

Posted Compensation Inequality

Xuanli Zhu
Keio University

May 21, 2025

The 8th Sophia Research Workshop in Economics

Outline

1. Introduction

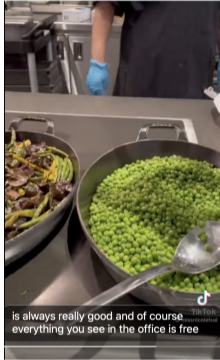
2. Empirical Facts

3. Discussion

4. A Simple Theory

5. Conclusion

A Vignette

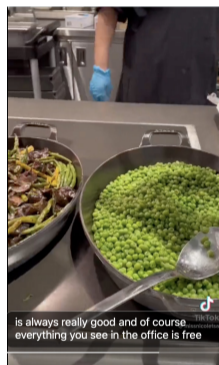
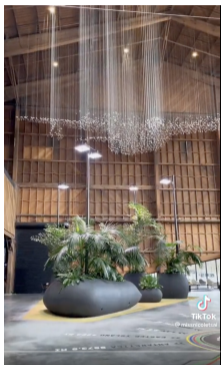
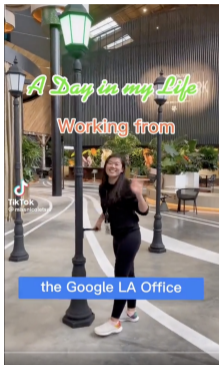


A Vignette



→ Compensating Differential?

A Vignette



→ Compensating Differential?
→ Workers valuing the amenity most matched with Firms of lowest provision costs?

Research Questions

- Empirical:

- (i) **What non-wage compensations** in today's office jobs? (No longer safety!)
- (ii) **What firms in what jobs** provide what amenities?
- (iii) How these relate to **wage inequality**?

Research Questions

- Empirical:

- (i) **What non-wage compensations** in today's office jobs? (No longer safety!)
- (ii) **What firms in what jobs** provide what amenities?
- (iii) How these relate to **wage inequality**?

- Theoretical:

- (i) What **else elements** can we use to explain firms' amenity provision patterns?
- (ii) Can it also **reconcile the empirical failures** of compensating differential?

What This Paper Does

1. Investigate the **patterns & wage effects of firm non-wage compensation provision** (both pecuniary & nonpecuniary) by using **job ads/vacancy data**
 - Difficult to observe in census/survey data
 - Study statistical patterns in the data
 - Discuss the inconsistency between findings and existing theories

What This Paper Does

1. Investigate the **patterns & wage effects of firm non-wage compensation provision** (both pecuniary & nonpecuniary) by using **job ads/vacancy data**
 - Difficult to observe in census/survey data
 - Study statistical patterns in the data
 - Discuss the inconsistency between findings and existing theories
2. Construct **a new & simple theory** of amenity provision and matching based on **productivity**
 - Orthogonal (not against!) to the classic channel of compensating differential
 - Reconcile the empirical findings and offer important implications

Outline

1. Introduction

2. Empirical Facts

3. Discussion

4. A Simple Theory

5. Conclusion

Data Source [▶ details](#)

- 4 million job postings in an online job board in China btw. 2013-2020
- 1/3 IT engineering jobs; 2/3 design, media, business operation, sales, admin, ..., jobs
- Each job declares job benefits along with job requirements and posted wages

A Sample Vacancy

Job Title
iOS开发工程师 **Wage**
18k-22k (该职位已下线)

深圳 / 经验1年以下 / 本科及以上学历 / web前端 / 全职

内容资讯 短视频 **Basic Job Info**

字节跳动 2018-09-10 发布于拉勾网 **Post Info**

收藏 已下线 完善在线简历 上传附件简历

查看原职位详情 -

职位诱惑: **Job Benefits**
六险一金, 弹性工作, 免费三餐, 餐补, 租房补贴, 带薪年假, 扁平管理, 晋升空间, 团队氛围好

职位描述: **Job Description and Requirement**

职位职责:

- 负责产品迭代改进及移动新产品的开发;
- 参与 APP 性能、体验优化及质量监控评估体系建设;
- 参与客户端基础组件及架构设计, 推进研发效率;
- 参与 hybrid 容器搭建, 插件、React Native 等动态技术调研。

职位要求:

- 本科及以上学历, 计算机相关专业;
- 热爱计算机科学和互联网技术, 对移动产品有浓厚兴趣;
- 扎实的数据结构和算法基础; 精通至少一门编程语言, 包括但不限于: Objective-C、Swift、C、C++、Java;
- 熟悉 iOS 平台原理, 具备将产品逻辑抽象为技术方案的能力;
- 关注用户体验, 能够积极把技术转化到用户体验改进上;
- 对新技术保持热情, 具备良好的分析、解决问题的能力。

工作地址

深圳 - 南山区 - 广东省深圳市南山区南海大道2163号来福士广场15层 **Work Address** 查看地图

Full
Vocabulary
Set
V

Amenity
Vocabulary
Subset

Vcomp C V

Firm Info

字节跳动

字节跳动

内容资讯, 短视频

D轮及以上

2000人以上

<http://www.bytedance.com>

Q1: What non-wage compensations firms use to attract workers?

Q2: Do high-skill jobs or high-wage firms provide non-wage compensations different from those lower ones?

Posted-Wage Regression

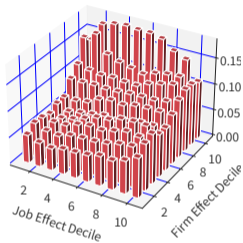
- Posted wage regression: $\ln w_{i,j,t} = \theta_i + \psi_j + \delta_i + \iota_t + \epsilon_i$
 - $\theta_i \equiv X_i\beta$ (job/worker effect), $X_i = \{\text{EDU}_i, \text{EXP}_i, V_{i,\setminus\text{comp}}\}$
 - ψ_j (firm fixed effect)
 - $\delta_i \equiv V_{i,\text{comp}}\gamma$ (compensation effect)
 - This posted wage regression does a similar job to the AKM framework (Zhu, 2022)

Fact 2: Systematic Differences in Compensation Provision Across Firm and Job Deciles

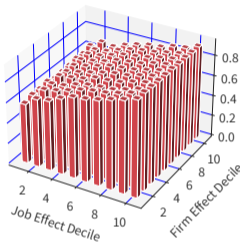
◀ details of selection ▶ more types

Occurrence Rate

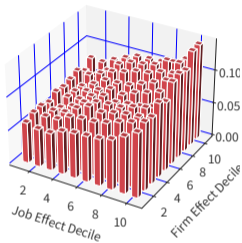
(a) Advanced Insurance



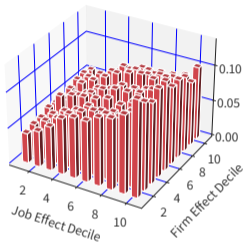
(b) Development



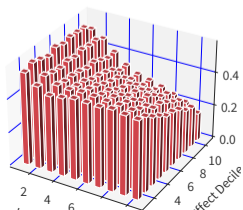
(c) Backloading Wage



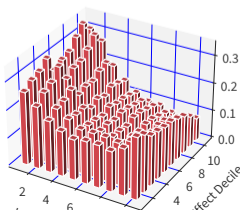
(d) Stock Option



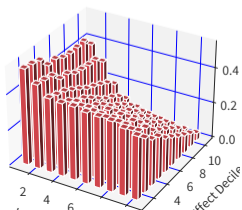
(e) Basic Insurance



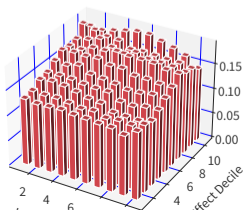
(f) Training



(g) Weekend, Holiday, Fixed Work-Time



(h) Work(-Time) Flexibility



Q3: How these non-wage compensations affect posted wages?

Fact 3: Hedonic Regression Results are Mixed but in A Systematic Way

	(1)	(2)	(3)	
Provided in High Skill/Premium Jobs/Firms	Advanced Insurance	.117** (.001)	.087** (.001)	.014** (.001)
	Backloading Wage	.054** (.001)	.030** (.001)	.010** (.001)
	Stock Option	.114** (.001)	.058** (.001)	.087** (.001)
	Coworker Quality	.140** (.001)	.059** (.001)	.024** (.001)
	Work-Flexibility	.046** (.001)	.032** (.001)	.010** (.001)
Provided in Low Skill/Premium Jobs/Firms	Basic Insurance	-.062** (.000)	-.046** (.000)	-.025** (.000)
	Training	-.057** (.001)	-.012** (.001)	-.003** (.001)
	Work-Time	-.113** (.001)	-.081** (.000)	-.021** (.000)
	Education FE	✓	✓	✓
	Experience FE	✓	✓	✓
	Year FE	✓	✓	✓
	C_{comp}		✓	✓
	Firm FE			✓
	Adj. R ²	.506	.633	.738
	No. Obs	3998840	3998840	3998840

Fact 3': Better Jobs/Firms Have Non-Wage Compensations with More Positive Wage Effects

◀ variance decomposition

Positively correlations between amenity (δ_i) and skill (θ_i) and firm FE (ψ_j) components

	With δ		Without δ	
	Comp.	Share	Comp.	Share
Var(ln w)	.362	-	.362	-
Var(θ_i)	.158	.437	.163	.450
Var(ψ_j)	.046	.128	.049	.136
Var(δ_i)	.002	.004		
Var(ϵ_i)	.097	.269	.098	.272
2 Cov(θ_i, ψ_j)	.049	.137	.052	.142
2 Cov(δ_i, θ_i)	.006	.017		
2 Cov(δ_i, ψ_j)	.003	.008		
Corr(θ_i, ψ_j)	.289		.288	
Corr(δ_i, θ_i)	.193			
Corr(δ_i, ψ_j)	.174			
Obs	3998840		3998840	
Firm	86165		86165	

Outline

1. Introduction

2. Empirical Facts

3. Discussion

4. A Simple Theory

5. Conclusion

Summary of Empirical Findings

1. Most non-wage compensations in the labor market are **common/standard ones**
→ unlikely many exogenous variations in costs
2. **High-paid firms in high-paid jobs** also in general provide **better non-wage compensations** (except for work-time)
→ compensation provision based on firm/worker productivity
3. However they are **not compensated from the posted wages**
→ at odds with the classic compensating differential theory

The Phantom of Unobserved Worker Ability

- Wait, isn't it just **unobserved worker ability** not-captured causing bias in above estimation? (Rosen, 1986; Hwang et al., 1992)

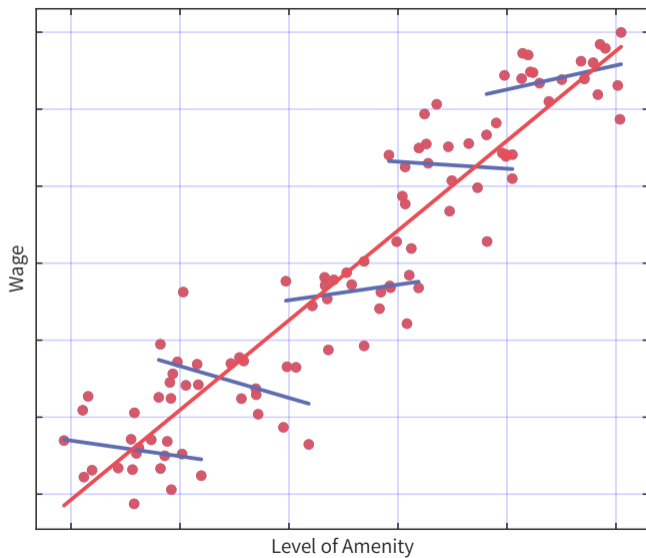
The Phantom of Unobserved Worker Ability

- Wait, isn't it just **unobserved worker ability** not-captured causing bias in above estimation? (Rosen, 1986; Hwang et al., 1992)

Two Counter-Arguments:

1. But **to what extent** would unobserved skill heterogeneity matter?
 - We have already added controls for **thousands of usually-unobserved job characteristics** beyond edu, exp, and occ
 - Unobserved job attributes is typically **positively correlated with observed job attributes**
2. Is compensation differential the most important force of **compensation inequality**?
 - The **empirical toughness of the omitted-variable problem** indicates other dominant mechanism of compensating dispersion

Unobserved Worker Ability \rightarrow Compensation Inequality?



Can Existing Theories Explain Positive Wage-Amenity Relationship?

- Hwang et al. (1992); Mortensen (2005): income effect
- Hwang et al. (1998): firms with low amenity-providing cost use both better amenity and higher wage to attract workers

Can Existing Theories Explain Positive Wage-Amenity Relationship?

- Hwang et al. (1992); Mortensen (2005): income effect
- Hwang et al. (1998): firms with low amenity-providing cost use both better amenity and higher wage to attract workers

Problems:

1. income effect cannot explain why it is low-pay firms provide **leisure** but not high-pay firms (e.g. notorious 996 working culture in Chinese IT industry)
 2. amenity-producing cost cannot explain why it is high-pay firms provide many superior amenities like **insurance or backloading wages**
 3. **sorting** is purely from exogenous heterogenous costs and preference or from wage-queue tradeoff
- Our model reconciles all these from a simple yet new angle

Outline

1. Introduction
2. Empirical Facts
3. Discussion
- 4. A Simple Theory**
5. Conclusion

Model Setting

- A continuum of worker with heterogenous productivity $q \in [0, 1]$ and additively separable (quasi-linear) utility function

$$U(C, a, h) = C + \underbrace{\phi_a a}_{\text{Amenity}} - \underbrace{\frac{h^{1+\phi_h}}{1+\phi_h}}_{\text{Disamenity}}$$

- $a \in \{0, 1\}$, e.g. advanced insurance
- $h \geq 0$, e.g. additional working hours
- Firms have O-Ring production function:

$$Y_j = AN_j^{1+\alpha} \prod_{i=1}^{N_j} q_i \cdot \underbrace{e(a, h)}_{\text{Efficiency Compensation}}$$

- **Compensations are (in)efficient:** $e(a, h) = 1 + \gamma_a a + \frac{h^{\gamma_h}}{\gamma_h}$
(microfoundations: e.g. less exogenous or endogenous exit (Hwang et al., 1998; Dey and Flinn, 2005); convexity in hour productivity (Goldin, 2014))

Equilibrium

- Firm problem:
$$\max_{\{q_i\}_{i=1}^N, a, h, w(q)} Y(q_1, \dots, q_N, a, h) - \sum_{i=1}^N w(q_i) - a\kappa N$$

s.t.
$$U(w(q), a, h) \geq u(q) \quad \forall q \in \{q_i\}_{i=1}^N$$

- $u(q)$ is the market equilibrium utility profile

- In the competitive equilibrium, a firm will employ workers with same q , and there is positive assortative matching (PAM)

Firms' Optimal Amenity Choices

- $a^* = \begin{cases} 1, & \text{if } q \geq q_a \\ 0, & \text{if } q < q_a \end{cases}$, and $\underbrace{AN^\alpha q_a^N \gamma_a + \phi_a}_{\text{marginal profit}} = \underbrace{\kappa}_{\text{marginal cost}}$
 - If a is not efficient, i.e. $\gamma_a = 0$, return back to the canonical compensating differential
 - If unit cost is $q\kappa$, higher q firms are still more likely to provide a
- $h^* = (AN^\alpha q^N)^{\frac{1}{1+\phi_h-\gamma_h}}$ increases in q
 - $h^*(q)$ will be fully compensated by $w(q)$, thus provision cost ex-post depends on q

Market Wage

$$- w(q) = \bar{A}q^N + \underbrace{\gamma_a \bar{A}q^N - \kappa}_{a \text{ effect}} + \underbrace{\frac{(\bar{A}q^N)^{1+\omega}}{(1+\omega)(1+\gamma_h)} + \frac{(\bar{A}q^N)^\omega}{1+\gamma_h}}_{h \text{ effect}}, \text{ if } q \geq q_a$$

- Recall $\gamma_a \bar{A}q^N - \kappa = -\phi_a$ when $q = q_a$ and can be positive when $q \uparrow$
→ **offsetting compensating differential** (for efficient amenity)
- $\frac{(\bar{A}q^N)^{1+\omega}}{(1+\omega)(1+\gamma_h)}$ is the efficiency gain from h ; $\frac{(\bar{A}q^N)^\omega}{1+\gamma_h}$ is the compensation for h
→ **magnifying compensating differential** (for efficient disamenity)

Model Implications

1. Results of empirical tests on compensating differential will depend on the targeted labor market
2. Compensation variations can be larger between-markets than within-markets
3. Efficiency compensations can enlarge both utility dispersion and wage dispersion
4. The set of non-wage compensations that can justify downward job moves is limited to inefficient amenities

Outline

1. Introduction
2. Empirical Facts
3. Discussion
4. A Simple Theory
- 5. Conclusion**

Take-Away Story

- Rather than saying something abstract and high-level, remember the initial story:
- Most large IT firms provide free and good meals from morning to evening
 - Not b.c. they have low costs to do so
 - Not b.c. their workers like free meals most
 - We argue b.c. they find it "efficient"!

Appendix

Preview of Empirical Findings

1. Firms use **common non-wage compensations** to **attract job seekers**:
 - insurance&fund package; leisure or work flexibility; bonus or backloading wage; growth potential; workplace environment/culture; fringe benefits; ...
 2. Info on non-wage compensations have **predictive power on posted wage** not because of their own variations, but of their **correlations with job/firm qualities**
 3. **Diff firms in diff jobs** have **distinct compensation-provision patterns**
 - High-wage firms w/ high-skill jobs: advanced insurance, backloading wage, ...
 - Low-wage firms w/ low-skill jobs: basic insurance, leisure, ...
 4. **Hedonic regression** shows **mixed results of compensating differential**
 - Compensations mainly posted by high-pay firms: counter evidences
 - Compensations mainly posted by low-pay firms: clear evidences
- Discuss how these findings are **inconsistent** with the views of existing theories

Preview of Theoretical Model

- We suggest a new theory that extends **compensating differential** with **efficiency compensation** and **productivity-based firm-worker sorting**
 - Key idea: **"Efficiency Compensation"**
 1. Many compensations observed in data are (in)efficiency compensation
 2. The level of efficiency depends on firm & worker productivity
 - Mechanism: **A new channel** works in addition to compensating differential
 1. When a compensation is **efficient**, it **counteracts** compensating differential effect
 2. When a compensation is **inefficient**, it **magnifies** compensating differential effect
 3. Extent of this (in)efficiency channel depends on firm-worker productivity sorting
- This simple modification **reconciles all findings** and generates many important **general implications**

Related Literature

1. Literature on Compensating Differential:

- Classic: Rosen (1974); Brown (1980); Rosen (1986); Hwang et al. (1992)
 - Recent: Mas and Pallais (2017); Maestas et al. (2018); Wissmann (2022) / Sorkin (2018); Taber and Vejlin (2020); Lamadon et al. (2022)
- New insights & New theory that reconciles existed empirical failures

2. Literature on Compensation Provision:

- Theory: Rosen (1974, 1986); Hwang et al. (1998); Hamermesh (1999); Mortensen (2005); Dey and Flinn (2005); Bonhomme and Jolivet (2009)
 - Empirical: Sockin (2022); Lachowska et al. (2022); Bana et al. (2022); Lamadon et al. (2022)
- New evidences & New theory that explains those new evidences

3. Literature on Efficiency Wage:

- Salop and Salop (1976); Shapiro and Stiglitz (1984); Katz (1986); Krueger and Summers (1988); Bloesch et al. (2021)
- Apply the insights to a more suitable place: "Efficiency Compensation"

Data Source

Lagou.com: the largest IT-centered online job board in **China**

- Over **6 million** job vacancies between 2013 and 2020 ▶ [vacancy trend](#)
- Mainly jobs in the **occupations** demanded by **IT-producing/using firms** and are **(routine or non-routine) cognitive**: Computer, Design & Media, Business Operation, Financial & Law, Sales, Admin ▶ [occupation classification](#)
- Like other vacancy data, **biased** to young/low-experienced & high education workers/jobs in large cities ▶ [summary statistics](#)
- Vacancy **information**: job name, **posted wage**, location, requirements on education and experience, **job task&skill description**, **job benefits**, firm name, ... ▶ [vacancy sample](#)
- **Final Sample** after cleaning: **4 million** job vacancies ▶ [sample cleaning](#)

Posted Compensation/Amenity Information

- Pros:

1. Hard to observe in census or survey data
2. Compensations or amenities that firms regard as important to attract workers
3. Also observe detailed job information

- Cons:

1. Not a full list of the compensations that a firm offer
2. Mainly amenities, rare disamenities (strategic hiding?)
3. Maybe cheap talk?

- Our empirical results will be mainly **descriptive & exploratory**

- No priori, let the data speak
- Find stylized facts of patterns & correlations in the data
- Shed new insights in thinking theories

Shortcomings & Some Reliefs

◀ Back Intro

◀ Back Data

- Vacancy data may be **selective or less representative**
 - Vacancy data is inclined to **young and more educated** workers, esp. here
 - **Not all jobs on the internet** or different post frequency than job composition

*(Valid issue for all vacancy data; Extent is an **empirical question**; With dev and structural transform, more and more likely to **become the dominant cases**; help to consider the aging worker cases)*

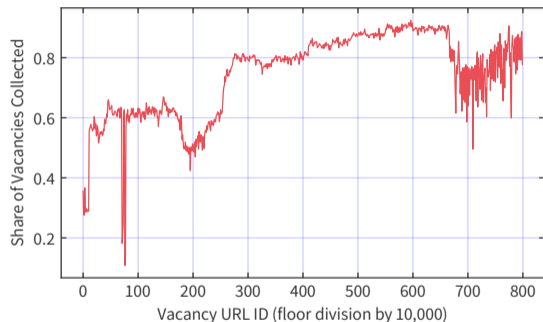
- Our wage measure incorporates **variation in hours**
 - One might worry that those efficient compensations are **solely compensating more working hours**

*(Often **additional pay** for overtime hours; **Variation is limited** comparing to wage; Inequality is often considered on **overall compensation level**; Need to think **hour and wage as a package**)*

Trends on Collected Vacancies [◀ Back](#)



of Vacancies by Posted Month



% of Vacancies by Post ID Chunks

Unstructured Text Data

- V : full vocabulary set with 110,000+ tokens/features (i.e. words or terms)
- $V_{\text{comp}} \subset V$: compensation vocabulary set with 13,000+ features
 - Not all uniques: synonyms, different versions, typos
 - Common words or stop words
 - Irrelevant texts
- $\mathbf{C}_{\text{comp}} \in \mathbb{R}^{N \times |V_{\text{comp}}|}$: an indicator matrix to run regression
- So, high-dimensional data \rightarrow (basic) Machine Learning methods

- Drop vacancies with not full-time jobs, outlier wages, job descriptions less than 20 words, nonChinese content
- Drop vacancies in 2013
- Drop vacancies from firms with less than 10 posts and from all the locations that have less than 1000 vacancies
- Drop duplicated vacancies based on job descriptions and education and experience requirements
- Drop vacancies with occupations not in selected major occupations

Data: Occupation Classification [◀ Back](#)

- No ready-for-use occupation classification
- Match to a set of selected 6-digit occupations ("minor") in six 2-digit occupations ("major") in U.S. SOC 2018
- Key idea: an occupation is defined by a bundle of skills and tasks
- 1st step: for each occupation choose several exclusive keywords, and find the set of just-match vacancies as the "learning" sample
- 2nd step: use the "learning" group to train a Naive Bayes classifier based on the job titles and job descriptions
- 3rd step: apply the trained classifier to both the "unknown" sample and the "learning" sample [▶ confusion matrix](#)

Data: Summary Statistics [◀ Back](#)

	Pooled		Major Occupation				
	-	Computer	Design_ Media	Business_ Operations	Financial_ Legal	Sales	Admin
Vacancy #	3,999,005	1,330,001	561,236	1,162,404	214,661	452,771	277,932
- share	1.00	.33	.14	.29	.05	.11	.07
Avg # Words	108.91	104.26	103.05	115.60	110.69	120.31	95.09
Wage (1k CNY):							
- Mean	13.64	17.38	10.68	14.19	11.95	10.21	6.32
- SD	9.24	9.79	6.31	9.52	9.19	6.53	3.90
Firm:							
- #	86,330	67,369	68,092	78,244	41,285	58,847	59,016
- Avg Posts	46.32	19.74	8.24	14.86	5.20	7.69	4.71
- Median Posts	20.0	9.0	4.0	6.0	2.0	3.0	2.0
Firm Size (share):							
- -15	.03	.03	.05	.02	.02	.03	.03
- 15-50	.18	.17	.25	.16	.15	.19	.20
- 50-150	.23	.21	.26	.22	.22	.23	.26
- 150-500	.21	.21	.21	.22	.23	.20	.23
- 500-2000	.15	.16	.12	.16	.18	.15	.14
- 2000+	.20	.23	.11	.22	.21	.19	.13
Education (share):							
- Vocational College	.33	.24	.38	.29	.27	.51	.52
- Bachelor	.54	.66	.47	.61	.63	.22	.24
- Master/Doctor	.01	.02	.00	.01	.03	.00	.00
- Not Specified	.12	.08	.15	.09	.07	.27	.23
Experience (share):							
- 0	.22	.12	.21	.16	.25	.48	.50
- 1-3	.37	.33	.48	.37	.36	.31	.38
- 3-5	.31	.41	.25	.33	.26	.16	.10
- 5-10	.11	.14	.05	.14	.13	.05	.03

What Are The Non-Wage Compensations That Firms Post?

[◀ Back](#)



insurance&fund; leisure; growth potential, bonus, environment, fringe benefits, ...

Lasso Regression using V_{comp} : Top Features (Frequency > 1%)

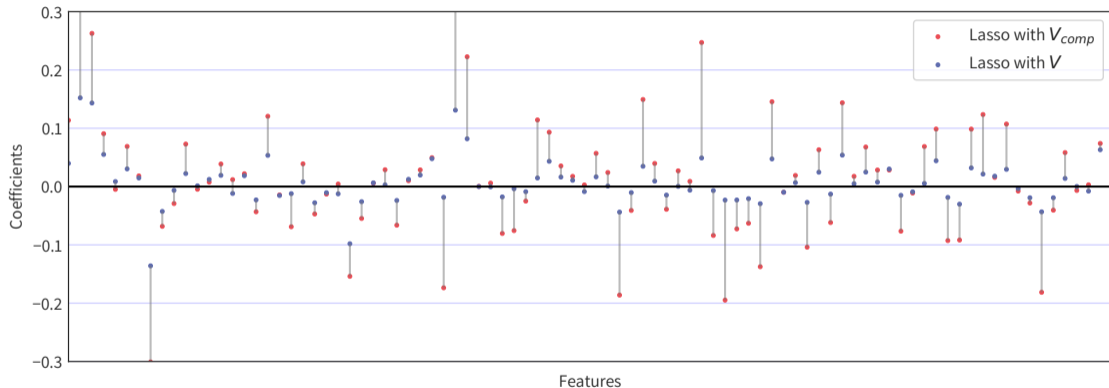
[lasso details](#)

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

Lasso Regression using V: Top Features (Frequency > 1%)

	Top Positive			Top Negative		
	token	coef	freq	token	coeff	freq
1	<u>14th month pay</u>	.152	.014	freshmen	-.155	.018
2	<u>three meals</u>	.143	.014	<u>five insurance</u>	-.136	.030
3	<u>large platform</u>	.131	.019	graduates	-.128	.033
4	master degree	.126	.015	vocational major	-.100	.036
5	lead	.107	.041	<u>two-day weekend</u>	-.098	.166
6	c++	.092	.051	vocational college	-.094	.148
7	algorithm	.082	.061	assistant	-.079	.011
8	<u>guru</u>	.082	.028	customer service	-.075	.030
9	famous	.079	.019	<u>social insurance</u>	-.073	.028
10	machine learning	.077	.016	accounting	-.071	.019
11	formation	.076	.013	<u>accommodation</u>	-.067	.016
12	undergraduate	.074	.319	administration	-.067	.027
13	overseas	.072	.026	commissioner	-.063	.011
14	react	.072	.020	taobao	-.059	.015
15	<u>development</u>	.071	.374	assistance	-.058	.164
16	undergraduate	.066	.029	ps	-.056	.029
17	<u>high salary</u>	.063	.028	ltd.	-.056	.012
18	landing	.060	.067	installation	-.055	.020
19	strategy	.057	.047	photoshop	-.052	.039

Fact 2a: Firm Non-Wage Compensations Correlated With Job Attributes that Can Explain Posted Wages



◀ All V'_{comp}

Lasso Regression

[◀ Back](#)

- Two purposes: (i) a first look of the wage-amenity relationship (ii) shrink features
- Run a **Lasso regression** of log posted wage $\ln w$ on an **indicator matrix** $\mathbf{C}_{\text{comp}} \in \mathbb{R}^{N \times |V_{\text{comp}}|}$
 - Use **BIC** to tune the Lasso penalization hyper-parameter [▶ lasso details](#)
- It shrinks V_{comp} to a **vocabulary subset** V'_{comp} with only 800 features (and \mathbf{C}_{comp} to $\mathbf{C}'_{\text{comp}}$)
- Inference & Robustness:
 - Coefficients are in general **not interpretable** due to multicollinearity & flexibility
 - Use **subsampling** to do inference, results are robust [▶ subsampling](#)
- Conduct **same Lasso regression** for $\mathbf{C} \in \mathbb{R}^{N \times |V|}$, and **inspect** top features & changes

- Lasso regression (L1 penalization):

$$\hat{\zeta} = \arg \min_{\zeta} \sum_{i=1}^N \left(\ln w_i - \sum_{k=1}^K c_{ik} \zeta_k \right)^2 + \lambda \sum_{k=1}^K |\zeta_k|$$

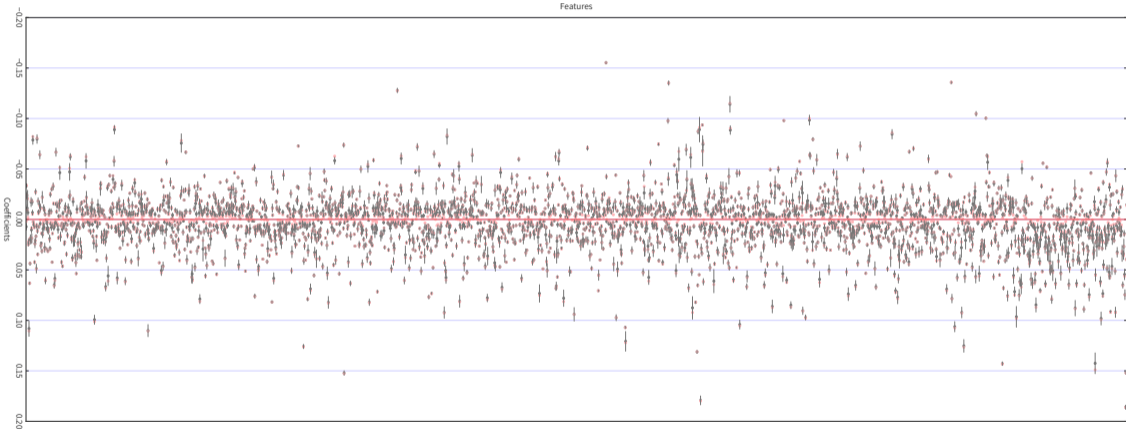
- BIC as the criterion to gauge the hyperparameter λ :

$$\min \text{BIC}(\lambda) = \frac{\| \ln \mathbf{w} - \mathbf{C} \hat{\zeta}_{\lambda} \|^2}{\sigma^2} + \hat{df}_{\lambda} \log N$$

- Inference via [subsampling](#) (10x10)

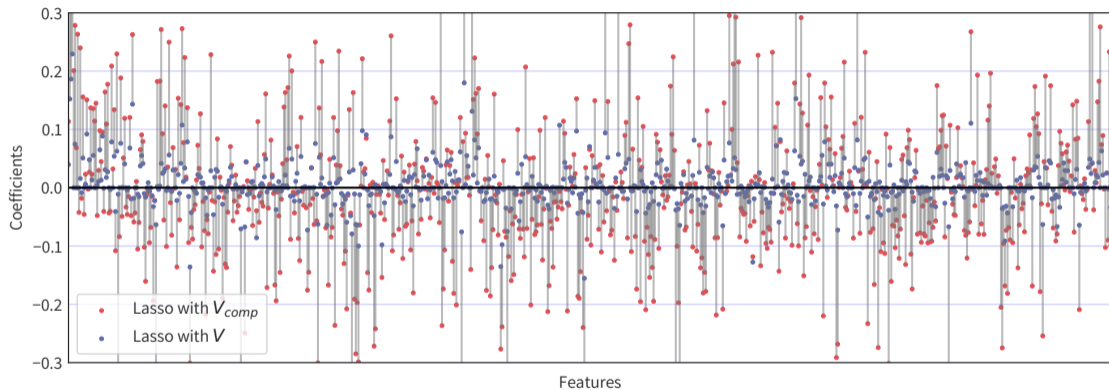
Confidence Intervals on Lasso Coefficients via Subsampling

[◀ Back](#)



Compare Lasso Coefficients

[◀ Back](#)



Posted-Wage Regression

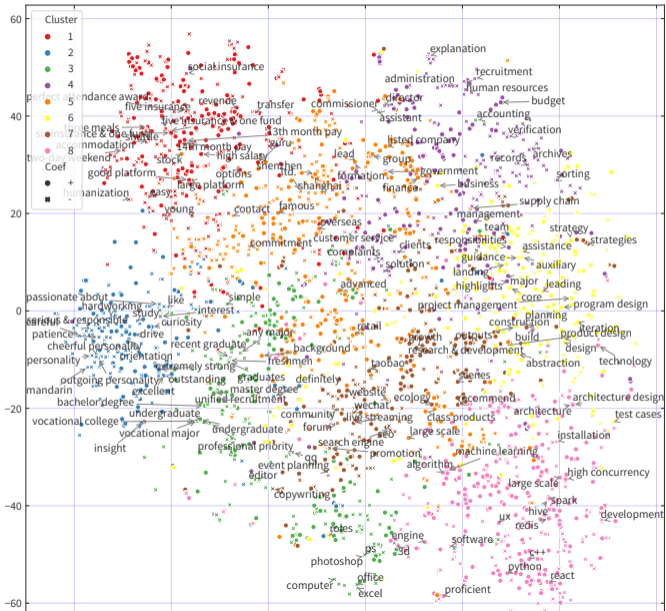
- Examine linkage of amenities with both **job skills/tasks** and **firm premiums**
- **Posted wage regression:** $\ln w_{i,j,t} = \theta_i + \psi_j + \delta_i + \iota_t + \epsilon_i$
 - $\theta_i \equiv X_i \beta$ (job/worker effect), $X_i = \{\text{EDU}_i, \text{EXP}_i, \mathbf{c}'_{i,\setminus \text{comp}}\}$
 - ψ_j (firm fixed effect)
 - $\delta_i \equiv \mathbf{c}'_{i,\text{comp}} \gamma$ (**compensation effect**)
 - ι_t (year fixed effect)
 - In practice, further dimensional reduction on $\mathbf{c}'_{i,\setminus \text{comp}}$ & $\mathbf{c}'_{i,\text{comp}}$ using PLS
 - This posted wage regression does a similar job to the AKM framework ([Zhu, 2022](#))
- **Variance decomposition:** $\text{var}(\ln w_i) =$
 $\text{var}(\theta_i) + \text{var}(\psi_j) + \text{var}(\delta_i) + 2 \text{cov}(\theta_i, \psi_j) + 2 \text{cov}(\theta_i, \delta_i) + 2 \text{cov}(\psi_j, \delta_i) + \text{var}(\epsilon_i)$

Gather Important Types and Check Occurrence

- We can take a direct look on if high wage firms & high jobs are accompanied with low valued amenities
- We do this by picking a set of **well-defined and well-studied compensation types/topics** from V_{comp} based on the frequencies & Lasso coefficients
- We **gather all relevant terms** by checking proximate terms in the **embedding space of a work-embedding model** trained on the whole job texts
- We then examine how the **occurrence ratio** for each type differ across different firms & jobs

Feature Clustering: Visualization

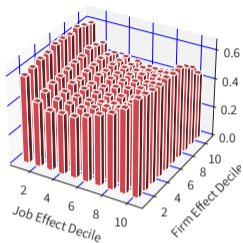
[◀ Back](#)



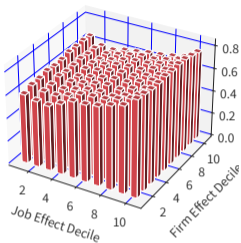
Compensation Occurrence (More)

[◀ Back](#)

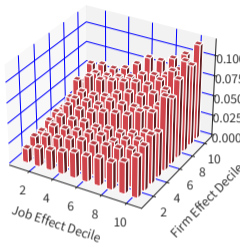
(a) Management



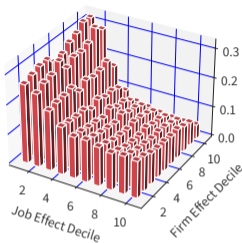
(b) Environment



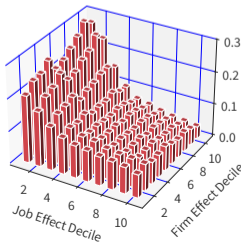
(c) Coworker Quality



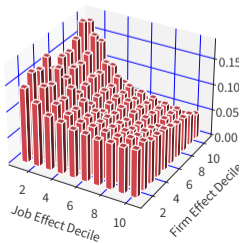
(d) Commission



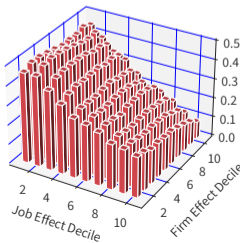
(e) Promotion



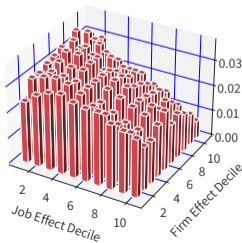
(f) Location



(g) Team Building



(h) Coworker Looking



Hedonic Regression

	Pooled	Computer	Design_ Media	Admin
	(1)	(2)	(3)	(4)
Advanced Insurance	.014** (.001)	.016** (.001)	.009** (.002)	.002 (.003)
Backloading Wage	.010** (.001)	.013** (.001)	.022** (.002)	.011** (.002)
Stock Option	.087** (.001)	.068** (.001)	.060** (.002)	.040** (.003)
Coworker Quality	.024** (.001)	.016** (.001)	.005* (.002)	.008+ (.004)
Work-Flexibility	.010** (.001)	.007** (.001)	.009** (.001)	.005** (.002)
Basic Insurance	-.025** (.000)	-.024** (.001)	-.017** (.001)	-.013** (.001)
Training	-.003** (.001)	-.019** (.001)	-.003 (.002)	.013** (.002)
Work-Time	-.021** (.000)	-.018** (.001)	-.020** (.001)	-.022** (.001)
Education FE	✓	✓	✓	✓
Experience FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Ξ_2, \dots, Ξ_8	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Adj. R ²	.738	.748	.730	.657
No. Obs	3998840	1325260	548808	260364

Model Overview

- **Workers:** heterogeneous in productivity; homogenous in preference
- **Firms:** heterogeneous in productivity; homogenous in (dis)amenity prod func
- **Firm-worker sorting** is thus solely based on productivity
 - In the classic Rosen model, sorting purely on worker preference & firm cost
 - Reality is likely a mix and depends on demographics ([Lentz et al., 2021](#))
 - We use O-Ring structure, so workers-sorting & only ex-post firm heterogeneity
- Key element: **Non-wage compensations can be "(in)efficient"**
 - Motivation: efficiency wage literature and the critiques; [Dey and Flinn \(2005\)](#)
 - Various micro-foundations: here the simplest way—"inducing effort"
 - Extra feature: the level of (in)efficiency depends on productivity sorting
 - We set one efficient amenity and one inefficient amenity for illustration

Competitive Equilibrium & Matching

- **Competitive equilibrium** in this economy is defined as an **assignment** of worker types to firms and a **utility schedule**, $u(q)$ such that
 - Firms maximize their profits
 - Labor market clears

- Firm problem:
$$\max_{\{q_i\}_{i=1}^N, a, h, w(q)} AN^{1+\alpha} \prod_{i=1}^N q_i e(a, h) - \sum_{i=1}^N w(q_i) - a\kappa N$$

s.t.
$$w(q) + \phi_a a - \frac{h^{1+\phi_h}}{1+\phi_h} \geq u(q) \quad \forall q \in \{q_i\}_{i=1}^N$$

- Complementary production function & additively separable utility function ensure **positive assortative matching (PAM)** even under imperfect transferable utility
 → a firm will employ workers with same q

Firms' Optimal Choices on Compensations

- Rewrite the firm problem:

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

$$- a^* = \begin{cases} 1, & \text{if } q \geq q_a \\ 0, & \text{if } q < q_a \end{cases}, \text{ and } \underbrace{AN^\alpha q_a^N \gamma_a + \phi_a}_{\text{mb}} = \underbrace{\kappa}_{\text{mc}}$$

- If a is not efficient, i.e. $\gamma_a = 0$, return back to the canonical compensating differential
- If unit cost is $q\kappa$, higher q firms are still more likely to provide a

$$- \text{FOCs: } \begin{aligned} AN^{1+\alpha}q^{N-1}e(a, h) &= u'(q) \\ AN^\alpha q^N h^{\gamma_h-1} &= h^{\phi_h} \end{aligned}$$

- $h^* = (AN^\alpha q^N)^{\frac{1}{1+\phi_h-\gamma_h}}$ increases in q
 - $h^*(q)$ will be fully compensated by $w(q)$, thus provision cost ex-post depends on q

Market Wage

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

- where $\bar{A} \equiv AN^\alpha$, $\omega = \frac{1+\gamma_h}{1+\phi_h-\gamma_h}$, $u_0 = 0$, and $u_a = \phi_a - \kappa$.

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

- Recall $\gamma_a\bar{A}q^N - \kappa = -\phi_a$ when $q = q_a$ and can be positive when $q \uparrow$

→ offsetting compensating differential

- $\frac{(\bar{A}q^N)^{1+\omega}}{(1+\omega)(1+\gamma_h)}$ is the efficiency gain from h ; $\frac{(\bar{A}q^N)^\omega}{1+\gamma_h}$ is the compensation for h

→ magnifying compensating differential

If Firm Size Is Endogenous (Typical O-Ring Results) [◀ Back](#)

- N is also a choice of the firm
- Additional FOC: $AN^\alpha q^N e(a, h) (1 + \alpha + N \ln(q)) = w + ac$
- Optimal choice on firm size: $N(q) = \frac{1+\alpha}{-\ln(q)}$
- Firm size increases in productivity q and is irrelevant to the choices of amenities
- All the relationships between productivity and amenity provision can be now directly translate to the firm size

Model Implications 1. Compensating Differential

1.1 The result of an empirical test on compensating differential depends on the targeted labor market

- If focusing on the low-end labor market (e.g. close to q_a or $q < q_a$ with imperfectly mandated policies), it is easy to find clear evidence of compensation differential
- If focusing on board or high-end labor market (and with heterogeneity in the use of efficiency compensation or imperfect match), such test are likely to fail

1.2 Compensating effects can be confounded with productivity effects

- Esp. for the up-end labor market where (in)efficiency forces are strong

1.3 Available variations for wage-amenity packages can be limited conditional on worker

- Depends on exogenous heterogeneity v.s. endogenous heterogeneity
- Constrains on both low-end and high-end markets

→ Those identified under field/choice experiments (WtP) or RCT-like experiments (exogenous variations) not necessarily capture how the labor market works

Model Implications 2. Labor Market Inequality

- 2 Efficiency compensations can **enlarge both utility dispersion & wage dispersion**
 - Ignoring non-wage compensations can underestimate labor market inequality
 - Moreover those compensations per se can actually be the drivers of wage inequality
- Increased sorting or better use of efficiency compensations increases wage inequality

Model Implications 3. Job Mobility & Choice

3.1 The set of non-wage compensations that can justify job moves to low wage-premium firms is likely **limited to inefficient amenities**

- Work-time/effort is the most likely culprit for moving downgrade

3.2 **Greater compensating** than just "compensating differential"

- A worker with a ϕ_h shock would suffer not only traditional compensation differential but also a worse matching & an inferior package of other compensations
- Again, available choices for wage-amenities packages are limited

→ Potential implications for gender wage gap and etc.

Take-Away Message

1. **Think explicitly** about non-wage compensations: insurance/fund, work-time, pay schemes, work environment, fringe benefits, ...
→ empirical focus & policy targets & intuition when back-out revealed preference
2. Different Firms in different jobs have **distinct provision patterns**
→ compensating differential \neq provision inequality
3. **(In)Efficiency compensations & productivity sorting** reconciles empirical findings and generates important implications
→ high-wage firms can also offer better compensations without wage discounts

Future Plan/Possibility

- Model the posted compensation as a discrete choice of firms?
- Interact/Distinguish with the income effect?
- Allow for heterogeneous preference and multi-dimensional sorting?
- Allow for search frictions and mismatch?
- Bring the model implications to the data?
- Combine with worker self-reporting data?
- Test if the similar empirical facts in the Japanese Data?