Beyond Compensating Differentials: Non-Wage Amenities, Sorting, and Efficiency

Xuanli Zhu

April 25, 2025
Latest Version Here

Abstract

This paper studies the supply of non-wage compensations across different types of firms and jobs, as well as their impact on wage determination. Taking advantage of the data from a Chinese online job board where most firms document pecuniary and nonpecuniary amenities to attract applicants, we extract a large set of non-wage compensations that employers and jobseekers recognize as essential for their matching in the labor market. We find that different firms in different jobs also provide different non-wage compensations in a systematic way. In particular, high wage-premium firms sorted with high skill jobs also more likely provide advanced insurance packages, backloading wage and stock options, professional coworkers, and flexible work-time, and such amenities are positively correlated with posted wage. In contrast, low wage-premium firms sorted with low skill jobs more likely to offer weekend, holiday, and regular work-time, and such amenities are strongly subject to compensating differential. To account for these findings, we propose a new theory that combines the compensating differential with two new forces: efficiency compensation and (productivity-based) firm-worker sorting. This framework not only reconciles the inconsistency between our data and the incumbent theories but also offers important general insights into the labor market inequalities.

Keywords: job amenities, compensating differential, wage inequality, firm-worker sorting

1 Introduction

Labor economists, over the years, have acknowledged the significant role of non-wage compensations play in wage determination within the labor market. The concept of "compensating differential" or "equalizing differences" encapsulates the simple yet profound idea that firms offering an amenity can provide lower wages than a comparable firm, while still attracting workers. The utility derived by workers from such an amenity can outweigh their monetary loss in the wage, rendering the firm a preferred choice. Conversely, the same logic applies for a disamenity. Classical models of compensating differential propose that variations in compensation provision arise from heterogeneity in both firms' compensation production functions and workers' utility functions concerning these compensations. Once we control for other workers' heterogeneity, which influences wage and may link with other compensations, we should observe the effect of equalizing differences from the wage disparities associated with varying compensations levels among marginal workers. Despite its intuitive reasoning, empirical validations of this theory have often led to mixed results. Labor economists typically find some coefficients in their hedonic regressions support the theory, while others exhibit zero or inverse direction, even when controlling for unobserved worker characteristics through worker fixed effects (Brown, 1980; Rosen, 1986). Early failures in empirical investigations are primarily attributed to the identification problem of unobserved confounding variables. Recently, two new strands of literature have shed fresh light on this topic. The first, through field experiments (Eriksson and Kristensen, 2014; Mas and Pallais, 2017; Maestas et al., 2018) or natural experiments (Wissmann, 2022), confirms the significance of compensating differentials by demonstrating workers' substantial willingness to pay for certain compensations offered by firms. The second strand (Sorkin, 2018; Taber and Vejlin, 2020; Lamadon et al., 2022) employs structural modelling on workers' job moves and wage changes, indirectly affirming the importance of compensating differentials through a revealed preference perspective, thereby emphasizing their role in wage determination and worker mobility. Despite these advancements, considerable gaps persist in our understanding of non-wage compensation provision and its labor market impact. Empirically, we lack direct observations of primary non-wage compensations in the labor market affecting wages and worker's job choices, and we don't fully understand if different types of compensations are systematically provided by different firms for different workers based on heterogeneity in costs or preferences. Theoretically, there remains open questions as to which theories can reconcile the early empirical inconsistencies with the newly founded evidence regarding firm compensation provisions, and if there exists other significant forces that may either compete or interplay with the compensating differential in shaping the wage-amenity relationship in the labor market.

In this study, we aim to fill these gaps first by leveraging a novel dataset to unearth fresh empirical facts regarding non-wage compensation provisions and their correlations with wage differentials. Concurrently, we propose a new theory that extends the classic compensating differential theory with new forces, enabling us to reconcile our findings with theoretical predictions. On the empirical side, we exploit online job vacancy data derived from a Chinese job board, extracting all non-wage compensation terms specified by employers in job advertisements. While the compensations in each job post may not constitute the complete list offered for that specific position, we contend that they are arguably the most important ones—leveraged by firms to attract potential workers or to justify their posted wages through compensating differential.

Consequently, the merit of our dataset is to allow us to discern, from the data itself, the key non-wage compensations that employers and job seekers deem vital, rather than predetermining what compensations to investigate, while preserving the transparency unattainable through many revealed preference methods. Our data reveals a broad variety of non-wage compensations, including both pecuniary and nonpecuniary elements. Another advantage of our dataset is the visibility it provides into detailed job attributes, such as posted wages and job requirements and descriptions, which assist in mitigating potential confounding factor bias. Given the unstructured, high-dimensional nature of our textual data, we employ an array of basic machine learning techniques (word cloud, Lasso regression, and Word-Embedding model), in addition to traditional econometric methods (wage regression and hedonic regression), to aid our analysis.[1] Recognizing potential data limitations and the lack of causal inference tools, our empirical results primarily adopt a descriptive and exploratory stance, drawing out stylized facts, patterns, and correlations, with an aim to foster new theoretical insights.[2]

Our empirical analysis unveils four key findings, which provide important insights on key aspects of theoretical models. First, our data reveal that firms predominantly offer standard non-wage compensations, such as insurance and fund packages, work-time arrangements, additional pay schemes, and workplace attributes. This suggests that the variations in firm amenity-provision cost functions are not as extensive or as exogenous as the compensating differential theories. Instead, these variations are likely endogenous, influenced by the quality/productivity of the worker they recruit. Second, non-wage compensations can predict posted wages, as demonstrated by a Lasso regression analysis, but this relationship appears to be more reflective of their correlation with other wage determinants such as job skills/tasks and firm wage premiums, as revealed by our posted wage regression and subsequent variance decomposition. Moreover, while productivity-based firm-worker sorting contributes significantly to the posted wage variance, the wage variations accounted for by non-wage compensations, potentially through the mechanism of equalizing differences, are marginal. This suggests that non-wage comepensations and wages are likely to be co-determined, influenced by job qualities and firm pay policies. Third, we find distinct compensation-provision patterns across different firms and jobs. Specifically, high wage-premium firms with high skilled jobs are more likely to offer of advanced insurance packages, backloading wages, high qualified coworkers, and flexible work-time arrangements, but less inclined to provide leisure such as weekend and holiday, or less overtime. Conversely, low wage-premium firms with low-skilled jobs tend to offer the

[1] We borrow several analysis procedures from our companion paper Zhu (2022), where we utilize the same dataset to investigate the determinants of posted wage inequality without considering the impact of non-wage compensations.

[2] The primary data limitation, as previously mentioned, is that non-wage compensations listed in job posts may not fully encompass the entire array of benefits or disamenities a firm might offer for a given position. Certain compensations may be omitted by firms if they are considered standard or trivial. Moreover, firms could strategically withhold information about job disamenities. This latter point, however, could be nuanced, as workers' preferences can vary significantly, and what some perceive as disamenities could be amenities for others, and vice versa. Even when there's consensus on what constitutes an amenity or disamenity, their impacts on utility can still differ greatly among workers. These variances could incentivize firms to disclose both types of information to facilitate optimal worker-firm matching. In our analysis, we find few systematic disamenities, which could also potentially be attributed to the nature of the office-based, cognitive jobs in our dataset. Another potential concern pertains to that firms possibly engaging in "cheap talk" in their job posts, claiming certain benefits without actual provision. However, such behavior would likely generate similar sets of amenities, while our results demonstrate distinct provision patterns among firms, thereby largely mitigating this concern.

opposite. This again hints at a pivotal mechanism of compensation provision associated with the quality/productivity of firms and workers . Lastly, our hedonic regression yields systematically mixed results of compensating differentials for compensations provided by different types of firms in different jobs in a systematic way. Specifically, compensations more likely provided by high-pay firms in high-skilled jobs are not compensated from, but rather positively correlated with, posted wages. In contrast, those more often provided by low-pay firms in low-skilled jobs show a negative correlation, thus being significantly compensated through the posted wage. This stylized finding underscores that the empirical inconsistencies commonly observed in empirical tests of compensating differential are linked to provision patterns and influenced by the heterogeneity in firm and work productivity. We discuss that these empirical findings mark a substantial deviation from the canonical compensating differential framework where sorting is entirely based on heterogeneity in firms' provision costs and workers' preference, and the wage discrepancies emerge solely from the force of equalizing differences.

The potential impact of unobserved worker abilities, a longstanding concern as we describe in the beginning, may understandably be suspected as a source of our mixed results on compensating differential. However, we suggest this concern might be disdirected due to two key reasons. First, our estimations account for the detailed job skills and tasks often unnoticed in standard datasets, reducing the impact of unobserved characteristics. Second, and perhaps more importantly, we argue that the persistent difficulty in finding empirical evidence of compensating differential actually points to important but overlooked wage-amenity variations in the labor market. These variations present wage-amenity dispersions that are orthogonal to the dispersions generated by the force of equalizing differences, suggesting another significant and competing force influencing the wage-amenity relationship. If this alternative force, which establishes a positive wage-amenity relationship, dominates in the market, it could substantially restrict workers' ability to equalize differences based on their preferences, and leads to large labor market inequality independent to the force of compensating differentials. Existing theories, such as an income effect reasoning Hwang et al. (1992); Mortensen (2005) or an extended Burdett-Mortensen model Hwang et al. (1998), can help to explain this positive wage-amenity correlation. However, these theories fall short in accounting for certain aspects of our empirical findings in the divergence in compensation provision, and they often abstract from productivity-based firm-worker sorting.

In light of our empirical findings, we propose a novel theory that marries the compensating differential with two additional elements, efficiency compensation and firm-worker productivity sorting, and show that it helps reconcile all stylized facts. Here by "efficiency compensation", we are borrowing the idea from the efficiency wage literature and refer to any efficiency (or inefficiency) effects pertaining to non-wage compensations, such as inducing workers' effort or reducing labor turnover costs.[3] This theory illustrates that if we allows that many non-wage compensations can either enhance or hinder production or firm operations, there will be an additional efficiency channel along with the traditional equalizing differential channel when firms decide their compensation levels. This new efficiency channel has the potential to either counteract or amplify the compensating differential mechanism, with its impact contingent

[3] The microfoundations underlying the (in)efficiency effects can be exactly the same as those posited in the efficiency wage theory, such as eliciting effort, reducing exogenous or endogenous labor turnover costs, etc. We argue that this is in fact a more inherent property of non-wage compensations than monetary wage, as suggested by some critiques on the efficiency wage theory.

upon the extent of productivity-based firm-worker sorting. For efficient compensations such as advanced insurance or backloading wages, the efficiency effect generates a productivity gain that offsets the force of equalizing differences. This efficiency effect is particularly pronounced for high-productivity firms matched with high-productivity workers, and when it dominates, it not only incentivizes those firms to provide superior compensations but also enhances, rather than reduces, their workers' wages. In contrast, for low-productivity firms matched with low-productivity workers, the efficiency effect is more subdued, discouraging the firm from offering efficiency compensations. In certain situations where an efficiency compensation is universally provided due to say regulatory requirements, its costs may not be fully offset by efficiency benefits in those low-productivity positions, leading to a wage reduction as being compensated, even though the provided amenity is inferior to those offered in high-productivity positions. On the other hand, the provision of inefficient compensations, such as leisure or light workloads, induces an efficiency loss through the (in)efficiency channel. This loss is considerably larger in high-productivity firm-worker matches, leading to the primary provision of these amenities by low-productivity firms in low-productivity jobs, as they are more easily to bear the cost and get compensated from wage discounts. Our new theory thus offers flexible patterns of compensation provision and its wage impact, yielding crucial implications for understanding labor market inequality. It elucidates why, in some cases, a well-identified hedonic wage regression could yield results inconsistent with the predictions of the classical compensating differential theory. It also provides insights into how non-wage compensation provision might generate important influence in the overall compensation inequalities in the labor market and helps identify key types of compensations that can justify certain job mobilities in the labor market. It underscores the wide-ranging impacts of non-wage compensations beyond traditional compensating differentials, and has potential implications for various labor market issues.

Related Literature. Our research builds on and contributes to the literature of compensating differentials and compensation provision, as exemplified in the classic paper by Rosen (1986). According to Rosen's framework, firms offer various levels of amenities or disamenities based on their cost functions to offset wage costs, while workers choose wage-amenity packages based on their preferences to maximize their utility. Despite the intuitive appeal and straightforward predictions of this theory, early empirical studies using hedonic wage regressions often produced mixed results, even when controlling for worker fixed effects (Brown, 1980; Rosen, 1986). This empirical failure has been attributed to the bias stemmed from unobserved confounding variables (Hwang et al., 1992; Bell, 2022) and been theoretically rationalized by introducing labor market frictions into job search and matching (Hwang et al., 1998; Mortensen, 2005; Lang and Majumdar, 2004; Bonhomme and Jolivet, 2009). Two recent lines of inquiry have revitalized this topic. The first involves empirical studies that use experimental or quasi-experimental methods to identify wage effects of certain types of compensations in specific contexts (Eriksson and Kristensen, 2014; Mas and Pallais, 2017; Maestas et al., 2018; Wissmann, 2022). The second approach studies the labor market by modelling unobserved compensation as a wedge to justify job moves and wage changes, acting as a justification for those moving to low wage-premium firms with wage loss (Card et al., 2018; Sorkin, 2018; Taber and Vejlin, 2020; Lamadon et al., 2022).

Our research contributes to this extensive literature through two aspects. Firstly, we pro-

vide new empirical evidence regarding firms' non-wage compensation practices and their impact on wage determination by leveraging a new data source—online vacancy data. To our knowledge, we are the first to mine the textual information provided in job advertisements to explore firms' behavior in the provision of non-wage compensation. We uncover a broad spectrum of pecuniary and nonpecuniary compensations, investigating their correlations with other textual job attributes as well as the the posted wage. In doing so, we adopt the estimation method developed in the wage differential literature Abowd et al. (1999); Kline et al. (2020); Lamadon et al. (2022), further enhanced by targeted techniques for job vacancy data introduced by Zhu (2022). Our investigation also reveals distinct compensation provision patterns among firms with varying wage premiums. Specifically, we observe that firms offering higher wage premiums are also more likely to provide an array of other compensations, including superior insurance and fund packages, backloading payments, and workplace with growth potential. In contrast, firms offering lower wage premiums are significantly more likely to offer leisure. These findings are consistent with recent studies by Sockin (2022); Bana et al. (2022) which also report that high-wage premium firms generally provide better amenities in many dimensions. Additionally, we suggest that this observed pattern is largely responsible for the mixed results typically found in hedonic regression testing for compensating differentials. In fact, we find a positive wage-amenity correlation only for those job amenities more frequently provided by high-wage premium firms, contradicting the predictions of compensating differential theory. We propose that this empirical pattern is more likely the result of an alternative force working in opposition to the equalizing differences force in shaping the wage-amenity relationship, rather than being a consequence of unobserved variable bias.

Secondly, we develop a novel theory that reconciles our empirical findings by integrating the elements of efficiency compensation and productivity-based firm-worker sorting with the traditional compensating differential mechanism. The idea of efficiency compensation stems from the classic efficiency wage literature, which proposes that increasing wage can have a direct effect in production efficiency Salop and Salop (1976); Shapiro and Stiglitz (1984); Katz (1986); Krueger and Summers (1988); Bloesch et al. (2021). Indeed, our transplantation is perhaps a more natural application than the original efficiency wage theory, as it echos one of its major critiques: that firms can employ alternative compensation schemes, rather than solely increasing wages, to achieve the same objective. The efficiency of alternative pay schemes has also long been discussed in the organizational literature (Lemieux et al., 2009). Concerning nonpecuniary compensations like fringe benefits, Dey and Flinn (2005) argue that offering health insurance can improve employer efficiency through reducing exogenous worker exit. We align with these earlier works in recognizing the efficiency nature of various non-wage compensations. Despite its simplicity, our model can can generate flexible results on firms' compensation provision and wage-amenity relationships, reconciling all empirical findings from our study and those we mentioned above. We thus complement to the rationalization through incorporating equalizing differences into a job search model with labor market frictions (Hwang et al., 1998; Mortensen, 2005; Lang and Majumdar, 2004; Bonhomme and Jolivet, 2009), which may find it challenging to rationalize some of our findings. Our model, with its efficiency channel that can potentially counteract the compensating differential force, suggests cautious policy implications for willingness-to-pay estimates derived from field and choice experiments(Eriksson and Kristensen, 2014; Mas and Pallais, 2017; Maestas et al., 2018), since provision of non-wage compensations could also create productivity effects,

and access to different wage-amenity packages may be limited depending on worker type and matching. Finally, we demonstrate that our new theory can generate important general implications for other topics of labor market inequality. For instance, it suggests that non-wage compensations could be itself a driver of increasing wage inequality, particularly in contexts of strong productivity-based sorting or between-firm productivity segregation. This resonates with recent studies identifying increased between-firm wage differentials as a significant factor in enlarged wage disparities (Card et al., 2013; Song et al., 2019). Furthermore, our model predicts larger compensating effects than the traditional equalizing differences channel when a worker experience a preference shock, as she might have to compensate not just through lower wages, but also through worse sorting and a poorer set of other non-wage compensations. This prediction could have important implications on wage inequality across different demographics, such as the gender wage gaps (Wiswall and Zafar, 2018; Le Barbanchon et al., 2021).

The remainder of this paper is structured as follows. In Section 2, we utilize vacancy data from China to illustrate the empirical realities concerning the patterns of non-wage compensation provisions present in our data, as well as their correlation with the posted wage. We discuss their implications and argues that the classical theory of compensating differential deviates substantially from these empirical observations. Subsequently, in Section 3, we propose a new theory that extends the traditional compensating differential mechanism by introducing two additional elements: efficiency compensations and firm-worker productivity sorting. We demonstrate that this enhanced theory can generate flexible provision patterns and wage impacts of compensations, thereby reconciling all the stylized facts uncovered in our data and generating important general implications on labor market inequality. We conclude in Section 4.

2 Empirical Facts From Vacancy Data

2.1 Data and Processing

Our main dataset used for the empirical analysis is derived from the online job vacancies posted from a leading Chinese online job board, Lagou.com. Lagou is the most popular online job board in the field of information technology (IT) job listings in China, and attracts a large and national user base, catering both IT-producing and IT-using firms across China. We collected a total of more than six million job posts on the site between 2013 and 2020 and recorded all the information in the job postings that job seekers can observe. Around one third of the job postings pertain to IT-specific occupations, such as IT engineers and programmers. The remaining two-thirds of the job vacancies encompass a wide array of other occupations, ranging from professional or high-skilled occupations like designers, writers, business operation specialists, and financial analysts to less professional or low-skilled occupations such as sales and marketing clerks, administrative staff, and customer service operators. Overall, our data primarily includes a large set of routine and non-routine cognitive jobs, with few manual jobs represented. Consequently, we shift the focus from many early studies in the compensating differential literature that examined disamenities related to harsh working conditions such as

job injury or mortality, to more general and common job amenities offered in modern office jobs. Given the secular trends of automation- and IT-based technological change and structural shifting from manufacturing to service industries all around the world, we believe that the labor market examined in this study is largely representative of present and future labor markets in many countries. After some basic data cleaning, our final sample for the empirical analysis in the next subsection comprises approximately four million job postings from over 86,000 firms. Further details on data collection and cleaning procedures can be found in our companion paper Zhu (2022).

Similar to other online job vacancy datasets that have been extensively used in the literature, our data contains detailed information about both the job and the hiring firm. The main information in the job posting that we will utilize is the full texts of the job title, job tasks, job requirements and other job characteristics, as well as the firm name and the posted wage.[4] Importantly for our purpose here is one type of often overlooked information in the analysis of online job vacancy data—the text information related to non-wage compensations and amenities that firms claim in the job postings to attract potential applicants. One major advantage of our data is that most job postings in our data include this information, partly due to a designated "job benefit" column for firms to complete when posting their job vacancies, and partly because many firms deem it important to write down such information for their recruitment. This direct observation of (at least part of) the non-wage compensations and job attributes offered can provide valuable insight into the composition of popular non-wage job compensations used by firms in the labor market. It thus complements the indirect approaches in the literature that rely on revealed preference to infer the integrated effect of various job compensations and helps to understand the nature and potential mechanisms behind these compensation provisions and their potential policy implciations. However, it is crucial to acknowledge that the compensations listed in a job posting may not fully reflect the entire package of non-wage benefits and amenities offered by the firm for that position. In particular, some compensations might not be mentioned if they are perceived as trivial or standard, and firms could strategically conceal information about job disamenities.[5] Despite fully acknowledging this limitation, we argue that the observed non-wage compensations in our data are likely the most important ones in the labor market, as perceived by the firms attempting to attract their potential workers or to justify their posted wages. Hence, the primary goal of our empirical analysis is to carefully utilize this observed information to explore the patterns of

[4]The availability of both posted wages and detailed information on job characteristics and requirements is crucial for studying wage determinants, as it allows us to study how firms determine their posted wage with minimal impact from unobserved confounding factors, as we suggested in Zhu (2022). This is also of particular importance here, as the unobserved worker abilities have been long regarded as a major concern in the studies of non-wage compensations.

[5]It may not be as straightforward as it seems for firms to strategically hide all the disamenities that could deter potential candidates. Firstly, for some job features, workers' idiosyncratic preferences might vary significantly, meaning that disamenities for some could be perceived as amenities for other, and vise versa. Additionally, even when all workers agree on the classification of certain job features as amenities or disamenities, their utility impacts could still differ considerably. As a result, firms might find both job amenities and disamenities valuable for matching and sorting with their ideal workers, providing an incentive for them to disclose even disamenity information. In fact, this is the sorting pattern that is suggested in the canonical framework of compensating differential in Rosen (1986), and such incentive would be even larger in the presence of search frictions and mismatching. Lastly, given that the context of the jobs in our dataset typically pertains to cognitive jobs in office, the types of systematic disamenities (e.g. job injury or work safety) are very likely to be rather limited.

firms' non-wage compensation provisions in the labor market and their potential relationship with wage determination and earning inequalities. We view our empirical results and findings in the following subsection as primarily descriptive rather than conclusive or casual, but we show that they turn out to be rather stylized and help to shed new lights on the development of theoretical models on the functions and mechanisms of non-wage compensations in the labor market.

Despite the richness of the information in job texts in our data, a significant challenge in our empirical analysis is extracting useful information into a tractable form from the raw texts of job postings, which describe non-wage compensations and other job attributes in natural language, so that we can conduct further analysis and estimations. To this end, we employ the typical method in the field of machine learning and textual analysis to tokenize the entire job texts and the job amenity texts, generating the full vocabulary set of job texts, V, and the vocabulary set of non-wage compensations, $V_{\text{comp}} \subset V$, respectively. The full vocabulary set, V, contains approximately 110,000 tokens or features (i.e. words or phrases, and we will use these words interchangeably hereafter), while the the compensation vocabulary, V_{comp}, includes over 13,000 tokens.[6] Using these tokenized datasets, we then construct the corresponding indicator matrices $\mathbf{C} \in \mathbb{R}^{N \times |V|}$ and $\mathbf{C}_{\text{comp}} \in \mathbb{R}^{N \times |V_{\text{comp}}|}$, where N represents the total number of job posts in our data sample. Each entry c_{ik} in $\mathbf{C}$ or $\mathbf{C}_{\text{comp}}$, with a value 0 or 1, indicates whetherBaB a job post i contains a certain token k in V or V_{comp}. We will use both $\mathbf{C}$ and $\mathbf{C}_{\text{comp}}$, as well as some transformations of them, to conduct the empirical analysis in the next subsection.

2.2 Stylized Facts

Our first stylized fact concerns the key types of non-wage compensations that firms in our data disclose to attract job applicants and/or justify posted wages. Given the large amount of the job postings and the extensive vocabulary V_{comp}, more efficient methods than merely eyeballing are necessary to examine these compensations. The perhaps most straightforward way to distill information from large text data is to create a word cloud based on the occurrence frequencies of tokens in the compensation vocabulary set V_{comp}, as shown in Figure 1. The tokens plotted represent the most frequently used by firms in our dataset, with the size of the tokens indicating their relative frequency. The most common non-wage compensations in our data is "five insurance and one fund," which represents the most common and standard package offered by firms within the Chinese social insurance system. "six insurance and one fund" also appears in the figure, albeit less frequently, and signifies an advanced package that includes an additional commercial insurance. [7] The prominence of insurance and fund

[6] We manually add a set of the common non-wage compensations that are compounded words into the dictionary of our tokenizer to ensure accurate tokenization in the context of Chinese job vacancy texts. Note that 13,000 tokens do not necessarily mean that we have 13,000 types of non-wage compensations, as firms can use different words or terms to indicate a certain type of compensation. However grouping features with similar meanings is not an easy task, and we will only conduct it for a selected set of non-wage compensations in our empirical analysis.

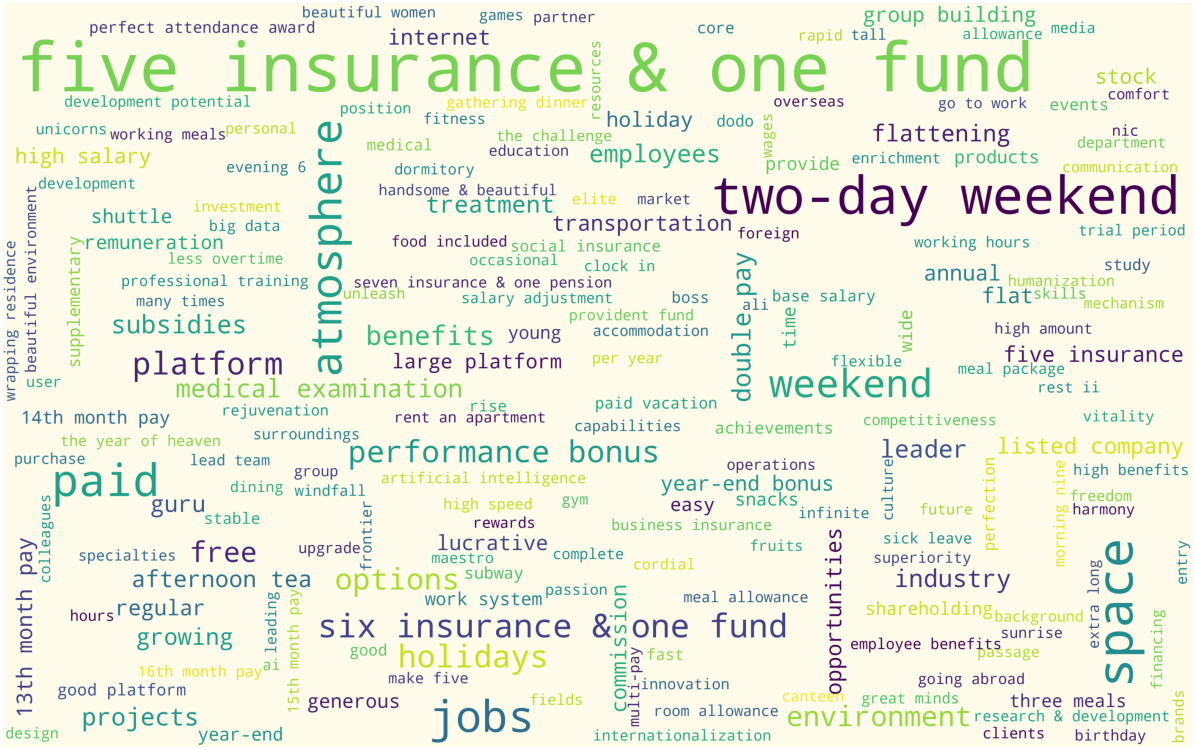
[7] The "five insurance" component comprises endowment insurance, medical insurance, employment insurance, employment injury insurance, and maternity insurance, which are mandated by law. However, enforcement is often limited, particularly in informal sectors and informal employment situations; though things have been improving in recent years. "One fund" refers to the housing provident fund, which is not legally compulsory but

packages is natural and unsurprising given their long-standing importance in the labor market as a key component of employee compensation. The second most frequently used compensation phrase in our data is the "two-day weekend", which is also unsurprising given that working hours and leisure time are among most critical concerns in labor supply decisions. We find other similar tokens related to work-time or leisure availability such as "weekend," "holiday," "flexible (worktime)". Furthermore, we observe frequently mentioned non-wage compensations related to career growth potential (e.g. "space"; "platform"; "opportunity") and workplace environment (e.g. "atmosphere"; "environment"; "flattening (organization)"). Another important group observed in the figure is monetary payments beyond regular pay, which primarily consist of various types of backloading payments such as "13 month pay", "double pay", "bonus", and "stock" or "options". Lastly, despite the relative lower frequency for individual tokens due to their variety, we can also see terms related to fringe benefits like "meal", "afternoon tea", or "fitness". Overall, a quick overview of the common job benefits in our dataset reveals that the types of non-wage compensations firms primarily offer include insurance and fund packages, work-time and leisure arrangements, growth opportunity and workplace environment, backloading payments, and fringe benefits. These findings align with what is generally observed in the Chinese labor market. This initial snapshot helps to provide a direct view of the non-wage compensations that play a crucial role in workers' job choices in the labor market and what might underlie the revealed preferences inferred from workers' job decisions.

Our next step involves running a least absolute shrinkage and selection operator (Lasso) regression of log posted wage on the indicator matrix $\mathbf{C}_{\text{comp}}$. This serves two purpose: (i) to shrink the vocabulary set to a more manageable size, allowing we to eliminate common and neutral words and focus on the most significant features; and (ii) to obtain a preliminary understanding of the correlation between wages and various non-wage compensations. Lasso regression is well-suited for achieving these goals, ad it adds a linear penalization cost function to the Ordinary Least Square (OLS) regression, enabling a tradeoff between minimizing residual sum of squares and the additional cost of incorporating more non-zero coefficients. The non-differentiable spike feature of the linear penalization cost function is particularly suitable for feature selection in text analysis, as it leads to sparse estimators. Following the suggestion for feature selection in the literature (see Gentzkow et al. (2019)), we adopt the Bayesian information criterion (BIC) to gauge the sole pre-determined hyperparameter in the Lasso regression, which governs the level of penalization. The turned Lasso regression compresses V_{comp} into a subset of features with nonzero coefficients estimated, donated as V'_{comp}, comprising a more reasonable size with approximately 800 tokens. In Table 1a, we display tokens with the largest absolute positive and negative coefficients and an occurrence frequency greater than 1 percent. These non-wage compensations thus hold strong predictive power on posted wages and occur quite frequently in our data. Most of the top tokens in Table 1a belong to the board categories that we summarized from our word cloud, and their coefficients are relatively large, indicating strong prediction power for posted wages. However, causal interpretations are prohibited, as these statistical correlations may arise from unobserved variable bias, i.e. there are

is provided by a large percent of formal firms, especially state-owned ones. That's why we can also observe a certain share of firms claiming only "five insurance" but no "one fund" in their job postings. "Six insurance" often includes the five basic insurance and an additional commercial supplementary medical insurance, which is typically offered only by a few "generous" or "high-welfare" firms. In relatively rare cases, we also observe "seven insurance" or "two fund", indicating more advanced insurance or fund support.

Figure 1: Word Cloud of Firm Posted Compensations



Notes. This word cloud figure simply illustrates the major non-wage compensations that firms in our data mention in their job postings. The size of each phrase represents the relative level of the frequencies of each phrases, and the position is just random.

uncontrolled job attributes that impact posted wages and are correlated with the presence of job amenities. In fact, we observe a large number of job amenity tokens with substantial positive coefficients, which contradicts the compensating differential theory and can be attributed to unobserved variable bias.

Table 1: Top Features Selected by Posted Wage Lasso Regression

(a) Lasso Regression on V_{comp}

	Top Positive			Top Negative		
	token	coef	freq	token	coeff	freq
1	14th month pay	.331	.013	five insurance	-.301	.020
2	large platform	.310	.016	commission	-.195	.022
3	three meals	.263	.013	young	-.186	.012
4	technology	.247	.025	easy	-.181	.014
5	guru	.223	.024	training	-.174	.018
6	flexibility	.149	.091	two-day weekend	-.154	.140
7	options	.146	.043	promotion	-.138	.068
8	shuttle	.144	.015	events	-.104	.010
9	remuneration	.124	.015	holiday	-.093	.017
10	six insurance & one fund	.121	.050	holidays	-.092	.046
11	platform	.114	.046	provide	-.084	.012
12	13th month pay	.114	.021	jobs	-.080	.097
13	supplementary	.107	.011	achievements	-.077	.010
14	stock	.099	.017	work system	-.076	.012
15	salary	.099	.025	travel	-.073	.058
16	good platform	.093	.010	entrepreneurship	-.069	.013
17	listed company	.091	.023	five insurance & one fund	-.068	.261
18	high salary	.074	.018	employees	-.066	.029
19	products	.073	.012	time	-.063	.012
20	lucrative	.069	.018	environment	-.062	.038
21	shareholding	.069	.012	double pay	-.055	.032
22	benefits	.068	.035	office	-.047	.018
23	motivation	.063	.016	company	-.043	.050
24	projects	.058	.030	wide	-.041	.012
25	year-end bonus	.057	.042	snacks	-.041	.013
26	team	.050	.108	growing	-.039	.025
27	treatment	.040	.027	transportation	-.029	.021
28	prospects	.039	.024	subsidies	-.028	.031
29	excellent	.039	.013	paid	-.025	.101
30	year-end	.035	.039	dividend	-.014	.010
31	development prospects	.029	.035	overtime	-.013	.014
32	group building	.029	.018	performance bonus	-.011	.044
33	space	.028	.128	opportunities	-.010	.022
34	management	.028	.051	subsidies	-.008	.019
35	flat	.027	.021	leader	-.007	.029
36	year-end bonus	.024	.018	afternoon tea	-.005	.024
37	free	.022	.038	enterprise	-.005	.010
38	atmosphere	.019	.092			
39	internet	.018	.019			
40	benefits	.018	.153			

(b) Lasso Regression on V

	Top Positive			Top Negative		
	token	coef	freq	token	coeff	freq
1	14th month pay	.152	.014	freshmen	-.155	.018
2	three meals	.143	.014	five insurance	-.136	.030
3	large platform	.131	.019	graduates	-.128	.033
4	master degree	.126	.015	vocational major	-.100	.036
5	lead	.107	.041	two-day weekend	-.098	.166
6	c++	.092	.051	vocational college	-.094	.148
7	algorithm	.082	.061	assistant	-.079	.011
8	guru	.082	.028	customer service	-.075	.030
9	famous	.079	.019	social insurance	-.073	.028
10	machine learning	.077	.016	accounting	-.071	.019
11	formation	.076	.013	accommodation	-.067	.016
12	undergraduate	.074	.319	administration	-.067	.027
13	overseas	.072	.026	commissioner	-.063	.011
14	react	.072	.020	taobao	-.059	.015
15	development	.071	.374	assistance	-.058	.164
16	undergraduate	.066	.029	ps	-.056	.029
17	high salary	.063	.028	ltd.	-.056	.012
18	landing	.060	.067	installation	-.055	.020
19	strategy	.057	.047	photoshop	-.052	.039
20	live streaming	.056	.014	careful	-.050	.032
21	listed company	.055	.027	hardworking	-.050	.032
22	large scale	.055	.072	verification	-.048	.011
23	responsibilities	.055	.048	human resources	-.047	.032
24	shuttle	.054	.018	website	-.047	.090
25	finance	.054	.070	any major	-.047	.020
26	six insurance & one fund	.053	.055	humanization	-.046	.012
27	python	.052	.066	excel	-.046	.047
28	director	.052	.022	mandarin	-.045	.027
29	unified recruitment	.051	.042	explanation	-.044	.013
30	hive	.051	.013	young	-.044	.025
31	technology	.049	.285	contact	-.044	.010
32	engine	.049	.017	easy	-.043	.027
33	team	.048	.552	commitment	-.043	.014
34	options	.047	.052	recent graduate	-.043	.029
35	revenue	.047	.019	five insurance & one fund	-.043	.294
36	group	.046	.022	editor	-.042	.042
37	ecology	.045	.012	recruitment	-.041	.057
38	leading	.045	.025	seo	-.041	.010
39	growth	.044	.021	established	-.041	.011
40	stock	.044	.022	computer	-.039	.014

Notes. Table 1a displays the tokens with the largest absolute values of positive and negative coefficients and with an occurrence frequency larger than 1 percent, derived from the Lasso regression using the V_{comp} feature set. Table 1b shows the top positive and negative features from the Lasso regression using instead the entire vocabulary set V, with tokens that are also present in V'comp underlined.

To address the potential bias due to unobserved job attributes, we perform a similar posted wage Lasso regression on the indicator matrix of the entire vocabulary set, **C**. In doing so, we essentially control for all job characteristics that firms document in their job postings, including not only education, experience, and occupation, but also other detailed skills and tasks often uncontrolled by econometricians. This Lasso regression reduces the entire vocabulary set V, containing over 100,000 tokens, into a subset V' with slightly over 3,000 tokens that have non-zero coefficients.[8] We again plot the top positive and negative features in Table 1b, with tokens that are also in V'_{comp} underlined. Although the derived coefficients are still not subject to casual interpretation due to multicollinearity and flexibility under high-dimensional space even without unobserved bias concerns, we can now apply domain knowledge—the knowledge that economists have about wage determinations, such as those from Mincer regressions—to examine whether the signs of some estimators are intuitively sensible. Our estimation results reveal that many non-compensation job attributes, especially those familiar within labor economics literature, exhibit intuitively pausible signs. For example, "master degree" and "undergraduate (degree)" are among the most positive tokens, and indicating a high-level education degree requirement is positively correlated with posted wages. In contrast, "vocational major" or "vocational college" and "freshmen" or "(new) graduates" are among the most native tokens, indicating that lower education and experience requirements are negatively correlated with posted wages. For other job attribute tokens that are not compensations, many are skills and tasks. In general, we observe high-level skills like "lead", "c++", and "machine learning" among the top negative tokens and low-level skills such as "assistant", "customer service", and "administration" among the top negative tokens. These findings suggest that our Lasso regression uncovers statistical correlations with some causal and economics significance.

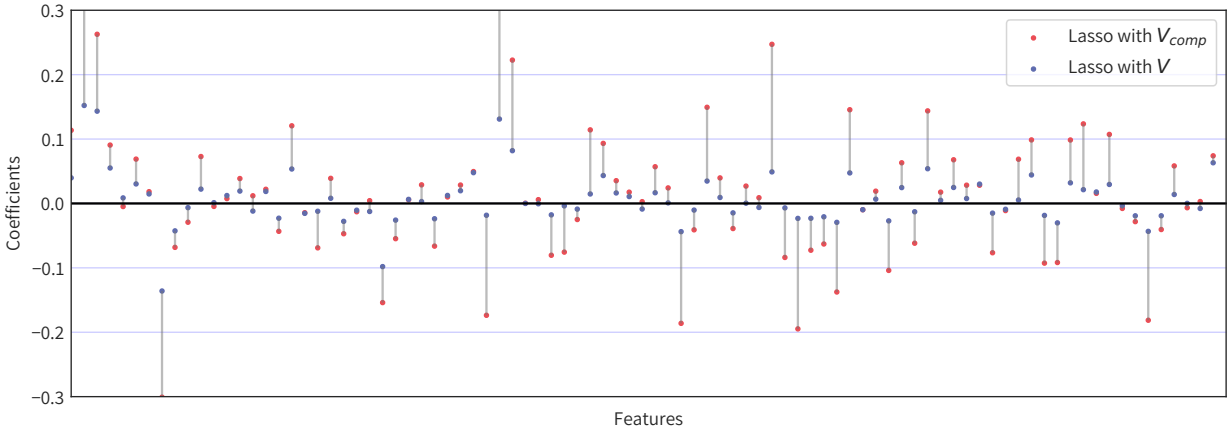
We now revisit our primary focus, those compensations terms captured in V'_{comp}. We find that many of the top tokens obtained from the early Lasso regression with features in V_{comp} also list as top tokens in the Lasso regression with the entire feature set V. Moreover, despite significant decreases in the absolute value of their coefficients, these top compensation features retain the same sign as the initial estimation. In fact, this pattern holds true for most of the nonzero tokens in V'_{comp}, as shown in Figure 2, except for those with coefficients very close to 0 in the first place.[9] These results suggest that the strong predictive power of non-wage compensation provisions on posted wages can be largely attributed to their systematic correlation with job skills and tasks that influence posted wage determination in a more direct way. The decision to provide job amenities, therefore, likely depends on the job type or quality, indicating a connection between firms' provision of job positions and (both wage and non-wage) compensations. Another potentially interesting patterns we observe in the Lasso top features is that some types of non-wage compensations seem to have systematically different predictive power on posted wage. For instance, while the advanced package of "six insurance and one fund" appears among the top positive tokens, basic or low level packages like "five insurance" or "social insurance" are among the top negative tokens. Additionally, tokens related to backloading pay and growth potential often appear among the top positive tokens, while tokens related

[8] Not all tokens in V'_{comp} are included in V'. In our empirical test, over 90 percent of the positive nonzero tokens in V'_{comp} remain in V', while for negative nonzero tokens, only around 70 percent maintain non-zero status.

[9] Figure 2 shows those tokens in V'_{comp} with frequencies greater than 1 percent, while in Figure A1 we show that the same pattern largely holds for all other low-frequency tokens in V'_{comp} as well.

to leisure or relaxation mainly appear in top negative tokens. This finding again suggests the possibility that firms systemically select different non-wage compensations for different jobs with different posted wages. The positive correlations between many non-wage compensations, even after controlling for all job attributes, continue to raise puzzles on why the force of compensating differential does not work for these important and material amenities, though we cannot rule out the possibility that these correlations are purely raised due to their linkage with other job skills and tasks within a high-dimensional context.

Figure 2: Compare Lasso Coefficients Under Different Specifications



Notes. Figure 2 compares the coefficients of tokens in V'comp with frequencies larger than 1 percent, between the Lasso regression using V_{comp} and the one using the entire feature set V. For those tokens in V'comp with lower frequencies see Figure A1.

The connection between the provision of non-wage compensations and job skill and task attributes suggests the possibility of correlations between other wage determinants and amenity provisions. One crucial determinant, as highlighted in the wage differential differential (see the survey in Card et al. (2018)), is firms' wage-setting policies. Briefly, Abowd et al. (1999) (hereafter AKM) pioneered an approach for studying the sources of labor market wage inequalities by incorporating both worker fixed effects and firm fixed effects in a wage estimation under panel data, decomposing entire wage variation into parts attributable to worker heterogeneity, firm heterogeneity, and their sorting. Subsequent estimation by AKM and other papers reveal the significant roles of firm (and firm-worker sorting, see Bonhomme et al. (2020)) in wage inequalities across labor markets in a large amount of countries. To investigate the potential link between firm wage premiums and non-wage compensation provisions, we adopt and extend the estimation approach developed in Zhu (2022). This approach is designed to estimate wage differentials in a manner similar to the AKM approach, but applied to job vacancy data, which lacks individual fixed effects but offers direct access to skills and tasks information. This information is used to generate a comprehensive set of controls for job heterogeneity, effectively mimicking worker heterogeniety. In this study, we extend the approach by incorporating a term representing the effects of non-wage compensations. Our estimation also explores the extent and channels through which non-wage compensation information can account for posted wage

variations. Specifically, our log posted wage regression is specified as

$$\ln w_{ijt} = X_i\beta + \psi_j + \mathbf{c}'_{i,\text{comp}}\gamma + \iota_t + \epsilon_i \tag{1}$$

, where w_i is the mean of the posted wage range for each vacancy i, ψ_i represents the firm fixed effects that capture time-invariant firm pay premiums, ι_t corresponds to the year effects, and ϵ_i is the error term. The term $\mathbf{c}'_{i,\text{comp}}$ refers to the row in the non-wage compensation indicator matrix $\mathbf{C}'_{\text{comp}}$ corresponding to job vacancy i, which contains non-wage compensation tokens in V'_{comp} that are selected by our Lasso regression. The vector X_i encompasses all job attributes, including $\{\text{EDU}_i, \text{EXP}_i, \mathbf{c}'_{i,\setminus\text{comp}}\}$, where EDU_i and EXP_i are dummy variables for the levels of the education and experience requirements in the job i, respectively, and $\mathbf{c}'_{i,\setminus\text{comp}}$ represents the corresponding row of all job attributes of job vacancy i in indicator matrix $\mathbf{C}'$ of Lasso selected tokens in V', excluding non-wage compensation tokens in V'_{comp}.[10] In Zhu (2022), we demonstrate that $\mathbf{c}'_{i,\setminus\text{comp}}$ captures a substantial amount of detailed and often unobserved job skills and tasks, which play an important role in wage differentials in addition to education and experience. Consequently, the entire job controls X can effectively replace the worker fixed effects used in the AKM approach, yielding estimation results well consistent with those found in the literature.

After estimating the posted wage regression in Equation (1), we can conduct the following decomposition on the posted wage variances:

$$\begin{aligned}\text{var}(\ln w_i) = &\text{var}(\theta_i) + \text{var}(\delta_i) + \text{var}\left(\psi_j\right) + \text{var}(\epsilon_i) \\ &+ 2\,\text{cov}\left(\theta_i, \psi_j\right) + 2\,\text{cov}(\theta_i, \delta_i) + 2\,\text{cov}\left(\psi_j, \delta_i\right)\end{aligned} \tag{2}$$

, where we denote $\theta_i \equiv X_i\beta$ and $\delta_i \equiv \mathbf{c}'_{i,\text{comp}}\gamma$ for notation convenience. In other words, we can statistically decompose the overall posted wage variations into primary components: variations in job qualities, variations in firm wage premiums, and variations of the compensation provisions, alongside all the covariance terms of these three heterogeneities and an additional error term. The results of the variance decomposition in Equation (2) on our data are presented in the first column of Table 2. Job heterogeneity in skills and tasks constitutes the largest source of posted wage differentials, accounting for 44 percent of the total wage variance. Firm heterogeneity in wage premiums also plays an important role, explaining 13 percent of posted wage variance, with sorting between job and firm further accounting for 14 percent. These figures align with estimation results from the set of recent studies that employ the AKM framework and linked employer-employee panel data across various OECD countries (see Bonhomme et al. (2020)). However, the variance of non-wage compensations, represented by $\text{Var}(\delta_i)$, explains a mere 0.4 percent of the total posted wage variance. In contrast, non-wage compensations contribute significantly more to the wage differentials through their covariances with the job and firm effects, though the levels remain relative low, with less than 2 percent and 1 percent, respectively. Consequently, the significant predictive power of non-wage compensations on

[10] In practice, we employ an additional dimensional reduction method to map the still-high-dimensional matrices of $\mathbf{C}'_{\text{comp}}$ and $\mathbf{C}'_{\setminus\text{comp}}$ into lo-dimensional representations to save memory and facilitate computation. However, the results remain qualitatively unchanged with or without this dimensional reduction procedure.

posted wages likely arises not only from their correlation with job attributes like job skills and tasks, as demonstrated in the earlier comparison of two Lasso regressions, but also from their correlation with firm wage premiums. In fact, when calculating correlations between three core terms, we find significant positive correlations between δ_i and job effect θ_i (0.19) and firm effect ψ_j (0.17), along with a substantial correlation between job effect and firm effect (0.29). This strong correlation is further evidenced by comparing the results when assuming a specification of posted wage regression without the non-wage compensation terms, as shown in the second column of Table 2. In the absence of the δ_i terms, the total explained wage variance declines by only 0.3 percent points, while the shares accounted for by the job effect and firm effect increase by 1.3 percent and 0.8 percent, respectively, largely replacing the portions attributed to the δ_i variances and covariances. Thus, there is overlapping between the variations in non-wage compensation provisions and those in job qualities and firm wage policies, making it statistically challenging to distinguish their impacts on the posted wage differentials. However, since the overall explanatory power of non-wage compensations on wages appears to be limited after accounting for job and firm characteristics, it suggests that non-wage compensations are likely to play a minor or indirect role in influencing wage differentials compared to those direct influencers. We will further discuss about the interpretations of our findings here, including the observed positive (rather than negative) correlations, from a theoretical perspective in Section 2.3.

Given our observation that non-wage compensations provisions as a whole are significantly correlated with job and firm qualities, it is natural to subsequently investigate how exactly specific sets of non-wage compensations are provided by different firms for different jobs. This analysis offers a direct and clear understanding of compensation provision patterns across firms and jobs ,and helps facilitate our later discussion on evaluating if various theoretical arguments hold intuitively in our data. To this end, we select a set of compensation topics and examine the occurrence patterns for across different firms and jobs types. Our selection is based on our earlier word cloud and Lasso regression results, the clarity of the related tokens in the natural language context (i.e. if well-defined and without multiple meanings), and the importance in labor economics literature. Specifically, we select eight types of non-wage compensations: basic insurance, advanced insurance, backloading wage, stock and options, coworker quality, training, leisure and fixed work-time, and work-time flexibility. These compensations are well-established job amenities in labor economics and possess clearly defined and commonly used categories, as observed in our earlier results. To identify all tokens related to each compensation type in our compensation vocabulary set V_{comp}, we train a natural language processing (NLP) model—the word embedding model—based on all the job vacancy texts. This model captures the in-text relationships among tokens and locates all the tokens in an embedded space with meaningful proximity metrics. We utilize this embedded space to identify all terms in V_{comp} that are relevant to each compensation type since related tokens tend to cluster together in the constructed embedded space.[11] With the token sets of the eight compensation

[11] In practice, this is done by selecting one typical token for a certain compensation type and then examing the cluster of tokens with a small Gaussian distance in the constructed embedding space. Within this cluster we often find synonyms or different versions for a certain compensation type. For examples, we can find all types of insurance and fund packages in the market, and then classify them into two types: those that state the basic or mandated level and those that state advanced or premium packages. Similarly, we can find all tokens related to leisure and fixed work-time, which turns out to encompass a large set of synonyms and similar terms. In contrast,

Table 2: Decomposition of Posted Wage Variance

	With δ		Without δ	
	Comp.	Share	Comp.	Share
$\text{Var}(\ln w)$	.362	-	.362	-
$\text{Var}(\theta_i)$	.158	.437	.163	.450
$\text{Var}(\psi_j)$	.046	.128	.049	.136
$\text{Var}(\delta_i)$	.002	.004		
$\text{Var}(\epsilon_i)$	.097	.269	.098	.272
$2\,\text{Cov}(\theta_i, \psi_j)$	.049	.137	.052	.142
$2\,\text{Cov}(\delta_i, \theta_i)$	.006	.017		
$2\,\text{Cov}(\delta_i, \psi_j)$	.003	.008		
$\text{Corr}(\theta_i, \psi_j)$	.289		.288	
$\text{Corr}(\delta_i, \theta_i)$	.193			
$\text{Corr}(\delta_i, \psi_j)$	.174			
Obs	3998840		3998840	
Firm	86165		86165	

Notes. The first column presents the variance decomposition (both components and shares) based on Equation (2), with the total posted wage variance being attributed to the variances of job effect θ_i, firm effect ψ_j, compensation effect δ_i, their covariances, and the variance of the error term. The second column displays the results obtained under the alternative specification of Equation (1) without the δ_i term. All variance and covariance terms are computed using the heteroscedasticity correction method proposed by Kline et al. (2020) to address potential limited sample bias arising from the high-dimensional fixed effects context. The findings presented in this table are also robust across subsamples featuring occupations with different skill levels.

types determined, we apply our estimation results from Equation (1) to partition all job vacancies into 10×10 job-firm joint decile cells based on the values of θ_i and ψ_i. We then calculate the occurrence ratio of each compensation type for each cell by determining how many job vacancies in a certain cell contain any terms in the token set of a certain compensation type.

The resulting occurrence distributions are illustrated in Figure 3, with the job and firm deciles ranking from the lowest (1) to the highest (10).[12] For all eight types, we see compensation occurrence rates systematically increase or decrease along either or both two axes of job and firm ranks, though there are distinct trend patterns among different types of compensations. In particular, for advanced insurance, backloading wage, stock and option, coworker quality, and work-time flexibility, we observe that the occurrence increases in both the level of job effect and the level of the firm effect, although the extent to which each effect matters varies across compensation types. Conversely, for basic insurance and leisure and fixed work-time, their occurrences in job vacancies decrease significantly with both firm effect and job effect, and for training, the occurrence reduces strongly with job effect but with ambiguous impact of firm effect.[13] In other words, our results suggest that high-pay firms in high-skill jobs are more likely to provide better insurance and fund package, non-wage pecuniary compensations like backloading wage and stock option, as well as nonpecuniary workplace amenities such as better coworkers and flexible worktime. On the other hand, low-pay firms in low-skill jobs more often mention basic insurance, leisure and fixed work-time, and training as the amenities.[14] Taken together, it is this distinct provision pattern, potentially also applied to other tokens in V'_{comp}, that enables the non-wage compensation provision information to predict posted wages through their association with job and firm attributes. A natural question that arises is why do different firms in different jobs provide diverse non-wage compensations. We will discuss this question by adopting theoretical perspectives in the next subsection.

Our final empirical analysis is to test the potential wage effects of the eight compensation

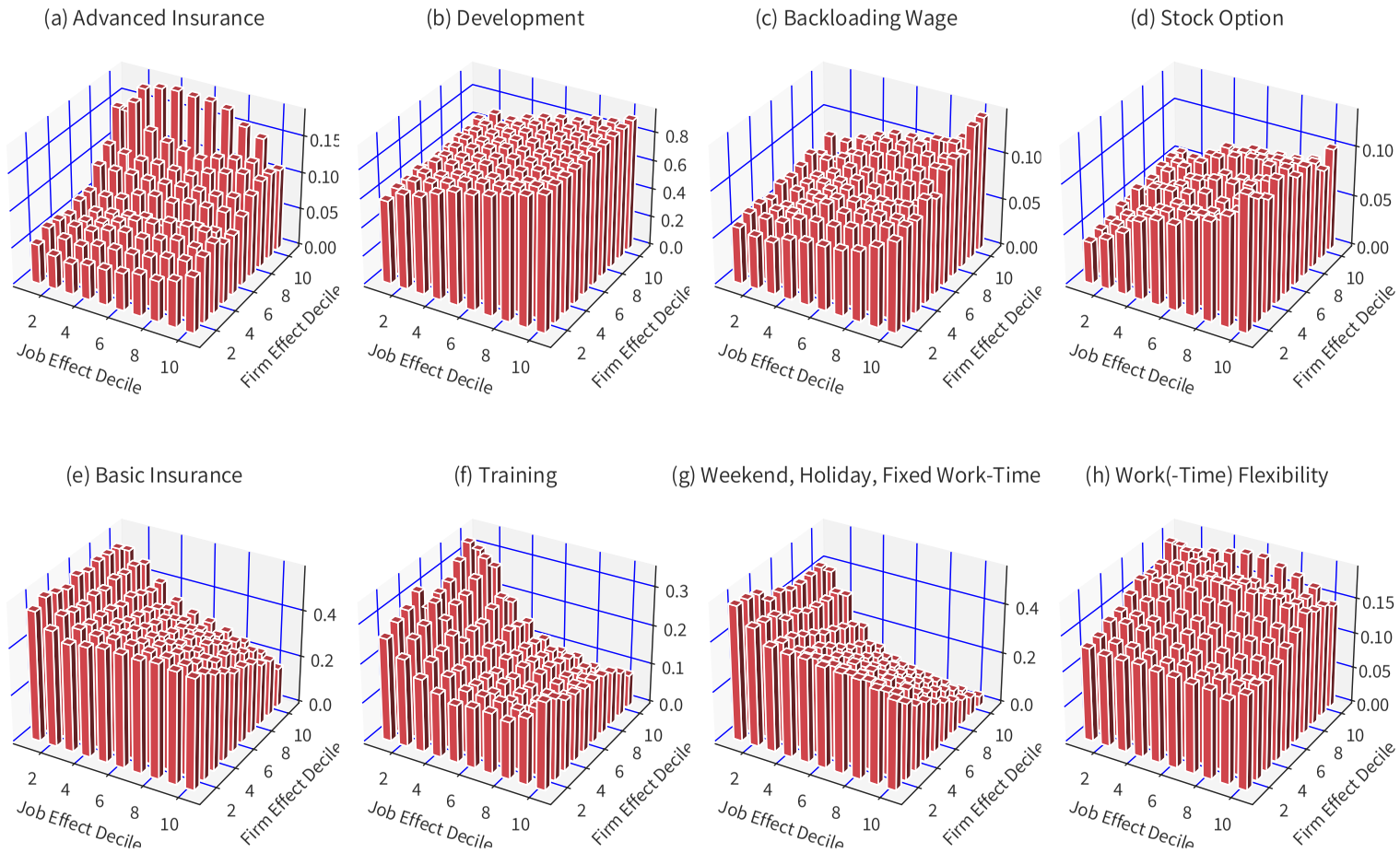
for some compensation types such as stock and options or work-time flexibility, the relevant set of tokens is rather limited and easily identified.

[12] The same patterns generally hold when we examine subsamples of board occupations with high, medium, or low skill levels, as shown in Figure A3. Additionally, Figure A2 displays the provision patterns for eight more compensation types within the major compensation genres observed in our data. Although these compensation types are relatively less well-defined in natural language contexts or of lesser economics importance (for example those amenities of fringe benefits) we still observe systematic differences in compensation provisions across firms and jobs.

[13] The non-monotone relationship between job effect and training occurrence is because our method cannot distinguish that if the training terms mentioned in job text indicate receiving training or offering training. Actually after checking the raw data we find that the increase in training occurrence in the top deciles of job effect is completely due to these high-skill jobs require tasks of offering training to other workers in the firm. Although we can resolve this problem by applying more advanced NLP model to our text data, we argue that such case is relatively rare in our vacancy text data, and thus we stick with simpler method.

[14] It is important to note that our results do not necessarily imply that better firms with better jobs are less likely to provide basic insurance and fund packages. This is because such firms are more likely to offer advanced insurance package and thus, will not mention the basic package in those case. Additionally, given that the basic insurance package is compulsory and well-implemented in high-welfare firms, those firms may not consider it as an attractive compensation for their potential workers, and thus may not mention it even when they are actually providing it. We don't think a similar argument applies to the work-life balance because there are a large number of anecdotes on long working hours in many big and well-paid firms, and the income effect would suggest that higher income workers should prefer at least not less, if not more, leisure. Consequently, high pay premium firms should find offering more generous leisure and work-time useful for attracting workers if they wish to do so.

Figure 3: Compensation Occurrence Across Job and Firm Ranks



Notes. Job effects and firm effects here are the ones estimated from Equation (1). The occurrence ratio is calculated as the percentage of job vacancies in each job-firm cell whose vacancy texts contain any of the terms related with a certain type of compensation. For each compensation type, we gather all the relevant tokens in V'_{comp} by training a word-embedding model on all job texts and checking for tokens cluster together in the embedding space. We show the patterns for some other compensation types in Figure A2.

types we selected earlier. Specifically, we run a hedonic regression by modifying the specification in Equation (1), with $\mathbf{C}'_{\text{comp}}$ in δ_i now replaced by the indicator matrix for the eight types of non-wage compensations. In doing this, we are effectively testing the idea of compensating differential, as has been done in the literature. While the literature often addresses concerns about unobserved worker ability bias by using worker fixed effects and panel data, we once again employ our full set of controls for documented skills and tasks $\mathbf{c}'_{i,\backslash\text{comp}}$ to account any job heterogeneity. In other words, our hedonic regression controls for nearly all the information documented in the job vacancy, and examines how the closure of different non-wage compensations might impact posted wages. We also present the results from three specifications of different sets of controls, allowing us to confirm robustness and compare coefficients estimated from different sources of data variations. The estimated results are shown in Table 3. All coefficients for the eight compensations are significant in all three specifications, suggesting that these compensations do play a role in determining posted wage determination. The estimated coefficients decline substantially after including controls for detailed job skills and tasks and introducing firm fixed effects, consistent with our earlier finding that their provisions are linked with job and firm quality or productivity. Perhaps more interestingly, we find that these coefficients exhibit mixed signs in a systematic manner. Specifically, for those compensations more likely to be provided by high wage-premium firms in high-quality jobs (i.e. advanced insurance, backloading wage, stock and option, coworker quality, and work-time flexibility), the coefficients are significantly positive across all three specifications. In contrast, for those compensations such as basic insurance and leisure and fixed work-time, which are more likely to be offered by low wage-premium firms in low-skilled jobs, their coefficients are significantly negative in all three specifications.Taken at the face value, these results indicate that amenities provided by high-wage-premium firms in high-skilled jobs not only are not compensated for, but actually result in increase in posted wage. On the other hand, amenities provided by low-pay-premium firms in low-skilled jobs are significantly offset by their posted wages. In summary, our hedonic regression, with detailed job controls, aligns with previous empirical studies: compensating differentials work in some cases, but in other cases, we see the exact opposite, creating puzzles for the theory. Notably, our study offer additional insights by demonstrating that these diverse outcomes may arise from the specific provision patterns across different jobs and firms.

2.3 Summary and Discussion

In this subsection, we first summarize our empirical findings and then discuss their implications on the literature of non-wage compensations in terms of both theories and empirics. Combining theory with empirics is crucial for interpreting our empirical findings, as any interpretation beyond statistical patterns and correlations depends on the economic model one has in mind regarding wage differentials, compensation provision, and worker-firm sorting. Our primary theoretical reference will be the classical framework of compensating differential in Rosen (1974, 1986), as it serves as the baseline model guiding both empirical examination and theoretical development. In this framework, compensating differential is the sole mechanism in the labor market for wage differentials, and workers and firms sort entirely on workers' heterogenous preferences for various job amenities and disamenities and firms' heterogenous

Table 3: Hedonic Regression on Selected Compensations

	(1)	(2)	(3)
Advanced Insurance	.117**	.087**	.014**
	(.001)	(.001)	(.001)
Backloading Wage	.054**	.030**	.010**
	(.001)	(.001)	(.001)
Stock Option	.114**	.058**	.087**
	(.001)	(.001)	(.001)
Coworker Quality	.140**	.059**	.024**
	(.001)	(.001)	(.001)
Work-Flexibility	.046**	.032**	.010**
	(.001)	(.001)	(.001)
Basic Insurance	-.062**	-.046**	-.025**
	(.000)	(.000)	(.000)
Training	-.057**	-.012**	-.003**
	(.001)	(.001)	(.001)
Work-Time	-.113**	-.081**	-.021**
	(.001)	(.000)	(.000)
Education FE	✓	✓	✓
Experience FE	✓	✓	✓
Year FE	✓	✓	✓
$\mathbf{C}_{\setminus comp}$		✓	✓
Firm FE			✓
Adj. R^2	.506	.633	.738
No. Obs	3998840	3998840	3998840

Notes. The specification of the hedonic regression is modified from the specification in Equation (1), with $\mathbf{C}'_{\text{comp}}$ in δ_i now replaced by the indicator matrix for the eight types of non-wage compensations. The indicator matrix for the compensation types are constructed in the same way as Figure 3.

costs of producing these non-wage job attributes. We will also discuss alternative frameworks such as the search models of Hwang et al. (1998) and Mortensen (2005) that feature other essential determinants and mechanisms. This discussion illustrates the inconsistencies between assumptions or predictions of existing theories and our empirical findings, emphasizing the need for new theories to reconcile these discrepancies, and also why the estimations of compensating differential might overlook vital aspects of compensation inequalities in the labor market.

To summarize our empirical results, we find that firms in our data document a variety of non-wage compensations and amenities in their job advertisement to attract workers. These include insurance and fund packages, leisure and work-time arrangements, work environment and workplace features, additional and backloading pay schemes, and fringe benefits, all of which are commonly encountered items in the labor market. Although the provision of these non-wage compensations emerges as a significant predicator for posted wages, we discover that this relationship is primarily driven by their strong correlation with other determinants of posted wage, such as job skills/tasks and firm wage policies. Once these factors are accounted for, non-wage compensations provisions only explain a small portion of posted wage differentials, with their impact mainly resulting from the covariances with job and firm effect. When we examine specific compensation types, we find distinct provision patterns across different jobs and firms. In particular, high wage-premium firms and high-skilled jobs are more likely to document the provision of advanced insurance and fund packages, backloading pays, high qualified colleagues, and flexible worktime. In contrast, low wage-premium firms and low-skilled jobs are more likely to document providing basic insurance, training, and leisure and fixed work-time as job amenities. Lastly, we find that in a hedonic regression, the first amenity group, those provided by high-pay firms in high-skilled jobs, are not compensated from posted wage but are positively correlated with them. In contrast, the second group, those provided by low-pay firms in low-skilled jobs, are significantly compensated through posted wage.

Our empirical results above reveal a world that deviates substantially in many aspects from the setting in the classical framework of compensating differential described in the canonical work by Rosen (1986). The first point concerns the assumed heterogeneity of market agents. In the framework of Rosen (1986), firms only differ in their costs of producing amenities, either in terms of direct cost or indirect cost on production. As a result, such cost heterogeneity is exogenous, meaning it is an endowment and unrelated to firms' labor choices or their productivity in production.[15] However, the compensations predominantly provided by firms in our data are those that either link with labor choices and worker wages, such as insurance & fund packages, leisure, and backloading wages, or with firm productivity like coworker quality, growth potential, and stock options, or are market goods like fringe benefits with limited cost variations across different firms. Similarly, given the commonality of these compensation goods, workers idiosyncratic preference are likely to be either quite homogenous, especially for pecuniary ones, or linked with their other characteristics that determine their labor market outcomes. The next two closely related points involve the sorting pattern and the pattern of compensation provision. In the canonical framework of compensating differential, sorting is solely based on firms' heterogeneous costs and workers' heterogeneous preferences on the

[15] In fact, the main example in Rosen (1986) is that two manufacturing factories have different cost functions for reducing the byproducts of pollution during production.

compensations. However, in our data, and in many other recent studies of wage differential (see the survey in Bonhomme et al. (2020)), there is clear evidence of productivity-based matching. In particular, in our posted wage decomposition results, we observe strong positive correlation between job quality and firm wage premiums. Job quality encompasses the requirements placed on potential workers, including education, experience, and specific skills and tasks, and firm wage premiums have been linked to firm productivity in previous studies (see e.g. Barth et al. (2016)). This feature aligns with a complementary production function on the inputs of firm and worker (or of multiple workers within a firm). Moreover, there is counter-evidence in our data against the notion that firms offer amenities that they have the lowest costs to those workers who value them the most. For amenities like insurance and fund packages or backloading wages, one might expect the costs to be higher in high-pay firms with high-skilled jobs and high-wage workers. However, we observe that these amenities are actually more likely to be provided by high wage-premium firms in high-skilled jobs than the inverse case. For amenities like leisure, high income workers are likely to have a higher preference if there exists a strong income effect, but we find that low wage-premium firms in low-skilled jobs are substantially more likely to provide them. Hence, the compensating differential theory show limited usage in accounting for the empirical findings presented here. Furthermore, our finding that the provisions of non-wage compensations are highly correlated with other fundamental determinants of wage suggests the potentially important impact of worker and firm productivities in firms' compensation provision decisions, which an aspect lacking in the perhaps oversimplified basic framework of compensating differential.

Our last point concerns the inconsistency between the theory and our empirical findings regarding the force of compensating differentials in impacting wage differentials. In the basic model of compensating differential, all wage variations arise from the fact that firms provide combinations of wage and amenity in different levels, equalizing wage differentials for marginal workers in equilibrium. However, our posted wage variance decomposition results suggest that most of the posted wage variations stem from job and firm characteristics, with a relatively small portion explained by information on non-wage compensation provision alone.This indicates that from firms' perspective, provisions of non-wage compensations either have little impact on wage decision or are, in fact, co-determined by those direct wage determinants. Furthermore, if we simply employ the logic of equalizing differential and abstract from worker preference heterogeneity, the estimated value of δ_i in Equation (1) should represent the wage value that is equalized in each job vacancy due to the compensation provision. Specifically, since firms in our study are generally more likely to document amenities rather than disamenities, the value of δ_i would indicate to extent to which the posted wage of a vacancy is discounted due to the non-wage compensations provided by this job. A low value of δ_i estimated in our case implies that the amenities provided by the employer in that job are highly valued by potential jobseekers, justifying a significant discount in the posted wage of the job. Conversely, a high estimated value of δ_i suggests that the job amenities provided yield only limited utility for potential workers, and thus cannot act as much compensation for the posted wage. As a result, the positive correlations between δ_i and both job effect θ_i and firm effect ψ_j imply that high wage premium firms and high-skilled jobs are accompanied by amenities with low values, and thus are less compensated. In contrast, low wage premium firms with low-skilled jobs are more likely to provide amenities with higher worth, leading to greater compensation from

posted wages.[16] Nonetheless, our analysis reveals an inverse pattern when examining the provision patterns of specific compensations. For instance, we find that high wage-premium firms in high-skilled jobs are more likely to provide advanced insurance and fund packages, while low-tier firms and jobs are more likely to be associated with basic packages. Moreover, critical compensations like backloading wage or growth potential are more commonly provided by high-pay firms in high-skilled jobs, though low-pay firms do offer more leisure and constrained work-time. Consequently, the observed patterns suggest that high wage premium firms tend to offer not inferior but superior non-wage compensations compared to their low-wage counterparts, a finding consistent with anecdotal evidences. This inconsistency between the prediction of compensating differential theory and our data can be observed more clearly in our hedonic regression results. Even after controlling for all job characteristics and firm fixed effects, we find that those compensations more likely to be provided by high wage-premium firms are not subject to equalizing differential, but instead are associated with increased posted wages. Conversely, compensations more likely to be provided by low wage-premium firms are significantly compensated. This stylized fact raises the question of why the equalizing differential force only operates for certain types of compensations but not for others, and suggests the possibility of other underlying factors at play.

One might question whether the persistent concern over the potential unobserved worker ability, which has long plagued the empirical literature on compensating differential, is responsible for our mixed results. This concern, which suggests that our failure to control for certain worker or job characteristics might skew the statistical relationships due to their correlation with non-wage compensation provisions and their direct impact on wages, was cogently addressed by Hwang et al. (1992) and remains even after efforts to address it using worker fixed effects (Brown, 1980). However, we propose two reasons why such concerns might be misplaced. First, our estimations control for detailed job skills and tasks that are documented in the job texts but are often unobserved by econometricians in standard datasets. As such, the remaining unobserved job and worker characteristics are likely to be minimal. Moreover, in Zhu (2022), we demonstrate that while these specific skills and tasks are crucial in explaining wage variations in our data, they only account for an additional 5 percent share of the total wage variations, largely because they are highly correlated with traditional worker controls such as education, experience, and granular occupation dummies. This implies that unobserved worker abilities, which are typically positively correlated with observed worker abilities, are less potent in inducing substantial estimation bias. Consequently, even if there is residual unobserved worker or job heterogeneity, its extent and impact are likely to be relatively limited.[17]

[16] As we presume no preference heterogeneity, these results can be further interpreted as high premium firms having larger costs in providing valuable amenities, whereas low premium firms have lower costs in providing amenities valued by workers. Alternatively, if we allow for large variations in idiosyncratic preference on non-wage compensations but disregard cost heterogeneity, we might interpret these results as high-skilled workers value firm-provided amenities less, while low-skilled workers value amenities more. Notably, all these interpretations rely on the implicit assumption that compensating differential is either the only source for wage differentials and firm-worker sorting, or orthogonal to other forces in the labor market. This assumption, though, generates predictions clearly inconsistent with our empirical findings. Furthermore, as we suggested earlier, the existence of substantial heterogeneity on the worker preference side and firm cost side is an empirical question with limited empirical support, especially for the latter.

[17] We refer to Hwang et al. (1992) for both a theoretical and empirical examination on how the extent and correlation between observed and unobserved worker characteristics can influence empirical testing of compensating

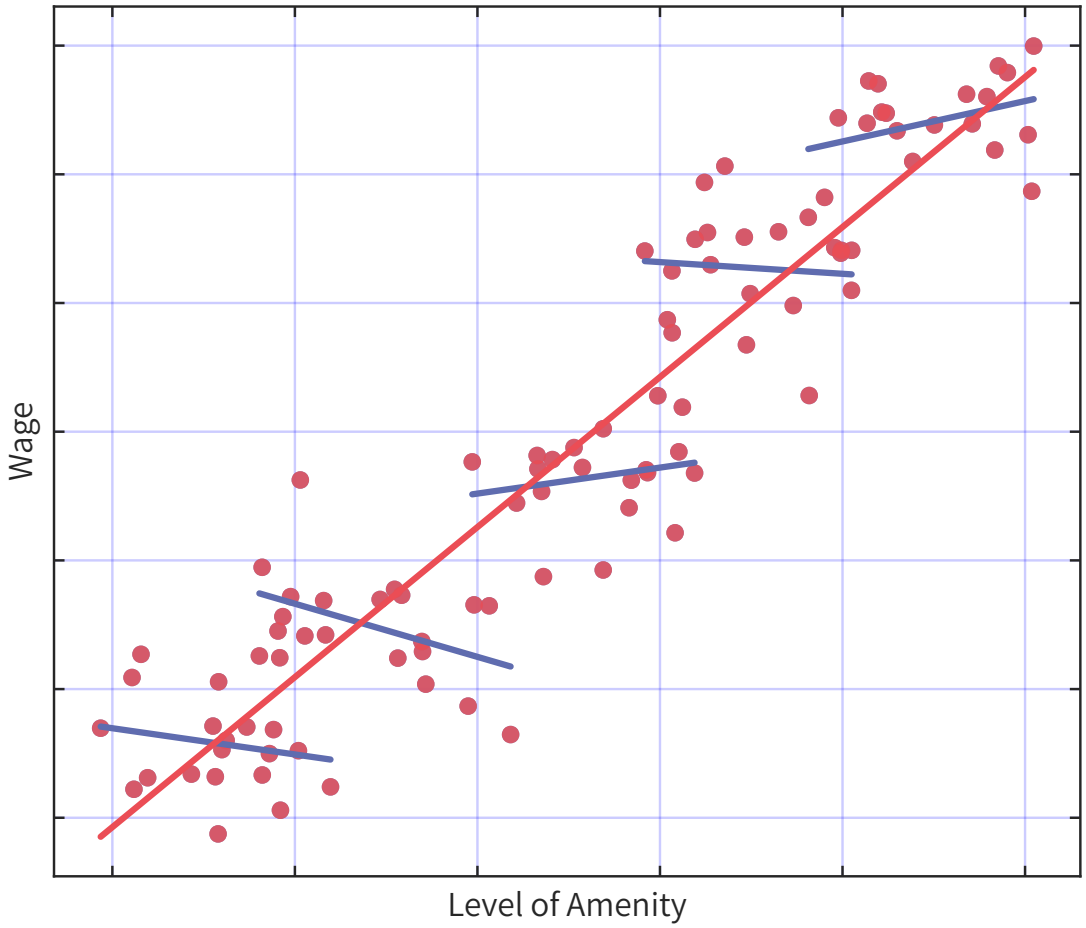
Our second reason for alleviating the concern over unobserved factor bias is more conceptual. Drawing on the idea in Hwang et al. (1992), we present in Figure 4 a typical figure used to illustrate the issue of unobserved worker abilities. The figure, while indicating a positive wage-amenity relationship in the data (the red line), suggests that such a correlation can be misleading unless we control for unobserved worker abilities and re-examine wage-amenity variations conditional on these worker differences (the blue lines). While we concur with this point qualitatively, we believe that there is another important insight to be noticed from this figure. Notably, the most significant dispersion in wage-amenity variations exists along the red line rather than the blue lines, suggesting the presence of another influential force affecting the wage-amenity relationship, one that operates counter to the force of compensating differential. The smaller the variations along the blue lines relative to those along the red line, the more significant is the competing force in labor market compensation determination that we are potentially overlooking, regardless of whether we ultimately find definitive evidence of compensating differentials. In fact, the difficulty in empirically substantiating the mechanism of compensating differential is itself indicative of other competing forces at work in the labor market. Understanding the potential competition or interplay between these forces is essential to fully grasp the market forces underpinning non-wage compensation provision and wage-amenity relationships, as well as their potential policy implications. For instance, if the alternative force represented by the red line dominates in the market, then workers' ability to equalize differential, i.e. their access to the compensation choices, based on their idiosyncratic preference becomes limited—a worker possessing the human capital at the bottom left corner can only be compensated within a limited scope and can never attain the amenity level in the top right corner. As such, in light of our empirical findings, it is equally, if not more important, to investigate this competing force as it is to continue the search for the evidence of compensating differentials.

Is there any existing theory that can explain the positive wage-amenity relationship represented by the red line? The answer is yes. The income effect reasoning adopted in Hwang et al. (1992) and Mortensen (2005), or the extended Burdett-Mortensen model used in Hwang et al. (1998) can both rationalize a positive wage-amenity correlation. The former achieves this by assuming that agents with higher income derive greater utility from other non-pecuniary amenities, while the latter suggests that those firms with low amenity-provision costs utilize both higher wages and better amenities to speed up hiring, thereby reducing search costs. However, both these two reasonings yield predictions that are inconsistent with our empirical findings. First, if the income effect is driving our findings, it becomes hard to explain why high wage premium firms that typically sort with high-skilled workers are substantially less likely to provide amenities such as weekends, holidays, and less overtime, whereas low wage firms often provide such leisure options to low income workers that they typically hire. It is almost implausible that high-income workers have a strong preference for many nonpecuniary amenities but not leisure, especially considering the recent widespread protests in China against the so-called "996 working hour system" (i.e. work from 9:00 am to 9:00 pm, 6 days per week) by employees of large, well-paid IT firms.[18] This suggests that it is not likely the

differentials.

[18]For more information, see the Wikipedia page and the news reports linked here: https://en.wikipedia.org/wiki/996_working_hour_system.

Figure 4: Illustration of Compensating Differential and Compensation Inequality



Notes. This figure is an illustration of the idea of omitted-ability bias in the estimating of compensating wage differential as well as of the dispersion of the compensation provision inequality in the labor market. The basic idea comes from Hwang et al. (1992) but here we use it to reveal additional implications.

worker preference heterogeneity driving the divergence in compensation provision. Second, the amenity-providing cost variations, as assumed in Hwang et al. (1998), fail to explain why high-pay firms often provide superior amenities such as insurance or backloading wages, bearing higher costs than their low-pay counterparts. While large, high-productivity firms might enjoy economics of scale in providing certain types of fringe benefits, such cost advantages do not apply to many of the significant non-wage compensations that we have identified in our data. Therefore, it is also difficult to use firm cost heterogeneity to explain the strong linkages between compensation provision and firm and job effects. Lastly, as we suggested earlier, there is substantial sorting based on firm and job productivity that is not accounted for in these models, which could have important impacts for firms' joint determination of wages and non-wage compensations. Given the challenge of reconciling our empirical evidences with existing theories, in the next section, we propose a simple yet novel theory that introduces a new force of efficiency compensation. This new force interacts with the compensation differential mechanism within the context of firm-worker productivity sorting. We will show that this theory can not only reconcile all our empirical findings but also generate important general implications for labor market inequalities.

3 A New Theory

3.1 The Basic Idea

In this subsection, we suggest that the puzzle in the compensating differential literature, which is also occurred in our results, is not a problem of identification but a problem of incomplete theory. In particular, we argue that as long as we combine two additional elements, which are also observed in the literature and in our data, with the canonical mechanism of compensating differential, we can then generate patterns of compensation provision and different levels of compensating differential that are consistent with our empirical findings. The first new element is efficiency compensation, i.e. non-wage compensation can be efficient in production or in firm-operation.[19] The second new element is firm and worker sorting, or the firm and job sorting in our case. While the existence of the second element, sorting, have been confirmed by the recent literature on wage inequality and by the results here, the first element, the efficiency function of compensation, is often dismissed when empirically testing the impact of compensations on wage. Here we argue that the level of efficiency (or inefficiency) is a general and important feature of non-wage compensations.[20] First, it is not difficult to

[19] We call it "efficiency compensation" because it is analogous to the idea of efficiency wage theory, which suggests firms pay wages higher than market clearing level for various efficiency reasons such that it is optimal for the production or profit maximization. Actually we think efficiency compensation is even a more nature idea because one key critique on the efficiency wage theory is that firms should be able to take advantage of other non-wage compensations to achieve the same efficiency aim (see Katz, 1986).

[20] To be clear, in the canonical compensation differential the provision of a compensation can be also efficient or inefficient in production. However, the theory of compensation differential assumes that the sign of the impact of the compensation on the production be must inverse to the sign of its impact on the workers' utility. In other words, an amenity for the workers must cause a reduction in production productivity or a direct cost in production. Here we relax this restriction and allow an amenity to be either efficient or inefficient or having no impact on production at all.

see the efficiency nature of those monetary compensations like backloading wage and stock option. In fact the literature have been long argued that alternative payment structure can help firm to improve efficiency through effort inducing, turnover reduction, and so on (see e.g. Lemieux et al., 2009). Similarly, it has been argued in the literature that amenities like health insurance and other insurance can reduce exogenous worker exit (Dey and Flinn, 2005) or facilitate efficient endogenous worker turnover (Hwang et al., 1998), and that amenities like better coworker quality improves both production productivity and on-the-job learning efficiency in a complementary production setting (Jarosch et al., 2021). In contrast, weekend, holiday, and less overtime or limited work duty are straightforward inefficient because they allow less work-time and effort. Other amenities like training or work-time flexibility are perhaps more unambiguous and if they are efficient or inefficient likely depends on the detailed cases.

A formal model setting and derivation of our new theory, which combines a simple framework of worker sorting with efficiency compensation, are documented in Section 3.2. For the rest of this subsection we briefly introduce the key ideas, intuitions, and implications of our new theory.[21] The key idea is that when an amenity is allowed to be efficient in production, then in additional to the wage saving benefit, firm will also take the marginal product of efficiency improvement into account when considering the provision of a certain compensation. With firm-work sorting in the labor market, the level of this marginal production benefits from offering efficiency compensation will be larger in high wage premium firms that are sorted with high productivity workers or jobs. In other words, the better the firm or the job, more efficient will be the compensations. As a result, higher wage premium firms and higher wage jobs are more likely to provide those efficiency compensations, and because increase in productivity will often at least partially translate into increase in wage, this efficiency gain act in counter to the classical compensating differential mechanism. And if the level of the compensation has a large span and the marginal product does not decline too fast, it is also possible that the efficiency channel dominates the compensating differential above some threshold of firm and worker level, generating positive wage effect in net, i.e. firms providing better compensations now cause wage increase rather than wage decrease. In contrast, firms with lower wage premiums and sorted with low productive workers or jobs are less likely to provide efficient compensations because the marginal production benefits are small. And when firms do provide such compensations in some cases, say basic insurance that is mandated by the government, their net loss between the provision cost and the efficiency effect, if any, will be

[21] In fact, the setting in Section 3.2 is one of the simplest way, but not the only way, to generate the desired results, and there are many potential or further extensions that can be added to the basic framework. To distinguish with the traditional compensating differential model and to clarify our new mechanisms, in our model we assume workers are homogenous in their preference on all non-wage compensations and firms have the same direct cost functions on providing all compensations. However, both firms and workers are heterogeneous in their productivity, and they form pairs endogenously, and the joint production function is assumed to be supermodular—a necessary condition to generate positively assortative matching between firms and workers in the economy. Compensations provided by firms are assumed to be either efficient or inefficient, i.e. they affect an efficiency terms of the firms' production which acts as another complementary input in the production function. We show that this simple and parsimonious setting that contains efficiency compensation and sorting is enough to generate rich features of compensations provision and different levels of compensating differential. More realistic models can be constructed by adding heterogeneous worker preference or search frictions so that the sorting becomes no longer monotone or perfect.

equalized through reduction in wage, and the lower is the rank and productivity of the firm and the work, the severe is the level of compensating differential. Therefore, our theory can generate the feature that while an amenity is significantly compensated from wage by low pay firms in low pay jobs, the same or even a superior amenity is not compensated from but actually positively correlated with wage in high wage firms and high skill jobs. The similar logic can be applied to compensations that are inefficient, say generous work-time or work-life balance.[22] Under complementarity, the higher the productivity and rank of the firm and worker, the larger the efficiency loss coming from the provision of such compensation.[23] Consequently, as long as the income effect on leisure is not too strong, high wage premium firms and jobs will not provide such compensations, but rather compensate workers for their utility loss with higher wage. On the other hand, such efficiency cost is small when the firm and the job have low rank and low productivity, and thus low wage premium firms with low skill jobs are more likely to document such inefficient compensations for attracting workers. In other words, now the efficiency channel is in the same direction as the compensating differential channel, and the impact of firm-worker sorting on the efficiency channel in fact act as an amplifier for compensating differential. Finally, when a compensation is neither efficient nor inefficient, the efficiency channel shuts down, and the model returns back to the traditional compensating differential model.

Our new theory have three implications that are important for understanding the labor market inequality in wage and non-wage compensation. First, the efficiency aspect of different pecuniary or nonpecuniary compensations could be the key to dissolve the puzzle that is brought by the mixed results found in the empirical tests for the theory of compensating differential. As our new theory shows, the efficiency effect can totally offset the effects of equalizing differential and generate results inverse to the predictions by the compensating differential theory. Our theory thus predicts that while it might be not difficult to find the clear evidences for compensating differential in the submarket with low-pay firms and low-skill workers, the similar evidences will be hard to find when targeting to the high-end labor market or the entire labor market. Also, directly adopting the estimation results and conclusions found from a particular compensation in a particular labor market to other compensations and other labor markets could be dangerous and misleading. Second, with firm-worker sorting and efficiency compensation, the labor market inequality could be underestimated by just looking at wage or monetary payments. The high-skill workers employed in high wage premium firms are likely to also enjoy the best non-wage compensations in many aspects, including both additional earnings from bonus and stock and nonpecuniary amenities like better insurance or fringe benefits, though at the expense of high effort. Perhaps more surprisingly, our theory suggest that the provision of compensations can not only generate inequality in non-wage compensation itself but also further enlarge the wage inequality. This is because efficiency compensations can simultaneously increase the workers' direct utility on non-wage compensations and increase workers' wage through a boost in their productivity. In other words, efficiency compensations

[22] It is arguable that in some cases generous rests like paid leave or maternity leave can be actually efficient if they help to retain workers and the turnover cost is very high. In fact Bana et al. (2022) find the in the U.S. high wage premium firms are more likely to participate in Paid Family Leave programs and have lower turnover rates. However, it could be a difficult empirical question to answer ex-ante that if an amenity like this is efficient or not.

[23] Note that in additional to the linkage with the firm-worker match productivity, such inefficient compensation also offsets the effect of other efficient compensations.

work as an amplifier for the labor market inequality at both observed wage level and observed utility level. Third, our theory suggest that the set of the unobserved non-wage compensations that drive the large amount workers' moving to low-wage premium firms will be rather limited (see Sorkin, 2018; Bonhomme et al., 2019). In fact, our theory suggest that these compensations must be inefficient ones like less work-time because high-wage premium firms will also provide better efficient compensations. Moreover, a worker that goes down the firm ladder due to some changes in preferences for certain amenities like leisure will suffer not only a worse matching but also a downgrading on many other efficient compensations, both of which will negatively affect the wage that the worker receive.[24]

3.2 A Simple Model

We now build a simple model with firm-worker sorting and efficiency compensation and show that the results derived can be consistent with the results we find in our data and other similar results found in the recent literature. In comparison to the canonical compensation differential model, in our model workers and firms are not heterogeneous in their tastes of and cost functions of various peculiarity and nonpecuniary compensations, rather they are heterogeneous in their productivity like the typical assignment model.[25] The variation of the firms' provision of a certain compensation comes from different efficiency (or inefficiency) levels of the compensation in different firm-worker pairs. In fact, the traditional mechanism of compensation differential still exists but is now offset by a new efficiency mechanism when the compensation is efficient or magnified when the compensation is inefficient. The level of this new efficiency channel depends on the level of firm-worker sorting. We show that this model can thus generate flexible results on the wage impact of compensation provision that could be both consistent and inconsistent with the predictions of the traditional compensating differential model. We also show that this new theory can generate important implications for understanding labor market inequalities.

In the economy there is a continuum of workers with same utility function $U(C, a, h)$ but heterogeneous productivity $q \in [0, 1]$, where C is the monetary consumption, $a \in \{0, 1\}$ is the indicator of if the worker receiving a nonpecuniary amenity, e.g. the firm's provision of a certain insurance, and h is the level of a nonpecuniary disamenity, e.g. the requirement on additional working hour. To simplify the analysis, we assume the utility function takes an

[24] In fact in the section 5.4 of Rosen (1986), Rosen suggests an application of the compensation differential theory as "hours of work (or work schedules more generally) may be formally treated as nonpecuniary aspects of jobs. Then the market transaction must be viewed as a tie-in in which a firm offers a fixed wage-hours package to workers, take it or leave it, with these package deals varying from firm to firm". He then suggests two sources for the equilibrium distribution of different packages generated in the labor market: coordination in production or set up costs. Our idea of labor market sorting as the source for heterogeneous provision of working hour and wage packages is close to the idea of coordinating production, but different from the classic compensation differential framework that Rosen suggest, in our argument the interpersonal differences in productivity affect the equilibrium allocation not only through the resulted heterogeneity in preference but also through firms' opportunity cost of offering such "inefficient" compensations. In addition, the nonpecuniary aspect of job we consider here can be more general and contains not only hours of work but also latent effort.

[25] A more general analysis with heterogeneous worker preferences (and heterogeneous firm cost functions of compensation provision) requires to extend the problem to multidimensional matching. However, this will significantly increase the analysis tractability, and thus we leave it for future research.

additively separable form,

$$U(C,a,h) = C + \phi_a a - \frac{h^{1+\phi_h}}{1+\phi_h} \tag{3}$$

, with $\phi_a, \phi_h > 0$. We assume that there is no income other than wage, thus $C = w(q)$, and the level of amenity a and disamenity h are decided by the firm that the worker matched in the equilibrium.

The economy also has a large continuum of potential firms that are ex-ante homogenous and want to hire workers. To facilitate sorting or separation, we assume that these firms face the same O-Ring type production technology, i.e. a production function with complementarity across all labor positions.[26] In addition, we extend the original O-Ring production function with efficiency compensation so that the labor productivity and the output Y_j of a firm j will also depend on the firm's choices on its supply of the compensations. In particular, the production function takes the form:

$$Y_j = A N_j^{1+\alpha} \prod_{i=1}^{N_j} q_i e(a,h) \tag{4}$$

, where A is a common productivity, N_j is the total number of worker the firm hires, and $e(\cdot)$ is an efficiency function which is increasing and strictly concave in both arguments. To simplify the exposition, we also assume the function $e(\cdot)$ takes an additively separable form:

$$e(a,h) = 1 + \gamma_a a + \frac{h^{\gamma_h}}{\gamma_h}$$

, where $\gamma_a, \gamma_h \in [0,1)$ control the decreasing return for these two efficiency compensation. This also helps us to illustrate that even when the decisions on the level of different amenities and disamenities are irrelevant by themselves, their occurrence could be still correlated through their relationships with the firm productivity. We further assume that for the amenity a, firm will pay the cost, which is a per-worker cost κ multiplied by the total number of worker N if $a = 1$. In comparison, for the disamenity h, a firm does not pay any direct cost but need to pay a higher wage w to compensate the loss in worker's utility. For analytical tractability, we also assume a fixed N for all firms but will show later that relaxing N to be another endogenous firm choice does not change our results.

The competitive equilibrium in this economy is defined as an assignment of worker types to firms and a utility schedule, $u(q)$, such that (i) given the utility schedule, all active firms maximize their profits by employing their workers in a way consistent with the assignment and by choosing the wages for their workers, $w(q)$, and the levels of two (dis)amenities, a and h, and that (ii) the labor market clears for workers of all productivity levels. The competitive equilibrium here coincides with the stable matching of the assignment problem: the workers'

[26] An alternative way to establish sorting is to generate two-dimensional matching by assuming an exogenous distribution of firms with heterogeneous productivity and assuming a pair-wise production function with complementarity. One advantage of our one-dimensional matching setting is that we can easily generate endogenous firm size and have the firm size correlated with firm level productivity, as what we find the data and in the literature. The choice of this alternative setting and our setting does not affect our main results on compensation provision. More generally one can have both two-side heterogeneity and endogenous firm size, see Eeckhout and Kircher (2018).

utility schedule and the firms' profit schedule is on the possibility frontiers and there does not exist other assignments that can generate larger payoffs.

We then characterize the competitive equilibrium by analyzing the firm's profit maximization problem:

$$\begin{aligned} \max_{\{q_i\}_{i=1}^N, a, h, w(q)} \; & AN^{1+\alpha} \prod_{i=1}^{N} q_i e(a,h) - \sum_{i=1}^{N} w(q_i) - a\kappa N \\ \text{s.t.} \quad & w(q) + \phi_a a - \frac{h^{1+\phi_h}}{1+\phi_h} \geq u(q) \quad \forall q \in \{q_i\}_{i=1}^N \end{aligned} \tag{5}$$

. The profit possibility frontier for the firm is $v(q_1, \ldots, q_N, u)$ with the utility compatibility constraint holding in equality and $a, h, w(q)$ chosen optimally. The derivative of v with respective to each single q_i is $\frac{dv}{dq_i} = AN^{1+\alpha} e(a,h) \prod_{i' \neq i}^{N} q_{i'} e(a,h) > 0$, and thus v is type increasing. It is then easy to see that the cross partial derivatives $\frac{d^2 v}{dq_i dq_{i'}} > 0$ and $\frac{d^2 v}{dq_i du} = 0$, and thus the equilibrium allocation in the economy satisfies positive assortative matching (PAM) and in our case this means all workers employed by any single firm will have the same type, q.[27] Under perfect segregation, the firm's problem in Equation (5) now can be written as

$$\max_{q,a,h} AN^{1+\alpha} q^N (1 + \gamma_a a + \frac{h^{\gamma_h}}{\gamma_h}) - N \left(u(q) - \phi_a a + \frac{h^{1+\phi_h}}{1+\phi_h} \right) - a\kappa N \tag{6}$$

. Because the amenity a is a discrete choice, there is a productivity threshold q_a such that the firm that hires workers with q_a will be indifferent between providing or not providing amenity a. In particular the optimal provision strategy will be

$$a = \begin{cases} 1, & \text{if } q \geq q_a \\ 0, & \text{if } q < q_a \end{cases}, \text{ and } AN^{\alpha} q_a^N \gamma_a + \phi_a = \kappa \tag{7}$$

This threshold is decided from the equation $AN^{\alpha} q_a^N \gamma_a + \phi_a = \kappa$, i.e. the marginal benefit of providing a equals the marginal cost of providing a. Note that if the amenity a is not efficient at all, i.e. $\gamma_a = 0$, then it can return back to the canonical compensation differential where the dispersion of preference ϕ_a and of the cost κ generate sorting between workers and firms. As a result, more productive firm with more productive workers are more likely to offer these efficient compensations. The importance of this channel can be seen more clearly when the cost of the amenity is increasing in the level of the worker, as the case of many insurance and fund. If we assume the per-worker cost of a is actually $q\kappa$, then the cost of providing such amenity increases in firms which employ high q workers and pay high wage w. As a result with the traditional compensating differential mechanism alone, high rank firms are less likely to pay for such amenity, which is inconsistent with the empirical facts found in the literature. How-

[27] In the original O-Ring mode there is perfect transferable utility, and thus the complementarity across different labor inputs in the O-Ring production function, i.e. $\frac{d^2 v}{dq_i dq_{i'}} > 0$, will be enough ensure that in the equilibrium assignment is PAM. However, the amenity and disamenity terms in the worker's utility function generate imperfectly transferable utility and as a result an additional condition $\frac{d^2 v}{dq_i du} \geq 0$ is required for segregation assignment to be achieved in the equilibrium. For more details about the sufficient conditions for monotone matching in an economy with assignment problem see Legros and Newman (2007).

ever, when a non-wage compensation is efficient and there is enough productivity dispersion across firms and workers, the differences in the efficiency effect could dominate and derive the discrepancy in compensation provision. However, with the efficiency channel, the increase in q has increased marginal benefit from efficiency effect more than covering the increased provision cost, generating positive relationship between firm productivity and compensation provision.

The first order conditions for the rest of two maximization choices q and h are

$$AN^{1+\alpha}q^{N-1}e(a,h) = u'(q), \tag{8}$$

$$AN^{\alpha}q^{N}h^{\gamma_h-1} = h^{\phi_h} \tag{9}$$

respectively. It's clear from Equation (9) that the optimal level of disamenity $h = \left(AN^{\alpha}q^{N}\right)^{\frac{1}{1+\phi_h-\gamma_h}}$ is also increasing in productivity q due to the same efficiency reason as amenity a. Therefore, more productive firm will also require high level of disamenity h. However, in this case, this disamenity will be fully compensated by the increase in the wage. To obtain the market wage, we first derive the market utility profile by replacing the optimal efficiency level $e(a,h)$ and then integrate Equation (8) over the entire distribution of worker productivity:

$$u(q) = \begin{cases} \frac{\left(\bar{A}q^N\right)^{1+\omega}}{(1+\omega)(1+\gamma_h)} + (1+\gamma_a)\bar{A}q^N + u_a, & \text{if } q \geq q_a \\ \frac{\left(\bar{A}q^N\right)^{1+\omega}}{(1+\omega)(1+\gamma_h)} + \bar{A}q^N + u_0, & \text{if } q < q_a \end{cases} \tag{10}$$

, where $\bar{A} \equiv AN^{\alpha}$, $\omega = \frac{1+\gamma_h}{1+\phi_h-\gamma_h}$, and u_0 and u_a is the constant of integration. In fact u_0 will be the utility or wage that workers of $p = 0$ obtain, and thus is pinned down by free-entry condition such that $u_0 = 0$. Similarly, u_a is pinned down by the firm indifference at p_a such that $u_a = \phi_a - \kappa$. Finally, the market wage profile can thus be derived from Equation (10) as

$$w(q) = \begin{cases} \bar{A}q^N + \gamma_a\bar{A}q^N - \kappa + \frac{\left(\bar{A}q^N\right)^{1+\omega}}{(1+\omega)(1+\gamma_h)} + \frac{\left(\bar{A}q^N\right)^{\omega}}{1+\gamma_h}, & \text{if } q \geq q_a \\ \bar{A}q^N + \frac{\left(\bar{A}q^N\right)^{1+\omega}}{(1+\omega)(1+\gamma_h)} + \frac{\left(\bar{A}q^N\right)^{\omega}}{1+\gamma_h}, & \text{if } q < q_a \end{cases} \tag{11}$$

. The first term $\bar{A}q^N$ is the wage function when there is no non-wage compensation, which, perhaps not surprisingly, increases in productivity p, indicating that firms with high q workers also provide higher wage. The second part $\gamma_a\bar{A}q^N - \kappa$ in the case $q \geq q_a$ is a combination of the increase in wage due to efficiency effect and the compensation differential on a. Note when $q = q_a$ this term is $-\phi_a$, i.e. the workers' utility benefits from firms' provision of a is fully compensated from the reduction in wage, and thus for some range of the productivity $p > p_a$, it will generate a wage plunge comparing to firms that have a close productivity of workers but do no provide amenity a. However, going up the firm rank, the efficiency effect will increase in productivity q, and it's possible that the provision of the amenity a are not compensated from wage reduction at all but actually generate wage gain for the workers. The last two terms, $\frac{\left(\bar{A}q^N\right)^{1+\omega}}{(1+\omega)(1+\phi_h)}$ and $\frac{\left(\bar{A}q^N\right)^{\omega}}{1+\phi_h}$, are the increase in wage due to efficiency effect and compensation

differential on h. In this case because h is an amenity and high rank firms demand more h, this generates a positive compensating differential on the wage. Therefore, our model shows that in a standard setting of firm-worker sorting or segregation, an efficient non-wage compensation, whether amenity or disamenity, could further enlarge market wage inequality, especially when the dispersion of productivity is large across firms. In other words, non-wage compensation can be not only an unobserved and overlooked labor market inequality, but also potential drivers behind the increasing in observed wage or earning inequality.

Next we discuss two more implications of our model on the empirical estimation of compensating differential. First, note for q close to q_a, one can find clear evidence of compensation differential by examining wage difference and controlling for worker characteristics. However, if a is a multiple discrete choice or close to an continuous choice, higher level of a will be correlated with higher wage, counteracting to the force of equalizing differential and thus confounding the estimation results. Also, if we assume that at some point in time the government mandate the firms' provision of amenity a (again for example the enforcement of a certain insurance) but the enforcement is not perfect. This will not change anything for all firms with worker productivity $q \geq q_a$, but it now requires all firms with $q < q_a$ also provide a. As a result, the compensation in wage reduction for these firms will be $\gamma_a \bar{A} q^N - \kappa$ and this reduction is larger for firms with lower q, providing larger incentive for these firms to circumvent the provision of a. The empirical estimation will thus find strong evidence of compensation differential for this firms, but again the higher the labor market with firms and workers with high productivity, the smaller will be the equalizing effect and in some cases be even negative. Second, given that firms gathered with high productivity workers are likely to both provide high levels of efficiency compensations and require high levels of efficiency disamenities, a worker deviated from in the common utility structure for certain amenities will have to also deviate from the current optimal matching and be subject to changes in receiving other compensations. For example, if a middle-age female worker with a high productivity q gives a birth and thus has a large increase in disamenity in working hour (i.e. an increase ϕ_h), she has to go down the job ladder and match to a firm in which workers have productivity lower than q. Moreover, the level of other non-wage compensations might be also downgraded if these compensations are efficient. Therefore, in additional to the part of usual equalizing differential, the cost of such compensation differential will also incorporate the wage decline due to a worse matching and the utility decline due to a less generous package of other compensations.

Finally, we consider the case when N is also a choice of the firm. The additional first order condition with respect to N in this case is

$$AN^{\alpha} q^N e(a,h)(1 + \alpha + N \ln(q)) = w + ac \tag{12}$$

. Further differentiating Equation (12) with respect to q and evaluated at the optimal level, we obtain the optimal choice on firm size:

$$N(q) = \frac{1+\alpha}{-\ln(q)} \tag{13}$$

. This result shows that the firm size increases in productivity q and is irrelevant to the choices of amenities. Therefore, all the relationships we have found between productivity and amenity

provision can be now directly translate to the firm size.

4 Conclusion

In this paper, we take advantage of online job advertisement data, where firms document their non-wage compensations and amenities to attract workers, to document several new empirical findings about firms' compensation provision behavior. Most importantly, we find while high wage premium firms sorted with high quality works or jobs are more likely to also provide many other non-wage compensations like advanced insurance or stock option, low wage premium firms sorted with low quality worker or jobs are more likely to provide weekend, holiday, and fixed work time. The classic compensating differential theory does not provide explanations for such distinguished behavior in compensation provision. We also find the puzzling results that the compensations that high wage premium firms provide are positively correlated with the posted wage, which is at odds with the prediction of compensating differential theory, although the inverse is true in for low wage premium firms, which supports the existence of equalizing differential. To reconcile these stylized facts, we suggest a new theory which extends the classic mechanism of compensating differential with an additional channel of efficiency compensation, of which the extent depends on the level of firm-worker sorting. We use a simple model to show that our new theory can not only reconcile all the empirical findings we find, but also have important implications on the labor market inequalities in terms of both wage and non-wage compensations. For the purpose of tractability, our model is rather stylized and perhaps over-simplified in that we assume homogenous worker preference in a quasi-linear form, homogenous firm production function, and perfect assortative matching. One potential future work is to further generalize the model so that we can bring the model to the data.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

Bana, S., K. Bedard, M. Rossin-Slater, and J. Stearns (2022). Unequal use of social insurance benefits: The role of employers. *Journal of Econometrics*.

Barth, E., A. Bryson, J. C. Davis, and R. Freeman (2016). It's where you work: Increases in the dispersion of earnings across establishments and individuals in the united states. *Journal of Labor Economics 34*(S2), S67–S97.

Bell, A. (2022). Job amenities and earnings inequality. *Working Paper*.

Bloesch, J., B. Larsen, and B. Taska (2021). Which workers earn more at productive firms? position specific skills and individual worker hold-up power. *Working Paper*.

Bonhomme, S., K. Holzheu, T. Lamadon, E. Manresa, M. Mogstad, and B. Setzler (2020). How much should we trust estimates of firm effects and worker sorting? Technical report, National Bureau of Economic Research.

Bonhomme, S. and G. Jolivet (2009). The pervasive absence of compensating differentials. *Journal of Applied Econometrics 24*(5), 763–795.

Bonhomme, S., T. Lamadon, and E. Manresa (2019). A distributional framework for matched employer employee data. *Econometrica 87*(3), 699–739.

Brown, C. (1980). Equalizing differences in the labor market. *The Quarterly Journal of Economics 94*(1), 113–134.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics 36*(S1), S13–S70.

Card, D., J. Heining, and P. Kline (2013). Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly journal of economics 128*(3), 967–1015.

Dey, M. S. and C. J. Flinn (2005). An equilibrium model of health insurance provision and wage determination. *Econometrica 73*(2), 571–627.

Eeckhout, J. and P. Kircher (2018). Assortative matching with large firms. *Econometrica 86*(1), 85–132.

Eriksson, T. and N. Kristensen (2014). Wages or fringes? some evidence on trade-offs and sorting. *Journal of Labor Economics 32*(4), 899–928.

Gentzkow, M., B. Kelly, and M. Taddy (2019). Text as data. *Journal of Economic Literature 57*(3), 535–74.

Hwang, H.-s., D. T. Mortensen, and W. R. Reed (1998). Hedonic wages and labor market search. *Journal of Labor Economics 16*(4), 815–847.

Hwang, H.-s., W. R. Reed, and C. Hubbard (1992). Compensating wage differentials and unobserved productivity. *Journal of Political Economy 100*(4), 835–858.

Jarosch, G., E. Oberfield, and E. Rossi-Hansberg (2021). Learning from coworkers. *Econometrica 89*(2), 647–676.

Katz, L. F. (1986). Efficiency wage theories: A partial evaluation. *NBER macroeconomics annual 1*, 235–276.

Kline, P., R. Saggio, and M. Sølvsten (2020). Leave-out estimation of variance components. *Econometrica 88*(5), 1859–1898.

Krueger, A. B. and L. H. Summers (1988). Efficiency wages and the inter-industry wage structure. *Econometrica: Journal of the Econometric Society*, 259–293.

Lamadon, T., M. Mogstad, and B. Setzler (2022). Imperfect competition, compensating differentials, and rent sharing in the us labor market. *American Economic Review 112*(1), 169–212.

Lang, K. and S. Majumdar (2004). The pricing of job characteristics when markets do not clear: theory and policy implications. *International Economic Review 45*(4), 1111–1128.

Le Barbanchon, T., R. Rathelot, and A. Roulet (2021). Gender differences in job search: Trading off commute against wage. *The Quarterly Journal of Economics 136*(1), 381–426.

Legros, P. and A. F. Newman (2007). Beauty is a beast, frog is a prince: Assortative matching with nontransferabilities. *Econometrica 75*(4), 1073–1102.

Lemieux, T., W. B. MacLeod, and D. Parent (2009). Performance pay and wage inequality. *The Quarterly Journal of Economics 124*(1), 1–49.

Maestas, N., K. J. Mullen, D. Powell, T. Von Wachter, and J. B. Wenger (2018). The value of working conditions in the united states and implications for the structure of wages. Technical report, National Bureau of Economic Research.

Mas, A. and A. Pallais (2017). Valuing alternative work arrangements. *American Economic Review 107*(12), 3722–59.

Mortensen, D. (2005). *Wage dispersion: why are similar workers paid differently?* MIT press.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy 82*(1), 34–55.

Rosen, S. (1986). The theory of equalizing differences. *Handbook of labor economics 1*, 641–692.

Salop, J. and S. Salop (1976). Self-selection and turnover in the labor market. *The Quarterly Journal of Economics 90*(4), 619–627.

Shapiro, C. and J. E. Stiglitz (1984). Equilibrium unemployment as a worker discipline device. *The American economic review 74*(3), 433–444.

Sockin, J. (2022). Show me the amenity: Are higher-paying firms better all around? *CESifo Working Paper*.

Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. Von Wachter (2019). Firming up inequality. *The Quarterly journal of economics 134*(1), 1–50.

Sorkin, I. (2018). Ranking firms using revealed preference. *The quarterly journal of economics 133*(3), 1331–1393.

Taber, C. and R. Vejlin (2020). Estimation of a roy/search/compensating differential model of the labor market. *Econometrica 88*(3), 1031–1069.

Wissmann, D. (2022). Finally a smoking gun? compensating differentials and the introduction of smoking bans. *American Economic Journal: Applied Economics 14*(1), 75–106.

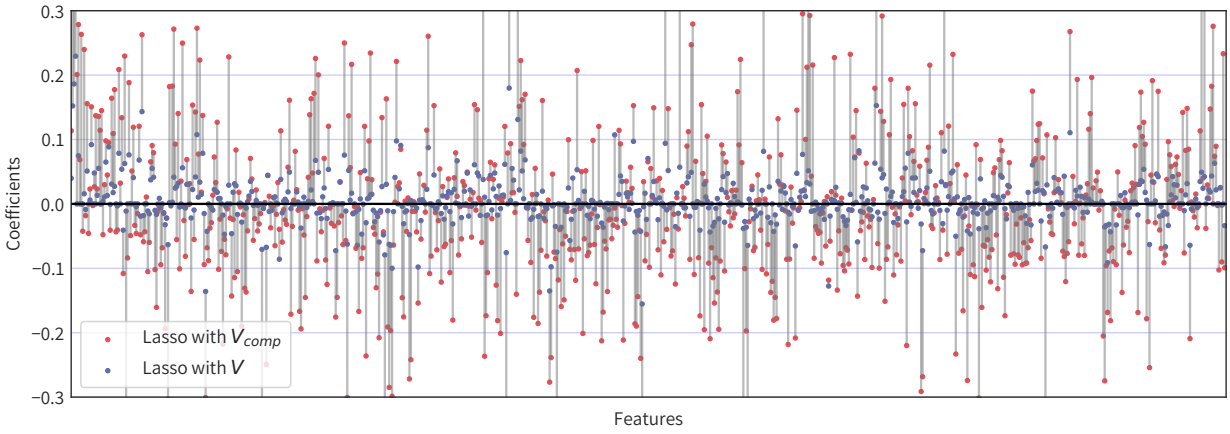
Wiswall, M. and B. Zafar (2018). Preference for the workplace, investment in human capital, and gender. *The Quarterly Journal of Economics 133*(1), 457–507.

Zhu, X. (2022). Posted wage inequality. *Working Paper*.

Appendices

Appendix A. Additional Tables And Figures

Figure A1: Lasso Coefficients Under Different Specifications



Notes. See note of Figure 2.

Figure A2: Compensation Occurrence of Other Compensation Types

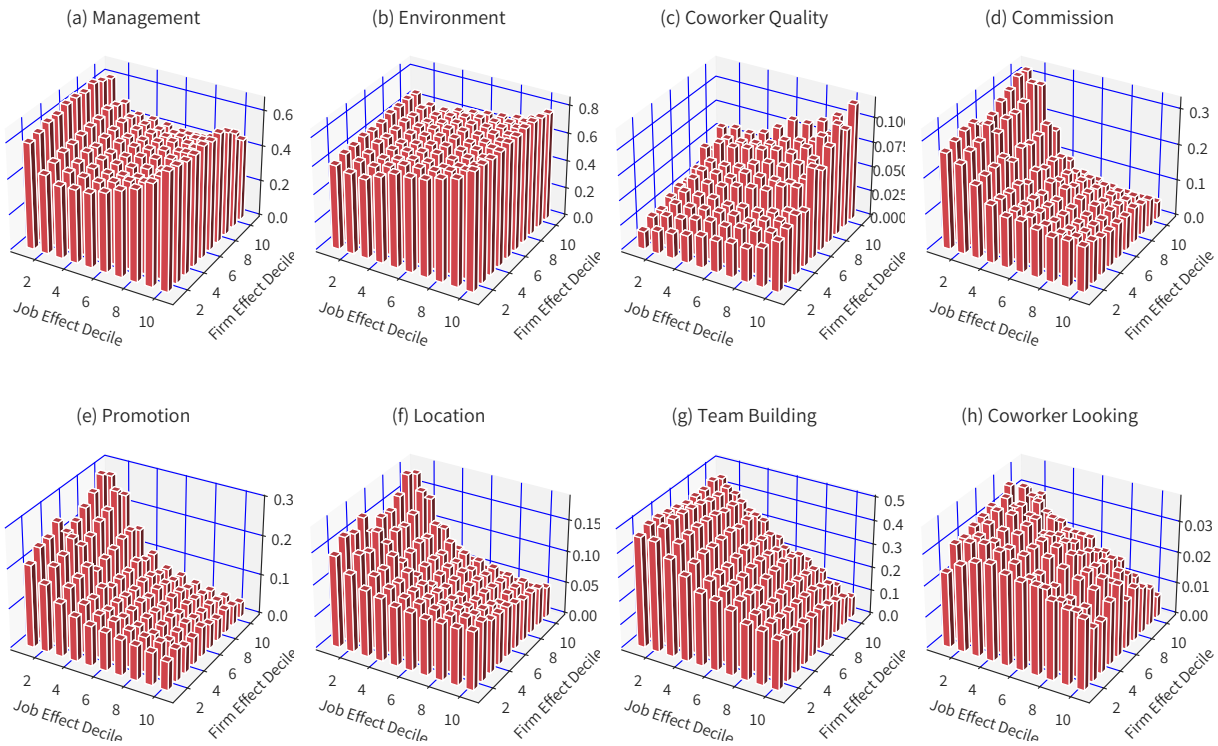
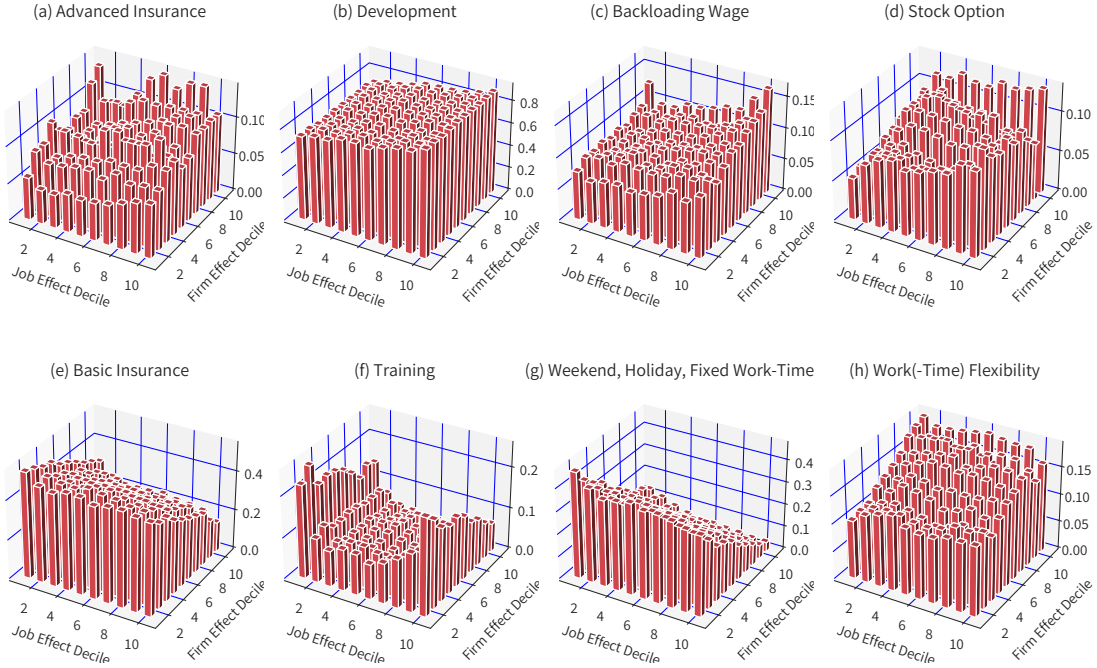
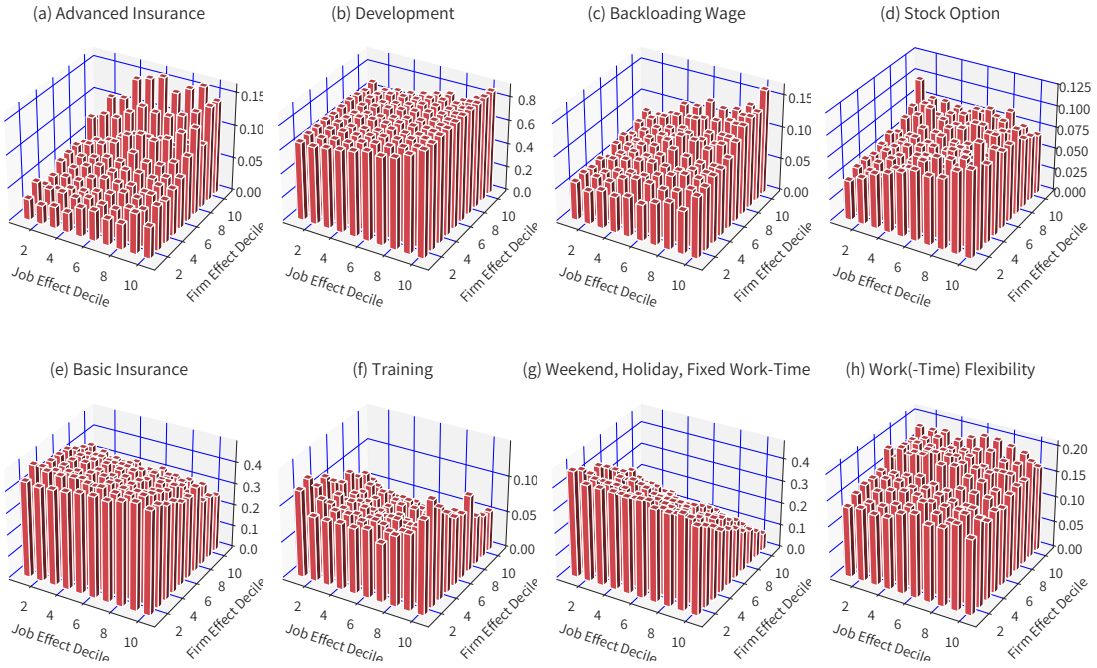


Figure A3: Compensation Occurrence by Occupation

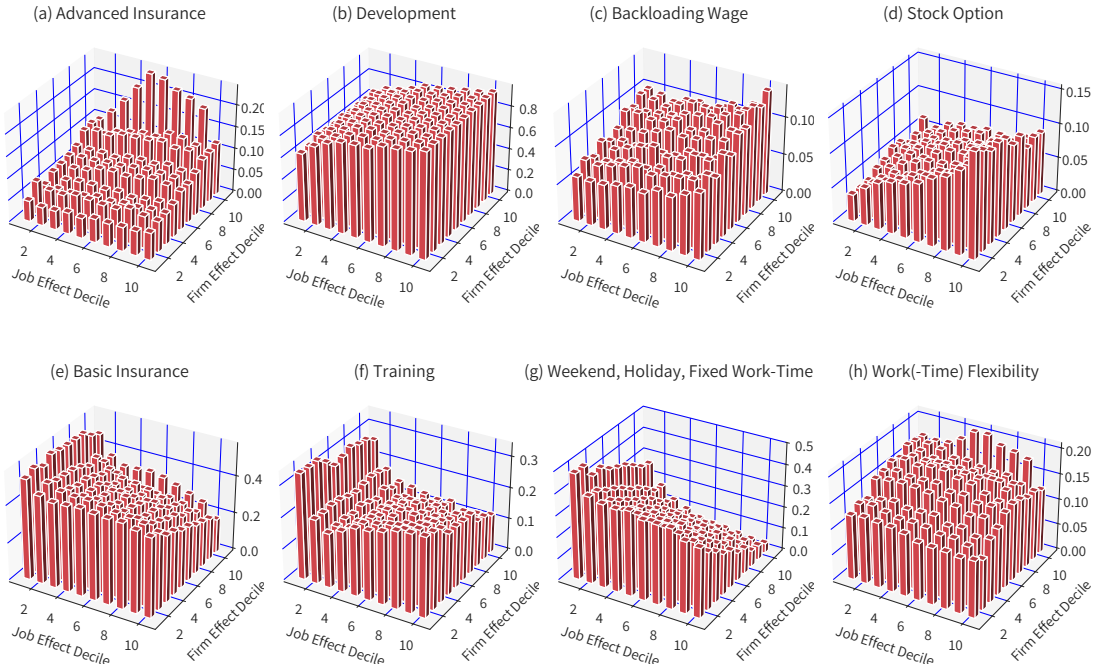
(a) Computer



(b) Design_Media



(c) Business_Operations



(d) Administrative