Learning about Pay and Non-pay from Vacancy Texts

Richard Audoly[1] Manudeep Bhuller[2] Tore Adam Reiremo[2]

[1]New York Fed

[2]University of Oslo (UiO)

August 30, 2023

Motivation

- Long tradition of seeing employers as more than pay [Rosen 1986]

  $\rightarrow$ "pay and non-pay", "amenity", "compensating differentials"

- Wide class of models where workers know the value of alternative employers j. For instance, [Card et al. 2018]

$$\max_{j \in \mathcal{J}} \; u_j = \beta \ln w_j + a_j + \varepsilon_{ij}$$

This paper: Is there information in vacancy texts?

What we do

1. Retrieve and categorize advertised job attributes from the text of 3+ million job ads in Norway

   $\rightarrow$ Detect up to 46 job attributes per ad

What we do

1. Retrieve and categorize advertised job attributes from the text of 3+ million job ads in Norway

   $\rightarrow$ Detect up to 46 job attributes per ad

2. Estimate value of employers using a revealed preference approach

   $\rightarrow$ Employers differ in pay, non-pay, and job stability

What we do

1. Retrieve and categorize advertised job attributes from the text of 3+ million job ads in Norway

   $\rightarrow$ Detect up to 46 job attributes per ad

2. Estimate value of employers using a revealed preference approach

   $\rightarrow$ Employers differ in pay, non-pay, and job stability

3. Quantify information value of publicly advertised attributes

   $\rightarrow$ Link extracted attributes to estimated value

Related Literature

- **Text of vacancies** Marinescu and Wolthoff [2020], Deming and Khan [2018], Deming and Noray [2020], . . .

  → Focus on "We offer" part of ads

- **Wage and non-wage attributes** Mas and Palais [2017], Card et al. [2018], Sorkin [2018], Lamadon et al. [2020], Morchio and Moser [2020], . . .

  → Estimates of value of employers in pay, nonpay, and job stability

- **Unpacking search frictions?** Belot et al. [2019], Jaeger et al [2021], . . .

  → Information featured on vacancies

Outline

The Norwegian vacancy data

- 3+ million job ads posted in Norway between 2002 and 2019
  - Most of the talk: 2015-2019 period
- Near-universe of publicly posted ads: legal requirement to report job openings (Labor Market Act §7)
- Main sources: online job boards, newspaper
- For each vacancy: industry, occupation, establishment ID, opening date, ...
- ... and full text of job ad

Goal: Extract information on job attributes in the text of job ads

A sample vacancy [English from Google Translate]

IT-consultant

About the job: Full-time, Permanent position

We are looking for a new team player for our IT department! As an IT consultant, you will be responsible for our IT environment. It also includes getting to know large parts of our organization well through having internal user support as well as employee training. You simply get a varied working day, with different tasks to brush up on. In the future, you will also assist in larger IT development projects [...]

Tasks:

- IT security.
- SharePoint.
- Data warehouse.
- SOX Compliance.

Required qualifications:

- Likes IT technical challenges.
- Sociable and likes to communicate with people.
- Has relevant education.
- Has experience from similar positions.

We can offer:

- A workplace with many varied tasks and a high focus on quality.
- Training in our key products/services.
- Good working environment with solid professional expertise.
- Flexible working hours.
- Competitive conditions.

[...]

A sample vacancy [English from Google Translate]

IT-consultant

About the job: Full-time, Permanent position

We are looking for a new team player for our IT department! As an IT consultant, you will be responsible for our IT environment. It also includes getting to know large parts of our organization well through having internal user support as well as employee training. You simply get a varied working day, with different tasks to brush up on. In the future, you will also assist in larger IT development projects [...]

Tasks:

- IT security.
- SharePoint.
- Data warehouse.
- SOX Compliance.

Required qualifications:

- Likes IT technical challenges.
- Sociable and likes to communicate with people.
- Has relevant education.
- Has experience from similar positions.

We can offer:

- A workplace with many varied tasks and a high focus on quality.
- Training in our key products/services.
- Good working environment with solid professional expertise.
- Flexible working hours.
- Competitive conditions.

[...]

Detecting job attributes in vacancy texts

1. Manually label 1,200 ads creating job attributes "labels" and noting key phrases

2. Unsupervised topic modelling using structured lists: (i) tasks, (ii) requirements, (iii) "We offer", (iv) residual [Blei et al., 2003]

   - "We offer" lists are additional source of phrases for attributes

3. Find closely related phrases for each attribute using Common Bag of Words [Mikolov et al., 2013]

   - Obtain full "dictionary" of <u>1,478 phrases for 46 attributes</u>

4. Identify attributes in the full vacancy corpus

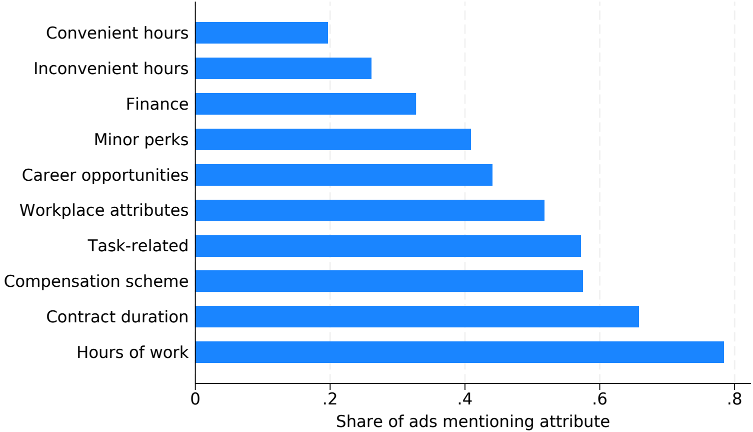
   - Evaluate performance on random sample with manual recognition

Summary of main extracted attributes

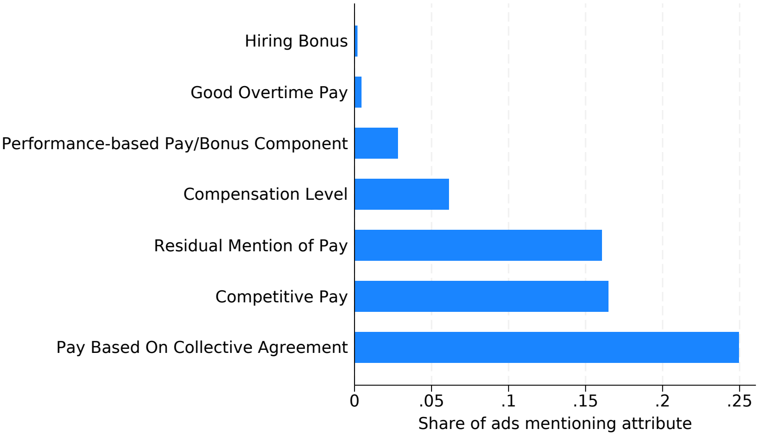
Main Category	Examples:
A. Pay Attributes	
– Compensation Scheme	Salary band; Bonuses; CBA-based pay
– Financial Attributes	Insurance scheme; Pension scheme
– Career Opportunities	On-the-job training; Career opportunities
B. Non-Pay Attributes	
– Hours of Work	Full-time; Part-time; On-call Job
– Convenient Hours	Flexible working hours
– Inconvenient Hours	Shift-work; Weekend/evening/nights
– Contract Duration	Permanent; Temporary
– Workplace Attributes	Social environment; Remote work
– Task-Related Attributes	Interesting tasks; Challenging tasks
– Other Minor Perks	Location; Company cabin; Company gym

Share of ads with aggregate attributes

- 95% of ads mention at least one attribute
- 6.6 attributes per ad on average

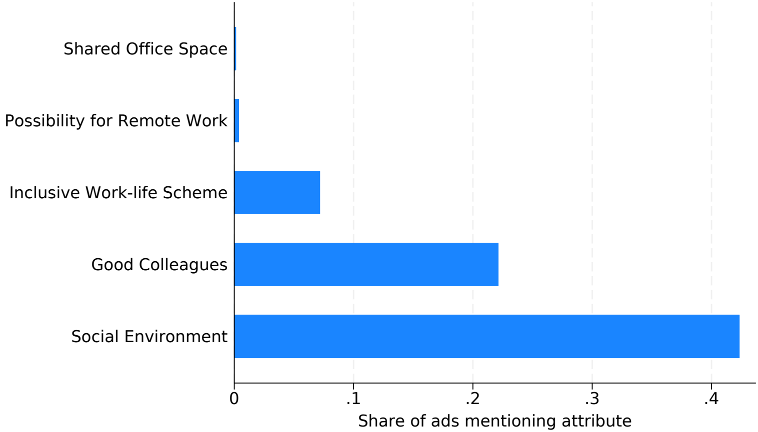


Share of ads with specific category: Pay-related



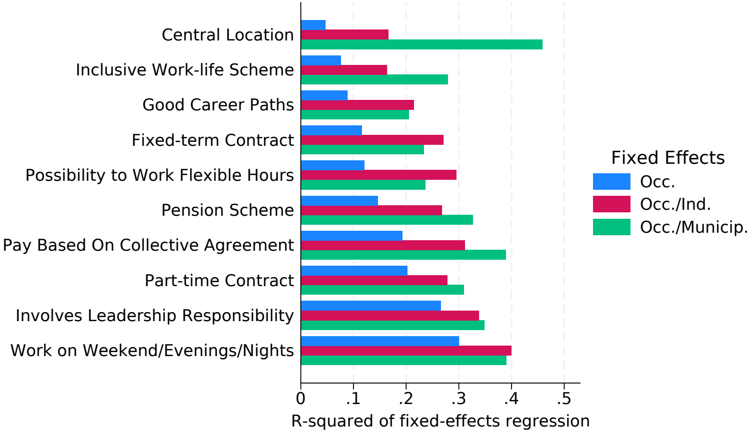
- Less than 10% of ads mention level of pay
- Additional pointers through collective agreements

Share of ads with specific category: Workplace



- Employers frequently use words to say: "it's nice to work here"

Is this all about occupation and industry?



- $R^2 < .4$ regressing attribute indicators on rich set of FEs

Outline

Objectives

- Estimate value of heterogenous employers using a revealed preference approach
- Multiple dimensions of this heterogeneity, such as pay and non-pay

Burdett Mortensen [1998] framework with "value posting" Sorkin [2018], Morchio and Moser [2020]

Value of working at employer j

$$
\begin{aligned}
V_j = {} & \ln W_j + a_j && \text{(flow utility: pay and non-pay)} \\
& + \beta\delta_j \mathbb{E}\Big[V_N + \varepsilon_N\Big] && \text{(destruction shock – } EN\text{)} \\
& + \beta\rho_j \sum_{k \neq j} f_k \mathbb{E}\Big[V_k + \varepsilon_k\Big] && \text{(relocation shock – forced } EE\text{)} \\
& + \beta(1-\delta_j-\rho_j)\lambda_1 \sum_{k \neq j} f_k \mathbb{E}\Big[\max\left\{V_j + \varepsilon_j, V_k + \varepsilon_k\right\}\Big] && \text{(offer)} \\
& + \beta(1-\delta_j-\rho_j)(1-\lambda_1)\mathbb{E}\Big[\max\left\{V_N + \varepsilon_N, V_j + \varepsilon_j\right\}\Big] && \\
& && \text{(no offer)}
\end{aligned}
$$

- All $\mathbb{E}[.]$ over $\varepsilon-$shocks $\sim$ i.i.d. Gumbel$(0, \sigma)$
- Essentially BM98 with idiosyncratic taste shocks

Decomposing value of firms V_j

1. Pay premium: $\ln W_j := \psi_j$
2. Non-pay/compensating differentials: a_j
3. Job security: $s_j := 1 - \delta_j - \rho_j$

Identification

1. EE, NE, EN flows identify $\{V_j, \rho_j, \delta_j, f_j, \lambda_1\}$ up to scale of taste shocks σ within set of strongly connected set of employers [Sorkin 2018]

2. Assuming piece-rate wage contracts [Barlevy 2008], AKM firm effects identify $\ln W_j$ (strongly connected $\subset$ connected set)

3. Pick a scale for σ to recover a_j

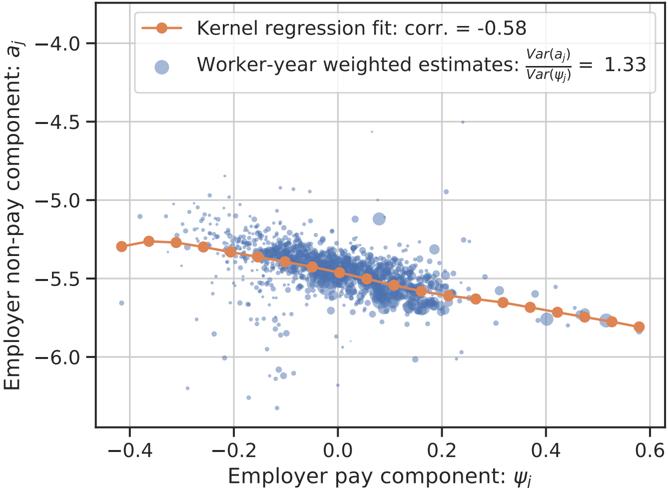
    - Assumption that utility flow has scale std($\ln W_j$)

Strongly selects larger, more stable establishments, but 70+ percent of employment retained (firm-size distribution)

Estimation

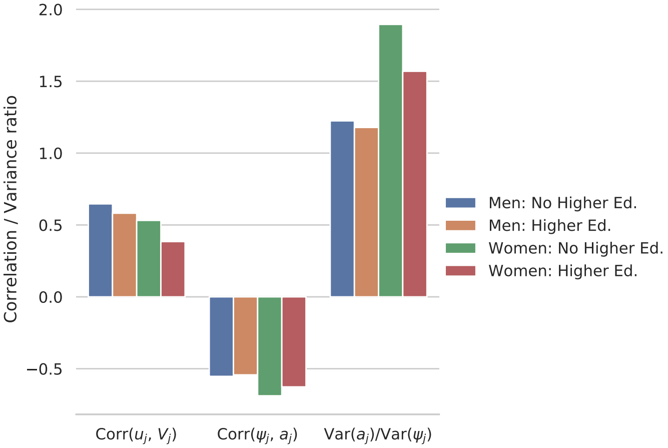
- Model estimated on Norwegian employer-employee data (start and end date of each spell, earnings and hours, some demographics)
- k-means to group establishment N_j into N_g groups (mapping $j \mapsto g$)
  - Grouping variables: (i) wage distribution, (ii) job creation/destruction rates, (iii) industry, occupation, location
  - Weighted by employment
  - Firm FEs computed using these clusters (address limited mobility bias)
- Split sample by gender and education to allow for (some) worker heterogeneity

Pay and non-pay in model



- Non-pay component matters

Heterogeneity by subgroup



- Non-pay seems to matter more for women

Outline

Advertised attributes and value of employers

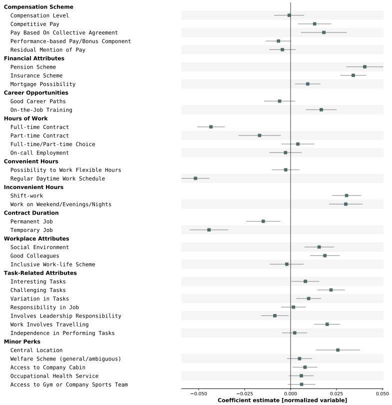
Quantify importance of advertised attributes with regression model

$$y_{g(j)} = \sum_k \beta_k \cdot \text{Attributes}_{k,g(j)} + \gamma \cdot X_{g(j)} + \epsilon_g$$

- $y_{g(j)}$: Estimated value of employer
- $\text{Attributes}_{k,g(j)}$: Share of ads mentioning extracted attribute k
- $X_{g(j)}$: occupation, industry, location proxy
- Stacked regressions, all weighted by employment

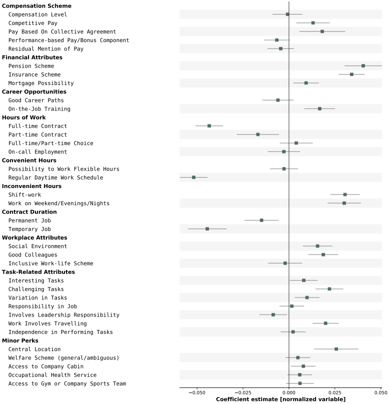
Separate regressions on attributes & controls

(a) Pay

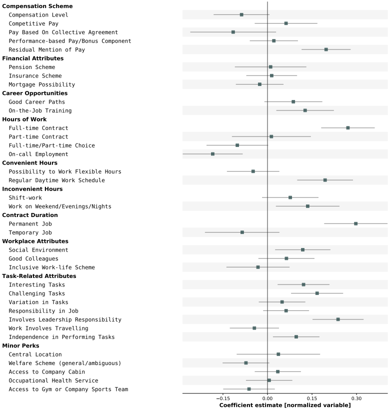


Separate regressions on attributes & controls

(a) Pay

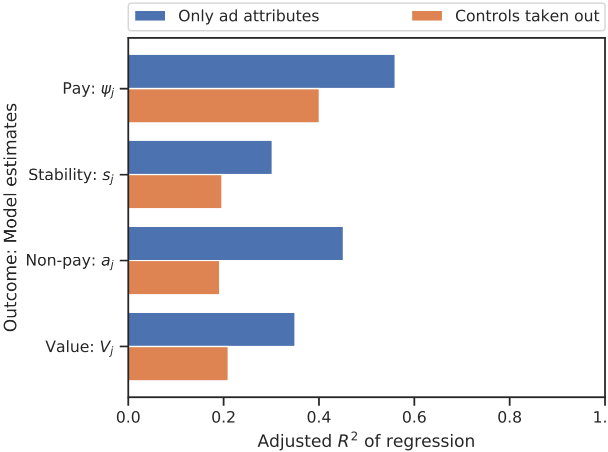


(b) Non-pay



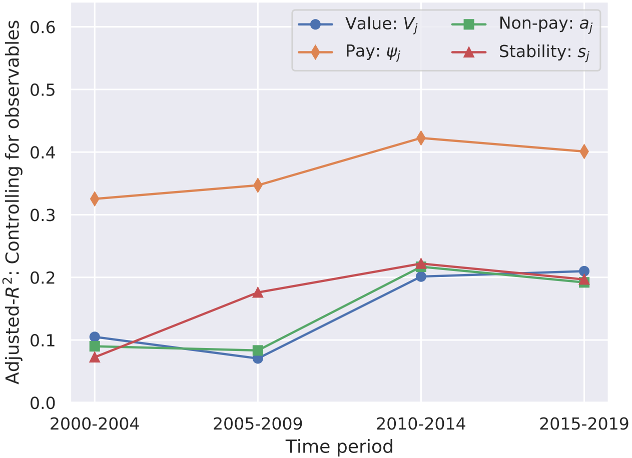
- High amenity employers advertise better hours/duration <u>and</u> some softer attributes (workplace, tasks)

Explanatory power of ad attributes



- Ad attributes have prediction power above controls

Trend in explanatory power of ad attributes



- Explanatory power of ad attributes shows slight increase after 2020

A mobility based measure of the information value of ads

Two alternative information set for workers:

1. Predict value of employers using all information, including ads:

$$\check{V}_{g(j)} = \sum_k \check{\beta}_k \cdot \text{Attributes}_{k,g(j)} + \check{\gamma} \cdot X_{g(j)}$$

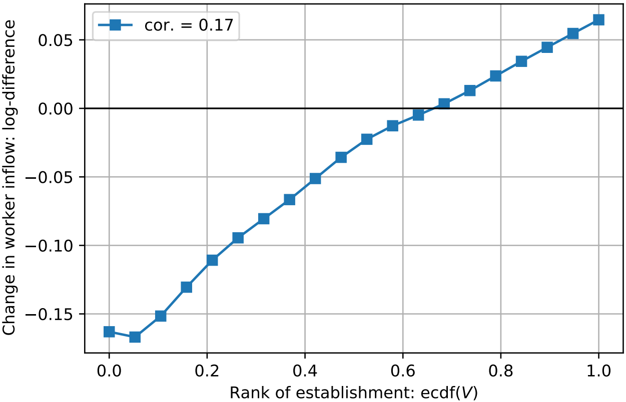
2. Predict value of employers <u>without</u> ad information:

$$\hat{V}_{g(j)} = \hat{\gamma} \cdot X_{g(j)}$$

Counterfactual (partial equilibrium) mobility where workers make decisions as:

$$\max\left\{V_j + \varepsilon_j, \hat{V}_k + \varepsilon_k\right\} \quad \text{vs.} \quad \max\left\{V_j + \varepsilon_j, \check{V}_k + \varepsilon_k\right\}$$

Counterfactual EE inflows



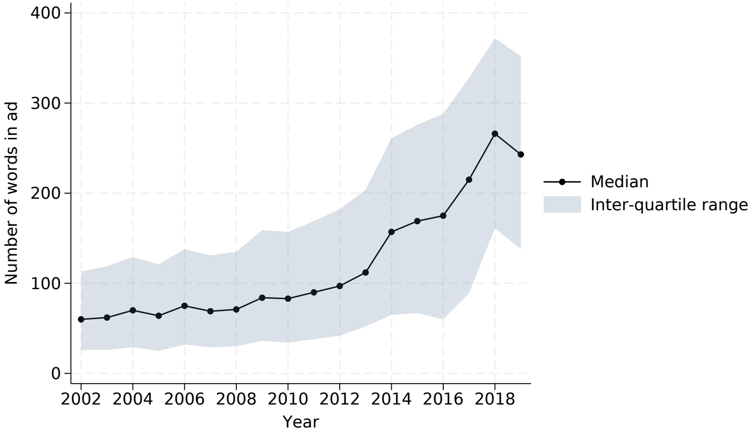
- Ad information improves mobility toward higher value employers
- Find correlation measure is relatively stable over time

Conclusion

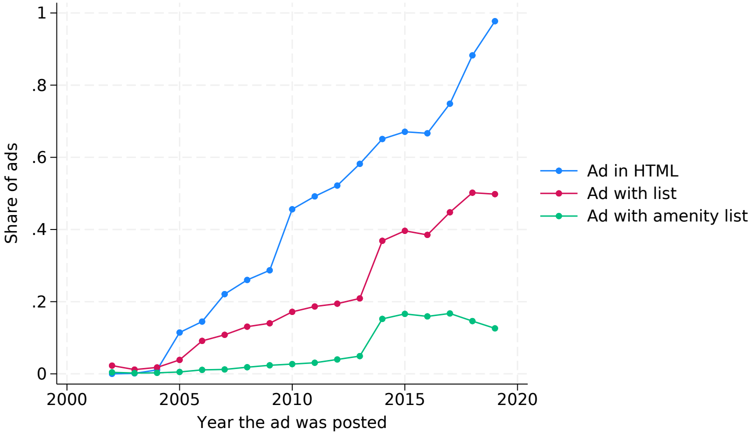
- We analyze the job attributes advertised by employers in vacancy texts
- We find that some advertised attributes correlate with specific dimension of the value of employer & predict value of employers in an economically meaningful way
- Open policy question: Is there room to make employers provide more information?

Appendix

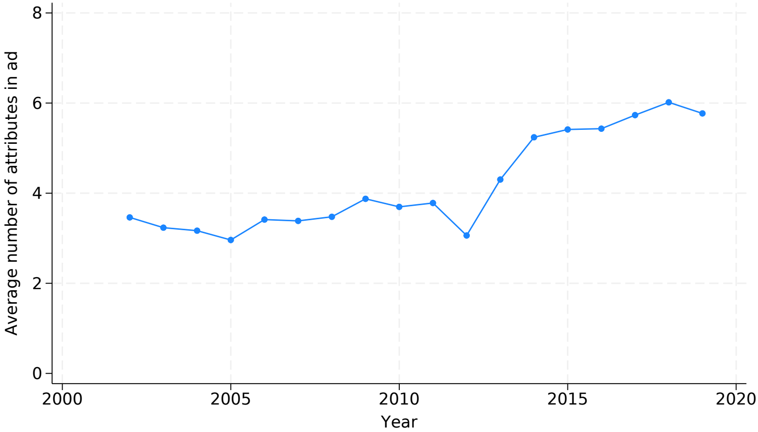
Job ad trends: word count



Job ad trends: html and lists



Job ad trends: number of amenities advertised



Validation of extracted attributes

	All:	Random Sample:		
	Text Analysis (1)	Text Analysis (2)	Manual Recognition (3)	Success Rate (4)
Any Attribute (%)	98.0	98.2	98.0	98.2
Number of Attributes	6.6	7.1	6.3	–
A. Pay Attributes (%)				
– Compensation Scheme	57.5	64.8	61.8	92.5
– Financial Attributes	32.8	37.8	34.2	96.5
– Career Opportunities	44.1	48.0	41.8	79.8
B. Non-Pay Attributes (%)				
– Hours of Work	78.4	77.2	72.2	92.0
– Convenient Hours	19.7	14.5	15.5	95.5
– Inconvenient Hours	26.1	27.8	21.8	92.0
– Contract Duration	65.8	67.2	66.8	88.0
– Workplace Attributes	51.8	55.0	51.0	82.0
– Task-Related Attributes	57.2	65.0	46.5	72.0
– Other Minor Perks	40.9	46.0	11.5	63.5
Number of Job Ads	858,745	400	400	400