i.e. perfectly interchangeable) job applicants compete for jobs.

$$\begin{aligned} c_q^e(\mathbf{a}_q, \mathbf{a}_k) &= \max I(S(\mathbf{a}_q, \mathbf{a}_k)), \forall k \\ c_q^i(c_q^e) &= 1 - c_q^e \end{aligned}$$

These two equations describe, respectively, the total percentage of jobs subject to external competition, and internal competition. Again, as applicants from other professional backgrounds are only eligible to



the same “not too specialized”, or “non canonical”, portion of job posts for a given occupation, they can *at most* compete for the same  $c_q^e$  jobs equal to the maximum interchangeability level for that occupation.

Of course, this is an oversimplification, and we’ll develop a more sophisticated model at the end of this Section. It is evident, for example, that not every single applicant is in fact eligible to the full  $c_q^e$  portion of “non canonical” job posts, but rather a more restricted  $c_q^e(\mathbf{a}_q, \mathbf{a}_j) \leq c_q^e(\mathbf{a}_q, \mathbf{a}_k)$  portion. In other words, job applicants from other fields overlap with each other from the smallest degree of interchangeability to the highest  $c_q^e(\mathbf{a}_q, \mathbf{a}_k) = \max I(S(\mathbf{a}_q, \mathbf{a}_k))$ ,  $\forall k$ , so that inter-occupation competitiveness progressively decreases as interchangeability increases. Put simply, the more a particular job resembles the “canonical” representation of an occupation, the less it is subject to external competition. So, the more competitive jobs are the less specialized ones (i.e. those that could be performed by any untrained person).

Furthermore, professional competitiveness is also a function of the number of eligible applicants. Thus, different quantities of applicants compete for each of these  $c_q^e$  and  $c_q^t$  portions of a  $n_q$  subset of the labor market. That is, while jobs subject to internal competition can only be performed by  $c_q^t n_q$  people, the remaining  $c_q^e$  jobs can attract any number of applicants from other fields, so that each of these interchangeable occupations can bring up to  $I(\mathbf{a}_q, \mathbf{a}_k) n_k$  people to compete for those job posts.

However, whether we’re talking about internal competition or external competition, the competitiveness coefficient in each case is determined by an applicants-to-jobs ratio:

$$c_q(n_q, n_q^*) = \frac{n_q^*}{n_q},$$

where  $n_q^*$  represents the sum of people already working a particular job post with people willing to apply for those same jobs (viz. the economic agents for which occupation  $q$  is an optimal alternative compared to their current positions).

According to this formulation, then, the more people apply for a constant number of job posts, the more competitive that position is. Of course, this is a partly bounded coefficient, as  $c_q(n_q, n_q^*) \in [0, +\infty) \forall n_q, n_q^* \geq 0$ . It is easy to see, for example, that competitiveness reaches its minimum when there is not a single applicant for a particular job post:

$$n_q^* = 0 \implies c_q(n_q, n_q^*) = \frac{n_q^*}{n_q} = \frac{0}{n_q} = 0$$

In this situation, any eligible applicant would be immediately hired. And, conversely, the competitiveness coefficient approaches infinity if the number of available jobs is zero:

$$n_q = 0 \implies c_q(n_q, n_q^*) = \frac{n_q^*}{n_q} = \frac{n_q^*}{0} = +\infty$$

Now, the problem of professional competitiveness becomes more complicated when we consider utility-consistent metrics. In this model, the number of eligible applicants is determined not only by interchangeability of professional profiles with a particular “canonical” occupation, but also by wages, and economic agents’s utility functions. Here we define the simplest utility-consistent model:

$$U(\mathbf{a}_q, \mathbf{a}_k, w_q, w_k) = \frac{I(\mathbf{a}_q, \mathbf{a}_k) w_k}{w_q}$$

In the model above, an economic agent compares the expected outcomes of working their current job  $q$  or switching over to job  $k$  to earn a fraction of the  $w_k$  wage proportional to their  $I(\mathbf{a}_q, \mathbf{a}_k)$  interchangeability with that occupation. This utility function, then, is evaluated  $\forall j$  available job posts, and the agent chooses to either stay on their current jobs, if no benefit is to be gained by switching jobs, or applies to the position which yields the highest utility.

### 3.2 Job Competitiveness in a Simulated Economy

These metrics become much more interesting in a simulated strategic economic context, where each agent can freely choose their preferred jobs based on a utility function, and therefore influence other agents's optimal choices, as well as the whole economy.

The motivation for such a model derives from the problems of endogeneity implied by the competitiveness metrics above. It is easy to see, for instance, there is no analytical solution to calculate equilibrium utility-consistent job applications. Thus, in order to estimate the most adequate model, we must iterate through a number of successive games of job trading, where each applicant either stays on their current position or applies to the job yielding the highest expect utility. Then, assuming no information asymmetry, as well as constant demand for each job, salaries are adjusted in order to keep industry values at the original levels, so that each new applicant reduces the wages of all workers in the field through supply and demand pressures. Finally, at the end of this economic strategic game, we measure the total  $n_k^*$  eligible workers willing to apply to each of  $k \in [1, j]$  occupations, and calculate the resulting  $c(n_k, n_k^*)$  competitiveness coefficients.

This model has the advantage of simulating the strategic behavior of rational economic agents and also of providing more precise competitiveness metrics, compared to the simplest models written above. It can be estimated through the following algorithm:

1. Get all the  $\mathbf{a}_k$  unique professional profiles in each occupation.
2. Calculate  $\max[U(a_k, a_j)] \forall k, j \in [1, j]$ .
3. If the highest utility-yielding job is not the current job, then switch jobs. Else, maintain the current position.
4. Edit the original vs current occupations record to keep track of economic agents's positions.
5. If agents switch jobs, add 1 to their job history counter to keep track of how many times they switched jobs.
6. Adjust employment levels:  $n_k^{t+1} = n_k^t + \Delta(n_k^t)$ , where  $\Delta(n_k^t)$  is the net flow of workers in a given occupation at time  $t$ .
7. Adjust wages:  $w_k^{t+1} = \frac{n_k^t w_k^t}{n_k^{t+1}}$ , in order to keep industry value constant.
8. Repeat steps (1)–(7) until no economic agent changes positions or the algorithm reaches an arbitrary maximum number of iterations.
9. Calculate the final competitiveness coefficient  $c(n_k^t, n_k^T) = \frac{n_k^T}{n_k^t} \forall k \in [1, j]$ , where  $t$  is the initial period of measurement, and  $T$  the last period.