Why is it so hard to find a job now? Enter Ghost Jobs*

Hunter Ng
Baruch College, City University of New York
hunterboonhian.ng@baruch.cuny.edu

24th May 2025

Abstract

This paper investigates "ghost jobs", which are vacancies posted without intent to hire, using a novel dataset of interview reviews from Glassdoor. Using a fine-tuned BERT model, I find that approximately 21% of job postings exhibit patterns consistent with ghost jobs. These are disproportionately concentrated in larger firms and high-skill industries, where firms may benefit from resume collection, market intelligence, or signaling. I also show that incorporating ghost job prevalence helps reconcile the recent disconnect in the Beveridge Curve between vacancy and hiring rates. The results highlight how ghost hiring imposes costs on job seekers, distorts labor market indicators, and warrants closer scrutiny from policymakers.

Keywords: Ghost jobs; Labor market dynamics; Job search fatigue; Glassdoor; ChatGPT; BERT; Beveridge Curve

JEL Classifications: J60, J23, J31, M51, C81, L22, D83, J45, J70, J11.

*I am grateful to Kevin Shih, Ruth Milkman, Sandra Sequeira and two anonymous referees for helpful comments. I am also grateful for helpful comments from seminar participants at Baruch College.

1. Introduction

Workers who apply for jobs assume that employers are looking to fill the jobs. That is not the case. Popular media has documented a new phenomenon of "ghost jobs"[1,2] (Paradis, 2024; Woods and Wong, 2024). The term ghost jobs is not formally defined[3,4], but instead, reflects a loosely termed phenomenon that is widely observed but not empirically analyzed. In this paper, I attempt to characterize this trend using a novel methodology.

To more properly define what a ghost job is, I categorize job postings into three types in Table 1. The key differentiating factor is the employer's intent. Type 1 refers to genuine openings, where there is a clear and immediate need to hire. Type 2 refers to pipeline postings, where the company intends to hire eventually but not right away, and often to build a pool of candidates for future openings. Type 3 refers to ghost jobs, where there is no real plan to hire in the foreseeable future. These may be posted for strategic reasons, such as collecting market information, signaling company growth, maintaining HR activity, or fulfilling compliance requirements. Employers often disguise Type 3 ghost postings as Type 2 pipeline roles, making it hard to distinguish them. Because both pipeline and ghost postings lack immediate hiring plans, employers often pool them, but only ghost jobs lack genuine intent to hire, making intent the critical but hidden differentiating factor. The ideal way to study this would be through internal company data such as management reports

[1] In an interview done for the Connecticut Public Radio, the authors interviewed Lisa Simon, chief economist of Revelio Labs, a jobs analytics firm and she gave statistics that in 2018, 1 in 5 job postings would not result in a hire but in 2024, 1 in 2 job postings do not result in a hire. They cited an example Allison Giddens, owner of a company called WinTech, which makes parts for planes and the military. Giddens noted that she kept an ad all year round because it's hard to find the highly skilled workers she needs.

[2] A ghost job is distinct from ghost work, a term popularized by Gray and Suri (2019) which refers to hidden labor behind digital tasks and which if often exploited due to a lack of formal employment status.

[3] Bienasz (2022) does not explicitly define the term but that it is a job opening that employers 'don't intend to hire for, much to the chagrin of job seekers'

[4] Wikipedia defines a ghost job as a position that is non-existent or has already been filled, though this is not a formal or textbook definition. I expand on the qualities of a ghost job in the paper.

on hiring plans and position performance but such data is sensitive and rarely available. Instead, I use an alternative approach where I analyze interview reviews from job-seekers.

I first hand-collect a dataset of job-seeker interview reviews from Glassdoor, a widely used platform where applicants report their experiences with employers. Glassdoor is an especially useful source because it captures firsthand accounts of the hiring process across thousands of companies. Nearly 70% of unemployed Americans use the internet to search for jobs (Socking and Sojourner, 2021), and Glassdoor has been cited as a credible information source in prior labor market studies (Kim and Ra, 2022; Kyiu et al., 2023; Symitsi et al., 2018).

I find that up to 21% of job postings could be classified as ghost jobs using a deep-learning, BERT-based classification model trained on interview review data. This suggests that a significant portion of the online job market may not reflect true hiring demand. In line with my hypotheses, I find that larger firms and those with dedicated HR departments are more likely to post ghost jobs. This is likely because they have more resources to maintain postings and benefit more from collecting applicant data. Interestingly, the highest concentration of ghost jobs occurs not at the largest firms, but among mid-sized companies, which may resort to strategic reasons to post ghost jobs, while lacking the scale to fill these roles rapidly. I also find that specialized and high-skill jobs, such as those in software, publishing, or technical consulting, are more likely to be flagged as ghost listings. While some of these may reflect genuine pipeline postings rather than ghost jobs, this finding also show the difficulty of distinguishing between legitimate hiring delays and ghost jobs.

To test whether ghost job signals in high-skill sectors simply reflect slow but genuine pipeline hiring, I compare reviews flagged by the BERT model across high- and low-skill industries. Even after accounting for differences in hiring timelines, high-skill reviews continue to show markers of disingenuous processes, such as vague interviews and lack of follow-up, unlike their low-skill counterparts. This supports the interpretation that these are not merely pipeline postings. Since employer intent is unobservable in vacancy data, interview reviews offer a useful alternative for identifying ghost jobs, especially in sectors where traditional methods fall short.

These results are robust to firm-level and year fixed effects, and they hold across a wide range of industries. By aggregating these micro-level patterns, I document a broader trend: the prevalence of ghost jobs has grown over time, coinciding with declining costs of job ad placement and greater strategic use of online postings. On a macroeconomic level, to test whether ghost jobs help explain puzzling labor market patterns, I integrate my estimates with JOLTS data and adjust the job openings rate accordingly. This adjustment helps reconcile the disconnect in the Beveridge curve highlighted by Mongey and Horwich (2023)[5], where official vacancy rates appear elevated even as hiring and quitting slow down. Figure 1 shows this increasing disconnect in the Beveridge curve in the last fifteen years. My findings suggest that ghost jobs may be a plausible explanation for this disconnect.

Why is it important to study ghost jobs? I offer two main reasons. Firstly, job-seekers bear the cost of applying, such as writing cover letters, preparing for interviews, and enduring long wait times. Over time, repeated experiences with non-responsive or misleading job ads can lead to job-search fatigue, which may worsen long-term unemployment (Lim et al., 2013). Secondly, widespread ghost postings can distort labor market signals. If many firms in an industry post roles they don't plan to fill, it creates a false impression of sectoral growth, which can lead to the misallocation of training, policy attention, and job-seeker effort (Goldfarb et al., 2007).

Why would firms post ghost jobs and continue to interview for them? I provide several strategic reasons in Table 1. One is that internal HR departments may want to appear active and essential. In organizations where headcount or budget depends on recruitment activity, keeping job postings live can help justify their role. This has been referred to as "productivity theater," where departments simulate busyness to show legitimacy (Sprague, 2023; Marchese, 2023). Alternatively, some firms post ghost jobs for informational or strategic advantage. By collecting resumes and gauging

[5]Economists Simon Mongey and Jeff Horwich provided a thought-provoking article in December, 2023 where they warned of a soft-landing using JOLTS data, contrary to what prominent economists such as Larry Summers and Olivier Blanchard have posited. They find that vacancy rates are not matching the other stock flow rates such as quitting rates and hiring rates, and they do not take a stand on the issue. However, they acknowledge that factors such as changes to'cost to employers of job posting, recruiting and evaluating candidates'

expected salary expectations, employers can benchmark labor market trends, keep internal wage growth in check, and maintain negotiating leverage over current staff. This type of opportunistic behavior is well documented in the corporate strategy and management literature (Chalmers et al., 2002; Nienhaus, 2022).

Another legitimate motive is operational risk management, particularly for highly specialized roles with long hiring lead times. In these cases, employers may post job ads even when there is no immediate vacancy, simply to build a pipeline of potential candidates in case of sudden turnover. This behavior comes under Type 2 pipeline postings, and is not inherently deceptive. However, from the outside, it can be difficult to separate these pipeline postings from Type 3 ghost jobs, which have no intention of being filled at all. The ambiguity in employer intent poses a serious challenge for labor market statistics. For instance, the Job Opening and Labor Turnover Survey (JOLTS), which is a key source of monthly vacancy data from the Bureau of Labor Statistics, does not account for ghost jobs under its current methodology.[6] According to the JOLTS definition, a job opening must meet several criteria, including that the job "could" begin within 30 days and that the employer is "actively recruiting." Both criteria are difficult to verify, and firms that post ghost or pipeline ads often meet them on paper. As a result, ghost jobs are counted in official vacancy statistics, potentially inflating vacancy rates and contributing to the apparent disconnect between job openings and actual hiring seen in the Beveridge curve.

That said, my approach comes with important caveats. Since it relies on interview reviews, it can only capture ghost jobs that proceed far enough to reach an interview stage. Listings that receive no follow-up or lead nowhere without ever scheduling an interview will not appear in this dataset. In other words, my definition of ghost jobs reflects the perceptions of applicants who were invited to participate in a hiring process that appeared performative. While this gives valuable insight into how ghost jobs are experienced by real job-seekers, it may underestimate jobs that never generate

[6]A JOLTS staff replied to my correspondence and indicated that BLS trains employees at the firms participating in the survey, on how to fill in the JOLTS surveys but admitted that it has no tangible way to verify whether the opening is legitimate or not.

any interaction, or overestimate cases where poor communication or poor fit may not reflect actual ghost jobs.

This approach also focuses on how ghost jobs are perceived and experienced by candidates, rather than on employer intent, which is fundamentally unobservable in large-scale data. In doing so, it sidesteps the measurement challenges that have limited prior research. Compared to alternative methods such as tracking job ads that remain posted for unusually long periods or that are recycled over time, this study uses rich textual narratives and large-language model (LLM) techniques to infer the sincerity of the hiring process itself.

A further limitation is that the sample of companies is not representative of the broader labor market. The dataset includes reviews from approximately 1,200 companies out of over 2.3 million listed on Glassdoor. Despite these limitations, the paper makes a methodological contribution by using natural language processing on interview reviews to identify ghost jobs from the perspective of job-seekers. In the absence of internal firm data, this offers a scalable and novel way to detect potentially insincere hiring practices.

This article contributes to two strands of literature. First, it speaks to the classical labor search framework, particularly models following Mortensen and Pissarides (1994), which have long served as the foundation for analyzing unemployment dynamics. These models typically assume that job vacancies reflect real hiring intent and do not account for job-seeker costs from applying to positions that may never be filled. Nor do they consider firms' incentives to acquire market information through job postings, even when not actively hiring. Recent work by Domash and Summers (2022) and Mongey and Horwich (2023) highlights growing mismatches between vacancy rates and unemployment, suggesting that traditional vacancy-based indicators may be less reliable in the modern labor market. This paper offers a novel explanation for this disconnect by introducing ghost jobs as a structural feature that distorts vacancy data and imposes search costs on workers.

Second, the paper contributes methodologically to the growing use of deep learning models in economics research. While prior studies have applied advanced text methods to political text classification (Jelveh et al., 2024), social behavior analysis (Ziems et al., 2024), and large-scale document processing (Dell, 2024), this study is among the first to apply deep learning techniques to job market microdata. By using BERT—a transformer-based text classification model—to analyze job-seeker interview reviews, I develop a scalable, applicant-centered approach to detecting ghost jobs. This methodology offers a new way to study hiring behavior in the absence of direct firm-level intent data.

The remainder of this paper is structured as follows: Section 2 reviews relevant literature on ghost jobs and develops the hypotheses. Section 3 describes the methodology and results. Section 4 concludes with a summary of my key insights and suggestions for future work.

2. Background and Data

2.1 Technology drives down the cost of posting job ads

The idea of jobs that were not intended to be filled can be roughly traced to an article in 2013 when the Wall Street Journal first highlighted the phenomenon of "phantom job listings" (Weber and Kwoh, 2013). Their article showed anecdotal cases where employers posted job ads for reasons unrelated to actual hiring, such as meeting formal requirements or preserving HR budgets. A separate survey of 1,045 hiring managers in 2022 showed that only 16% of jobs posted plan to be filled in one month (Clarify Capital, 2023). While this does not in itself imply that the remaining 84% are ghost jobs, it shows that most job postings remain open for extended periods, making it difficult to differentiate between slow, legitimate hiring processes and listings with no real hiring intent. This ambiguity points to a measurement gap in jobs data. For instance, the JOLTS survey, a monthly survey conducted by the Bureau of Labor Statistics (BLS), does not verify whether a job ad reflects actual hiring activity. Instead, all qualifying postings are considered legitimate openings. This has led to a widening disconnect and irrelevance in the Beveridge curve (Mongey and Horwich, 2023), where unemployment rate is inversely proportional to job vacancies.

A contemporaneous reason driving the increase in ghost jobs is the minimal marginal cost to employers in posting a new job on top of existing job postings (Bhuller et al. 2023)[7,8]. This is different from the hiring costs in classical labor economics literature, which entails other costs such as interview costs, writing the job description, processing the salary and benefits, on-boarding costs, etc. As shown in Table 1, the most popular job platforms now offer free or very low-cost job postings, including Indeed, LinkedIn (first job), Glassdoor, and others. This cost structure makes it increasingly feasible for firms to post listings for strategic reasons, even in the absence of immediate hiring needs.

After I released a preprint of this study, a discussion appeared on Hacker News, a popular technology and work forum. Many users, some of them hiring managers, shared firsthand experiences that supported the paper's main idea. One user explained, "We keep a few job listings open year-round just to have resumes on hand in case someone quits unexpectedly." Others admitted they posted jobs just to test the market or to make the company appear active. A few mentioned that jobs were posted to fulfill formal requirements, such as when applying for federal grants or visa sponsorships, even though there was no plan to hire. These anecdotes show how low job posting costs and policy incentives can encourage firms to keep ads open regardless of real hiring intent (Hacker News, 2024).

While the phenomenon of ghost jobs has been showcased through anecdotes and surveys, empirical evidence is lacking due to the difficulty in verifying whether an opening is a ghost job or not. While one could get data on how many jobs are posted and left for a period of more than 30 days or are continuously posted throughout the year, these cannot robustly be categorized as ghost jobs because the hiring firm can argue that they have difficulty finding good candidates or simply

[7] In the same interview in Footnote 1, Giddens noted that job listing companies gave her an incentive to put up more job listings than are immediately required. She cited ZipRecruiter, a recruiting company charging her $400 a month for three job listings, even though she only needed one hire. The rational mindset was thus - since the job ad was already paid for, there was no marginal cost to posting two more, even though she was not looking to fill them.$

[8] Appcast, which is a subsidiary of Stepstone Group, the 3rd largest jobs business in the world behind Indeed and LinkedIn, documented in its annual white paper that cost per click of hiring has been consistently decreasing since 2021.

disguise a Type 3 ghost job as a Type 2 pipeline posting or even a Type 1 genuine posting. In the next section, I introduce the novel dataset and a deep learning method to analyze the data.

2.2 A BERT approach to classifying reviews indicating ghost jobs

Glassdoor is a popular online platform founded in 2007 to allow employees to anonymously share reviews about employers. Over time, it has become a major resource for job seekers and researchers. In this paper, I focus on the pre-employment interview section that allows job-seekers to post their interview experiences. A job seeker can submit one interview review per employer per year on Glassdoor, which they can update as needed. However, they can review the interviews of multiple employers within the same year. Each review pertains to a specific job at a particular firm. Each review has an associated volunteer (the reviewer) and a creation time.

Figure 1a shows the survey users fill out when submitting an employer interview review to Glassdoor. Fields marked with an asterisk are required. The overall rating is restricted to whether it was positive, neutral or negative. They then add the job title, interview difficulty, process and other interview questions.

To create accountability, Glassdoor requires reviewers to have a verified active email address or a valid social network account. Each submitted review is assessed for compliance with content guidelines, and those that do not meet the standards are not displayed. Every interview review posted is anonymous but their accounts are verified. The review also shows whether the job-seeker was employed or not, and there is also an option to indicate the location of the interview. Figure 4B shows such an example.

Once an interview review is posted, visitors to the website can see it, vote it as helpful, and answer or comment about the interview questions posted. If they find that the interview review is problematic, they can also report it to Glassdoor, which will trigger an internal investigation.

When a job-seeker reads a review, they can click and label whether the review was helpful or refrain entirely from expressing an opinion. Glassdoor also provides additional information about reviewed firms through its list of companies page. I collect this information and the list of variables is shown in the Appendix.

This paper relies on the *"Interviews"* section, which is displayed on each company's landing page in Glassdoor. A screenshot of the web interface is given in Appendix A1. I determine that a robust way to check whether a company is putting out a ghost job is to rely on word-of-mouth from past interviewees. While it is possible that some interviewees may simply be upset with a rejection from the company and write "sour grapes" comments about the company, the findings in this paper show that there is a distinction between such "sour grapes" comments versus evidence of ghost jobs. The "sour grapes" comments typically show displeasure with the process. For example, on 5th August 2022, a review posted for the position of *Sales Associate* at Robert Half stated -

"An informational interview is followed by a video interview. I had multiple rounds of video interviews and I thought I did pretty well. As I am an immigrant, I guess they had some issues with our communication skills".

Contrast this with a review posted on 29th April 2016 for the position of *Marketing Associate* at HDR -

"It was a pretty slow process overall. I would say very slow. I mean they didnt call me back for 3 months. It's probably the worst HR ive ever encountered (I mean Omaha HR not hiring administrative staff). All the hiring is centered in Omaha. ... They didn't ask any cases, any "what would you do" questions, they didn't ask teamwork questions either. Pretty confusing overall. I initiated the salary negotiation myself because at that point I felt like they were going to hire me, but no one raised that topic."

that may indicate that this was a ghost job, where there were no questions asked on competence, nor clarification on expected salary. More examples can be found in Appendix A1. There is a distinct difference in the content of "sour grapes" review versus "ghost job" reviews, which lends credibility to the dataset. Specifically, I analyze jobseekers' posts about their interview experiences through an advanced large-language model (LLM) textual analysis with AI-assisted techniques.

To build my dataset, I extracted the first 1,203 companies listed by relevance on Glassdoor. These company listings are randomly sorted by the platform's algorithm and refreshed each time the browser is reloaded, providing a broad and quasi-random sample of firms rather than a static or curated list. As shown in Appendix A1, I use the filter where overall rating is 1.0 and above to capture as many companies as possible. I extract up to 30 pages of interview reviews from each company's Glassdoor page. This limit reflects the structure of the web-scraping script, which was designed to balance coverage and efficiency. Some companies have reviews dating back many years, and retrieving all of them would disproportionately weight older reviews while increasing runtime significantly. As such, this introduces a limitation: for firms with long histories on Glassdoor, the data may not fully capture more recent hiring practices.

The web-scraping is done in the month of July 2024 using a server and takes an entire month to complete. I retrieve a total of 29,294 pages from 1,203 companies. Some companies are smaller or not as popular and may not have up to 30 pages of interview reviews. There are also some reviews that are not in English as Glassdoor is an exhaustive dataset of global companies and MNCs. I use a deep-learning, language detector library in Python and identify that there are 269,347 english reviews, 4,315 portugese reviews, 3,593 french reviews and etc. The full list is given in Appendix A2. Table 2 shows the descriptive statistics of the variables. I use only the english reviews resulting in 1,194 valid companies across 97 industries.

To identify whether an interview review is indicative of a ghost job, I primarily rely on a deep learning–based classification approach. Specifically, I use ChatGPT-4o to label a random subset of 2,000 reviews, and then fine-tune a BERT model to classify the full dataset based on these labels. This model captures contextual cues in the interview narratives that are often too subtle for rule-based methods to detect. Because the dataset is large (over 269,000 English-language reviews), this two-step process balances scalability with classification accuracy.

As a supplementary check, I also implement a classic keyword search method, drawing from a hand-curated list of terms and phrases that indicated insincere hiring pipelines. Reviews containing these keywords are flagged as ghost jobs. This method is more limited as it lacks contextual understanding and risks false positives when keywords appear outside of ghost job contexts. I therefore use keyword search as a diagnostic tool to compare against the BERT model's predictions.

2.3 Highly-skilled jobs are more likely to be ghost jobs

Prior literature suggests that highly-skilled positions may be particularly prone to prolonged hiring processes due to their specialized skill-set and narrower applicant pools. Unlike lowly-skilled positions that have a larger number of potential applicants, highly-skilled roles often require specialized expertise, higher education qualifications, and extensive experience. This limited availability of qualified candidates means that employers face more significant challenges in finding suitable hires quickly. By advertising these roles even when there is no immediate need, firms can conduct preliminary screenings and build a database of potential hires. This approach minimizes the downtime associated with the hiring process, especially for jobs in certain environments (Moretti, 2013).

This hiring strategy is consistent with Type 2 pipeline postings, which are common in high-skill industries. However, the high prevalence of such postings makes it easier for Type 3 ghost jobs to blend in, complicating detection. This ambiguity motivates the following hypothesis:

H1: Highly-skilled industries are more likely to have ghost job postings

2.4 Bigger firms are more likely to have ghost jobs

Bigger firms are also more likely to have ghost jobs due to their extensive resources and complex organizational structures. Large companies often have multiple departments and a constant need to fill a variety of positions, making it more feasible for them to keep job advertisements open even when there are no immediate vacancies. This maintains a robust talent pipeline, ensuring that they can quickly respond to unexpected departures or sudden increases in demand. Additionally, bigger firms typically have dedicated HR departments with the capacity to manage ongoing recruitment processes, including screening and maintaining databases of potential candidates. There is also literature on how operational factors can affect stock returns, and the high vacancy rate may positively signal to shareholders that the company is thriving (Campbell, 1996; Chen and Li, 2023).

Moreover, larger organizations benefit from economies of scale in their hiring processes, making the marginal cost of maintaining job ads relatively low. The job portals often offer bulk advertising deals or subscription-based models that allow companies to post numerous job listings at a reduced cost. This low marginal advertising cost makes it practical for big firms to advertise positions continuously, even if they are not actively looking to hire immediately. The presence of ghost jobs in larger firms can also be attributed to strategic workforce planning, where companies aim to gather data on the labor market, including salary trends and skill availability. By keeping job ads active, they can continuously monitor market conditions. Thus, I hypothesize the following in alternate form.

H2: Bigger firms are more likely to have ghost job postings

3. Results

3.1 How Effectively Does BERT Identify Ghost Jobs?

To build my dataset, I extracted the first 1,203 companies listed by relevance on Glassdoor. These company listings are randomly sorted by the platform's algorithm and refreshed with each browser reload, providing a broad and quasi-random sample. From each company's Glassdoor page, I extract up to 30 pages of interview reviews. I remove all non-English reviews from the initial sample. After restricting the sample to companies with more than 10 interview reviews and ensuring that the firms have non-missing control variable data, I end up with 1,194 firms in the final dataset.

After cleaning the text, I do a Latent Dirichlet Allocation (LDA) topic analysis. I tokenize reviews using a bag-of-words approach and use the *gensim*[9] package to estimate the LDA model. Each LDA topic is a distribution over words, and each review is assigned a mixture of topics. Table 4 shows that the extracted topics, such as HR communication, hiring timelines, and interview difficulty—confirm that the Glassdoor interview data reflects genuine interview experiences.

To classify whether a review is indicative of a ghost job, I primarily rely on a fine-tuned BERT model. I begin by using ChatGPT-4o to label 2,000 randomly selected reviews for ghost job relevance, and then use these labels to fine-tune the BERT model for classification across the full dataset. Appendix A2 provides the prompt structure and training details. As a supplementary comparison, I also implement a keyword search method using a hand-picked list of ghost job indicators derived from news coverage and manual review of the text.

The BERT model serves as the primary method for identifying ghost job indicators in interview reviews, classifying 21% of reviews as indicative of ghost jobs, as shown in Table 5 Panel A. In contrast, a basic keyword-based method flags only 1.6% of reviews. This difference highlights the limitations of relying solely on text matching. The BERT model is better equipped to capture implied meanings (Gonzalez-Carvajal and Garrido-Merchan, 2023), context, and patterns of evasive or performative interview behavior that simple keyword filters often miss (see examples in Appendix A1). Moreover, the 21% estimate is more conservative than the expert's assessment that up to 50% of job postings may be ghost jobs (Woods and Wong, 2024). While some false positives may arise from rejected candidates misinterpreting their experience, this upward bias is likely counterbalanced by a downward bias from candidates who never leave a review at all.

I next repeat the LDA analysis on the subset of reviews classified as ghost jobs. Table 6 Panel A shows dominant topics include multiple interview rounds, long delays, lack of feedback, and no salary discussion. These are signals consistent with insincere hiring processes. Panel B confirms

[9] gensim is a Python library designed for scalable topic modeling and natural language processing. It implements Latent Dirichlet Allocation (LDA) using an online variational Bayes algorithm, which is well-suited for economic research involving unstructured text.

that ghost jobs are more common in negative experiences: 74.8% of BERT-classified ghost jobs are labeled negative, compared to just 6.2% under the keyword method.

To address the concern that the BERT model is merely detecting negative sentiment and not ghost jobs, I run a logistic regression predicting ghost job classification from job-seeker's sentiment. While sentiment is statistically significant, it explains only 37% of the variation with pseudo $r^2 = 0.370$), implying that BERT captures additional features beyond sentiment alone. This supports the interpretation that ghost job classification reflects structural cues such as repeated interviews, lack of feedback, or unclear role definitions, and not just the job-seeker's sentiment.

3.2 Can Ghost Job Adjustments Help Reconcile the Beveridge Curve?

Next, I explore whether the ghost job trend helps explain labor market dynamics by comparing my data to the JOLTS series. Fig 2a shows the annual share of reviews classified as ghost jobs. There is a discrete jump in 2015, which is likely driven by Glassdoor's $70 million funding round and expanded U.S. presence (Cook, 2015). Aside from that, the trend tracks labor market disruptions like the COVID-19 shock. This suggests that while Glassdoor is not a perfect sample, its signals track broader shifts in job search dynamics.

I adjust the JOLTS job openings rate using a rolling average of the estimated share of ghost jobs based on my Glassdoor classification. The adjustment applies a dynamic scaling factor based on the historical gap between JOLTS openings and quits rates, smoothing short-term fluctuations while preserving longer-term trends. In Fig 2b, the gap between openings and quits is narrowest during the COVID-19 period (2020–2021). This narrowing is likely driven by the decline in hiring intentions during the early pandemic, when many firms froze hiring and posted speculative openings to monitor the labor market, or maintained postings for signaling or compliance reasons. At the same time, quits sharply declined due to uncertainty and risk aversion among workers. The combination of inflated job openings and lower quits widened the Beveridge gap in raw JOLTS data. By accounting for ghost jobs, which correspondingly surged during this period, the adjustment possibly offered a more accurate picture.

That said, I acknowledge that this adjustment is based on data from 1,194 firms and is not representative of the U.S. economy as a whole. I present the Beveridge Curve adjustment in Fig 2b not as a definitive macroeconomic correction, but more as an illustrative example to show how accounting for the quality of job postings could potentially reconcile labor market indicators.

Using the Beveridge curve reconciliation, I can address the concern that the BERT model is merely identifying failed interviews rather than genuine ghost jobs. During the 2020–2021 period, the ghost jobs correction corrects for the disconnect to the largest degree. I use this period as a benchmark, as it most clearly reflects ghost job activity, and test whether the content in the reviews from this period differs from those in other periods where the reconciliation is not as effective.

If the NLP classification is simply identifying failed interviews rather than genuine ghost jobs, then the thematic content of ghost job reviews in 2020–2021 should resemble those in other periods, and also be indistinguishable from non-ghost (placebo) reviews. In contrast, if ghost jobs represent a stable and distinct employer behavior, then reviews classified as ghost jobs from adjacent periods should exhibit more similarity to the benchmark period than non-ghost reviews do.

To evaluate this, I apply Latent Dirichlet Allocation (LDA) to extract a single dominant topic from ghost job reviews in three time windows: 2018–2019, 2020–2021 (benchmark), and 2022–2023. I repeat the same procedure for reviews classified as non-ghost. I then compare the resulting topics using Jensen-Shannon Divergence (JSD), a symmetric and bounded metric of dissimilarity between word distributions. A lower JSD implies more topical similarity.

As shown in Table 7, the average JSD between ghost job topics and the benchmark period is 0.318, with a minimum of 0.263, indicating strong internal consistency in ghost job themes across time. In contrast, the average JSD between non-ghost topics and the benchmark is 0.645, with a minimum also at 0.645, demonstrating that non-ghost topics are far less similar to the benchmark. These findings suggest that the BERT model is capturing a consistent and recognizable structure associated with ghost jobs and not merely failed interviews or one-off hiring anomalies.

3.3 Why Do Firms Conduct Costly Interviews Without Hiring Intent?

While interviews impose real costs such as requiring time, staffing, and coordination, firms may still choose to conduct them even when they have no intention of hiring. This creates a puzzle: if ghost jobs are not meant to result in actual hires, why incur the added expense of interviews at all? As shown in Table 1, firms may use interviews to gather competitive intelligence, benchmark compensation expectations, signal company growth to investors or analysts, or satisfy internal HR activity metrics. In these cases, the interview itself, not the hire, is the objective. Thus, the interview process gives the company the benefits of ghost jobs, not just posting the ad itself.

To detect such behavior at scale, I use the interview-to-job ratio, defined as the ratio of Interviews_Count to Jobs_Count. In typical hiring processes, this ratio should remain relatively stable as most firms interview a modest number of candidates per job. When this ratio is unusually high, it suggests an imbalance where there is intensive interviewing without corresponding job openings, which may indicate performative behavior.

I estimate separate regressions for firms in the top 10% and bottom 90% of the interview-to-job ratio distribution, with industry fixed effects in Table 8. This directly tests the concern that ghost jobs should not involve interviews due to their cost. If firms would not incur real costs—such as interview coordination and staff time—without genuine hiring intent, then we should not observe significant interview activity among firms posting ghost jobs. However, the results show otherwise. Among the bottom 90% (typical firms), interview reviews scale modestly with job postings (β = 0.055) and strongly with salaries (β = 0.074), consistent with standard hiring behavior. But among the top 10%, each job posting is associated with over 18 interview reviews (β = 18.75), and the model intercept is statistically insignificant, suggesting that interviews occur regardless of actual hiring outcomes.

This suggests that interviews can occur even in the absence of hiring intent. Firms with extreme interview-to-job ratios could be interviewing for strategic purposes. If the interview itself provides value, the cost becomes justifiable, even without a hire.

3.4 Do Larger Firms Post More Ghost Jobs?

Next, I examine how the likelihood of ghost job postings varies with firm size. I hypothesize that larger firms are more likely to post ghost jobs due to slack in their HR operations or a greater need to gather labor market intelligence—such as benchmarking salaries or gauging skill availability. These firms may also have more unused job ad credits or internal incentives to maintain hiring visibility.

To test this, I split the companies into seven employee-size categories, ranging from 1–50 employees to over 10,000. For each group, I calculate the number and percentage of ghost jobs, defined as the number of interview reviews classified as indicative of ghost jobs divided by the total number of interview reviews in that bracket. Table 9 presents these results.

The descriptive findings show that the highest ghost job percentage occurs in mid-to-large firms. Specifically, the keyword-based method reveals that companies with 5,001 to 10,000 employees have the highest ghost job share at 1.58%, while the BERT model finds the peak among firms with 1,001 to 5,000 employees. This aligns with the hypothesis that mid-sized and large firms may have more institutional inertia or resources that enable ghost job posting. In contrast, the smallest and largest firms tend to operate leaner.

To explain the pattern in Table 9, I develop a simple model (Appendix C) treating ghost job postings as a strategy for acquiring labor market information. Drawing on the costly information acquisition literature (Grossman and Stiglitz, 1980; Morris and Shin, 2002; Sims, 2003; Veldkamp, 2023), the model predicts an inverted-U relationship between firm size and ghost job activity. Small firms lack HR capacity, large firms face diminishing informational returns, and mid-sized firms strike the optimal balance between cost and informational gain. This suggests that ghost job posting is most prevalent where institutional capacity and informational incentives are best aligned.

3.5 Are High-Skill Industries More Prone to Ghost Jobs?

Lastly, I examine whether ghost job prevalence varies by industry, particularly in high-skill sectors. These industries often use ghost postings strategically to build talent pipelines, benchmark pay, or signal hiring in competitive markets. In contrast, low-skill sectors face tighter budgets, more predictable staffing needs, and lower informational value from excess applicants, making speculative postings less common and less beneficial.

Table 10 shows that ghost job rates are highest in high-skill sectors. Industries such as publishing, internet and web services, software development, and commercial equipment services, all reliant on technical talent, exhibit ghost job shares above 29%. These patterns align closely with Type 2 pipeline postings and Type 3 ghost jobs, where ads are used not to fill immediate vacancies. These sectors often face volatile demand or evolving skill requirements, making speculative posting a rational strategy. Charitable foundations also rank high, likely due to grant-contingent hiring, where openings are posted in anticipation of future funding. By contrast, routine or low-skill sectors such as restaurants, accounting, and government-linked institutions report ghost job rates near or below 13%. These organizations typically have more stable hiring pipelines and tighter resource constraints.

To test whether job skill drives these patterns, I estimate a fixed-effects regression using BERT-classified ghost job indicators as the outcome in Table 11. High-skill roles are significantly more likely to be ghost jobs across all specifications. This supports the interpretation that firms disproportionately use ghost job listings for professional and managerial positions. Importantly, this effect remains robust even after controlling for sentiment, review length, and firm size.

Table 11 also reinforces the inverted-U relationship between firm size and ghost job prevalence, consistent with the theory developed in section 3.4 and Appendix A3. Size is significantly positive, while the square of size is negative, indicating that ghost jobs are most common among mid-sized firms. Given that high-skill jobs are less frequent overall (BLS, 2024), the fact that they are more likely to be ghost jobs even after adjusting for firm size strengthens the claim that ghost postings are a strategic feature of high-skill labor markets.

4. Conclusion

Ghost jobs are a growing feature of the online labor market. Using a BERT-based classification on Glassdoor interview reviews, I find that roughly one in five job postings may be ghost jobs. These postings are especially common in high-skill industries and mid-sized firms, where firms face low marginal posting costs and strategic incentives to collect labor market data or maintain a perception of growth.

The costs of ghost jobs are not borne by firms but by job seekers, who invest time and emotional effort pursuing roles that do not exist. This mismatch can increase job search fatigue, mislead policy and training priorities, and contribute to distorted macro indicators such as the Beveridge Curve. In this paper, I illustratively show that adjusting job openings for estimated ghost job trend reduces the observed disconnect between vacancies and quits during the COVID-19 period.

While not all ghost jobs are deceptive, the opacity of employer intent presents real challenges for labor economists and policymakers. This study offers a scalable approach to detecting ghost hiring and contributes to a broader understanding of how firms interact with the labor market under low-cost job advertising regimes.

References

Appcast. (2024). 2024 Recruitment Marketing Benchmark Report (US Edition). https://whitepaperseries.com/wp-content/uploads/2024/03/135886_ML-2024-Recruitment-Marketing-Benchmark-Report.pdf#page=1.00

Bhuller, M., Ferraro, D., Kostøl, A. R., & Vigtel, T. C. (2023). The Internet, Search Frictions and Aggregate Unemployment (Working Paper 30911). National Bureau of Economic Research. https://doi.org/10.3386/w30911

Bienasz, G. (2022, September 20). What Are "Ghost Jobs" and Why Are They Everywhere Now? Entrepreneur. https://www.entrepreneur.com/business-news/what-are-ghost-jobs-and-why-are-they-everywhere-now/435666

BLS. (2024). BLS: U.S. Bureau of Labor Statistics. Bureau of Labor Statistics. https://www.bls.gov/oes/tables.htm

Campbell, J. Y. (1996). Understanding Risk and Return. Journal of Political Economy, 104(2), 298–345. https://doi.org/10.1086/262026

Chalmers, J. M. R., Dann, L. Y., & Harford, J. (2002). Managerial Opportunism? Evidence from Directors' and Officers' Insurance Purchases. The Journal of Finance, 57(2), 609–636. https://doi.org/10.1111/1540-6261.00436

Chen, C.-W., & Li, L. Y. (2023). Is hiring fast a good sign? The informativeness of job vacancy duration for future firm profitability. Review of Accounting Studies, 28(3), 1316–1353. https://doi.org/10.1007/s11142-023-09797-2

Clarify Capital. (2023, October 23). Survey: Job Seekers Beware of Ghost Jobs. Clarify Capital. https://clarifycapital.com/job-seekers-beware-of-ghost-jobs-survey

Cook, J. (2015, January 6). Rich Barton-backed career site Glassdoor picks up 70 million in funding. GeekWire. https://www.geekwire.com/2015/rich-barton-backed-career-site-glassdoor-picks-70-million-funding/

Dell, M. (2024). Deep Learning for Economists (Working Paper 32768). National Bureau of Economic Research. https://doi.org/10.3386/w32768

Domash, A., & Summers, L. H. (2022). How Tight are U.S. Labor Markets? (Working Paper 29739). National Bureau of Economic Research. https://doi.org/10.3386/w29739

González-Carvajal, S., & Garrido-Merchán, E. C. (2023). Comparing BERT against traditional machine learning text classification. Journal of Computational and Cognitive Engineering, 2(4), 352–356. https://doi.org/10.47852/bonviewJCCE3202838

Goldfarb, B., Kirsch, D., & Miller, D. A. (2007). Was there too little entry during the Dot Com Era? Journal of Financial Economics, 86(1), 100–144. https://doi.org/10.1016/j.jfineco.2006.03.009

Gray, M. L., & Suri, S. (2019). Ghost work: How to stop Silicon Valley from building a new global underclass. Eamon Dolan Books.

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. The American Economic Review, 70(3), 393–408.

Hacker News. (2024, May 13). Why is it so hard to find a job now? Enter Ghost Jobs | Hacker News. https://news.ycombinator.com/item?id=42136469

Jelveh, Z., Kogut, B., & Naidu, S. (2024). Political Language in Economics. The Economic Journal, 134(662), 2439–2469. https://doi.org/10.1093/ej/ueae026

Kim, J. (Simon), & Ra, K. (2022). Employee satisfaction and asymmetric cost behavior: Evidence from Glassdoor. Economics Letters, 219, 110829. https://doi.org/10.1016/j.econlet.2022.110829

Kyiu, A., Tawiah, B., & Boamah, E. O. (2023). Employees' reviews and stock price informativeness. Economics Letters, 233, 111406. https://doi.org/10.1016/j.econlet.2023.111406

Lim, V. K., Chen, D. J. Q., & Tan, M. (2013). Unemployed and Exhausted? Fatigue During Job-Search and Its Impact on Reemployment Quality. Academy of Management Proceedings, 2013(1), 13399. https://doi.org/10.5465/ambpp.2013.13399abstract

Marchese, D. (2023, January 23). The Digital Workplace Is Designed to Bring You Down. The New York Times. https://www.nytimes.com/interactive/2023/01/23/magazine/cal-newport-interview.html

Mongey, S., & Horwich, J. (2023). Are job vacancies still as plentiful as they appear? Implications for the "soft landing" | Federal Reserve Bank of Minneapolis. Federal Reserve Bank of Minneapolis. https://www.minneapolisfed.org/article/2023/are-job-vacancies-still-as-plentiful-as-they-appear-implications-for-the-soft-landing

Moretti, E. (2013). Real Wage Inequality. American Economic Journal: Applied Economics, 5(1), 65–103. https://doi.org/10.1257/app.5.1.65

Morris, S., & Shin, H. S. (2002). Social value of public information. American Economic Review, 92(5), 1521–1534.

Nienhaus, M. (2022). Executive equity incentives and opportunistic manager behavior: New evidence from a quasi-natural experiment. Review of Accounting Studies, 27(4), 1276–1318. https://doi.org/10.1007/s11142-021-09633-5

Paradis, T. (2024). It's wild how many job listings might be fake. Business Insider. https://www.businessinsider.com/companies-posting-fake-job-listings-search-resume-2024-6

Sims, C. A. (2003). Implications of rational inattention. Journal of Monetary Economics, 50(3), 665–690.

Sockin, J., & Sojourner, A. (2023). What's the Inside Scoop? Challenges in the Supply and Demand for Information on Employers. Journal of Labor Economics, 41(4), 1041–1079. https://doi.org/10.1086/721701

Sprague, R. (2023). Privacy Self-Management: A Strategy to Protect Worker Privacy from Excessive Employer Surveillance in Light of Scant Legal Protections. American Business Law Journal, 60(4), 793–836.

Symitsi, E., Stamolampros, P., & Daskalakis, G. (2018). Employees' online reviews and equity prices. Economics Letters, 162, 53–55. https://doi.org/10.1016/j.econlet.2017.10.027

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30. https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

Veldkamp, L. (2023). Information choice in macroeconomics and finance. Princeton University Press. https://books.google.com/books?hl=en&lr=&id=qlSIEAAAQBAJ&oi=fnd&pg=PP9&dq=Information+Choice+in+Macroeconomics+and+Finance&ots=DNuPDbVFEL&sig=tA8ia0BC7g1DB5pVABPnD-auDBs

Weber, L., & Kwoh, L. (2016, May 12). Want a New Job? Beware “Phantom” Postings - WSJ. Archive.Ph. https://archive.ph/kNYmI

Wikipedia. (2025). Ghost job. In Wikipedia. https://en.wikipedia.org/w/index.php?title=Ghost_job&oldid=1290926967

Woods, D., & Wong, W. (2024). What are “ghost jobs”? Connecticut Public Radio. https://www.ctpublic.org/2024-06-14/what-are-ghost-jobs

Ziems, C., Held, W., Shaikh, O., Chen, J., Zhang, Z., & Yang, D. (2024). Can Large Language Models Transform Computational Social Science? Computational Linguistics, 50(1), 237–291. https://doi.org/10.1162/coli_a_00502

Tables and Figures

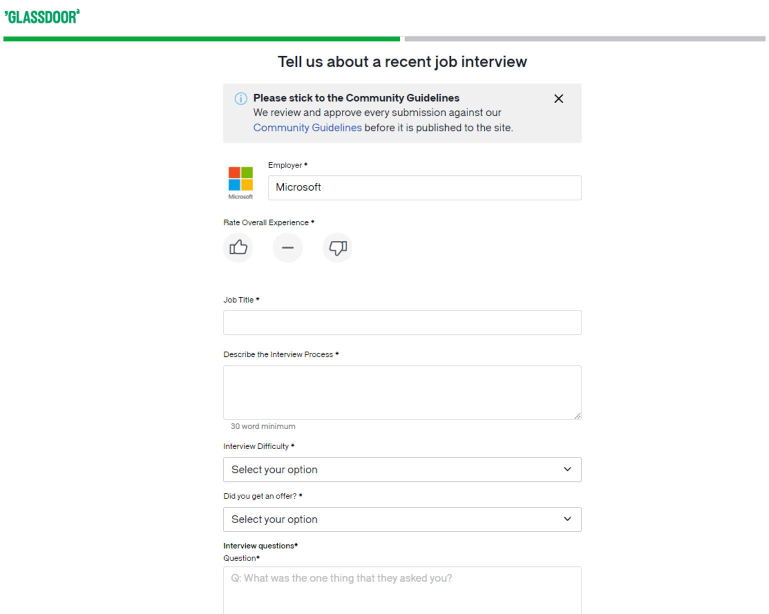

Fig 1a. Screenshot of the Glassdoor Interview Review Section, where job-seekers can post their experiences.

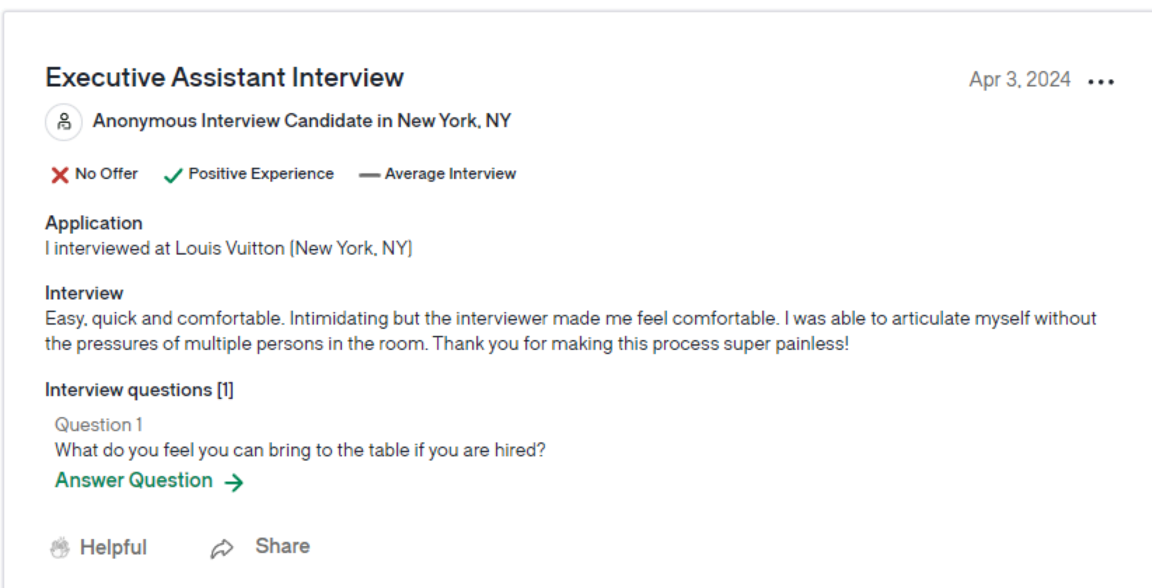

Fig 1b. Screenshot of an interview review posted for Louis Vuitton in New York on 3th April 2024

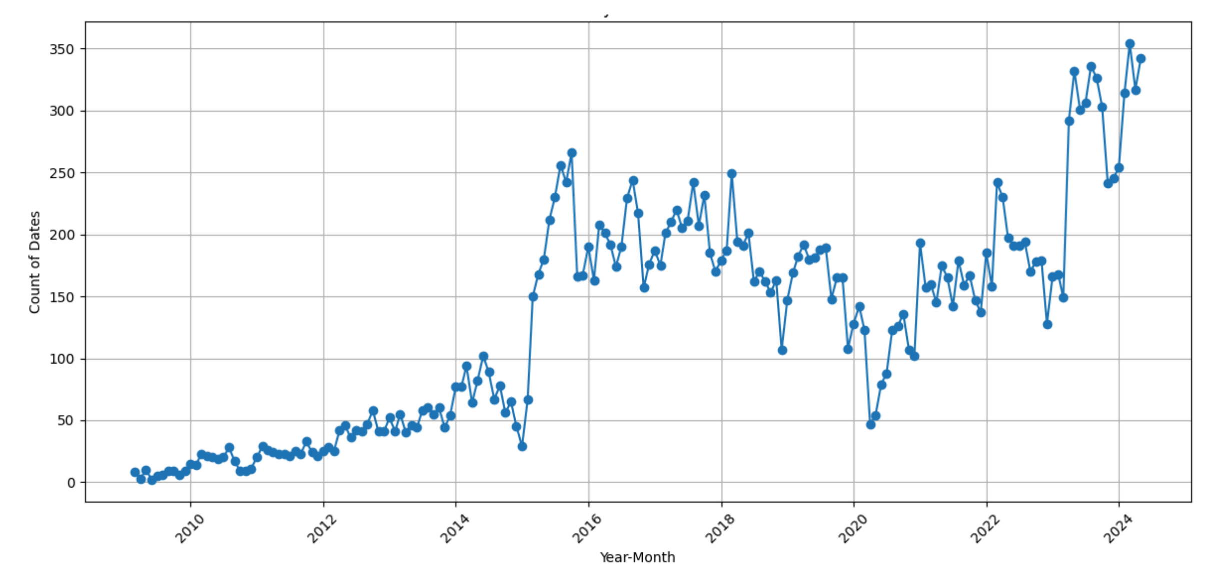

Fig 2a. Number of ghost jobs classified by BERT for each month. Only interview reviews with a US location are taken.

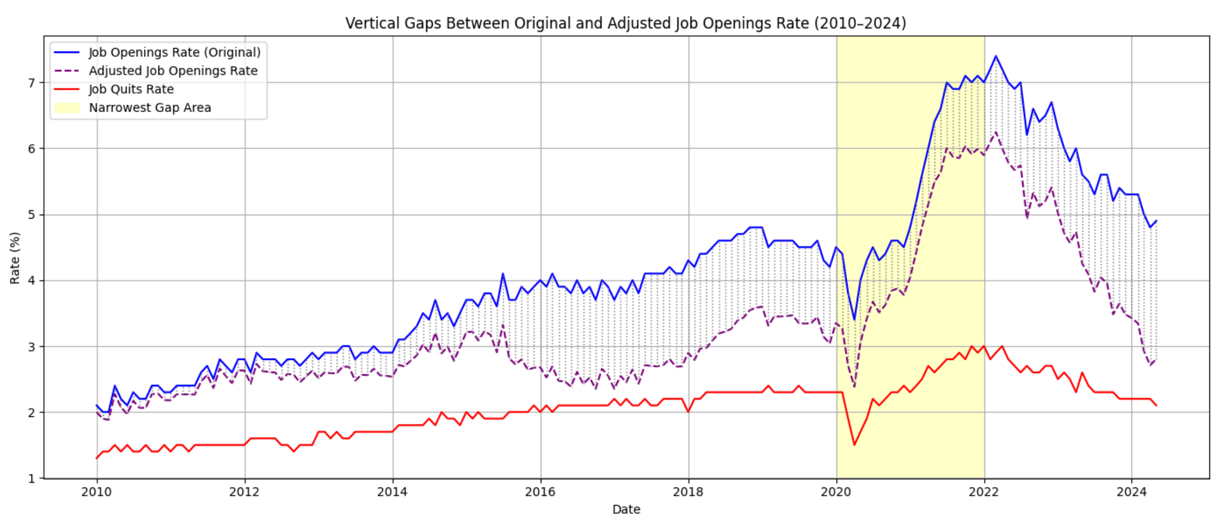

Fig 2b. Rolling average-adjusted *job opening rate* after taking into account the number of ghost job reviews posted, and the *job quits rate* in the US. This line graph shows that the disconnect closes back up after the Covid pandemic. The formula is given by

$$\overline{G}(t) = \left(\frac{1}{k}\right)\sum_{i=0}^{k-1} G(t-i)$$

$$J_t^{\text{adj}} = J(t)*\left(1 - J_t^{\text{adj}}(t)\right)$$

where J(t) = raw JOLTS job openings rate at time t , G(t) = estimated ghost job share at time t , $\overline{G}(t)$ = rolling average of G(t) over a window of length k , J_t^{adj} = adjusted job openings rate at time t

Table 1. Typology of Ghost Jobs by Hiring Intent. This table outlines three categories of job postings based on employers' hiring intent. Ghost jobs are defined in this paper as postings in Category 3, where there is no immediate or eventual plan to hire. Categories 1 and 2, while not deceptive per se, may still distort labor market signals.

Type	Hiring Intent	Common Features and Interpretation
Type 1: Genuine Openings	Immediate and Real	The firm has a real and immediate need to hire. Interviews are structured, questions are role-specific, and timelines are clear. There is typically a budgeted vacancy, with salary discussions and decision-making processes in place. These ads reflect a genuine intent to fill a defined role.
Type 2: Pipeline Postings	Delayed but Real	There is eventual intent to hire, but no current opening or budgeted role. Firms may use these listings to build a steady flow of résumés for hard-to-fill or specialized positions. Interview processes may be slower or less urgent, and timelines vague. These ads are not deceptive but can still mislead applicants.
Type 3: Ghost Jobs	None	No intention to hire in the near or foreseeable future. Could be used for market intelligence, employee signaling to keep staff on edge, investor signaling, HR budget justification, compliance or regulatory theatre for grants, employer branding, or quota optics to show "active engagement" with DEI or other goals. Ads may or may not be genuine and can persist all year round. Interviews may be short or prolonged depending on internal strategy.

Table 2. How much does it cost to post a job ad in 2024? Table 2 describes the cost of posting a job ad on the most popular online recruitment websites. There are many free options, which also feature job postings by SnP500 companies. This provides evidence that job ad posting costs has become marginal today. Source - various websites

Jobs Portal	Cost	Details
Indeed	0	Indeed doesn't charge for posting ads, only for showcasing your post
LinkedIn	0	First job posting is free, subsequent jobs charged based on cost-per-click
Glassdoor	0	Glassdoor lets companies post up to 10 jobs free for 7 days
Handshake	0	Unlimited posting for free
SimplyHired	0	Unlimited posting for free
Guru	0	Unlimited posting for free
HubStaffTalent	0	Unlimited posting for free
Illfound	0	Unlimited posting for free
ZipRecruiter	$16/day	Charged for jobs posting per day
Monster Jobs	$12/day	Charged for jobs posting per day

Table X. Descriptive Statistics of the dataset used in this study

	count	mean	min	25%	50%	75%	max	std
word_count	269,330	56.64	1	33	40	56	967	53.42
skill	199,186	0.642	0	0	1	1	1	0.479
date	269,330		3/11/2009	9/13/2017	6/25/2021	5/17/2023	7/16/2024	
Jobs_Count	1,144	1,256.86	1	94.75	326	1,000	84,000	4,892.35
Reviews_Count	1,194	4,607.97	123	1,400	1,800	2,600	207,000	14,103.68
Interviews_Count	1,194	837.80	19	190.25	290.5	517	54,000	2,692.46

Table 4 Latent Dirichlet Analysis on the interview reviews. Sample size = 269,347.

Topic	Keywords	Interpretation
0	0.019*time, 0.014*would, 0.013*recruiter, 0.012*company, 0.011*candidate, 0.010*even, 0.009*interview, 0.009*process, 0.008*role, 0.008*like	Focuses on the candidate's perspective, including interactions with recruiters and company roles in the interview process.
1	0.069*round, 0.046*interview, 0.045*question, 0.037*technical, 0.020*hr, 0.013*asked, 0.013*first, 0.013*one, 0.012*2, 0.011*coding	Focuses on the different rounds of interviews, particularly technical and HR interviews, and specific questions such as coding
2	0.034*interview, 0.030*question, 0.021*test, 0.018*online, 0.018*assessment, 0.015*take, 0.015*answer, 0.013*time, 0.012*video, 0.011*minute	Focuses on the structure and components of the interview process, including online assessments and video interviews.
3	0.111*interview, 0.038*manager, 0.027*phone, 0.027*call, 0.024*process, 0.022*hiring, 0.021*recruiter, 0.020*week, 0.016*team, 0.015*day	Focuses on the managers and recruiters, phone interviews, and the overall hiring timeline.
4	0.050*interview, 0.044*question, 0.025*process, 0.024*experience, 0.019*asked, 0.017*good, 0.016*easy, 0.013*job, 0.013*work, 0.012*interviewer	Focuses on the overall experience of the interview process, question difficulty and interviewer interaction.

Table 5 Panel A. BERT Matching to check whether interview review indicative of ghost jobs

Category	Count
Number of True values in ‘keyword-match’ column for English reviews	56,559
Number of False values in ‘keyword-match’ column for English reviews	212,788
% of Ghost Jobs	21%

Table 5 Panel B. Keyword Matching to check whether interview review indicative of ghost jobs.

Category	Count
Number of reviews matched for any of the keyword in the list	3,624
Number of reviews not matched for any of the keyword in the list	265,723
% of Ghost Jobs	1.3%

Table 6 Panel A. Latent Dirichlet Analysis on the interview reviews which have been classified as ghost jobs using the BERT model. Sample size = 56,559.

Topic	Keywords	Interpretation
1	interview, round, process, hr, time, 2, feedback, 3, final, month	Focuses on multiple rounds, HR involvement, feedback, and the time frame of the interview stages.
2	interview, job, day, manager, time, one, told, get, minute, would	Focuses on individual interview experiences, including interactions with managers, the time taken, and what candidates were told during the process.
3	interview, recruiter, call, email, week, back, would, never, time, phone	Focuses on interactions with recruiters, follow-up calls and emails, and instances of delayed or no responses.
4	company, like, time, work, know, candidate, even, get, people, want	Focuses on perceptions and experiences with the company, including the work environment and how they felt about the company's communication and treatment.
5	offer, salary, job, pay, process, position, long, low, offered, get	Focuses on job offers, including salary negotiations, the overall offer process, and the positions offered to candidates.

Table 6 Panel B. Classification of the BERT search based on whether the sentiment of the interview was negative, positive or neutral (code as neg exp, pos exp and neutral exp respectively).

	(1) Total	(2) Neg	(3) Pos	(4) Neutral
Number of Ghost Jobs	56559	36262	7896	12103
% of Ghost Jobs	21	74.8	4.64	24.8

Table 6 Panel C Logit regression of whether an interview review is indicative of a ghost job. The dependent variable is a binary indicator for ghost jobs as classified by the BERT model. The independent variable is the sentiment of the interview experience, coded as 1 for positive, 0 for neutral, and −1 for negative.

	Logit Coef.
Constant	−1.0192***

	(0.006)
Sentiment	−2.0631***
	(0.008)
Pseudo R^2: 0.3700, N = 267,632	

*, **, *** indicate significance at the 10%, 5%, and 1% levels respectively.

Table 7. Jensen-Shannon Divergence (JSD) between ghost job topics and non-ghost (placebo) topics relative to the benchmark period (2020–2021). Lower values indicate greater similarity in topic content.

Comparison	Average JSD	Minimum JSD
Ghost Topics vs 2020–2021	0.318	0.263
	(0.054)	
Placebo Topics vs 2020–2021	0.645	0.645
	(0.0)	

Table 8. OLS regression coefficients for interview volume, stratified by firms' interview-to-job ratios. The dependent variable is the number of interview reviews. All models include industry fixed effects.

Bottom 90% of Interview-to-Job Ratio	
Variable	Coef.
Intercept	1328.77***
	(298.81)
Jobs_Count	0.055***
	(0.0061)
Salaries_Count	0.074***
	(0.0009)
Top 10% of Interview-to-Job Ratio	
Variable	Coef.
Intercept	−190.98
	(617.51)
Jobs_Count	18.75***
	(2.58)
Salaries_Count	0.054***
	(0.0045)

*, **, *** represent significance at the 10%, 5%, and 1% level.

Table 9. Number and Percentages of ghost jobs using the BERT model, sorted by how many employees the company has. The *% of ghost jobs* is calculated by taking the number of interview reviews indicative of ghost jobs divided by total number of interview reviews of companies in that employee number bracket. Variable definitions are detailed in Table 2.

	Number of Ghost Jobs	% of Ghost Jobs
BERT (1 to 50 Employees)	314	14.5
BERT (51 to 200 Employees)	480	16.8
BERT (201 to 500 Employees)	216	23.9
BERT (501 to 1000 Employees)	756	23
BERT (1001 to 5000 Employees)	15200	**24.8**
BERT (5001 to 10000 Employees)	13000	22.7
BERT (10000+ Employees)	25800	18.7
BERT (Unknown)	857	19.8

Table 10 Ranking of the Industries - the top 5 industries with the highest percentage of interview reviews indicative of ghost jobs and the industries with the lowest percentage of interview reviews indicative of ghost jobs, using both keyword search and BERT models. The upper panel shows the top industries and the bottom panel, the industries with the lowest percentages, and % of ghost jobs refer to the number of interview reviews indicative of ghost jobs divided by the total number of interview reviews for companies in that particular industry. Variable definitions are detailed in Table 2.

Top Industries (BERT)	Number of Ghost Jobs (BERT)	% of Ghost Jobs (BERT)
Publishing	572	34.4
Internet & Web Services	2200	30.1
Grantmaking & Charitable Foundations	170	29.9
Commercial Equipment Services	102	29.8
Software Development	1830	29.8
Bottom Industries (BERT)	**Number of Ghost Jobs (BERT)**	**% of Ghost Jobs (BERT)**
Restaurants & Cafes	1320	13.6
National Agencies	550	13.2
Accounting & Tax	252	12.5
Religious Institutions	26	12.4
Commercial Printing	25	12.2

Table 11. Does skill of the job affect whether the job ad is more likely to be a ghost job? This table shows a fixed effect OLS regression on the dependent variable of whether the review is indicative of a ghost job using the BERT model. Word Count refers to the word count of the interview review. Sentiment refers to the job-seeker's experience of the interview, which can be either positive, neutral or negative, with positive being 1, neutral being 0 and negative being -1. Size refers to the 5 levels of the company sizes. Unknown company sizes are discarded. Skill refers to whether the job is high-skilled or low-skilled. Standard errors, shown in parentheses, are clustered at the year level. Variable definitions are detailed in Table 2.

	(1)	(2)	(3)
Skill	0.030***	0.029***	0.029***
	(0.002)	(0.002)	(0.002)
Size	0.019***	0.000	0.000
	(0.006)	(.)	(.)
Size^2	−0.003***	0.000	0.000
	(0.001)	(.)	(.)
Word Count	0.002***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)
Sentiment	−0.302***	−0.298***	−0.298***
	(0.006)	(0.007)	(0.007)
Company FE	No	Yes	Yes
Year Fe	No	No	Yes
adj r^2	0.458	0.467	0.467
N	194809	194807	194807

*, **, *** represent significance at the 10%, 5%, and 1% level.

Appendix

A1. Examples of interview reviews that may indicate "ghost jobs"

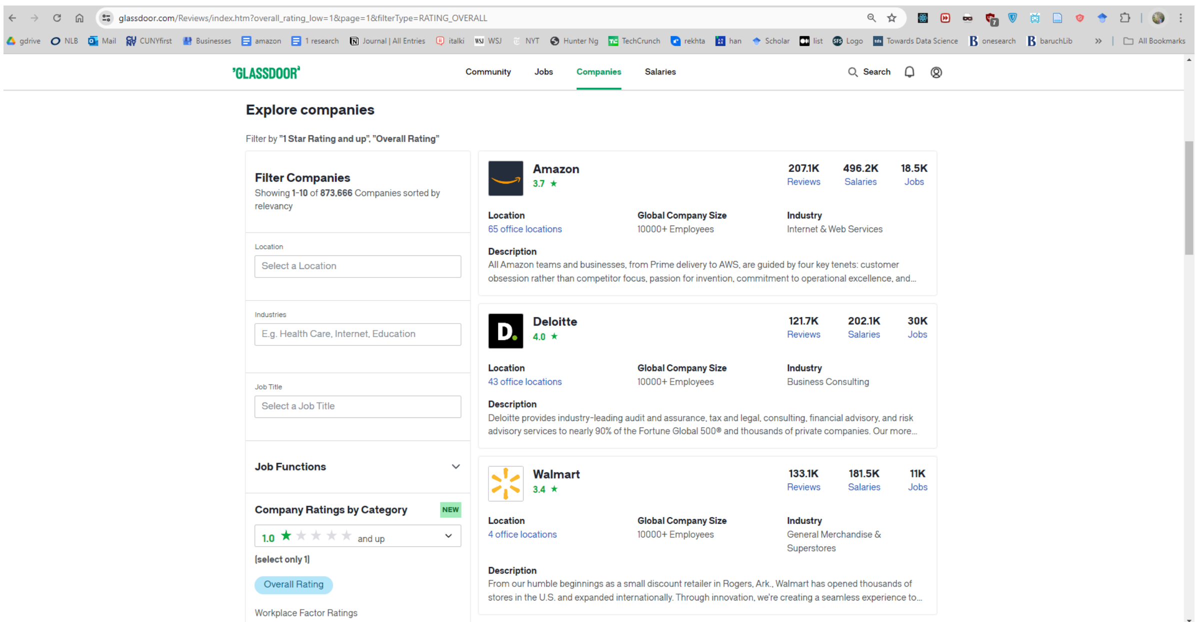

Figure A1. A screenshot of the *Explore Companies* section, where all companies that are listed on Glassdoor are displayed here. There are 873,666 companies that have a overall 1 start or up rating. These are sorted by relevance to the general user and randomly displayed across all industries and company size. I retrieve randomly a representative sample of 1,203 companies. *glassdoor.com*

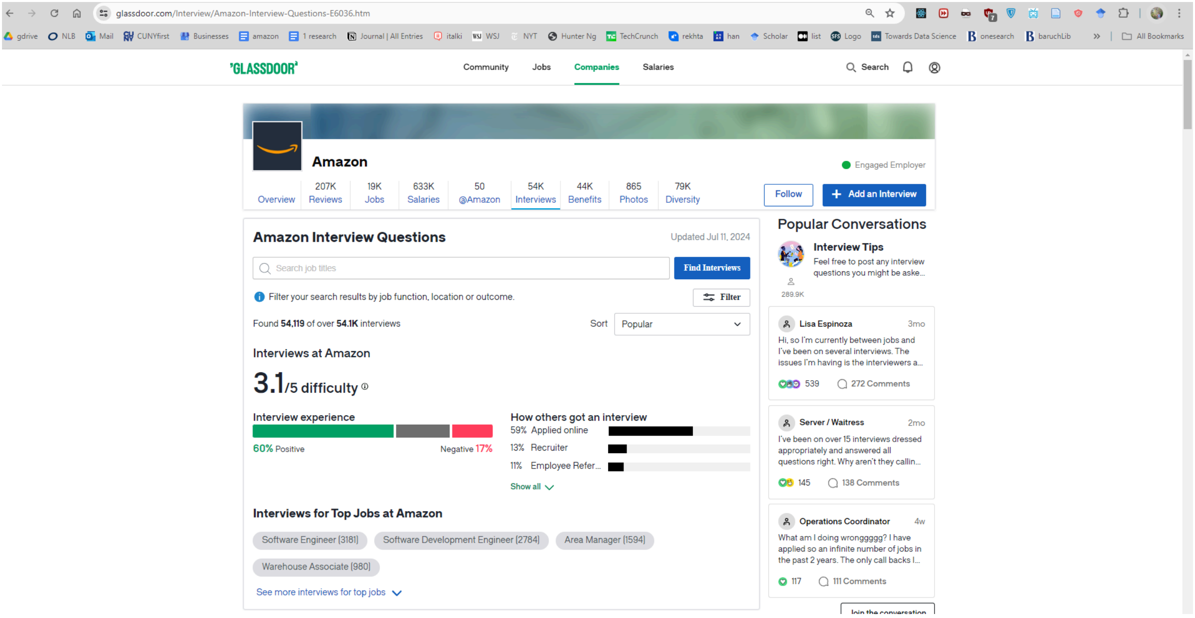

Figure A2. A screenshot of the *Interviews* section, where job-seekers can post their experiences about their job interviews. Glassdoor has stated that these interviews will never be taken down on request by the implicated firm, to signal their commitment to transparency on Glassdoor. *glassdoor.com*

"The recruiter didn't call at our scheduled time, I emailed her asking if that was still a good time for her and no response. 15 minutes later she emails me saying she couldn't reach me, BS. I reschedule for later that day and it was clear why she avoided calling earlier. She was not prepared and had almost no info on the role. She proceeded to tell me that the hiring manager picked our resume out of the bunch, blah blah, all to keep me interested after I told her I didn't want to move forward with more interviews unless it was serious (***they tried filling this role last year and put it on hold bc "they didn't know what they were looking for*** *"-recruiter. Also, I'd been interviewing for months after being laid off and was exhausted). I then met with the hiring manager, I had a great convo and she seemed very excited. I completed the final round and met with the team. I waited a couple days and then reached out to the recruiter to check for updates, and surprise she ghosted me. Not even the decency to respond to a candidate that completed their effing interviews, disgusting."*

Figure A3. An interview experience posted on 12th April 2023 for the position of Product Designer at 2U. The bolded words indicate that this could be a "ghost job" as they could be gathering a pipeline of candidates in anticipation of current employees quitting their jobs.

*"The interview process was pretty quick because at the time I was actively interviewing with another company and was pretty far into interviews. I really appreciated that they came together to get me in quick. As part of their interview process, they DID ask me to spend almost an entire day interviewing with several different people on the potential team and hiring manager. In some ways, it is nice to get to know several people, but on the other hand, as a candidate, **it's exhausting to have interviews every half hour for 5 hours**. That to me seems unnecessary. They did also give me a surprise scenario to test whether or not I knew what I was talking about. Would have appreciated more feedback about how interviews Int. In the end, I was a bit ghosted and **they didn't even update me on having cancelled their position until I reached out to them**."*

Figure A4. An interview experience posted on 10th June 2019 for the position of Change Management Support at ADM. The bolded words indicate that this could be a "ghost job" and HR is getting other departments to interview candidates to possibly get a sense of the market.

B1. Textual Analysis

I first use the *langdetect* library in Python, which is a port of Google's language-detection library and uses statistical methods to identify the language of a given text. It relies on pre-trained language models that use n-grams to make predictions about which language the text is in. Table A1 shows the breakdown.

Appendix Table B1. Breakdown of the number of reviews initially in each language in the dataset. I use only the English reviews - "en" for the rest of the article.

en	269347	pt	4315
fr	3593	es	3148
de	1554	it	1451
nl	373	unknown	292
af	190	ca	36
da	13	cy	11
ro	8	tl	8
so	8	sv	5
no	5	id	5
pl	4	sl	3
fi	3	sk	2
et	2	hr	1
tr	1	vi	1

For the keyword search, I used the following list of words compiled through manual observation and research from the news articles on this phenomenon. I tested with a final list of 192 phrases which could represent a review that was indicative of a ghost job.

Appendix Table B2. List of keywords used for the keyword search to check whether an interview review is indicative of a ghost job.

weren't hiring	position was put on hold	no intention of hiring
position had been filled	no positions were available	collect resumes
not actually hiring	no actual vacancy	just interviewing
fake job post	no real job	not actively hiring
position canceled	position withdrawn	never intended to hire
not filling the position	job ad was a sham	not recruiting

advertisement only	job on hold	no open positions
recruiting freeze	job freeze	no real vacancies
ghosted	always accepting applications	placeholder job
fake listing	fake vacancy	interview without intent to hire
no intent	no intention	job posting expired
phantom job	phony job ad	resume collector
job no longer available	fake hiring process	dummy position
not a real opening	job ad scam	no job to fill
job not real	position never existed	job posting hoax
interviewed but no job	position indefinitely on hold	fake opportunity
never contacted back	no follow-up after interview	interview for non-existent job
job ad fraud	position already filled	job ad deception
no intent to recruit	position no longer open	no job offer
ghosted by employer	interviewed for closed position	interview with no feedback
no vacancies at this time	job listing scam	interviewed but position not available
job application black hole	interviewed but no response	no job at the end
interviewed but never hired	recruiting for non-existent job	job offer never materialized
job application trap	no job interviews	position disappeared
resume gathering	interviewed with no intent to hire	bogus job ad
false job advertisement	fake recruitment process	interviewing for show
job no longer exists	resume solicitation	position on indefinite hold
job ad just for publicity	fake job vacancy	interviewing without hiring
interview for nonexistent position	no plans to hire	not actually recruiting
job interview deception	phantom hiring	position doesnt exist
no position available	interview with no job	no real job offer
interview for phantom position	no plans to recruit	job application scam
no recruitment	dummy vacancy	not recruiting actively
position permanently closed	no active hiring	position not open
sham job ad	no longer hiring	ad with no hiring intention
always hiring without positions	no active vacancies	job freeze in place
recruitment stopped	job ad with no positions	position not existent

application trap	interviewed but nothing	position already taken
job freeze announcement	job ad without positions	no real hiring process
phantom job ad	position withdrawn permanently	ad without actual hiring
interview without positions	application but no job	no active recruitment
interviewed but no roles	no real job posting	position unavailable
resume collection ad	interviewed for non-existent role	job ad permanently closed
position unavailable permanently	no follow-up interview	interview without job offer
no follow-up contact	resume collection without jobs	position filled but still interviewing
no recruitment intention	no roles available	fake job advertisement
no vacancies advertised	ghosted job application	job ad without hiring
phantom vacancy	interview with no open positions	no positions open
dummy job advertisement	job not open	position freeze
no hiring intentions	position closed	interview with no result
no hiring after interview	interviewed without jobs	no recruitment process
no available positions	sham vacancy	fake recruitment
position fake	no follow-up hiring	resume collection for no job
no job positions	vacancy not real	interview but no jobs
interview without vacancy	position on hold indefinitely	job posting scam
no hiring after application	ghost application process	no recruitment after interview
fake job listing	phantom job process	no job hiring
interviewed with no result	no recruitment intention	position closed indefinitely
resume collection scam	fake application process	no follow-up after application
interviewed without vacancy	no job recruitment	application without jobs
fake job role	position not open for hiring	application but no hiring
resume collection for no vacancy		

For the classification of whether a interview review was indicative of a ghost job or not, I employed the keyword search in Table A2, and I also trained a BERT model to the ghost jobs. Since there may be a lot of nuances in the language that could indicate whether a review was a ghost job or not, a deep-learning model may be a better approach. I used ChatGPT-4o, the most advanced AI model currently. I used the prompt -

"Ghost jobs refer to the phenomenon where employers put out job ads but are not actually hiring. I will give a list of interview reviews with numbers. If you are 90% sure the interview review indicates a ghost job, say yes; otherwise, say no. No explanation needed."

on 2,000 randomly selected interview reviews, and then I trained a BERT model to . The specifications for the BERT model are listed in Table A3.

Appendix Table B3. List of specifications for the training of the BERT model

num_train_epochs	3
per_device_train_batch_size	16
per_device_eval_batch_size	16
warmup_steps	500
Iight_decay	0.01
save_steps	1
evaluation_strategy	epoch

B4. Technical Notes on Latent Dirichlet Allocation and AI technique

Latent Dirichlet Allocation (LDA) is used to determine the topics inherent in the interview reviews. LDA is a generative probabilistic model used for topic modeling in natural language processing (NLP). LDA assumes that each interview review is a mixture of several topics, and each topic is characterized by a distribution of words. The generative process involves sampling a topic distribution for each review, then sampling a topic for each word in the review, and finally sampling the word from the topic-specific distribution. The model parameters are then inferred using variational Bayes to reduce the Kullback-Leibler (KL) distance among topics.

ChatGPT-4o achieves accurate text classification through its sophisticated architecture, primarily built on the transformer model introduced by Vaswani et al. (2017). The transformer architecture consists of multiple layers of self-attention mechanisms and feed-forward neural networks, which enable the model to capture intricate relationships and dependencies within the given job interview review.

C1. Informational Gain through Ghost Jobs

1.1 Model Setup

To set up my model, I rely on the costly information acquisition framework (Grossman and Stiglitz, 1980; Morris and Shin, 2002; Sims, 2003; Veldkamp, 2023). I first define several key variables that represent the post of ghost jobs in my empirical study.

1. Company Size: Let a set of companies C be differentiated by size S. The size of a company is denoted by S_i, where S_i can be small S, medium M, or large L.

2. HR Department Size: Let each company have a HR department size denoted by H_i, which scales with company size S_i. Assume $H_i = \alpha S_i$, where α is a positive constant. Thus, larger companies tend to have larger HR departments.

3. Ghost Jobs: Let G_i represent the number of ghost jobs posted by company i. Companies post ghost jobs to acquire information about the job market, denoted by $I(G_i)$, a function of G_i.

4. Cost of Posting Ghost Jobs: The cost function for posting ghost jobs is denoted by $C(G_i, H_i)$. The cost increases with the number of ghost jobs and decreases with a larger HR department. Assume the cost function is given by:

$$C(G_i, H_i) = \beta G_i^2 - \gamma H_i G_i$$

where β and γ are positive constants. βG_i^2 represents the increasing marginal cost of handling more applications, and $\gamma H_i G_i$ represents efficiency gained from a larger HR department.

5. Information Value: The value of information obtained from ghost jobs is denoted by $V(I(G_i))$, which is an increasing but concave function:

$$V(I(G_i)) = \delta \ln(1 + G_i)$$

where δ is a positive constant representing the value derived from the information gathered.

1.2 Firm's Optimization Problem

Each company seeks to maximize its net benefit from posting ghost jobs. The net benefit function $B(G_i)$ for company i is:

$$B(G_i) = V(I(G_i)) - C(G_i, H_i)$$

Substituting the functions, we get:

$$B(G_i) = \delta \ln(1 + G_i) - (\beta G_i^2 - \gamma H_i G_i)$$

To find the optimal number of ghost jobs G_i^* that maximizes $B(G_i)$, I take the first-order condition by setting the derivative of $B(G_i)$ with respect to G_i to zero:

$$\frac{dB}{dG_i} = \frac{\delta}{1 + G_i} - 2\beta G_i + \gamma H_i = 0$$

$$\delta = (1 + G_i)(2\beta G_i - \gamma H_i)$$

Solving,

$$G_i^* = \frac{-(\gamma H_i - 2\beta) + \sqrt{(\gamma H_i - 2\beta)^2 + (8\beta(\delta + \gamma H_i))}}{4\beta}$$

1.3 How does number of ghost jobs posting varies with company size?

Small Companies S: Small companies have a smaller H_i. Thus,

- γH_i is relatively low.
- $\sqrt{(\gamma H_i - 2\beta)^2 + (8\beta(\delta + \gamma H_i))}$ is dominated by 2β.
- Smaller G_i^* because the marginal cost of handling additional ghost jobs outweighs the benefits.

Medium-Sized Companies M: Medium-sized companies have a moderate H_i. Thus,

- γH_i is larger than for small companies but not as large as for big companies.
- The term $(\gamma H_i - 2\beta)^2$ may be close to zero, reducing the negative impact in the quadratic solution, thus increasing G_i^*.
- The balance between information acquisition benefits and handling costs is optimal, resulting in a higher G_i^*.

Large Companies L: Large companies have a large H_ii. Thus,

- γH_i is large.
- $\sqrt{(\gamma H_i - 2\beta)^2 + (8\beta(\delta + \gamma H_i))}$ is affected more by γH_i.
- Marginal value of additional information is low, leading to a smaller increase in G_i^* compared to medium-sized companies.

1.4 Comparative Statics

Effect of HR Department Size H_i:

- When H_i is small (small companies), the term γHigamma H_iγHi is small, making G_i^* low.
- When H_i is moderate (medium-sized companies),G_i^* is maximized because costs are balanced, and information value remains significant.
- When H_i is large (large companies), the costs are lowered, but diminishing returns on information value cap the increase in G_i^*.

1.5 FOC Conclusion

By solving the first-order condition, I show that:

- Medium-sized companies post more ghost jobs than both small and large companies. This is because they have sufficient HR capacity to handle the costs associated with ghost jobs but still derive significant value from the additional information.
- Small companies post fewer ghost jobs due to limited HR resources and high marginal cost
- Large companies also post fewer ghost jobs because they face diminishing marginal returns on information due to other ways of acquiring market intelligence

D1. Variable Definitions

Appendix Table D1. Variable Definitions. This table presents definitions of the variables used in the paper.		
Variables	Definition	Source
$review_{i,t}$	Interview Review, based on company i and time t	glassdoor.com
$experience_{i,t}$	Sentiment of the Interview Review, whether it is good, neutral of negative, based on company i and time t	glassdoor.com
$word\_count_{i,t}$	Length of words of Interview Review, based on company i and time t	glassdoor.com
$industry_{i,t}$	Industry of the company revieId for the job interview based on company i and time t	glassdoor.com
$skill_{i,t}$	Whether the job is a highly skilled, professional or managerial role. Takes a boolean of 1 for highly skilled and 0 otherwise based on company i and time t	glassdoor.com