Faculty Hiring Network Reveals Possible Decision-Making Mechanism

Sunjing Zheng 1, Nan Jiang 1,2, Xiaomeng Li 1, Mingzhong Xiao 3,* and Qinghua Chen 1,*

1 School of Systems Science, Beijing Normal University, Beijing 100875, China;
sunjingzheng@mail.bnu.edu.cn (S.Z.); nanjiang@mail.bnu.edu.cn (N.J.); lixiaomeng@bnu.edu.cn (X.L.)
2 School of Science, Wuhan University of Science and Technology, Wuhan 430081, China
3 School of Artificial Intelligence, Beijing Normal University, Beijing 100875, China
* Correspondence: xmz@bnu.edu.cn (M.X.); qinghuachen@bnu.edu.cn (Q.C.);
Tel.: +86-10-58807943 (M.X.); +86-10-58807085 (Q.C.)

Abstract: Social physics (or sociophysics) offers new research perspectives for addressing social issues in various domains. In this study, we explore the decision-making process of doctoral graduates during their transition from graduation to employment, drawing on the ideas of sociophysics. We divide the process into two decision steps and propose a generative model based on appropriate assumptions. This model effectively reproduces empirical data, allowing us to derive essential parameters that influence the decision-making process from empirical observations. Through a comparison of the best-fit parameters, we discover that doctoral graduates in business disciplines tend to exhibit more concentrated employment choices, while those in computer science and history disciplines demonstrate a greater diversity of options. Furthermore, we observe that universities consider factors beyond rankings when selecting doctoral graduates.

Keywords: sociophysics; faculty hiring network; decision-making behavior; generative model

1. Introduction

The exploration of the laws governing social dynamics is a complex and diverse long-term process for humanity, involving numerous disciplines and fields. Contributions to the study of social dynamics have been made by various disciplines, including economics, political science, sociology, psychology, physics, and other disciplines. Despite many disciplines attempting to understand the dynamics of social systems, progress in understanding society has been slow for a considerable period. The main reason for this is the presence of human factors within social systems. The decision-making behavior of individuals and their interactions with others contribute to the complexity of social systems. In situations where traditional data are difficult to obtain, many analyses remain qualitative in nature [1,2].

There is a new idea utilizing statistical physics that considers individuals as entities and accounts for human behavior to understand the underlying mechanisms of human social behavior [3,4]. Human activities are a simple extension of the “natural order”, and applying natural laws can uncover the intrinsic rules followed by human activities. Building upon the groundbreaking contributions of scientists such as Auguste Comte, George Kingsley Zipf, and John Quincy Stewart, Serge Galam introduced the concept of “sociophysics” [5]. Galam proposed a theoretical framework called “social models” to explain the formation and evolution of social phenomena and collective behavior [6]. He used physics to study hierarchical systems, voting in group decisions, stability and fragmentation of alliances between nations, minority opinion propagation, and rumor phenomena. Over the past few years, sociophysics has evolved into a multidisciplinary field. It provides a relatively objective and normative epistemological path and principles in studying social science issues such as political science, economics, and communication, revealing the structure
and behavioral patterns of social systems at various scales (from individuals to the societal level) and capturing the regularities of long-term societal evolution \[7,8\].

The complexity of social systems arises from the decision-making behaviors of individuals and the interactions among people. Understanding the underlying mechanisms of human decision-making behavior becomes a powerful tool for expanding thinking and guiding practice. We focus on the social issue of employment decisions for doctoral graduates. Scientific discovery requires the capacity to seek, nurture, and combine internal and external sources of knowledge. Universities, in particular, serve as vital “containers” for the advancement and integration of this knowledge \[9\], the employment of doctoral graduates as faculty members is a crucial aspect of this continuous cycle. Faculty hiring decisions determine the composition of the academic workforce and directly impact educational outcomes, career trajectories, and the development and dissemination of ideas \[10\].

Based on principles from sociophysics, we propose a generative model to simulate the decision-making process of doctoral graduates’ employment in higher education institutions. We investigate this decision-making process by dividing it into two stages: the application process, where graduates submit resumes to universities, and the recruitment process, where universities consider and hire applicants. Since individual behaviors are mostly determined by their local contexts and internal psychology \[11\], we analyze the model’s parameters to understand the psychological characteristics of doctoral graduates’ job-seeking and universities’ recruitment processes. The remaining sections of this paper are organized as follows. In Section 2, we first introduce the data source and distribution characteristics, followed by the construction of the generative model. In Section 3, we estimate and compare the parameters of the generative model. Finally, we summarize and discuss the paper in Section 4.

2. Materials and Methods

2.1. Data Sources

In this paper, we use the university faculty hiring data \[12\] of PhD graduates collected in Ref. \[13\] covering the three disciplines of computer science, business, and history. The relevant data information of various disciplines is shown in Table 1, including hundreds of universities in each discipline and the flow of tenure or tenure-track faculty between them, as well as the number of universities ranked in U.S. News & World Report and USA National Research Council (NRC) for each discipline. In detail, each tenure-track faculty was composed of a pair of graduating colleges and in-service colleges in the U.S. universities or institutions.

<table>
<thead>
<tr>
<th></th>
<th>Computer Science</th>
<th>Business</th>
<th>History</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of institutions</td>
<td>205</td>
<td>112</td>
<td>144</td>
</tr>
<tr>
<td>No. of regular faculty</td>
<td>5032</td>
<td>9336</td>
<td>4556</td>
</tr>
<tr>
<td>Mean size of faculty</td>
<td>25</td>
<td>83</td>
<td>32</td>
</tr>
<tr>
<td>No. of ranking university by U.S. News &amp; World Report</td>
<td>153</td>
<td>111</td>
<td>143</td>
</tr>
<tr>
<td>No. of ranking university by USA National Research Council (NRC)</td>
<td>103</td>
<td>-</td>
<td>122</td>
</tr>
</tbody>
</table>

2.2. Empirical Evidence of Asymmetric Distribution Jump

To obtain a distribution of ranking differences, one first needs a university ranking. Unfortunately, a universally recognized objective and fair ranking does not currently exist. In the present study, we use rankings from some authoritative organizations for the following reasons: Authoritative organizations invest a significant amount of funds and resources in ranking processes to ensure the accuracy and objectivity of the results \[14\]. The rankings generated by these organizations are widely trusted and used by students and
individuals in the education field. Students often rely on these rankings to make decisions regarding their choice of universities, highlighting the influential role that authoritative rankings play in the decision-making process [15]. In this study, we adopted the rankings provided by the U.S. News & World Report and NRC as the reference rankings. In addition, these two organizations’ rankings are considered relatively objective, and the Pearson correlation coefficient between the two rankings exceeds 0.8.

Based on the collected authoritative rankings, we analyze the distribution characteristics of ranking differences in the empirical data. \( u \) represents the authoritative ranking of the university where the doctoral graduate obtained their doctoral degree, and \( v \) represents the authoritative ranking of the university where the doctoral graduate is employed. \( r = v - u \) indicates the difference in rankings between the university where the doctoral graduate is employed and the university from which they obtained their doctoral degree. It is worth noting that smaller values of \( u \) and \( v \) correspond to higher rankings, indicating that the university is better, while larger values correspond to lower rankings, indicating that the university is worse. The function \( p(r) \) represents the probability of a rank difference of \( r \).

As shown in Figure 1a–e, the distribution of ranking differences exhibits consistent patterns across different disciplines and authoritative rankings. The ranking differences are concentrated around 0, indicating that doctoral graduates tend to choose institutions for employment that are close in ranking to their own university. This trend is further highlighted by all the data exhibiting a prominent peak at \( r = 0 \), with 60% of doctoral graduates in the history field, 77% in the business field, and 44% in the computer science field being attributed to them remaining at their alma mater (the university they graduated from) for employment. The distribution also displays asymmetry, with lower frequencies on the left side and higher frequencies on the right side. Positive ranking differences on the right side indicate that doctoral graduates from higher-ranked universities tend to work in lower-ranked universities, while negative ranking differences on the left side represent the opposite situation.

Figure 1f reveals that over 60% of graduates go to universities that are lower in ranking than their alma mater \( (r > 0) \), while only a small number of graduates are able to secure employment in universities that are higher in ranking \( (r < 0) \). This empirical evidence is not an isolated finding; similar characteristics have been observed in Ref. [10] across the entire academic community and in eight different domains. The Minimum Violation Rankings (MVR) model developed in Ref. [13] is based on the core assumption that doctoral graduates tend to move more easily from higher-ranked universities to lower-ranked universities, while the reverse movement is less likely. The model introduces the concept of “reverse arcs”, which are instances where doctoral graduates move from lower-ranked universities to higher-ranked universities for employment, contrary to the general trend. The goal of the MVR model is to minimize the number of these “reverse arcs” in a stable doctoral graduate transition-to-employment network, thereby providing rankings of university prestige that align with the observed data. We utilize this characteristic as the basis for our subsequent modeling.

2.3. Generative Model Based on the Process of Doctoral Employment

We decompose the faculty recruitment process into two consecutive processes: graduates selecting their employer and schools selecting candidates. By setting the selection probabilities for these two processes, we obtain the final joint probability density distribution and generate flow data accordingly.

2.3.1. Two Decision-Making Processes in the Graduation-Employment Process

By observing the distribution of empirical data, one can see that the ranking differences of doctoral graduates’ employment institutions exhibit asymmetric distribution characteristics. Upon closer analysis, we determined that the mobility of doctoral graduates is not solely determined by one party, but rather the result of a mutual decision-making process involving both parties. As shown in Figure 2, a transition from \( u \) to \( v \) indicates the
following: First, a doctoral graduate who graduated from university $u$ has applied for a job to university $v$, and subsequently, university $v$ has approved the student's job application.

Figure 1. Asymmetry in ranking differences among disciplines in authoritative rankings. $r = v - u$ indicates the difference in rankings between the university where the doctoral graduate is employed and the university from which they obtained their doctoral degree. (a–c) The distribution of ranking differences between doctoral graduates’ working institutions and their alma maters in the fields of history, computer, and business, respectively, based on the authoritative ranking by U.S. News & World Report. (d,e) The distribution of ranking differences in the fields of history and computer science, respectively, based on the authoritative ranking by the USA National Research Council (NRC). (f) The proportion of hierarchy transitions in real data across different disciplines under various authoritative rankings. The data is obtained from Ref. [12].
Figure 2. The process of a PhD graduate flowing from $u$ to $v$ can be divided into two subprocesses.

When doctoral graduates choose their place of employment, they pay attention to the differences between their alma mater and the prospective employer university. Due to the broad impact of prestige, better universities have access to more resources, are more likely to gain recognition in the academic community, and often have higher paper acceptance rates [10]. As a result, doctoral graduates tend to prefer employment opportunities at these higher-ranked universities, that is, universities with smaller values of $v$.

However, doctoral graduates do not rashly apply to top-tier universities considering the costs involved in the application process. When deciding which university to apply to, the graduates speculate on the response they might receive. The graduates are aware that universities generally prefer candidates from higher-ranked institutions. If there is a significant disparity between their alma mater and the target university, there is a higher likelihood of rejection. In light of this, graduates tend to take a conservative approach and apply to universities similar in ranking to their own. The graduates are unwilling to apply to lower-ranked institutions, while they recognize the difficulty of securing a position at higher-ranked institutions. Here, we assume that the ranking difference, $r$, between the graduate’s employment institution and their alma mater follows a normal distribution, $N(0, \sigma^2)$, with a probability density given by

$$p_1(r) \propto \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{r^2}{2\sigma^2}}.$$  \hspace{1cm} (1)

To facilitate our analysis, we set the selection standard deviation for all students as $\sigma$. Notably, as the selection standard deviation increases, the perceived differences between universities diminish in the eyes of doctoral graduates, and their preferences for specific universities become less pronounced. Conversely, when $\sigma$ is relatively small, PhD graduates exhibit clear preferences and tend to select universities with rankings similar to their alma mater.

From the perspective of universities, it is indeed true that employers believe that universities with higher prestige have access to better resources to nurture their students. Moreover, doctoral graduates from more prestigious universities are perceived to possess a higher quality. Therefore, universities naturally tend to prioritize applications from students coming from more prestigious institutions. This preference for prestigious universities reinforces their central role in disseminating ideas, shaping academic norms,
and influencing broader culture [10]. Therefore, we assume that the selection of doctoral graduates by universities follows an “S”-shaped function, which can be specified as

\[ p_2(r) \propto \frac{1}{1 + e^{-1/2 \Delta T}}. \] (2)

Similarly, \( r \) represents the rank difference between the university with recruitment needs and the graduate’s graduation university. \( \Delta \) is an adjustable displacement parameter, and \( T \) represents the degree of irrationality. \( \Delta \) and \( T \) are used to adjust the degree of influence of the rank difference on the probability of selection. When \( r + \Delta > 0 \), the university considers the student to be from a better university and is more willing to offer to them, resulting in a higher selection probability. Conversely, when \( r + \Delta < 0 \), the university perceives the student to be from a lower-ranked university, leading to a lower selection probability. As soon as \( T \) represents the degree of irrationality, a smaller value of \( T \) indicates greater rationality of the universities. In this case, universities are almost unwilling to select students from lower-ranked universities. Conversely, a larger value of \( T \) suggests lower rationality of the universities. Although universities still tend to hire doctoral graduates from higher-ranked universities, they also consider exceptional doctoral graduates from slightly lower-ranked universities.

2.3.2. Simulation Process of the Generative Model

We decompose the faculty recruitment process into two consecutive processes: student selection of schools for job applications and school selection of whether to give an offer to the student. After both students and universities have gained knowledge of their own and other institutions and have transformed this knowledge into selection probabilities, we generate flow data through a generative mechanism. As shown in Figure 3, the specific generation process is as follows:

(1) We assume there are \( M \) schools ranked from 1 to \( M \) based on their prestige, with lower ranking values indicating better schools. We generate \( N \) graduates for each school.

(2) We randomly select a student who chooses a school to submit their resume using a roulette wheel selection method based on the probability \( p_1(r) \), where the standard deviation is a tunable parameter.

(3) After receiving the student’s resume, the school decides whether to give an offer to the student based on the probability \( p_2(r) \).

(4) If the university extends an offer, we record the ranking of the student’s alma mater \( u \) and the ranking of the university that extended the offer \( v \). We repeat steps (2) and (3) until all students have made their choices.

The simulated data obtained from the generative model are presented in Section 3.

2.3.3. Theoretical Analysis of the Model

The two decision-making behaviors in the above simulation process come from different entities, assuming that these two decision processes are independent of each other. When there are large values of \( M \) and \( N \), one can use mean-field analysis to obtain an expected value of the simulation result, denoted as \( p(r) \):

\[ p(r) = \frac{p_1(r) \cdot p_2(r)}{\sum p_1(r) \cdot p_2(r)} = \frac{\frac{1}{\sqrt{2\pi e^2}} \cdot \frac{1}{1 + e^{-1/2 \Delta T}}}{\sum \left( \frac{1}{\sqrt{2\pi e^2}} \cdot \frac{1}{1 + e^{-1/2 \Delta T}} \right)}. \] (3)

When \( \Delta = 0 \), based on Equation (3) one can see that as \( r \) approaches zero, the joint probability, \( p(r) \), increases, indicating that doctoral graduates are more likely to enter schools with rankings similar to their alma mater. Conversely, as \( r \) deviates from zero, \( p(r) \) decreases, suggesting that the probability of doctoral graduates entering schools with a large ranking difference from their alma mater decreases. The distribution of \( p(r) \) is
asymmetric and exhibits a characteristic of being high in the middle and low on both sides. The specific impact of the parameters on $p(r)$ are discussed in Section 3.

![Flowchart of the generation model](Image)

**Figure 3.** Flowchart of the generation model.

### 3. Results

#### 3.1. Impact of Parameters on the Generative Model

In this Section, we discuss the effects of three parameters in the model by examining the changes in the joint probability distribution. For the purpose of comparison, $p_1(r)$, $p_2(r)$, and $p(r)$ are plotted together to observe the relationship between the probabilities of individual decisions and the overall process, as shown in Figure 4.

Here, $\sigma$ represents the concentration level in student choices. When $\sigma$ is small, students tend to prefer schools with rankings close to their own alma mater. When $\sigma$ is large, students tend to consider rankings more extensively when selecting the schools to which they want to apply. In Figure 4b, schools dominate the entire recruitment process.
The parameter $\Delta$ represents the tolerance level of schools when making offers, which describes the degree to which schools are willing to give offers to doctoral graduates from lower-ranked institutions when selecting students. Smaller values of $\Delta$ result in a right-skewed probability density of school choices, leading to a right-skewed joint probability density. As shown in Figure 4c, most students end up in schools with lower rankings than that of their own alma mater. Conversely, when $\Delta$ is large, schools have a higher tolerance for doctoral graduates from lower-ranked institutions, as depicted in Figure 4d, and almost give offers to all students. In this case, students become the dominant factor in the entire recruitment process.

Figure 4. The impact of parameters $\sigma$ (a,b), $\Delta$ (c,d), and $T$ (e,f) on $p_1(r)$, $p_2(r)$, and $p(r)$. See text for details.

The parameter $T$ reflects the degree of importance that universities place on school rankings when making job offers to applicants. As illustrated in Figure 4e, when the irrationality level, $T$, is small, this indicates a higher level of rationality, with schools attaching more importance to rankings and rarely giving offer to students from lower-ranked institutions. Once the rank difference becomes negative, the selection probability experiences a sharp decline, and schools tend to choose students from higher-ranked alma maters. On the other hand, when the irrationality level $T$ is large, as shown in Figure 4f, schools prioritize other factors beyond rankings, such as students’ comprehensive qualities and actual abilities. Importantly, the term “irrationality” here does not imply...
that schools’ decisions are irrational. Rather, it suggests that schools are more likely to consider nonranking factors when making decisions, such as students’ overall qualities and practical abilities.

3.2. Parameter Estimation for Empirical Data

We identify the optimal fit parameters for \( p(r) \) by adjusting the parameters to minimize the Kolmogorov–Smirnov (KS) statistic between the actual ranking difference distribution and \( p(r) \). Then, we use these best-fitting parameters to generate simulation data by using the generative model.

Figure 5 illustrates the actual ranking difference distribution, the joint probability distribution, and the distribution of ranking differences in the simulation data. The blue solid line represents the joint probability density distribution curve, while the red circle represents simulation data generated based on the joint probability density. The joint probability density curve captures the similar distribution characteristics of the actual data rank differences: high probability in the middle and low probability on both sides, with a faster decrease in probability on the left and a slower decrease on the right.

Table 2 displays the optimal fit parameters for the three domains across various ranking authorities. By comparing the best-fit parameters obtained from two authoritative rankings in each domain, we observe a consistent pattern. These parameters can to some extent reflect the decision-making process of students in job seeking and schools in offering admissions.

As shown in Table 2, from the perspective of students, the standard deviation, \( \sigma \), of the selection distribution is relatively consistent for computer science and history majors, and they are noticeably larger compared to that for business. This indicates that doctoral graduates in the field of business tend to have more concentrated employment choices. In other words, doctoral graduates in the business field prefer to choose universities with smaller ranking differences from their alma mater, while the choices of doctoral graduates in computer science and history are more diverse.

From the perspective of universities, the offset values, \( \Delta \), in all three fields are positive. A positive \( \Delta \) indicates that universities have a certain level of tolerance for graduates from slightly lower-ranked institutions. This means that universities consider the actual abilities of doctoral graduates when selecting them, rather than solely focusing on the ranking of their alma mater. The appropriate value of \( T \) also suggests that universities consider factors other than just rankings.

<table>
<thead>
<tr>
<th>Authoritative Ranking</th>
<th>Parameter</th>
<th>Computer Science</th>
<th>Business</th>
<th>History</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S. News &amp; World Report</td>
<td>( \sigma )</td>
<td>71</td>
<td>45</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>( \Delta )</td>
<td>12</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>( T )</td>
<td>35</td>
<td>51</td>
<td>31</td>
</tr>
<tr>
<td>NRC</td>
<td>( \sigma )</td>
<td>58</td>
<td>-</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>( \Delta )</td>
<td>18</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>( T )</td>
<td>13</td>
<td>-</td>
<td>41</td>
</tr>
</tbody>
</table>
Figure 5. Comparison between the generative model and real data: (a–c) the fields of history, computer, and business, when ranked by U.S. News & World Report; (d,e) the fields of history and computer, when ranked by NRC; and (f) the proportion of hierarchy transitions of the simulation data across the various disciplines under different authoritative rankings.

4. Discussion

This study is an attempt to analyze the decision-making process of doctoral graduates in job searching based on the perspective of sociophysics. The process of graduates transitioning from graduation to employment is divided into two steps, and a generative model is proposed based on appropriate assumptions. This model effectively reproduces empirical data and allows for the estimation of important parameters that influence the decision-making process from the empirical data. Among these parameters, $\sigma$ is related to the diversity of students’ resume submissions, $\Delta$ reflects the tolerance of universities toward the ranking of graduates’ alma mater, and $T$ represents the degree of irrationality in university decision-making. By comparing the best-fitting parameters, we found that doctoral graduates in business fields have more concentrated employment choices, while graduates in computer science and history fields have broader preferences. Furthermore, universities consider factors beyond rankings when selecting doctoral graduates.

Understanding the mechanisms behind human decision-making can broaden our thinking and can guide practical applications. This research provides an initial exploration by proposing a generative model, aiming to inspire further research on decision-making problems in other social systems. Furthermore, with the extended application of this model,
it is possible for us to directly obtain the public perception of the excellence levels of various universities from graduation-to-employment data, even in the absence of authoritative ranking information. This approach can also be applied to other domains, such as national residency rates or corporate competitiveness.

The analysis presented here, included faculty members from different ranks, which indeed may have some influence on our conclusions. However, we strongly believe that the choice of doctoral institution reflects graduates’ academic achievements and capabilities, playing a significant role in determining their subsequent employment destinations. This belief is further validated by the study in Ref. [13], where the MVR model predicts university rankings and demonstrates a strong correlation with authoritative rankings, highlighting the significance of doctoral graduates’ backgrounds in shaping their career paths. Further detailed discussions on this topic remain open for researchers to explore.

Author Contributions: Conceptualization, S.Z.; Methodology, X.L.; Writing—original draft, S.Z.; Writing—review and editing, M.X.; Visualization, N.J.; Supervision, Q.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the National Natural Science Foundation of China under Grant 72274020.

Data Availability Statement: The data used in this study can be obtained from the website http://santafe.edu/~aaronc/facultyhiring/ (accessed on 31 July 2023).

Conflicts of Interest: The authors declare no conflicts of interest.

References

2. Weiwei, L. The integration trend and its characteristics of social physics and big data technology. *J. Dialect. Nat.* 2019, 41, 80–86. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.