The Value of Personality: Exploring the Wage Penalty Paradox in the UK

Matías Golman*

This version: January 9, 2026

Click here for the most recent version

Abstract

This study examines the relationship between personality trait and wage offers. Using text analysis of 11.7 million UK online job postings from 2017 to 2022, I identify Big Five personality traits in vacancy descriptions. Jobs mentioning personality traits offer 4.6 per cent lower wages on average, creating a paradox: why do employers advertise for traits if these are penalised? I test four explanations: correlation with low-pay job characteristics, compensating differentials, linguistic intensity effects, and demographic targeting. Trait penalties are significantly reduced by the inclusion of occupation fixed effects. Penalties are larger in high-skill occupations but smaller in female-dominated roles. Trait penalties decrease when combined with traditional benefits and workplace culture signals, but increase with remote work. Linguistic intensity analysis shows that low-intensity mentions carry minimal penalties, but high-intensity mentions show substantial penalties. Machine learning analysis of gender targeting contradicts simple demographic targeting predictions. The findings suggest employers use personality-traits for multiple purposes rather than a single mechanism: signalling requirements at low intensity, facilitating compensating differentials with appropriate amenity bundles, and serving organisational signalling functions at high intensity.

Keywords: personality traits, Big Five, job vacancy data, wage determination, compensating differentials, UK labour market, text analysis

JEL Codes: J31, J23, J24, C55

*University of Nottingham. Email: matias.golman@gmail.com.

Contents

1 Introduction

Individuals' personality traits are highly demanded in UK labour markets. Most British employers use personality tests during recruitment, showing that these characteristics matter in employment decisions (McGee and McGee, 2020). Individual-level studies consistently find that people with higher scores on traits like Conscientiousness, Openness, Extraversion, and Emotional Stability typically earn more, although high Agreeableness is associated with a wage penalty (Alderotti et al., 2023). Yet, recent research on online job advertisements reveals a puzzling contradiction: positions mentioning personality traits or character attributes are associated with *lower* wages (Acemoglu et al., 2022; Brenčič and McGee, 2023b; Deming and Kahn, 2018).

This negative association contradicts human capital theory. Skills and attributes that enhance productivity should be rewarded in the labour market. When employers explicitly request certain characteristics, economic theory suggests they value these attributes enough to pay for them. This expectation is strengthened by research on other job requirements such as education, experience, and technical skills, which show positive wage premiums when mentioned in advertisements (Alekseeva et al., 2021; Deming and Kahn, 2018; Hershbein and Kahn, 2018). Why would employers emphasise personality traits in job advertisements if these are not economically rewarded?

Analysing personality traits through job advertisements captures employers' strategies to attract candidates rather than workers' actual endowments, offering a distinct perspective from individual-level data. I identify four potential explanations from labour and personnel economics literature which can be tested using this data.

Correlation with other low-pay characteristics: Personality trait mentions might cluster systematically in lower-paying contexts. Trait mentions could correlate with job, occupation, or firm characteristics that determine wages (Brenčič and McGee, 2023b; Deming and Kahn, 2018; Woodcock, 2008). For instance, customer service or care-related occupations may mention more traits and simultaneously pay lower wages due to low market value. I first test this by including high-dimensional fixed effects. Then I examine how trait-wage relationships vary across occupations and interact traits with job and firm characteristics.

Compensating differentials through person-organisation fit: Following the theory of equalising differences, workers might trade monetary compensation for non-monetary benefits (Rosen, 1986). When jobs signal rewarding features such as flexible work, supportive culture, or meaningful impact, employers can offer lower wages whilst still attracting qualified candidates (Burbano, 2016; Escudero et al., 2024). Similarly, environments emphasising personality traits may attract individuals who value being valued for their personality (Cable and Judge, 1996; Kristof-Brown et al., 2005; Stevens and Szmerekovsky, 2010). I test this by examining whether amenity provision explains trait penalties and whether traits interact with specific amenities.

Demographic targeting and statistical discrimination: Drawing on models of statistical discrimination (Aigner and Cain, 1977; Arrow, 1971), employers might use personality trait language to target demographic groups they believe will accept lower wages. If employers associate certain traits with specific demographics, they might use trait language strategically to induce self-selection (Gaucher et al., 2011; Kuhn et al., 2020). Recent experimental evidence shows that trait language in advertisements affects application patterns differently across demographic groups (Abraham et al., 2024; Flory et al., 2015; Oldford and Fiset, 2021). I test this using machine learning to classify implicit gender targeting in advertisements.

Linguistic intensity and signalling: The intensity and placement of trait mentions may alter their interpretation and value signalling. Information structure theory suggests that message frequency, prominence, and contextual placement shape how receivers process and value information (Koch and Zerback, 2013; Lambrecht, 1996; Rosenfeld et al., 2015). I test this by measuring trait mention intensity using TF-IDF text analysis and examining whether high-intensity mentions show different wage effects than low-intensity mentions.

I begin by establishing a baseline model documenting the negative relationship between personality trait mentions and wages. Mentioning personality-related language in advertisements is associated with 4.6 per cent lower wages on average. The penalty is largest for Extraversion (3.2 per cent) and Emotional Stability (3.1 per cent), followed by Conscientiousness (2.0 per cent) and Agreeableness (1.4 per cent). Only Openness shows a small positive effect (0.7 per cent).

I then test each hypothesis using complementary methodological approaches. For correlation with low-pay characteristics, I use occupation, job, and firm fixed effects and interaction terms. For compensating differentials, I examine amenity controls and trait-amenity interactions. For linguistic intensity, I use TF-IDF text analysis. For demographic targeting, I employ BERT machine learning models to classify implicit gender preferences in advertisements.

The findings reveal that the relationship between personality trait mentions and wages reflects multiple overlapping employer strategies rather than a single mechanism. Trait language varies by intensity, context, and organisational characteristics. The persistence of penalties even with extensive controls, combined with context-dependent effects, suggests employers strategically deploy personality trait language for different purposes depending on job characteristics, linguistic emphasis, amenity bundles, and organisational positioning.

This paper proceeds as follows. Section 2 reviews the literature on personality traits' effect on wages. Section 3 describes the Adzuna job vacancy data. Section 4 presents the baseline analysis of the personality trait wage penalty. Sections 5–8 examine each of the four explanations presented in turn, providing both methodological approaches and results. Section 9 concludes.

2 Literature Review

2.1 Individual-Level Evidence: Personality Traits and Wage Premiums

Economists recognise that personality traits explain differences in economic outcomes (Borghans et al., 2008; Heckman and Kautz, 2012). However, the literature is not clear on the distinction about personality traits, soft skills, social skills or non-cognitive attributes, with authors often mixing their meanings and scope (Woessmann, 2024). I define personality traits as stable individual characteristics that remain relatively fixed after early adulthood, following psychology literature (Cobb-Clark and Schurer, 2012; McCrae and John, 1992).

Unlike soft skills or social skills, which are dynamic and learnable, personality traits are largely uncorrelated with cognitive abilities and difficult to modify through training (Grugulis and Vincent, 2009; Matteson et al., 2016). This stability gives employers strong incentives to select for required traits during hiring rather than developing them later. I use the Big Five framework (Openness, Conscientiousness, Extraversion, Agreeableness, and Emotional Stability), which has become the standard taxonomy in economics.

2.1.1 Direct Wage Effects

Individual-level studies consistently document positive wage returns to personality traits, with the exception of Agreeableness. Flinn et al. (2025) show that higher Conscientiousness and Emotional Stability, and lower Agreeableness increase earnings and job stability, with effects ranging between 1-5 per cent according to trait and gender. Mueller and Plug (2006) find that low Agreeableness increases earnings by 10 per cent for men and 5 per cent for women. Nyhus and Pons (2005) show that Emotional Stability and Extraversion correlate with higher wages, while Agreeableness reduces wages, particularly for women. Heineck (2011) reports 2-3 per cent wage premiums for Openness but similar penalties for Agreeableness in the UK.

The most comprehensive evidence comes from Alderotti et al. (2023)'s meta-analysis of 62 studies. They find consistent patterns across countries: Conscientiousness, Openness, and Extraversion show positive wage effects averaging 2-3 per cent, while Agreeableness and Neuroticism[1] show negative effects of similar magnitude. These effects persist after controlling for education, experience, and cognitive ability.

2.1.2 Mechanisms and Heterogeneity

Individual-level evidence mostly uses Bowles et al. (2001) "incentive-enhancing preferences" to explain wage effects. This theory posits that workers with certain characteristics, particularly Conscientiousness, require less monitoring and show greater reliability, making them more productive. Empirical support comes from several studies. Cubel et al. (2016) show in a laboratory experiment that conscientious individuals are more productive when completing tasks. Judge and Cable (1997) and Barrick and Mount (1991) find that Conscientiousness predicts job performance across occupations, while Extraversion particularly benefits management and sales roles.

Returns to personality traits vary across contexts. Gensowski (2018) shows that personality effects peak between ages 40-60, with larger returns for highly-educated workers. Collischon (2020) uses quantile regressions that show increased returns at higher wage levels. In contrast, Nyhus and Pons (2005) find that Conscientiousness affects wages primarily at career start, suggesting that employers reward the trait to increase the probability of recruiting and retaining the conscientious worker, but stop this once they learn about the workers' productivity.

Personality differences can also partly explain gender inequalities. Using German data, Braakmann (2009) finds that personality differences account for 18 per cent of the gender wage gaps and 7 per cent of full-time employment gaps. Nyhus and Pons (2012) show that Agreeableness and Intellect explain 11.5 per cent of the Dutch gender wage gap. Risse et al. (2018) find that men's lower Agreeableness and higher confidence contribute to Australian gender wage gap.

Occupation-specific returns matter considerably. John and Thomsen (2014) show identical traits have different returns across occupations. Cobb-Clark and Tan (2011) find that personality affects both occupation selection and wages within occupations. Denissen et al. (2018) show that personality-job fit, the congruence between individual traits and job demands, influences income.

Recent evidence suggests these patterns evolve with labour market changes. Deming (2017) and Weinberger (2014) document increasing returns to social skills and character attributes over time, particularly in team-based work. Rohrbach-Schmidt et al. (2023) find that trait-wage relationships depend on job task requirements, with stronger effects for non-routine and interactive tasks.

[1] The opposite of Emotional Stability

2.2 Job Advertisement Evidence: Personality Traits and Wage Penalties

The availability of online job posting data enables direct observation of employer demands. However, this literature finds the opposite pattern: personality trait mentions correlate with *lower* wages.

2.2.1 The Demand-Side Paradox

Deming and Kahn (2018) pioneered using job postings to measure job requirements across firms and labour markets, finding that social skills are positively correlated with wages and firm performance, particularly when combined with cognitive skills. The authors use a *character* feature inspired by broad personality attributes from Heckman and Kautz (2012). Similar to later studies by Alekseeva et al. (2021) and Acemoglu et al. (2022), they find negative wage correlations between this *character* control variable and posted wages.

Only two studies examine wage effects from Big Five traits specifically. Brenčič and McGee (2023b) analyse 140,193 job advertisements from Monster.com posted over a two-week period in 2006, using dictionaries based on trait-descriptive adjectives from Goldberg (1981) and John (1990). They find that 54 per cent of advertisements mention at least one personality trait, with employers primarily demanding Extroversion (31 per cent), Conscientiousness (26 per cent), and Openness (21 per cent).

Their analysis reveals negative wage penalties for most traits. Among advertisements posting hourly wages, they find negative coefficients for Extroversion (4.2 per cent with firm fixed effects), Agreeableness (6.3 per cent), and Emotional Stability (6.3 per cent), while only Openness shows positive effects (5 per cent). These findings establish a clear wage penalty paradox: despite employers explicitly demanding personality traits in advertisements, most traits are associated with lower posted wages.

Beyond wage effects, Brenčič and McGee (2023b) make a key theoretical contribution by examining how personality traits relate to compensation structure rather than wage levels. They develop a model where workers differ in intrinsic motivation and firms choose between fixed wages and incentive-based compensation. In their framework, conscientious workers supply effort even without extrinsic incentives, making fixed wage contracts more attractive to firms hiring such workers. Consistent with this model, advertisements demanding conscientious workers are 3.7 percentage points less likely to offer incentive pay and 2.7 percentage points less likely to reference promotion opportunities. These findings suggest firms use personality requirements as substitutes for extrinsic motivation mechanisms when selecting workers who are intrinsically motivated. While this compensation structure mechanism differs from the compensating differentials and demographic targeting hypotheses examined in the present study, it provides important context for understanding how personality trait language functions in job advertisements beyond simple wage determination.

In a subsequent study, Brenčič and McGee (2023a) investigate whether task requirements can explain the personality-wage relationship using the same Monster.com dataset. They examine employers' demands for personality traits across different job tasks and test for task-specific wage returns using the National Longitudinal Survey of Youth 1997. Their job advertisement analysis shows that trait demands correlate with general tasks (communication, teamwork, problem-solving) but are negatively correlated with routine, manual, and mathematics tasks.

Using longitudinal data with respondent and occupation fixed effects, they find limited evidence of task-specific wage returns to personality traits. Most interaction coefficients between traits and tasks are insignificant. The few significant results include a 1.2 per cent wage increase for Extroversion in teamwork-intensive occupations and a 2.3 per cent increase for Openness in communication-intensive roles. However, these effects are modest and inconsistent. Brenčič and McGee (2023a) conclude that task-specific wage returns do not provide compelling economic incentives for personality-based sorting, suggesting the wage penalties observed in job advertisements cannot be explained by differential returns to personality traits across tasks.

The present study builds on Brenčič and McGee (2023b)'s establishment of the wage penalty paradox but advances the literature in important ways. While their analysis uses a two-week sample from 2006, this research examines 11.7 million job postings spanning 2017-2022, providing contemporary relevance and statistical power to investigate underlying mechanisms. Where Brenčič and McGee (2023b) identify the paradox and test one potential explanation (incentive pay substitution for conscientiousness) and Brenčič and McGee (2023a) examine task-specific returns, this study provides the first systematic investigation of multiple competing explanations: correlation with occupation characteristics through extensive fixed effects and interaction analysis, compensating differentials through comprehensive amenity provision measures, linguistic intensity effects through TF-IDF analysis, and demographic targeting through machine learning classification of implicit gender preferences. Additionally, this study advances methodologically by implementing TF-IDF measures to capture trait mention intensity and employs BERT models to identify implicit demographic targeting patterns not examined in prior work.

3 Data

This study uses the Adzuna job vacancy dataset employed in Chapter 1. This chapter focuses on wage-related aspects and firm characteristics that influence wage determination. For details on general data characteristics and personality trait identification, see Chapter 1.

3.1 Data Source and Construction Process

The data construction process began with 276.3 million raw job advertisements from Adzuna, an aggregator whose algorithms continuously collect information on open vacancies from online job boards. The Office for National Statistics (ONS) of the United Kingdom uses Adzuna to derive its weekly vacancy indicators, reflecting the dataset's broad coverage of the UK labour market.

In the raw dataset, approximately 65 per cent of vacancies contained wage information, higher than other job vacancy sources used in the literature, which typically report wage information in only 13-25 per cent of postings (Banfi and Villena-Roldan, 2019; Brenčič, 2012; Kuhn and Shen, 2023; Marinescu and Wolthoff, 2020). This high wage coverage is a result of Adzuna's efforts in identifying wages from the different advertisement sections. It makes the dataset suitable for analysing the relationship between job requirements and offered wages.

To create a dataset appropriate for wage analysis, I applied a filtering and cleaning process shown in Table 1. First, I removed advertisements posted by recruitment agencies (45 per cent of vacancies), as these intermediaries make it impossible to identify the actual employing firm or industry sector. Without industry information, advertisements could not be accurately classified into occupational categories.

Table 1: Cleaning of the database

Process	**N vacancies**	**N no missing W***	**Mean W***	**SD W***	**Δ**
Adzuna raw database	276,325,386	178,229,590	31,120	18,524	-
Without recruitment agencies	150,918,209	83,727,680	29,882	18,558	-45%
Without missing companies match	66,833,773	32,647,950	28,365	18,392	-56%
Keep only Private companies with SIC	64,642,205	31,619,913	28,303	18,468	-3%
Remove duplicates vacancy-id, source-id, same description-location	20,783,111	10,490,006	28,342	18,830	-68%
Remove salaries missing	9,657,106	9,657,106	29,454	18,964	-54%
Incorporate new salaries from cleaned description	12,103,405	11,866,183	31,930	18,873	25%
Remove outliers in wages and text size	11,659,923	11,430,457	32,018	18,944	-4%

Notes: Author's creation based on Adzuna. $*$W = Salary considered after trimming 1% tails.

Next, I matched company names from the remaining advertisements with the Companies House database, a registry of UK businesses. This matching process required text normalisation to account for variations in company name formats and allowed me to identify Standard Industrial Classification (SIC) codes for each vacancy. After restricting to private companies with valid SIC codes and removing duplicates through multiple deduplication strategies, 20.8 million unique vacancies remained.

From this set, 9.7 million vacancies had structured wage fields in Adzuna's data. To expand wage information coverage, I implemented a text analysis procedure to extract additional wage information from unstructured job descriptions. This process involved:

1. Standardising salary formats through regular expressions, addressing variations such as annual figures (e.g., "£35,000"), ranges (e.g., "£30,000-£40,000"), abbreviations (e.g., "£35k"), and hourly/weekly/monthly rates.
2. For range values, calculating the mean of the lower and upper bounds.
3. Converting non-annual rates to annual equivalents based on the job's full-time/part-time status.
4. Applying a hierarchical decision rule that prioritised annual figures from the text, followed by structured wage fields after verification.

This wage extraction process recovered an additional 2.4 million vacancies with wage information not considered in Adzuna's wage fields. After removing outliers in both wages and text length, the final dataset contained 11.7 million vacancies, all with valid wage information, advertised by 128,559 companies.

The final average salary in the final dataset is approximately 15 per cent lower than in the UK, according to external sources. This difference partly reflects the potential selection bias noted by Batra et al. (2023) that posted wages are more common in advertisements seeking less-skilled workers, as well as the overrepresentation of certain lower-paying occupations like Caring Personal Services and Customer Service roles in online job postings.

3.2 Occupation and Firm Characteristics

Since Adzuna does not provide standard occupational classifications, I implemented a machine learning algorithm developed by Turrell et al. (2019) to classify advertisements into 3-digit Standard Occupation Codes (SOC). This algorithm assigns SOC codes based on job title, description, and industry information derived from the Companies House matching, with a reported accuracy rate of 76 per cent.

To investigate mechanisms through which personality trait demands might affect wages, I complemented the vacancy data with information at both occupation and firm levels.

At the occupation level, I incorporated data from the UK Household Longitudinal Study (UKHLS) on gender composition, educational attainment, migrant proportion, and average worker age for each SOC code. The UKHLS also includes self-reported Big Five personality assessments based on a 7-point scale, enabling comparison of employer demands for traits with actual trait distributions across occupations. These occupation-level characteristics are essential for testing whether the trait-wage relationship reflects sorting or demographic targeting mechanisms.

At the firm level, I matched companies with the Financial Analysis Made Easy (FAME) database, which provided information on firm size (employee count), annual profit, public listing status, and directors' demographic characteristics. This firm-level information allows testing whether organisational characteristics influence the relationship between personality traits and wages.

Together, these data linkages enable analysis of how personality trait demands relate to wages across different types of jobs, occupations, and firms, providing the foundation for testing the four mechanisms regarding why employers use personality-related language despite potential wage penalties.

3.3 Retrieving Personality Traits and Job Characteristics

This paper builds on the personality trait identification methodology developed in Chapter 1. I use the established Big Five personality framework widely used in economic and psychological research (Almlund et al., 2011; Borghans et al., 2008; Heckman and Kautz, 2012). This taxonomy categorises personality into five dimensions: Conscientiousness, Openness to experience, Extraversion, Agreeableness, and Emotional Stability.

Using a knowledge engineering approach, I identify personality traits through dictionaries compiled from authoritative sources including Goldberg (1990), Costa Jr and McCrae (1992), Goldberg (1992), McCrae and John (1992), John (1990), Saucier (1994) and Heckman and Kautz (2012). This approach allows for direct application of established taxonomic work while avoiding ambiguous or context-dependent terms that might lead to false positives. This process identified words with multiple meanings, such as *flexible* or *trust*, which could refer to work schedules rather than personality traits.

To minimise this risk, words were only included after being validated through context analysis: 8 samples of 100 randomly selected sentences containing each potential trait word were examined to ensure the term genuinely referred to a personality characteristic in the context of job requirements. I modified terms in contexts not referring to worker or team characteristics before dictionary matching. For instance, replace "flexible hours" with "flx hours". I removed words that, regardless of context, did not refer to personality traits in job vacancies, such as *modest*, *conventional*, *deliberate*, *order*, *aesthetic*, *deep*, *relaxed* or *comfortable*.

A job posting is labelled as requiring a specific personality trait if it includes at least one keyword from the corresponding trait dictionary in its job requirements or job title.[2] This boolean approach generates binary indicators for each of the five personality traits, which are then related to wage offers to test the hypothesis that trait mentions correlate with lower wages even after controlling for various job, occupation, and firm characteristics.

In addition to personality traits and wages, I created several other variables through text analysis of job advertisements. For location identification, I mapped Adzuna's location variables to the Nomenclature of Territorial Units for Statistics (NUTS), achieving 96 per cent coverage at the one-digit level and 85 per cent at the two-digit level. Education requirements were classified into five categories following Nandi and Nicoletti (2014), though notably only 12 per cent of vacancies explicitly mentioned education requirements.

I created indicators for hard and soft skills based on expanded dictionaries from Deming and Kahn (2018); Khaouja et al. (2021); Lyu and Liu (2021). I built 16 categories of amenities, inspired by the methodology and dictionaries from Escudero et al. (2024) and Boschetti Adamczyk et al. (2025). Remote work indicators were based on Draca et al. (2022)'s dictionary, capturing the increase in remote work mentions following the COVID-19 pandemic. Experience requirements were coded as binary indicators, while seniority levels (entry, mid, senior) were identified following hierarchical descriptions from previous literature. Job arrangement variables (part-time/full-time, temporary/permanent) combined Adzuna's structured fields with text-derived information when the former were missing. All dictionaries were subjected to validation through manual checks of word context to minimise false positives, with over 1,700 specific word replacements implemented to ensure accuracy.

For details on the dictionaries and validation process, see Chapter 1.

4 Baseline Analysis: The Personality Traits Wage Penalty

This section establishes the baseline relationship between personality trait mentions in job advertisements and offered wages. I first present the empirical methodology for identifying this relationship, then document the evidence of the wage penalty across different personality traits.

[2] Marinescu and Wolthoff (2020) show that job titles usually contain relevant hierarchical, experience and specialisation information which can be omitted in the body of the advertisement. Both were merged for the text analysis retrieval.

4.1 Empirical Methodology

To identify the relationship between personality traits and wages, I employ a high-dimensional fixed-effects (HDFE) model that controls for multiple sources of unobserved heterogeneity across job characteristics, occupations, and firms. This approach isolates the relationship between personality traits and wages while addressing a large set of observable and fixed unobservable factors. The baseline model is:

$$\ln(wage_i) = \beta_0 + \sum_{t=1}^{5} \beta_t Trait_{ti} + \beta_6 X_i + \delta_l + \theta_m + \epsilon_i \quad (1)$$

where $\ln(wage_i)$ is the log of the posted wage for vacancy i, $Trait_{ti}$ represents binary indicators for each of the five personality traits (Conscientiousness, Openness, Extraversion, Agreeableness, and Emotional Stability), and X_i is a vector of job posting characteristics including education requirements, experience, seniority, contract type, and other job attributes. The terms δ_l and θ_m capture location at two digits of NUTS and month-year fixed effects, respectively.

I progressively augment this specification by adding industry fixed effects (α_s), occupation fixed effects (γ_k), and finally firm fixed effects (ϕ_j):

$$\ln(wage_i) = \beta_0 + \sum_{t=1}^{5} \beta_t Trait_{ti} + \beta_6 X_i + \delta_l + \theta_m + \alpha_s + \gamma_k + \phi_j + \epsilon_i \quad (2)$$

This stepwise approach allows me to assess how much of the observed wage penalty for personality traits can be explained by within-industry, within-occupation, and within-firm variation. If the trait coefficients (β_t) remain overall negative after including these fixed effects, this suggests the wage penalty persists even within narrowly defined industries, occupations, and firms, pointing to a relationship that requires further investigation.

To further explore how the trait-wage relationship varies across different occupations, I also estimate the baseline specification separately for each occupation at 1-digit, 2-digit, and 3-digit SOC levels, excluding the occupation fixed effects from these models.

4.2 Results

4.2.1 Evidence of the Wage Penalty

The median nominal salary in online advertisements posted on Adzuna for the period 2017-2022 was £25,960, which is 15 per cent lower than the median income reported by the Annual Survey of Hours and Earnings (ASHE) for the UK during the same period. There are three potential explanations. First, Adzuna reports starting offered wages rather than realised wages. Second, these posted salaries do not include bonuses, performance-related rewards, overtime or other variable income. Third, the Adzuna dataset has more low-paid occupations compared to the wider labour market. For example, Caring Personal Services and Customer Services jobs, which typically offer below-median wages, make up about 18 per cent of vacancies in the Adzuna database but only 10 per cent of occupations in ASHE.

Table 2 shows the relationship between personality trait mentions and posted wages. Panel A presents the effect of mentioning any personality trait, while Panel B shows the coefficients for each of the Big Five traits separately. The first column in both panels gives the raw relationship without controls, showing a wage penalty of 17.6 per cent for job advertisements that mention any personality trait. Looking at individual traits in Panel B, all traits except Openness have large wage penalties in the raw data. Agreeableness has the largest penalty (21.4 per cent), followed by Extraversion (18.5 per cent), Emotional Stability (16.8 per cent), and Conscientiousness (6.7 per cent). Openness-to-experience has a small positive association with wages (1.6 per cent). On average, jobs mentioning personality traits offer lower wages than those that do not.

Table 2: Log posted wage and Personality Traits

	(1) Baseline	(2) +Controls	(3) +Month	(4) +Location	(5) +Industry	(6) +Occupation	(7) +Firm
Panel A: Effect of Any Personality Trait							
Any Personality Trait	-0.176***	-0.157***	-0.157***	-0.151***	-0.126***	-0.069***	-0.045***
	(0.000)	(0.000)	(0.000)	(0.005)	(0.009)	(0.010)	(0.007)
R^2	0.024	0.187	0.191	0.212	0.267	0.450	0.600
Panel B: Effects of Individual Personality Traits							
Conscientiousness	-0.067***	-0.034***	-0.034***	-0.030***	-0.029***	-0.022**	-0.020***
	(0.000)	(0.000)	(0.000)	(0.004)	(0.009)	(0.009)	(0.007)
Openness	0.016***	0.026***	0.027***	0.017***	0.021**	0.005	0.007*
	(0.000)	(0.000)	(0.000)	(0.003)	(0.008)	(0.007)	(0.004)
Extraversion	-0.185***	-0.140***	-0.140***	-0.135***	-0.117***	-0.060***	-0.032***
	(0.000)	(0.000)	(0.000)	(0.004)	(0.007)	(0.009)	(0.007)
Agreeableness	-0.214***	-0.106***	-0.109***	-0.104***	-0.083***	-0.022***	-0.014*
	(0.000)	(0.000)	(0.000)	(0.005)	(0.015)	(0.008)	(0.008)
Emotional Stability	-0.168***	-0.091***	-0.092***	-0.086***	-0.079***	-0.043***	-0.031***
	(0.000)	(0.000)	(0.000)	(0.004)	(0.013)	(0.010)	(0.008)
Observations	11662124	11662124	11662124	9,923,701	9,923,697	9,923,697	9,888,315
R^2	0.040	0.201	0.206	0.226	0.277	0.452	0.601
Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	No	No	Yes	Yes	Yes	Yes	Yes
Region FE	No	No	No	Yes	Yes	Yes	Yes
Industry FE	No	No	No	No	Yes	Yes	No
Occupation FE	No	No	No	No	No	Yes	Yes
Firm FE	No	No	No	No	No	No	Yes
N firms	128,559	128,559	128,559	128,559	128,559	128,559	81,843

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Each column represents a different specification with progressively added controls and fixed effects. All regressions with controls include hard and soft skills, 16 sets of amenities, advertisement length, contract type, work schedule, seniority level, education, experience, bonus availability, and negotiable salary indicators. Standard errors are clustered at the region level in models (4)-(5), and at the occupation and firm level in models (6)-(7).

4.2.2 Controlling for Job Characteristics and Fixed Effects

To examine whether these penalties might be due to other factors, the subsequent columns add controls for job characteristics (column 2), location (column 3), industry (column 4), occupation (column 5), and firm fixed effects (column 6). The covariates include education, experience, seniority, contract types, skill requirements, and 16 categories of amenities that capture workplace benefits and non-wage compensation[3].

Even with the most extensive set of controls (column 7), the penalty for mentioning any personality trait remains statistically significant at 4.5 per cent. When examining individual traits, Conscientiousness (2.0 per cent), Extraversion (3.2 per cent), Agreeableness (1.4 per cent), and Emotional Stability (3.1 per cent) all have statistically significant negative coefficients. Openness-to-experience shows a small positive coefficient (0.7 per cent), though it is only significant at a 10 per cent confidence level.

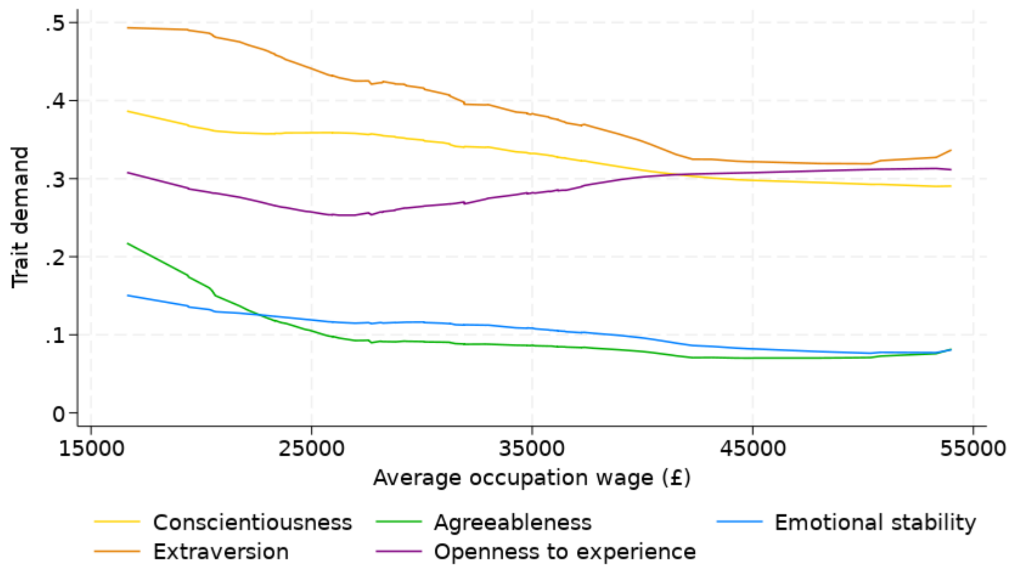
The wage penalties decrease when fixed effects are added, especially occupation fixed effects (column 6), which reduce the penalty for any trait from 12.6 to 6.9 per cent. This indicates that part of the observed penalty comes from selection into occupations—personality traits appear more often in advertisements for lower-paying occupations. However, because significant penalties remain even after controlling for occupation and firm characteristics, selection alone cannot fully explain this relationship. The inclusion of job requirements and job amenities controls also reduce the size of the wage penalty in the baseline specification (column 1), but these differences are ultimately captured by the subsequent fixed effects.

To illustrate how trait demand varies across occupation-level average wages, Figure 1 plots, at the occupation level, the proportion of advertisements which mention a trait against the average wage in that occupation. In total, there are 90 occupations at the 3-digit level of the SOC. Extraversion and Conscientiousness show strong negative trends across the wage distribution, with mentions decreasing as average occupation wages increase. For example, Extraversion appears in nearly 50 per cent of advertisements for occupations at the lowest end of the wage distribution but only in about 32 per cent for the highest-paying occupations. Agreeableness and Emotional Stability follow similar but less pronounced trends, with both traits mentioned more often in lower-wage occupations. Openness to experience shows a different trend, with lower mention rates in the middle of the wage distribution and slight increases at both extremes. This pattern aligns with findings from Chapter 1, where Openness was most prevalent in professional and technical occupations, but also in sales and customer service.

These differences across the wage distribution help explain why occupation fixed

[3] Amenities specifications are listed in Appendix B and the rest are thoroughly described in Chapter 1.

Figure 1: Personality trait demand across the wage distribution of occupations



Notes: 90 Occupations considered at the 3-digits level of the SOC.

effects reduce the wage penalties seen in Table 2. Personality traits appear more often in lower-paying occupations, creating a negative correlation between trait mentions and wages in the overall data. However, significant negative coefficients remain even after controlling for occupation, so this selection mechanism does not fully explain the wage penalties.

4.2.3 Occupation-Specific Trait Effects

To assess occupational differences, I estimated the baseline specification for each occupation at 1, 2 and 3 digits, excluding the occupation fixed effects from the model. Figure 2 presents the coefficient estimates and 95 per cent confidence intervals for each personality trait across the nine major occupation groups. The results show substantial heterogeneity in how personality traits affect wages across occupation categories.

Extraversion has negative wage effects across almost all occupation types, with especially large penalties in Administrative and Secretarial roles (5.7 per cent), Associate Professional and Technical occupations (4.9 per cent), and Sales and Customer Service roles (3.6 per cent). Conscientiousness also has negative coefficients across most occupation groups, with the largest penalties in Managers, Directors and Senior Officials (3 per cent) and Associate Professional and Technical occupations (3 per cent).

Openness to experience stands out as the only trait with predominantly positive effects. It has positive and significant coefficients in several occupation groups, including Skilled Trades (2.1 per cent), Sales and Customer Service (2.2 per cent), and Caring, Leisure and Other Service occupations (1.8 per cent). This confirms that the small positive effect of Openness in the overall analysis represents positive valuation of this trait in specific occupations.

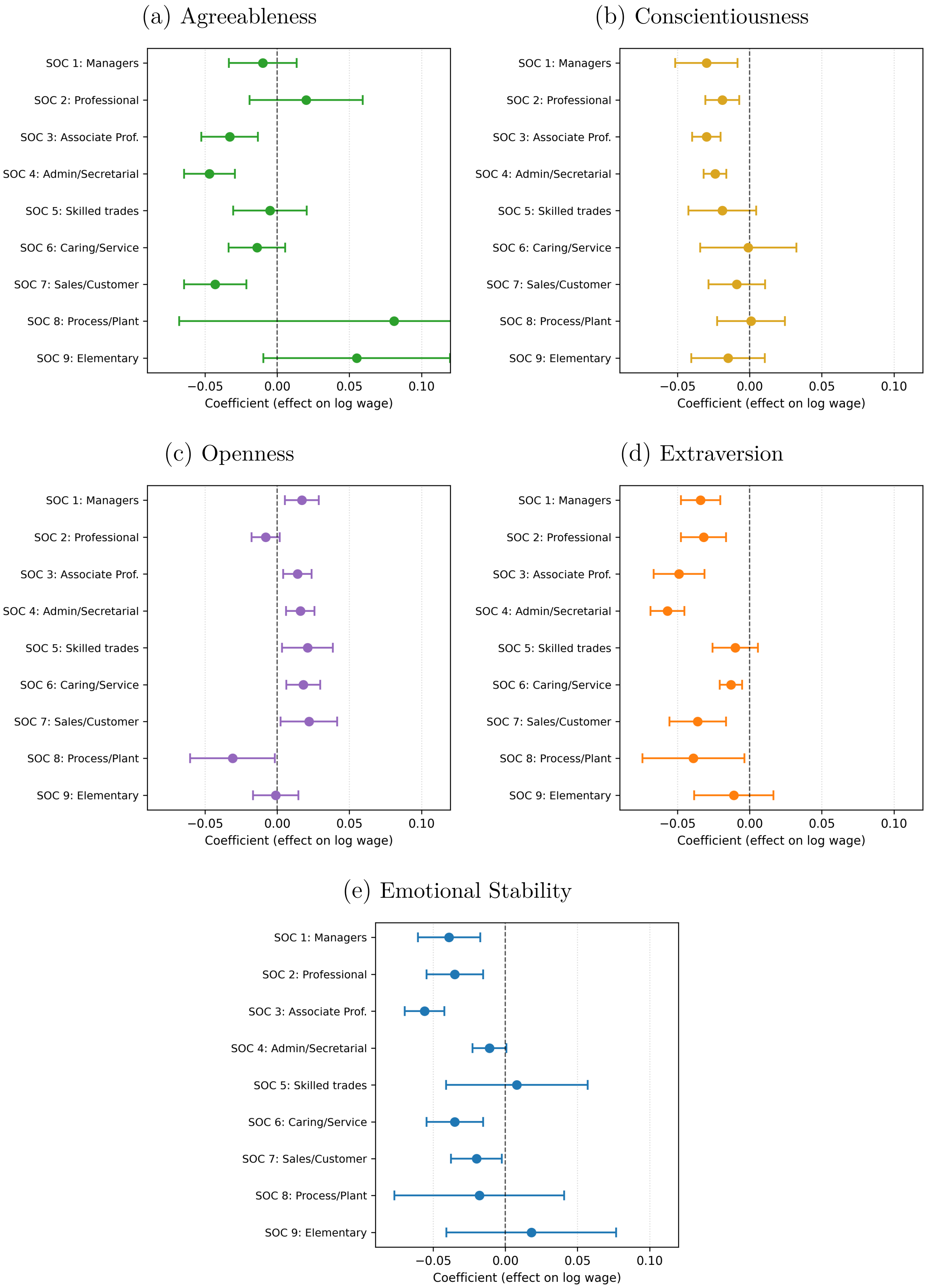
Agreeableness and Emotional Stability show markedly different effects across occupations. Agreeableness has strong negative effects in Administrative and Secretarial roles (4.7 per cent) and Sales and Customer Service positions (4.3 per cent), but shows a large positive effect in Elementary occupations (5.5 per cent), although this latter effect has a wide confidence interval. Emotional Stability shows the most pronounced variation, with substantial penalties in Associate Professional and Technical occupations (5.6 per cent) and Managers (3.9 per cent), but no significant penalty in Skilled Trades or Elementary occupations.

These patterns suggest that the wage penalty for personality traits varies with the specific requirements and characteristics of different types of work. The positive effects of Agreeableness in Elementary occupations and the absence of penalties for most traits in Skilled Trades may reflect productivity benefits in these contexts. Conversely, the consistent penalties for Extraversion across high-skill occupations suggest that this trait may serve different signalling functions in professional and managerial roles than in service-oriented positions.

More detailed analyses at the 2-digit and 3-digit SOC level in Table A.2 and Table A.3 in Appendix A show even greater differences, indicating complex relationships between personality traits and wages that cannot be explained by simple occupational sorting. The granular occupation-specific patterns present heterogeneity that strengthens the argument against uniform explanations for trait penalties. At the 2-digit level, Extraversion penalties range from non-significant effects in Teaching and Educational Professionals to substantial penalties of 6.1 per cent in Science, Research, Engineering and Technology Professionals. Similarly, Conscientiousness shows positive effects in Caring Personal Service Occupations (1.7 per cent) but large negative effects in Business, Media and Public Service Professionals (3.5 per cent).

This fine-grained heterogeneity suggests that personality trait language serves fundamentally different functions across occupational contexts, varying not just by broad occupational categories but by specific job roles within them. The variation in trait effects across the occupational hierarchy also provides insight into potential mechanisms beyond those directly tested—the differential effects of traits in entry-level elementary occupations versus senior managerial positions may reflect different stages of career development where personality characteristics are valued differently, relating to theories of employer learning about worker productivity over time.

Figure 2: Personality Trait Coefficients on Log Posted Wage, by 1-digit SOC Occupation



Notes: Each panel shows coefficients from separate regressions for each 1-digit SOC occupation. Points represent coefficient estimates; horizontal bars show 95 per cent confidence intervals. Effects crossing zero are not statistically significant at the 5 per cent level. All regressions include controls for soft skills, hard skills, remote work, advertisement length, contract type, work schedule, seniority level, education, experience, amenities availability, and negotiable salary. Fixed effects for location, month-year, and firm are included. Standard errors are clustered at the firm level. Estimations are reported in Table A.1 in Appendix A.

5 Correlation with occupation-level characteristics

In the previous subsection, I showed that wage penalties varied between occupations. In this section, I consider whether the variation in wage penalties can be explained by specific occupational characteristics, such as workers' average demography, education level, or income. In addition, I estimate whether wage penalties are moderated or increased when vacancies demand different skills, seniority levels, or types of contracts. Finally, I consider whether there are different effects by type of firms. The baseline model is extended by incorporating interaction terms between each personality trait and occupation, job, or firm characteristics.

For occupation-level characteristics, I estimate separate regressions for each trait t while controlling for the other four traits:

$$\begin{aligned} \ln(wage_i) = \beta_0 + \beta_t Trait_{ti} + \sum_{j \neq t}^{5} \beta_j Trait_{ji} + \gamma_c Occ\_Char_{kc} \\ + \delta_{tc}(Trait_{ti} \times Occ\_Char_{kc}) + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i \end{aligned} \tag{3}$$

where $Trait_{ji}$ represents the other four traits included as controls, Occ_Char_{kc} represents standardised values of occupation k's characteristics including women's share, educational attainment, migrants' share, mean worker age, mean hourly wage and share of managerial positions. By standardising the occupation characteristics, I interpret the interaction coefficients δ_{tc} as the change in the trait effect associated with a one standard deviation increase in the characteristic.

The results are presented in three components: (1) the effect when the characteristic is at its mean, (2) the change in effect for a one standard deviation increase, and (3) the total effect when the characteristic is one standard deviation above the mean. These regressions control for occupation-level variables while including occupation fixed effects, so the effect accounts for the trait-occupation-characteristic variation after controlling for occupation-specific features.

For job-requirement interactions, I estimate separate regressions for each trait t and job requirement r:

$$\begin{aligned} \ln(wage_i) = \beta_0 + \beta_t Trait_{ti} + \sum_{j \neq t}^{5} \beta_j Trait_{ji} + \gamma_r Job\_Req_{ir} \\ + \lambda_{tr}(Trait_{ti} \times Job\_Req_{ir}) + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i \end{aligned} \tag{4}$$

where Job_Req_{ir} includes binary indicators for job requirements such as soft skills, hard skills, amenities, negotiable salary, seniority levels, work schedule, contract period, experience, and education. The coefficients λ_{tr} show how the wage effect of trait t differs when job requirement r is present. For these analyses, I present the baseline and the interaction effect for each trait-requirement combination. These regressions include occupation fixed effects (γ_k) and firm fixed effects (ϕ_j).

For firm-level characteristics, I estimate separate regressions for each trait t and firm characteristic f:

$$\begin{gathered}\ln(wage_i) = \beta_0 + \beta_t Trait_{ti} + \sum_{j \neq t}^{5} \beta_j Trait_{ji} + \gamma_f Firm\_Char_{jf} \\ + \mu_{tf}(Trait_{ti} \times Firm\_Char_{jf}) + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i\end{gathered} \tag{5}$$

where $Firm\_Char_{jf}$ includes standardised continuous characteristics of firm j such as size (employee count), age, and profit, as well as binary indicators for public listing status. Similar to the occupation characteristics, the standardisation of continuous firm variables allows for direct interpretation of how a one standard deviation increase affects trait wage premiums or penalties. These regressions include occupation fixed effects (γ_k) and firm fixed effects (ϕ_j), allowing for an interpretation of the trait-firm-characteristic variation after controlling for firm-specific features.

If the wage penalty for personality traits is primarily driven by correlation with these observable characteristics, we would expect the interaction terms to be significant and the main effects of traits to become smaller or insignificant once these interactions are included. Alternatively, finding persistent main effects alongside significant interaction terms would suggest that while context matters, there is still an underlying relationship between personality trait mentions and wages that cannot be fully explained by these observable factors.

5.1 Results

5.1.1 Occupation-level heterogeneity

To test whether personality trait penalties reflect correlation with other occupation characteristics, I interact each trait with standardised occupation-level variables from the UK Household Longitudinal Study (UKHLS). Table 3 presents these interaction effects, showing how trait penalties vary with occupation characteristics.

The demographic composition of occupations influences trait-wage relationships. In female-dominated occupations, personality trait penalties are generally smaller or even become premiums. For instance, as women's representation increases by one standard deviation, the Conscientiousness penalty (2.1 per cent) becomes statistically insignificant (0.8 per cent), while Openness shifts from a marginally significant effect to a significant positive premium (1.6 per cent). Similarly, the Extraversion penalty is reduced from

3.3 to 2.3 per cent in occupations with higher female representation. The same pattern appears in occupations with higher migrant shares, where penalties for Conscientiousness, Extraversion, and Emotional Stability all diminish. Conscientiousness shows the largest shift, with its penalty (2.0 per cent) becoming statistically insignificant (0.8 per cent) as migrant representation increases by one standard deviation.

In contrast, increases in the human capital associated with an occupation generally amplifies trait penalties. As the proportion of degree holders increases, the penalty for Extraversion worsens (3.1 to 4.9 per cent), and Openness loses its positive effect, shifting from 0.8 per cent to becoming statistically insignificant (0.2 per cent). Only Agreeableness shows a different pattern, with its penalty remaining relatively stable around 1.6 to 2.3 per cent across education levels, suggesting this trait may operate through different mechanisms. Similarly, in higher-paying occupations, most traits face stronger penalties. This is pronounced for Extraversion, where the penalty increases from 3.2 to 4.6 per cent in occupations with wages one standard deviation above the mean. Emotional Stability also shows amplified penalties, from 3.3 to 4.5 per cent in higher-paying occupations.

Occupational age structure has limited effects on most traits, with only Emotional Stability showing a larger penalty (3.1 to 4.1 per cent) in occupations with older workers. However, managerial content has larger effects on traits. In management-intensive occupations, the Extraversion penalty increases (3.2 to 4.9 per cent), as does the Emotional Stability penalty (3.2 to 4.2 per cent). These results contradict traditional expectations that leadership roles would value traits like Extraversion, suggesting that high-level positions may use personality-traits for different purposes beyond identifying productive characteristics. Similarly, occupations with higher hourly wages tend to penalise all traits but Agreeableness more.

These findings provide partial support for the correlation with low-pay characteristics hypothesis. The smaller penalties in female-dominated and migrant-heavy occupations suggest that demographic composition partly explains trait penalties. However, the amplified penalties in high-skill, high-wage occupations indicate that correlation with observable low-pay characteristics cannot fully account for the wage penalty paradox. The persistence and even intensification of penalties in contexts where these traits should be most valued suggests additional mechanisms beyond simple correlation with occupation characteristics operate in determining the relationship between personality trait mentions and wages.

Table 3: Effects of personality traits with standardised occupation characteristics

Occ. Characteristic		Consc.	Open.	Extra.	Agree.	Emot.
Age (years)	At mean (0 SD)	-0.020***	0.007*	-0.032***	-0.014*	-0.031***
		(0.007)	(0.004)	(0.007)	(0.008)	(0.008)
	Per 1 SD increase	-0.005	0.002	-0.003	0.007	-0.010*
		(0.005)	(0.003)	(0.005)	(0.006)	(0.005)
	Total effect at +1 SD	-0.025**	0.009*	-0.035***	-0.006	-0.041***
		(0.010)	(0.006)	(0.009)	(0.012)	(0.010)
Degree holder share	At mean (0 SD)	-0.020***	0.008**	-0.031***	-0.016**	-0.033***
		(0.006)	(0.004)	(0.006)	(0.008)	(0.008)
	Per 1 SD increase	-0.008	-0.010**	-0.018***	-0.007	-0.020***
		(0.006)	(0.004)	(0.006)	(0.009)	(0.007)
	Total effect at +1 SD	-0.029***	-0.002	-0.049***	-0.023**	-0.053***
		(0.005)	(0.005)	(0.008)	(0.011)	(0.010)
Female share	At mean (0 SD)	-0.021***	0.007*	-0.033***	-0.010	-0.028***
		(0.006)	(0.004)	(0.006)	(0.011)	(0.009)
	Per 1 SD increase	0.013**	0.009**	0.010	-0.009	-0.010
		(0.005)	(0.004)	(0.006)	(0.011)	(0.008)
	Total effect at +1 SD	-0.008	0.016***	-0.023***	-0.019**	-0.038***
		(0.008)	(0.005)	(0.009)	(0.007)	(0.007)
Hourly income	At mean (0 SD)	-0.020***	0.008**	-0.032***	-0.013	-0.033***
		(0.006)	(0.004)	(0.006)	(0.008)	(0.008)
	Per 1 SD increase	-0.009*	-0.008*	-0.015***	0.006	-0.012*
		(0.005)	(0.005)	(0.006)	(0.006)	(0.007)
	Total effect at +1 SD	-0.030***	-0.001	-0.046***	-0.007	-0.045***
		(0.005)	(0.006)	(0.008)	(0.011)	(0.012)
Managerial share	At mean (0 SD)	-0.020***	0.007*	-0.032***	-0.014*	-0.032***
		(0.006)	(0.004)	(0.006)	(0.008)	(0.008)
	Per 1 SD increase	-0.009*	-0.001	-0.018***	-0.002	-0.010**
		(0.005)	(0.005)	(0.005)	(0.008)	(0.005)
	Total effect at +1 SD	-0.029***	0.006	-0.049***	-0.016	-0.042***
		(0.005)	(0.007)	(0.007)	(0.012)	(0.009)
Migrant share	At mean (0 SD)	-0.020***	0.007*	-0.032***	-0.015*	-0.032***
		(0.006)	(0.004)	(0.006)	(0.008)	(0.008)
	Per 1 SD increase	0.012**	0.002	0.012**	0.006	0.004
		(0.006)	(0.004)	(0.005)	(0.005)	(0.007)
	Total effect at +1 SD	-0.008	0.009	-0.021**	-0.009	-0.028***
		(0.008)	(0.007)	(0.010)	(0.009)	(0.010)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. "At mean" shows the effect when the standardised occupation characteristic equals zero (at its mean). "Per 1 SD" shows the change in effect for a one standard deviation increase in the occupation characteristic. "Total effect" shows the combined effect when the Regressions control for job requirements, location, time-month, occupation and firm fixed effects. 9888315 observations.

5.1.2 Job-requirement heterogeneity

Job-specific attributes also modify how personality traits affect wages. Table 4 presents interaction coefficients showing how trait-wage relationships change when specific job requirements are mentioned, with the baseline representing jobs that do not mention these requirements.

When soft skills are mentioned alongside personality traits, the interaction effects are generally positive, indicating that soft skills reduce trait penalties. The interaction is statistically significant for Extraversion (1.2 per cent) and Agreeableness (1.8 per cent), suggesting these traits are viewed more favourably when soft interpersonal skills are explicitly valued. This aligns with research showing personality traits and soft skills are complementary (Deming, 2017; Heckman and Kautz, 2012).

Conversely, hard skills requirements create negative interactions with personality traits. The interaction coefficients are negative for Openness (1.7 per cent) and Extraversion (3.3 per cent), indicating that technical skill requirements amplify penalties for these traits. This suggests employers may view personality characteristics as less valuable when technical competencies are the primary focus.

Low seniority roles have a significant positive interaction with Conscientiousness (2.1 per cent), suggesting this trait may entail productivity-enhancing demand in junior positions, as Bowles et al. (2001) theory predicts. Similarly, when no experience is needed, Conscientiousness shows a large positive interaction (8.5 per cent), higher than when experience is explicitly required (0.9 per cent). Interestingly, the negative effect of Agreeableness seems to be driven by its interaction with Experience needed (2.7 per cent).

Work schedule requirements create notable interaction effects. Full-time positions show positive interactions with Openness (1.8 per cent) and Extraversion (1.3 per cent). However, part-time positions have even stronger positive interactions with Openness (4.7 per cent) and Agreeableness (5.0 per cent), indicating these may be more valued in non-standard employment arrangements. Results for contract types produce negative effects for temporary positions, with a significant interaction for Emotional Stability (4.2 per cent). Only Openness shows a positive interaction with permanent positions (1.4 per cent).

Negotiable salary arrangements create a positive interaction with Openness (1.4 per cent) but a negative interaction with Agreeableness (3.1 per cent). The latter aligns with individual-level data findings, which argue that individuals with high Agreeableness scores may have lower risk preferences and bargaining power (Flinn et al., 2025).

Education requirements generally show positive interactions, particularly for A-Level requirements, which have significant positive interactions with Conscientiousness (2.0

per cent) and Emotional Stability (4.1 per cent). Bachelor's degree requirements show a positive interaction with Agreeableness (3.1 per cent) and GCSE with Extraversion (2.8 per cent).

These interaction patterns provide mixed evidence for the correlation with low-pay characteristics hypothesis. The positive interactions between personality traits and entry-level positions (low seniority, no experience required) suggest that trait penalties may partly reflect concentration in junior roles that typically offer lower wages. Similarly, the positive interactions with educational requirements indicate that when jobs explicitly require qualifications, personality traits face smaller penalties or even premiums. However, the negative interactions with hard skills and temporary contracts suggest that certain job characteristics amplify rather than explain trait penalties. The heterogeneity across different job contexts indicates that while some trait penalties may reflect correlation with low-paying job characteristics, this mechanism cannot fully account for the wage penalty across all job types.

Table 4: Interaction term coefficients and SE by Job characteristics

		Cons.	Open.	Extra.	Agree.	Emot.
Soft skills	Baseline	-0.027***	0.006	-0.040***	-0.027***	-0.033**
		(0.007)	(0.007)	(0.008)	(0.006)	(0.013)
	Interaction	0.010	0.002	0.012*	0.018**	0.002
		(0.007)	(0.007)	(0.006)	(0.009)	(0.015)
Hard skills	Baseline	-0.016	0.017***	-0.013	-0.004	-0.029**
		(0.013)	(0.005)	(0.010)	(0.013)	(0.013)
	Interaction	-0.006	-0.017***	-0.033***	-0.021	-0.005
		(0.012)	(0.006)	(0.009)	(0.014)	(0.013)
Negotiable	Baseline	-0.026**	0.000	-0.026***	-0.001	-0.025**
		(0.010)	(0.005)	(0.009)	(0.012)	(0.010)
	Interaction	0.012	0.014*	-0.013	-0.031**	-0.013
		(0.009)	(0.008)	(0.008)	(0.013)	(0.013)
Seniority	Baseline	-0.032***	0.004	-0.022**	-0.007	-0.048**
		(0.011)	(0.011)	(0.011)	(0.018)	(0.024)
	I. High	0.066	-0.009	-0.018	0.167	-0.001
		(0.042)	(0.035)	(0.035)	(0.105)	(0.030)
	I. Middle	-0.002	-0.013	-0.019	-0.023	0.010
		(0.013)	(0.018)	(0.012)	(0.018)	(0.024)
	I. Low	0.021*	0.009	-0.011	-0.007	0.025
		(0.010)	(0.010)	(0.012)	(0.020)	(0.025)

Continued on next page

Table 4 – Continued from previous page

		Cons.	Open.	Extra.	Agree.	Emot.
Work schedule	Baseline	-0.021***	-0.003	-0.036***	-0.022***	-0.035***
		(0.007)	(0.005)	(0.007)	(0.008)	(0.008)
	I. Full-time	0.010	0.018***	0.013**	0.005	-0.001
		(0.007)	(0.006)	(0.006)	(0.009)	(0.013)
	I. Part-time	-0.030	0.047**	0.008	0.050*	0.045
		(0.027)	(0.023)	(0.017)	(0.028)	(0.039)
Contract period	Baseline	-0.022**	0.004	-0.031***	-0.008	-0.025***
		(0.009)	(0.005)	(0.009)	(0.010)	(0.009)
	I. Temporary	-0.014	-0.016	-0.012	-0.023	-0.042*
		(0.010)	(0.010)	(0.012)	(0.016)	(0.024)
	I. Permanent	0.009	0.014***	0.002	-0.007	-0.002
		(0.008)	(0.005)	(0.006)	(0.011)	(0.009)
Experience	Baseline	-0.035***	0.001	-0.032***	0.005	-0.042*
		(0.011)	(0.011)	(0.011)	(0.016)	(0.023)
	I. Needed	0.019*	0.009	-0.000	-0.027*	0.013
		(0.010)	(0.013)	(0.009)	(0.014)	(0.024)
	I. Not needed	0.085***	0.015	0.029	-0.006	0.057
		(0.031)	(0.037)	(0.036)	(0.033)	(0.035)
Education	Baseline	-0.022***	0.007*	-0.032***	-0.017*	-0.035***
		(0.007)	(0.004)	(0.007)	(0.009)	(0.008)
	GCSE	0.005	-0.003	0.028*	0.030	0.023
		(0.013)	(0.014)	(0.015)	(0.025)	(0.017)
	A-Level	0.020*	-0.011	0.000	0.025	0.041***
		(0.010)	(0.010)	(0.012)	(0.019)	(0.010)
	Bachelor	0.012	0.003	-0.014	0.031*	0.031
		(0.018)	(0.014)	(0.014)	(0.018)	(0.026)
	Postgraduate	0.003	0.010	0.001	0.005	-0.025
		(0.018)	(0.018)	(0.015)	(0.012)	(0.016)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. Regressions control for job requirements, location, time-month, occupation and firm fixed effects. 9,888,315 Observations. *I.* correspond to Interaction terms (Baseline + Interaction).

5.1.3 Correlation with firm characteristics

This section examines how firm characteristics affect the relationship between personality traits and wages. Three mechanisms may explain these effects. First, according to Bowles et al. (2001), larger firms with higher monitoring costs may benefit more from employees with traits like Conscientiousness and Emotional Stability. Second, firm reputation may matter: established firms can rely on their reputation, while smaller or younger firms might use personality language to attract candidates, possibly offering lower wages but better perceived fit (Kristof, 1996; Rosen, 1986). Third, different organisational structures may value certain traits based on their business models (Nordman et al., 2019).

To test these mechanisms, I matched Adzuna data with the FAME database to get information on firm age, size, profit, and public listing status. Table 5 shows how each firm characteristic affects the relationship between personality traits and wages. These regressions include firm fixed effects, meaning the coefficients capture how trait-firm characteristic interactions vary within firms after controlling for firm-specific factors.

Firm age shows significant effects only for Extraversion, where penalties increase from 3.0 per cent at mean firm age to 4.7 per cent in older firms. This suggests that established firms may use Extraversion-related language more selectively or for different signalling purposes than younger firms.

Firm size and profitability show no statistically significant interaction effects for any personality trait. This contradicts monitoring cost theories that predict larger firms should value traits like Conscientiousness and Emotional Stability more highly due to greater supervision challenges. The absence of size effects also indicates that firm scale does not explain trait penalties through correlation with lower or higher wages.

Public listing status has significant effects for Agreeableness, where publicly quoted firms show larger penalties (7.5 per cent total effect) compared to other private firms (1.3 per cent). The effect is also larger for the rest of the traits when publicly quoted. This represents the strongest firm-level interaction observed, suggesting that organisational governance structures may influence how interpersonal traits are valued in compensation decisions. However, if we consider publicly quoted firms as more widely known, the results contradict the argument that firms with little reputation use traits to attract individuals while paying lower wages.

These results provide limited support for firm-based explanations of personality trait wage penalties within the correlation with low-pay characteristics hypothesis. The absence of significant interactions for firm size and profitability indicates that trait penalties do not simply reflect concentration in smaller, less profitable firms. The monitoring cost mechanism receives no empirical support, as larger firms do not show reduced penalties

Table 5: Effects of personality traits with standardised firm characteristics

Firm Characteristic		Consc.	Open.	Extra.	Agree.	Emot.
Firm Age (years)	At mean (0 SD)	-0.024**	0.009*	-0.030***	-0.006	-0.017
		(0.011)	(0.005)	(0.010)	(0.014)	(0.012)
	Per 1 SD increase	-0.001	0.002	-0.017***	-0.013	0.004
		(0.009)	(0.005)	(0.006)	(0.011)	(0.007)
	Total effect at +1 SD	-0.026*	0.011	-0.047***	-0.019	-0.014
		(0.015)	(0.008)	(0.011)	(0.018)	(0.009)
Firm Size (employees)	At mean (0 SD)	-0.024**	0.009	-0.031***	-0.002	-0.013
		(0.011)	(0.006)	(0.010)	(0.014)	(0.011)
	Per 1 SD increase	-0.006	0.001	0.008	0.050	0.018
		(0.008)	(0.003)	(0.009)	(0.033)	(0.018)
	Total effect at +1 SD	-0.029*	0.009	-0.023	0.048	0.005
		(0.016)	(0.007)	(0.016)	(0.043)	(0.026)
Profit (£)	At mean (0 SD)	-0.016**	0.002	-0.029***	-0.011	-0.015
		(0.008)	(0.006)	(0.010)	(0.010)	(0.010)
	Per 1 SD increase	0.000	0.005	0.001	0.014	0.004
		(0.007)	(0.005)	(0.008)	(0.012)	(0.010)
	Total effect at +1 SD	-0.016	0.007	-0.028**	0.003	-0.012
		(0.011)	(0.008)	(0.011)	(0.018)	(0.014)
Publicly Quoted	At mean (0 SD)	-0.020***	0.006	-0.031***	-0.013	-0.031***
		(0.007)	(0.004)	(0.007)	(0.008)	(0.008)
	When publicly quoted = 1	-0.007	0.026	-0.010	-0.062***	-0.007
		(0.015)	(0.032)	(0.018)	(0.018)	(0.021)
	Total effect when quoted	-0.027*	0.032	-0.042**	-0.075***	-0.038*
		(0.014)	(0.032)	(0.018)	(0.017)	(0.021)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. "At mean" shows the effect when the standardised firm characteristic equals zero (at its mean). "Per 1 SD increase" shows the combined effect when the firm characteristic is one standard deviation above its mean. Regressions control for job requirements, location, time-month, occupation and firm fixed effects. Sample size: 3619726 observations for Firm Age estimations, 1897228 observations for Profit, 3009216 observations for Firm Size and 9888315 observations for Publicly Quoted.

for traits that should facilitate supervision. Only public listing status shows meaningful interaction effects, and these amplify rather than explain trait penalties. Overall, these findings suggest that firm characteristics cannot account for the personality trait wage penalty through correlation with low-pay firm types, indicating that other mechanisms must explain this relationship.

5.1.4 Assessment of the Low-Pay Correlation Hypothesis

The correlation with low-pay characteristics hypothesis receives partial support but cannot fully explain the personality trait wage penalty. Three key findings emerge.

First, demographic composition at the occupation level matters. Personality traits have smaller penalties or even premiums in occupations with higher proportions of women and migrants, suggesting that some observed trait penalties reflect sorting into lower-paying occupational contexts. However, trait penalties generally increase with occupation education level, wage level, and managerial content—precisely the contexts where human capital theory would predict stronger returns to productivity-enhancing traits.

Second, job-level characteristics provide mixed evidence. The positive interactions between personality traits and entry-level positions (low seniority, no experience required) suggest that trait penalties partly reflect concentration in junior roles. However, some job characteristics amplify rather than explain trait penalties, particularly hard skills and temporary contracts.

Third, firm-level characteristics provide limited explanatory power. The absence of significant interactions for firm size and profitability indicates that trait penalties do not simply reflect concentration in smaller, less profitable firms. Only public listing status shows meaningful effects, and these amplify rather than explain trait penalties.

Overall, while demographic composition and career stage characteristics provide some explanatory power, the persistence and amplification of penalties in high-skill, high-wage contexts indicates that additional mechanisms beyond correlation with observable low-pay characteristics must account for this relationship.

6 Compensating Differentials

A second explanation proposes that employers use personality-traits to signal non-pay amenities valued by workers, which may result in lower posted wages through a compensating differentials mechanism. According to Rosen (1986)'s equalising differences theory, when jobs offer attractive non-wage characteristics, employers can offer lower wages while still attracting qualified candidates. Personality-traits in job advertisements may signal a workplace where employees will experience better person-organisation fit, which candidates may value alongside monetary compensation.

Firms benefit from attracting workers whose personalities align with their organisational identity (Kristof, 1996; Ployhart et al., 2006; Roberson et al., 2005). This alignment can enhance job satisfaction, organisational commitment, team cohesion, and reduce turnover (Judge and Cable, 1997; Krueger and Schkade, 2008). By signalling these benefits through personality-related language, employers may effectively lower their wage offers while maintaining their ability to attract suitable candidates.

Building on signalling theory (Spence, 1973), research shows that job-seekers evaluate not just compensation but also fit with organisational culture when making application decisions (Breaugh, 2013; Chapman et al., 2005; Stevens and Szmerekovsky, 2010). This is supported by person-organisation fit theory, which suggests applicants are more likely to pursue positions aligned with their values (Judge and Cable, 1997; Kristof, 1996; Roberson et al., 2005). In this framework, job seekers may prioritise intrinsic rewards over purely monetary compensation (Akerlof and Kranton, 2005; Cassar and Meier, 2018).

This compensating differentials approach parallels Handy and Katz (1998)'s model of non-profit compensation, where firms offer lower monetary compensation but attract workers who value the organisation's mission. Indeed, Burbano (2016) found that information about an employer's social responsibility reduces candidates' wage requirements, providing direct evidence for how non-monetary aspects can allow firms to offer lower wages.

Empirical studies confirm that workers are attracted to organisations that align with their personality traits. Stevens and Szmerekovsky (2010) demonstrated that students with high extraversion scores showed greater interest in job advertisements seeking outgoing candidates. Similarly, Van Hoye and Turban (2015) found extraverted applicants were more attracted to recruiters displaying similar traits, while Judge and Cable (1997) observed that conscientious individuals preferred organisations with achievement-oriented cultures. Conversely, candidates who perceive a poor fit with an organisation's identity are less likely to apply (Breaugh, 2013; Chapman et al., 2005; Wille and Derous, 2018).

6.1 Empirical Methodology

To test whether personality-traits serves as a signal for job quality and person-organisation fit, I examine the relationship between personality traits, advertised amenities, and wages. The analysis proceeds in three steps: first examining correlations between traits and amenities, then testing whether amenity controls explain trait penalties, and finally investigating interaction effects.

6.1.1 Identifying Advertised Amenities

The identification of job amenities builds on the established methodology developed by Escudero et al. (2024) and Boschetti Adamczyk et al. (2025), which provides a taxonomical foundation validated across multiple labour market contexts. Their approach initially drew from Maestas et al. (2023)'s nine job attributes and was expanded using Sockin (2022)'s categorisation of 48 non-wage amenities derived through topic-modelling machine learning algorithms.

I adapted this framework to suit the UK labour market context and the Adzuna dataset characteristics. Following Escudero et al. (2024)'s approach, I reorganised and expanded their keyword dictionaries by conducting extensive context analysis of the terms they used, adding synonyms where appropriate, and implementing over 600 specific word replacements to eliminate false positives identified through manual verification of randomly sampled advertisements. This process involved examining multiple instances of each keyword to ensure it genuinely referred to the intended amenity rather than unrelated concepts that might share similar terminology.

The final taxonomy incorporates several modifications to better capture the UK labour market context. I excluded the physical effort and pace of work category used in the original frameworks, as preliminary analysis suggested they may not be considered as positive amenities. I split the original flexible work arrangements category into two distinct measures: flexible work arrangements (covering flexible hours, part-time options, and work-life balance) and remote work possibilities (capturing teleworking, home-based work, and hybrid arrangements), reflecting the increased prominence and distinct nature of remote work following the COVID-19 pandemic. Additionally, I introduced a general benefits category to capture miscellaneous perks and benefits that did not fit clearly into the established categories but were frequently mentioned in UK job advertisements. This adapted framework resulted in 16 distinct amenity categories, each validated through the rule-based natural language processing approach described in the source methodology, with binary indicators created for each category based on the presence of relevant keywords in job descriptions.

For each category, I expand Escudero et al. (2024)'s keyword lists through contextual analysis, examining 100 randomly selected sentences containing each potential keyword to verify usage and identify synonyms. The final dictionaries contain between 6 keywords (overtime paid) and 145 keywords (food subsidies and discounts) per category. A full list of keywords and replacements done to the advertisements to facilitate amenities retrieval is available in Appendix B

Table 6 provides definitions of each amenity category, organised into four broad groups reflecting different aspects of compensation and working conditions.

Table 6: Categorisation of amenities and definitions

Panel (a): Variable Earnings	
Subcategory	**Definition**
A01–General benefits	Encompasses reward schemes, flexible benefit platforms, comprehensive benefit packages, staff benefits, recognition schemes, referral programmes, tax benefits, relocation support, and seasonal benefits offered by employers.
A02–Bonuses and commissions	Encompasses various forms of financial incentives and rewards aimed at motivating and compensating employees based on their performance, achievements, or specific goals within an organisation, including equity compensation, profit sharing, performance pay, sales compensation, and premium pay arrangements.
Panel (b): Fringe Benefits	
A03–Overtime paid	Encompasses aspects related to flexible earnings, reflecting the compensation and conditions associated with working beyond regular hours, including extra shifts, double pay, and unsocial hour premiums within an employment arrangement.
A04–Paid time off	Reflects provisions for employees to take time away from work while receiving compensation in specific circumstances, including vacation and holidays, sick leave, medical leave, maternity and paternity benefits, family leave, personal time, special leave, volunteer time, and wellbeing time arrangements.
A05–Health insurance	Includes provisions offered by employers to support employees' healthcare needs, ensuring access to medical services and providing financial protection, including private health insurance, dental and vision coverage, life insurance, accident insurance, free healthcare, health savings plans, wellness programmes, childcare support, eye care, and specialized coverage.
A06–Retirement contributions	Encompasses provisions designed to assist employees in saving for their retirement and securing financial stability during their later years, including pension schemes, employer contribution matching, savings accounts, and life benefits.
A07–Subsidies and discounts	Encompasses benefits related to food, housing, transportation, and various subsidies or discounts offered to employees, including meal provisions, drinks and beverages, fruit and snacks, gym and fitness benefits, transport benefits, discount schemes, clothing and uniform provision, accommodation support, vouchers and rewards, wellness perks, vehicle benefits, platform benefits, leisure entertainment, and insurance extras.

Continued on next page

Table 6 – continued from previous page

Subcategory	Definition
	Panel (c): Non-Wage Job Attributes
A08–Office space and amenities	Covers workplace-related benefits, including facilities and amenities provided by the employer such as fitness facilities, recreational spaces, dining facilities, modern office environments, specialised rooms, work facilities, outdoor spaces, service facilities, and storage amenities.
A09–Location and commuting	Focuses on factors related to the workplace's geographical location and how employees commute to and from work, including location quality, city center access, accessibility, commuting arrangements, parking facilities, public transport links, walking distance convenience, cycling facilities, local transport options, and nearby amenities.
A10–Work equipment and allowances	Sheds light on how the organisation assists employees in ensuring they have the necessary tools and technology, including remote work support, mobile devices, computing equipment, company vehicles, vehicle expenses, company cards, clothing and protective equipment, and general work equipment provision.
	Panel (d): Working Conditions
A11–Work schedule flexibility	Includes various aspects related to the flexibility of work schedules and arrangements, such as flexible shifts and hours, options for better work-life balance, and modern working practices that support family-friendly policies and employee autonomy over working time.
A12–Workplace safety and ergonomics	Pertains to all aspects related to ensuring a safe working environment for employees, including workplace safety culture, security measures, health and safety compliance, safety regulations and policies, and ergonomic office arrangements that prioritise employee wellbeing.
A13–Job security	Encompasses all aspects related to ensuring job security, stability, and financial protection for employees, including company stability indicators, severance benefits, unemployment support, and income protection measures in various employment scenarios.
A14–Remote work opportunities	Encompasses various arrangements that allow employees to work outside traditional office settings, including fully remote positions, home-based work, hybrid arrangements, telecommuting, and virtual job opportunities that became particularly prominent following the COVID-19 pandemic and changing work patterns.
A15–Work environment and impact on society	Provides insights into the organization's commitment to creating a positive workplace environment and contributing positively to the community and society, including workplace culture, work satisfaction, company values, social activities, company reputation, employee support, work quality, impact and purpose, diversity and inclusion, and team environment characteristics.
A16–Human capital development	Assesses the opportunities for personal and professional growth and development within the organisation, including career growth prospects, skill development opportunities, learning programmes, company growth potential, training and education provision, promotion and advancement pathways, merit-based progression, and personal growth support.

6.1.2 Empirical Models

The analysis employs three complementary approaches to test the compensating differentials hypothesis.

Step 1: Trait-Amenity Correlations First, I examine whether amenities predict personality trait mentions, which would suggest these characteristics cluster together in job advertisements. For each personality trait t, I estimate:

$$Trait_{ti} = \beta_0 + \sum_{a=1}^{16} \beta_a Amenity_{ai} + \sum_{j \neq t}^{5} \gamma_j Trait_{ji} + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i \tag{6}$$

where $Trait_{ti}$ is a binary indicator for whether advertisement i mentions personality trait t, $Amenity_{ai}$ represents the 16 amenity categories, and $\sum_{j \neq t}^{5} \gamma_j Trait_{ji}$ controls for the other four personality traits. Positive coefficients β_a indicate that advertisements mentioning amenity a are more likely to also mention trait t. If employers use personality language alongside amenities to signal attractive job characteristics, we would expect positive correlations between traits and amenities that facilitate compensating differential mechanisms.

Step 2: Amenity Controls and Trait Penalties Second, I test whether controlling for amenities reduces personality trait wage penalties. I compare two specifications:

Model without amenity controls:

$$\ln(wage_i) = \beta_0 + \sum_{t=1}^{5} \beta_t Trait_{ti} + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i \tag{7}$$

Model with amenity controls:

$$\ln(wage_i) = \beta_0 + \sum_{t=1}^{5} \beta_t^a Trait_{ti} + \sum_{a=1}^{16} \lambda_a Amenity_{ai} + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i \tag{8}$$

If personality traits serve as proxies for non-monetary compensation, we would expect $|\beta_t^a| < |\beta_t|$, indicating that amenity controls explain part of the trait penalty through systematic correlation with amenity provision.

Step 3: Trait-Amenity Interactions Third, I examine whether personality traits and amenities interact in determining wages, testing whether employers strategically combine trait requirements with specific amenity packages. This approach moves beyond testing simple correlation between traits and amenities (captured in Steps 1–2) to examine whether the wage effect of trait mentions depends on the presence of specific amenities.

For each trait t and amenity category a, I estimate:

$$\begin{aligned}\ln(wage_i) = \beta_0 + \beta_t Trait_{ti} + \lambda_a Amenity_{ai} + \eta_{ta}(Trait_{ti} \times Amenity_{ai}) \\ + \sum_{b \neq a}^{16} \lambda_b Amenity_{bi} + \sum_{j \neq t}^{5} \gamma_j Trait_{ji} + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i\end{aligned} \tag{9}$$

The interaction coefficient η_{ta} captures how the wage effect of trait t changes when amenity a is present. The economic interpretation depends on both the amenity main effect (λ_a) and the interaction effect (η_{ta}). Under pure compensating differentials theory, amenities should reduce wages ($\lambda_a < 0$) as employers offer lower monetary compensation when providing valuable non-wage benefits. However, some amenities may command wage premiums ($\lambda_a > 0$) due to productivity effects, selection requirements, or signalling functions. For instance, Escudero et al. (2024) document that bonuses and commissions are associated with lower posted wages, as is schedule flexibility, while working in teams and human capital development opportunities show positive wage associations.

When $\lambda_a < 0$ (compensating differentials), negative interactions ($\eta_{ta} < 0$) suggest trait requirements amplify the compensating differential effect, whilst positive interactions ($\eta_{ta} > 0$) suggest traits reduce or reverse the compensating differential mechanism. When $\lambda_a > 0$ (amenity premiums), negative interactions ($\eta_{ta} < 0$) suggest trait requirements eliminate the amenity premium, possibly indicating employers use trait language to justify lower wages despite offering attractive amenities.

The total wage effect when both trait and amenity are present equals $\beta_t + \lambda_a + \eta_{ta}$. If the baseline trait effect β_t becomes statistically insignificant whilst the interaction η_{ta} remains significant, this suggests trait language primarily functions as part of strategic compensation packages rather than reflecting standalone productivity requirements.

6.2 Results

6.2.1 Trait-Amenity Correlations

To test whether employers use personality trait language alongside amenity provision to signal attractive job characteristics, I first examine whether amenity mentions predict personality traits in job advertisements. Table 7 presents results from linear probability models where each personality trait serves as the dependent variable, with all 16 amenity categories as predictors.

Table 7: Amenities and personality trait demand

Amenity Category	Consc.	Open.	Extra.	Agree.	Emot.
A01 General Benefits	0.001	0.001	0.020*	0.018	0.005
	(0.013)	(0.009)	(0.011)	(0.011)	(0.009)
A02 Bonuses and Comm.	-0.006	-0.017**	0.015	-0.001	0.006
	(0.012)	(0.007)	(0.010)	(0.008)	(0.007)
A03 Overtime	-0.021**	-0.001	-0.014*	0.007	0.007
	(0.009)	(0.018)	(0.008)	(0.011)	(0.014)
A04 Paid Time Off	0.006	0.021**	-0.006	0.002	-0.005
	(0.009)	(0.009)	(0.011)	(0.006)	(0.007)
A05 Health Insurance	0.003	-0.000	-0.015	0.011	-0.004
	(0.011)	(0.012)	(0.013)	(0.008)	(0.011)
A06 Retirement	-0.030***	-0.015	0.001	-0.005	0.009
	(0.008)	(0.009)	(0.014)	(0.009)	(0.009)
A07 Subsidies and discounts	-0.002	-0.008	0.037***	-0.014	-0.023**
	(0.014)	(0.011)	(0.012)	(0.009)	(0.010)
A08 Office Space	-0.006	0.003	0.016*	-0.002	-0.016**
	(0.015)	(0.013)	(0.008)	(0.007)	(0.007)
A09 Location Amenities	-0.003	-0.006	0.016*	-0.005	-0.001
	(0.007)	(0.008)	(0.008)	(0.005)	(0.007)
A10 Work Equipment	0.018*	-0.025**	-0.000	-0.006	0.006
	(0.010)	(0.012)	(0.013)	(0.007)	(0.012)
A11 Work Schedule Flex	0.027	0.053***	0.023	-0.021	-0.032**
	(0.018)	(0.011)	(0.016)	(0.013)	(0.014)
A12 Workplace Safety	0.037***	-0.001	-0.058***	-0.001	0.018
	(0.011)	(0.011)	(0.013)	(0.010)	(0.012)
A13 Job Security	0.023	-0.028	0.023	0.021	0.019
	(0.036)	(0.024)	(0.031)	(0.023)	(0.023)
A14 Remote Work	-0.015	0.006	-0.011	-0.005	-0.002
	(0.014)	(0.016)	(0.012)	(0.006)	(0.005)
A15 Work Environment	0.007	0.025***	0.038***	0.021***	-0.003
	(0.008)	(0.007)	(0.008)	(0.007)	(0.007)
A16 Human Capital Dev	0.019**	-0.004	0.036***	-0.006	-0.010
	(0.008)	(0.005)	(0.007)	(0.004)	(0.006)

Notes: Net coefficients show coefficient β_a from Eqn (6) of each amenity category on trait demand, controlling for all other amenities. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. Regressions control for other traits, job requirements, location, time-month, occupation and firm fixed effects. Number of observations 9,888,315.

The results show selective clustering between personality traits and amenities that varies across trait types. Extraversion has the strongest positive associations with amenities, showing significant correlations with food subsidies and discounts (3.7 per cent), work environment features (3.8 per cent), human capital development (3.6 per cent), and location amenities (1.6 per cent). This clustering suggests that employers seeking extraverted workers bundle these positions with workplace amenities, which could facilitate compensating differential mechanisms.

Conscientiousness displays contradictory associations: positive correlations with workplace safety (3.7 per cent), work equipment (1.8 per cent), and human capital development (1.9 per cent), but negative correlations with overtime pay (2.1 per cent) and retirement benefits (3.0 per cent). These opposing patterns suggest that Conscientiousness requirements appear in different job contexts—some emphasising safety and development, others involving shift work with limited traditional benefits.

Emotional Stability shows predominantly negative associations with amenities, correlating negatively with food subsidies (2.3 per cent), office space (1.6 per cent), and work schedule flexibility (3.2 per cent). Openness has selective positive correlations with amenities that enhance work flexibility: paid time off (2.1 per cent), work environment (2.5 per cent), and flexible work schedules (5.3 per cent), but negative associations with bonuses and commissions (1.7 per cent) and work equipment (2.5 per cent). Agreeableness shows the most limited amenity associations, with only work environment showing a significant positive correlation (2.1 per cent).

These patterns indicate that trait-amenity clustering is trait-specific rather than uniform, suggesting that compensating differentials mechanisms, if they operate, must function differently across personality characteristics and job contexts.

6.2.2 Amenity Controls and Trait Wage Penalties

The second test examines whether controlling for amenities as additive terms reduces personality trait wage penalties, which would indicate that part of the penalty operates through systematic correlation with amenity provision. Table 8 compares trait coefficients with and without amenity controls in the full specification including all fixed effects.

The results provide limited evidence that amenities explain trait wage penalties through additive mechanisms. Controlling for amenities has minimal impact on trait coefficients: Agreeableness shows identical penalties (1.4 per cent), Conscientiousness penalties decrease marginally from 2.1 to 2.0 per cent, Emotional Stability penalties remain stable (3.0 to 3.1 per cent), Extraversion penalties decrease slightly from 3.4 to 3.2 per cent, and Openness effects remain small and positive (0.6 to 0.7 per cent). The R^2 increases minimally from 0.598 to 0.601, indicating that amenities explain very little additional variation in wages.

This stability suggests that the baseline wage penalties for personality traits do not

Table 8: Personality trait wage effects: with vs without amenities controls

Personality Trait	Without Amenities	With Amenities
Agreeableness	-0.014*	-0.014*
	(0.008)	(0.008)
Conscientiousness	-0.021***	-0.020***
	(0.007)	(0.007)
Emotional Stability	-0.030***	-0.031***
	(0.008)	(0.008)
Extraversion	-0.034***	-0.032***
	(0.008)	(0.007)
Openness	0.006	0.007*
	(0.004)	(0.004)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Both specifications include all personality traits together, plus job requirements, location, time-month, occupation and firm fixed effects. Without amenities: N= 9,888,315, R^2=0.598. With amenities: N= 9,888,315, R^2=0.601.

primarily operate through employers offering different amenity packages in trait-demanding jobs. The absence of additive effects indicates that compensating differentials, if they exist, must operate through interactive rather than additive mechanisms.

6.2.3 Interaction between Amenities and Traits

The interaction analysis tests whether employers strategically combine personality trait requirements with specific amenity packages, examining whether the wage effect of trait mentions varies depending on the compensation context. Table 9 presents both the baseline effects (when amenities are absent) and the total effects (baseline plus interaction, when amenities are present).

The analysis reveals evidence for context-dependent employer strategies operating through multiple mechanisms beyond pure compensating differentials. Amenities modify trait wage effects, with patterns varying across trait-amenity combinations, indicating that employers use personality trait language differently depending on the broader compensation package offered.

Table 9: Baseline and total wage effects of personality traits by amenity presence

		Consc.	Open.	Extra.	Agree.	Emot. Stab.
A01 General benefits	Baseline	-0.024***	0.004	-0.036***	-0.023***	-0.043***
		(0.007)	(0.006)	(0.007)	(0.008)	(0.009)
	Total	-0.013	0.014	-0.022**	-0.002	-0.013
		(0.011)	(0.009)	(0.010)	(0.012)	(0.015)

Continued on next page

Table 9 – continued from previous page

		Consc.	Open.	Extra.	Agree.	Emot. Stab.
A02 Bonuses and commissions	Baseline	-0.021**	0.003	-0.032***	-0.011	-0.026***
		(0.008)	(0.005)	(0.007)	(0.009)	(0.007)
	Total	-0.018**	0.018**	-0.030***	-0.023**	-0.044***
		(0.008)	(0.008)	(0.010)	(0.011)	(0.016)
A03 Overtime	Baseline	-0.019***	0.008*	-0.035***	-0.019***	-0.037***
		(0.006)	(0.004)	(0.007)	(0.006)	(0.006)
	Total	-0.051	-0.002	0.041	0.063	0.060
		(0.042)	(0.016)	(0.037)	(0.067)	(0.059)
A04 Paid time off	Baseline	-0.014*	0.006	-0.037***	-0.018**	-0.027***
		(0.007)	(0.005)	(0.007)	(0.007)	(0.005)
	Total	-0.037***	0.012	-0.016	-0.006	-0.041*
		(0.014)	(0.008)	(0.012)	(0.016)	(0.021)
A05 Health insurance	Baseline	-0.018***	0.005	-0.038***	-0.017***	-0.032***
		(0.006)	(0.005)	(0.006)	(0.006)	(0.007)
	Total	-0.034	0.019*	0.005	-0.003	-0.029
		(0.021)	(0.010)	(0.020)	(0.027)	(0.029)
A06 Retirement contributions	Baseline	-0.020***	0.003	-0.039***	-0.020***	-0.029***
		(0.007)	(0.005)	(0.006)	(0.007)	(0.008)
	Total	-0.021*	0.014**	-0.017	-0.005	-0.035**
		(0.012)	(0.005)	(0.011)	(0.015)	(0.016)
A07 Subsidies and discounts	Baseline	-0.020***	0.005	-0.038***	-0.022***	-0.032***
		(0.006)	(0.004)	(0.007)	(0.007)	(0.005)
	Total	-0.021	0.015*	-0.007	0.008	-0.029
		(0.016)	(0.008)	(0.015)	(0.020)	(0.022)
A08 Office space	Baseline	-0.021***	0.008*	-0.032***	-0.014*	-0.032***
		(0.007)	(0.004)	(0.007)	(0.008)	(0.008)
	Total	0.003	-0.005	-0.016	-0.024	-0.011
		(0.011)	(0.015)	(0.012)	(0.023)	(0.012)
A09 Location and commuting	Baseline	-0.024***	0.006	-0.034***	-0.017*	-0.032***
		(0.007)	(0.005)	(0.007)	(0.009)	(0.009)
	Total	0.004	0.012	-0.020**	-0.001	-0.030***
		(0.009)	(0.008)	(0.010)	(0.014)	(0.010)
A10 Work equipment	Baseline	-0.016**	0.005	-0.034***	-0.023***	-0.028***
		(0.006)	(0.005)	(0.006)	(0.006)	(0.006)
	Total	-0.053*	0.024**	-0.017	0.029	-0.047

Continued on next page

		Consc.	Open.	Extra.	Agree.	Emot. Stab.
		(0.027)	(0.009)	(0.022)	(0.027)	(0.039)
A11 Work schedule flexibility	Baseline	-0.018***	0.011**	-0.033***	-0.014*	-0.032***
		(0.007)	(0.005)	(0.007)	(0.008)	(0.007)
	Total	-0.035*	-0.013	-0.025*	-0.016	-0.027*
		(0.018)	(0.013)	(0.015)	(0.012)	(0.014)
A12 Workplace safety	Baseline	-0.022***	0.007	-0.033***	-0.012	-0.031***
		(0.008)	(0.005)	(0.008)	(0.008)	(0.008)
	Total	-0.008	0.008	-0.021**	-0.035***	-0.036***
		(0.008)	(0.010)	(0.010)	(0.013)	(0.008)
A13 Job security	Baseline	-0.021***	0.007	-0.032***	-0.015*	-0.032***
		(0.007)	(0.004)	(0.007)	(0.008)	(0.008)
	Total	0.010	0.016	-0.038*	0.030	-0.022
		(0.024)	(0.022)	(0.021)	(0.020)	(0.018)
A14 Remote work	Baseline	-0.020***	0.007	-0.031***	-0.013*	-0.030***
		(0.007)	(0.004)	(0.007)	(0.008)	(0.008)
	Total	-0.033***	0.024	-0.052***	-0.045***	-0.057***
		(0.008)	(0.016)	(0.013)	(0.015)	(0.011)
A15 Work environment	Baseline	-0.030***	0.003	-0.045***	-0.039***	-0.046***
		(0.006)	(0.006)	(0.006)	(0.008)	(0.011)
	Total	-0.008	0.012*	-0.015*	0.002	-0.020*
		(0.011)	(0.006)	(0.009)	(0.012)	(0.011)
A16 Human capital dev.	Baseline	-0.028***	0.004	-0.060***	-0.038***	-0.054***
		(0.007)	(0.006)	(0.008)	(0.005)	(0.006)
	Total	-0.017**	0.009*	-0.019**	-0.007	-0.024**
		(0.008)	(0.005)	(0.007)	(0.010)	(0.010)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. Baseline effects when amenity is absent vs total effects (baseline + interaction) when amenity is present. Regressions control for other traits, job requirements, location, time-month, occupation and firm fixed effects. R^2 ranges from 0.601 to 0.601. Some amenities may be omitted if not present in sample.

Traditional Benefits and Compensating Differentials

Traditional fringe benefits show patterns most consistent with compensating differentials theory. When health insurance is advertised, most trait penalties are eliminated: Conscientiousness penalties shift from 1.8 per cent to a non-significant 3.4 per cent, Extraversion penalties disappear entirely (from 3.8 per cent to a non-significant 0.5 per cent), and Agreeableness penalties are eliminated (from 1.7 per cent to a non-significant 0.3 per cent).

Retirement contributions similarly moderate penalties for most traits.

These patterns suggest that when employers offer traditional benefits packages, personality trait requirements function as part of compensating differentials packages that allow lower monetary compensation while maintaining candidate attraction through enhanced total compensation.

Workplace Culture and Development Amenities

Workplace culture and development amenities show the strongest compensating differentials effects. Work environment mentions reduce penalties across all traits: Conscientiousness penalties decline from 3.0 per cent to a non-significant 0.8 per cent, Agreeableness shifts from a penalty (3.9 per cent) to no significant effect (0.2 per cent), and Emotional Stability penalties are reduced (4.6 to 2.0 per cent). Human capital development opportunities create similar moderating effects.

These findings provide strong support for compensating differentials mechanisms when personality trait language is combined with signals of positive organisational culture and employee development.

Amenities with Wage Premiums and Amplified Trait Penalties

However, several amenities show patterns inconsistent with simple compensating differentials theory. Remote work creates uniformly larger penalties for most traits: Conscientiousness penalties increase from 2.0 per cent to 3.3 per cent, Extraversion penalties amplify from 3.1 per cent to 5.2 per cent, and Emotional Stability penalties nearly double from 3.0 per cent to 5.7 per cent. This finding is unexpected: standard economic models suggest employers should value conscientious workers more when direct monitoring is limited, yet the data show the opposite pattern. The mechanisms behind this counterintuitive result remain unclear and warrant further investigation.

Variable compensation arrangements show more expected patterns for some traits. When bonuses are offered, the Conscientiousness penalty moderates (from 2.1 per cent to 1.8 per cent) and Openness shifts from a non-significant baseline effect to a positive premium (1.8 per cent). This pattern for Conscientiousness aligns with theoretical predictions: performance-based pay and conscientiousness may serve as substitutes in incentivizing effort, as Brenčič and McGee (2023b) suggest in their analysis. However, Agreeableness and Emotional Stability show larger penalties when bonuses are offered, indicating that variable compensation does not universally moderate trait penalties.

Physical Workplace Amenities

Physical workplace amenities generally reduce trait penalties in ways consistent with compensating differentials. Office space amenities eliminate penalties for most traits, with Conscientiousness shifting from a penalty (2.1 per cent) to a non-significant positive effect (0.3 per cent). Location and commuting benefits similarly moderate penalties, particularly for Conscientiousness and Emotional Stability.

6.2.4 Assessment of the Compensating Differentials Hypothesis

The compensating differentials hypothesis receives substantial but selective support. Three key findings emerge.

First, trait-amenity clustering is selective rather than uniform. Extraversion shows the most consistent amenity associations, while other traits show contradictory or negative associations. This indicates that compensating differentials cannot operate through simple uniform bundling of traits with amenities.

Second, additive amenity controls provide limited explanatory power. The stability of trait coefficients when amenity controls are added suggests that baseline wage penalties do not primarily operate through employers offering different amenity packages. This indicates that compensating differentials must operate through interactive rather than additive mechanisms.

Third, the interaction analysis provides strong evidence for context-dependent compensating differentials mechanisms. When personality trait language is combined with traditional benefits (health insurance, retirement contributions), workplace culture signals (work environment, human capital development), and quality physical amenities, trait penalties are reduced or eliminated. However, some amenity combinations amplify rather than moderate trait penalties, particularly remote work and variable compensation.

The compensating differentials hypothesis receives substantial but selective support. The mechanism operates effectively when personality trait language is combined with traditional benefits and positive workplace culture signals, allowing employers to offer lower monetary compensation while still attracting suitable candidates through enhanced person-organisation fit. However, compensating differentials cannot explain all instances of trait penalties, particularly when traits are combined with amenities that command market premiums.

7 Language Intensity

The baseline analysis established that advertisements mentioning personality traits offer lower wages on average. However, this binary measure treats all trait mentions equally, regardless of how much emphasis employers place on these characteristics. This section examines whether the intensity of personality trait language relates to wage offers.

The relationship between trait intensity and wages can be understood through information processing and signalling frameworks. Information architecture theory suggests that the frequency and prominence of content elements shape their interpretation (Rosenfeld et al., 2015). When personality traits dominate job descriptions, they may trigger information saturation effects, where excessive repetition diminishes message value (Malhotra, 1982), potentially crowding out discussion of technical requirements. Conversely, Mahjoub and Kruyen (2021) note that detailed and specified job advertisements can positively influence perceptions about job appropriateness and truthfulness, helping potential applicants reduce search costs and better understand person-organisational fit (Wei et al., 2016).

I first present the methodology for measuring trait intensity, then examine how different levels of emphasis on personality traits and their placement in advertisements relate to wages.

7.1 Empirical Methodology

While the boolean approach identifies the presence of personality traits, it does not capture the emphasis or intensity with which traits are mentioned. Term frequency-inverse document frequency (TF-IDF) provides a methodological framework for capturing these linguistic patterns by measuring both local term prominence (how frequently traits appear within a single advertisement) and global distinctiveness (how unique these mentions are across all advertisements) (Salton and Buckley, 1988; Sparck Jones, 1972). High TF-IDF scores indicate advertisements where personality traits are both frequently mentioned and distinctive, suggesting deliberate emphasis beyond standard practice.

Technically, the TF-IDF method quantifies the importance of a word within a document by considering both its frequency in a specific document (Term Frequency) and its rarity across all documents (Inverse Document Frequency). For a term t in document d within a corpus of N documents:

$$\text{TF-IDF}(t, d) = \text{TF}(t, d) \times \text{IDF}(t)$$

where the Term Frequency is:

$$\text{TF}(t, d) = \text{frequency of term } t \text{ in document } d$$

and the Inverse Document Frequency is:

$$\text{IDF}(t) = \log\left(\frac{N}{\text{number of documents containing term } t}\right)$$

To measure trait intensity, I transform both job descriptions and trait-related keywords into TF-IDF vectors. I then calculate the cosine similarity between these vectors, which measures the cosine of the angle between them:

$$\text{similarity}(d, \text{trait}) = \frac{\vec{d} \cdot \vec{\text{trait}}}{||\vec{d}|| \times ||\vec{\text{trait}}||}$$

where $\vec{d}$ is the TF-IDF vector of the job description and $\vec{\text{trait}}$ is the TF-IDF vector of the trait keywords.

This approach generates a continuous *trait intensity score* between -1 and 1 for each trait in each advertisement. A higher score indicates that the job posting places greater emphasis on the trait, accounting for both the frequency of trait mentions and their relative importance in the broader corpus. A detailed mathematical description of the TF-IDF calculation and cosine similarity measure is provided in Appendix C.

7.1.1 Empirical models

I investigate the intensity hypothesis through three complementary approaches. First, I examine the linear relationship between trait intensity and wages. For each trait t, I estimate:

$$\begin{aligned} \ln(wage_i) = \beta_0 + \beta_t Trait\_Score_{ti} + \sum_{j \neq t}^{5} \beta_j Trait\_Score_{ji} \\ + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i \end{aligned} \tag{10}$$

where $Trait\_Score_{ti}$ is the standardised intensity score for trait t in advertisement i, derived from the cosine similarity between the advertisement's TF-IDF vector and the trait's TF-IDF vector. Crucially, this specification includes all advertisements in the sample: those not mentioning the trait receive a score of zero, while those mentioning it receive positive intensity scores. This approach captures the combined effect of the decision to mention traits and the intensity chosen. The other traits ($Trait\_Score_{ji}$) are also included as standardised scores.

Second, to examine potential non-linearities and separate low-intensity from high-intensity mentions, I divide advertisements into quintiles based on their trait intensity scores. For each trait t, I estimate:

$$\begin{aligned} \ln(wage_i) = \beta_0 + \sum_{q=1}^{5} \beta_{tq}(Trait_{ti} \times Q_{qi}) + \sum_{j \neq t}^{5} \beta_j Trait_{ji} \\ + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i \end{aligned} \tag{11}$$

where Q_{qi} is a dummy variable indicating whether advertisement i belongs to the q-th quintile of trait t intensity. The baseline category is advertisements that do not mention the trait at all. In this specification, the other traits ($Trait_{ji}$) are included as binary indicators.

The coefficients β_{tq} measure the wage effect for each quintile of trait intensity compared to not mentioning the trait, thus capturing the combined effect of mentioning and intensity level. However, the pattern across quintiles provides information about the intensive margin: if $\beta_{t5} > \beta_{t4} > ... > \beta_{t1}$, this indicates that higher intensity has additional wage effects beyond the initial decision to mention traits. This approach allows me to identify whether low-intensity mentions (Q1) have different wage effects compared to high-intensity mentions (Q5), and how these compare to not mentioning traits at all. I also provide a robustness exercise using simple frequency counts.

Third, I examine how the relationship between trait intensity and wages varies with firm characteristics by estimating, for each trait t and firm characteristic f:

$$\begin{aligned}\ln(wage_i) = \beta_0 + \sum_{q=1}^{5} \beta_{tq}(Trait_{ti} \times Q_{qi}) + \sum_{j \neq t}^{5} \beta_j Trait_{ji} \\ + \gamma_f Firm\_Char_{jf} + \sum_{q=1}^{5} \nu_{tfq}(Trait_{ti} \times Q_{qi} \times Firm\_Char_{jf}) \\ + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \epsilon_i\end{aligned} \tag{12}$$

where $Firm\_Char_{jf}$ represents standardised firm characteristics such as size, age, profit, and a dummy for public listing status. The coefficients ν_{tfq} capture how the wage effect of trait intensity at each quintile q varies with firm characteristic f. This approach tests whether certain types of firms are more likely to use trait-related language for different purposes. For instance, smaller or younger firms might have stronger incentives to use trait-related language as a compensating differential, resulting in larger wage penalties at high intensity levels.

If the linguistic intensity hypothesis is correct, we would expect different wage effects across quintiles. Furthermore, these patterns might vary systematically across different types of firms.

7.2 Results

7.2.1 Intensity scores and linear effects

Table 10 shows the effects of a one standard deviation increase in trait scores on wages. The linear intensity analysis reveals that stronger emphasis on personality traits is associated with wage penalties for most traits. Extraversion and Conscientiousness show the largest intensity effects (-0.9 percentage points each), followed by Emotional Stability (-0.5

percentage points). Agreeableness and Openness show no significant linear relationship between intensity and wages. These results suggest that the wage penalty for personality traits marginally increases with the emphasis placed on them in job advertisements.

Table 10: Effects of standardized personality trait intensity on wages

Agreeableness	Conscientiousness	Emotional Stability	Extraversion	Openness
-0.002	-0.009**	-0.005**	-0.009***	0.001
(0.009)	(0.003)	(0.002)	(0.003)	(0.003)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Coefficients show the effect of a one standard deviation increase in trait intensity on log wages. Regressions control for binary trait indicators, job requirements, location, time-month, occupation and firm fixed effects. Sample size: 2921390 observations.

7.2.2 Non-linear effects across intensity quintiles

To examine non-linear effects, I divide advertisements into quintiles based on trait intensity and estimate the wage effects for each quintile, controlling for location, month-year, occupation, and firm fixed effects. Figure 3 shows these results for each of the Big Five traits, with advertisements that do not mention the trait as the baseline.[4], with advertisements that do not mention the trait as the baseline.

The results show clear non-linear patterns that provide important insights into the wage penalty paradox. For most traits, low-intensity mentions (first quintile) have small or negligible wage effects, while high-intensity mentions (fourth and fifth quintiles) show larger wage penalties. This pattern is most pronounced for Agreeableness, which shows positive coefficients in the first quintile that become increasingly negative through higher quintiles, reaching significant penalties in the fourth quintile. Conscientiousness follows a similar trajectory, with near-zero effects in the first quintile and progressively more negative effects in higher quintiles.

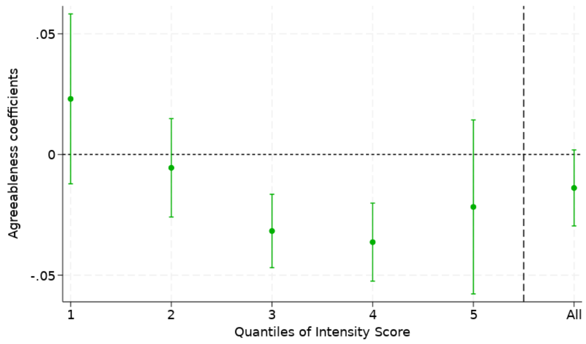
Extraversion has consistent penalties across all intensity levels, though these become more pronounced at higher intensities. Emotional Stability shows significant penalties beginning in the second quintile and maintaining these through higher intensity levels. Openness stands apart from other traits, maintaining relatively stable, small positive coefficients across most quintiles.

These intensity gradients suggest that employers use trait mentions strategically depending on emphasis level. When traits are mentioned with low intensity, they likely signal job requirements without triggering wage penalties. When traits receive high emphasis, they may serve different strategic purposes—potentially as compensating differential signals or demographic targeting mechanisms—that justify lower wage offers.

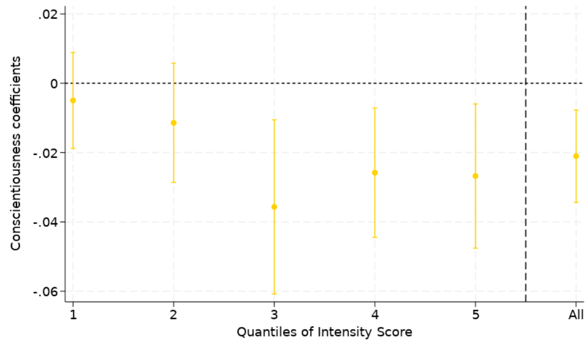
[4] Table A.9 in Appendix A presents the corresponding estimates

Figure 3: Binary Trait to Wage coefficients at different quantiles of the intensity score.

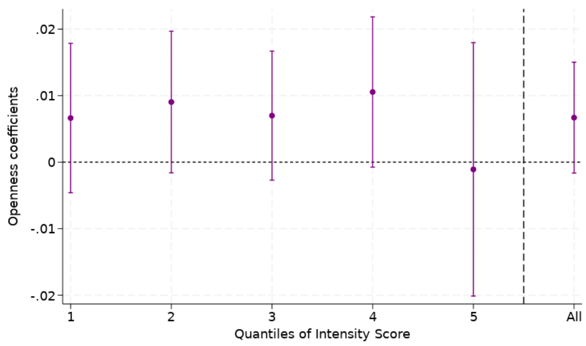
(a) Agreeableness



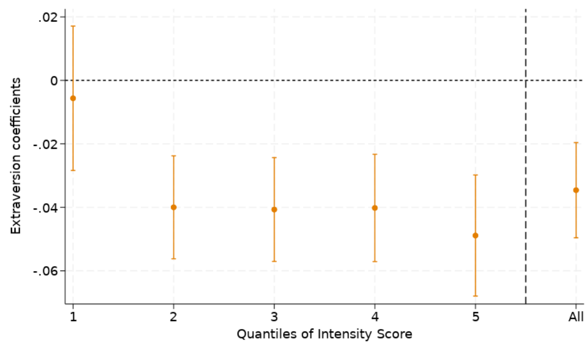
(b) Conscientiousness



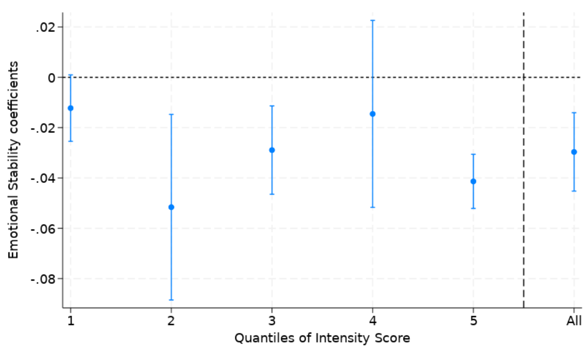
(c) Openness



(d) Extraversion



(e) Emotional Stability



Notes: Author's creation using Adzuna. Baseline is *no trait demand.*

7.2.3 Interaction with firm characteristics

The firm interaction analysis reveals that intensity effects vary systematically across different types of organisations, providing further evidence for strategic use of personality trait language. Tables A.5 through A.8 present these interaction effects across firm characteristics.

Larger firms demonstrate markedly different intensity patterns compared to average-sized firms. For Openness, large firms show significant premiums for low-intensity mentions (2.7 per cent in Q1) and high-intensity mentions (3.5 per cent in Q5), suggesting these firms may value this trait across intensity levels. Similarly, large firms show premiums for low-intensity Conscientiousness mentions (3.0 per cent in Q1), indicating that established firms with greater resources may use modest trait language to signal requirements.

Publicly quoted firms show dramatically different patterns, particularly for high-intensity trait mentions. For Conscientiousness, public firms impose severe penalties for high-intensity mentions, reaching -16.3 per cent in Q3 and -10.0 per cent in Q4, compared to much smaller penalties in non-public firms. Similarly, public firms penalize high-intensity Agreeableness mentions much more severely (-9.6 per cent in Q4) than private firms (-1.4 per cent). These patterns suggest that public firms may face different pressures that lead them to use high-intensity trait language for purposes other than identifying productive characteristics.

Firm age and profitability show more modest interaction effects, though older and more profitable firms generally show larger penalties for high-intensity trait mentions, particularly for Emotional Stability.

7.2.4 Frequency-Based Intensity Analysis

As a robustness check on the TF-IDF intensity analysis, I examine trait wage effects using simple frequency counts of trait mentions within advertisements. This approach provides a transparent measure of emphasis by counting raw mentions (1, 2, 3, 4, or 5+ times) without the contextual weighting inherent in TF-IDF measures. Table A.10 in Appendix A presents the wage effects for each frequency category, with advertisements not mentioning the trait as the baseline.

The frequency analysis confirms the non-linear intensity patterns observed in the TF-IDF analysis while providing additional insights into how raw repetition affects wages. Most traits show escalating penalties with increased mention frequency. Conscientiousness demonstrates a clear frequency gradient, with penalties increasing from -1.7 per cent for single mentions to -5.8 per cent for five or more mentions. Extraversion shows the most consistent escalation pattern, with penalties rising from -2.8 per cent (single mention) to

-9.2 per cent (five or more mentions).

Emotional Stability exhibits similar patterns, though with some variation at intermediate frequencies, ranging from -3.2 per cent for single mentions to -10.1 per cent at the highest frequency. Agreeableness shows more erratic patterns across frequency levels, with single mentions carrying -1.4 per cent penalties but higher frequencies showing mixed effects, including some positive but non-significant coefficients.

Openness maintains its exceptional status in the frequency analysis, showing positive premiums for moderate repetition (1.8 to 2.7 per cent for 2-4 mentions) before turning negative only at very high frequencies (-4.1 per cent for 5+ mentions). This pattern reinforces that Openness functions differently from other personality traits, with employers willing to pay premiums for moderate emphasis on this characteristic.

The consistency between TF-IDF quintile patterns and simple frequency counts provides strong evidence that the intensity effects are robust to different measurement approaches. Both methods demonstrate that low-emphasis trait mentions often carry minimal penalties while high-emphasis mentions show substantial penalties, supporting the hypothesis that employers use trait language strategically depending on the level of emphasis they wish to convey.

7.2.5 Trait Position Effects

I also examine whether the position of personality trait mentions within job advertisements affects their wage impact. Table 11 presents the wage effects across position quintiles, with advertisements not mentioning the trait as the baseline.

Table 11: Effects of trait position quintiles on wages

Trait	1-20% (Start)	21-40%	41-60%	61-80%	81-100% (End)
Agreeableness	-0.023***	-0.006	-0.013	-0.020**	-0.006
	(0.008)	(0.020)	(0.015)	(0.009)	(0.006)
Conscientiousness	-0.014**	-0.028***	-0.020	-0.015*	-0.024***
	(0.006)	(0.009)	(0.013)	(0.008)	(0.008)
Emotional Stability	-0.032**	-0.022	-0.040***	-0.031***	-0.031***
	(0.016)	(0.020)	(0.010)	(0.010)	(0.011)
Extraversion	-0.035***	-0.036***	-0.031***	-0.031***	-0.025***
	(0.010)	(0.010)	(0.008)	(0.007)	(0.008)
Openness	0.018*	0.003	0.005	0.001	0.008
	(0.009)	(0.006)	(0.005)	(0.011)	(0.006)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ Baseline is no trait mention. Position quintiles based on trait's first appearance as a percentage of advertisement length. Regression control for job requirements, location, time-month, occupation and firm fixed effects. Sample size: 9 888 315 observations.

The position analysis provides complementary evidence that linguistic presentation influences wage effects. Agreeableness shows the largest penalty (-2.3 per cent) when mentioned early in advertisements, with penalties diminishing when mentioned later in the text. Conscientiousness and Emotional Stability show penalties across most positions, ranging from -1.4 to -2.8 per cent for Conscientiousness and -2.2 to -4.0 per cent for Emotional Stability. Extraversion demonstrates consistent penalties of approximately -3.1 to -3.6 per cent across all positions, suggesting this trait carries wage penalties regardless of placement.

Only Openness shows a positive coefficient (1.8 per cent) when mentioned early, becoming non-significant when mentioned later. This pattern aligns with the intensity analysis, confirming that Openness differs fundamentally from other traits in its relationship with wages.

7.2.6 Assessment of the Linguistic Intensity Hypothesis

The linguistic intensity analysis provides evidence that employers use personality trait language strategically with different wage implications depending on the emphasis level. Three key findings emerge.

First, the non-linear relationship between trait intensity and wages shows that low-intensity mentions often carry minimal or no wage penalties, while high-intensity mentions consistently show larger penalties. This pattern suggests that moderate use of personality trait language likely reflects job requirements, while excessive emphasis serves different strategic purposes that justify lower wage offers.

Second, the firm interaction analysis reveals that established, larger firms show more favourable wage effects for low-intensity trait mentions, suggesting these organisations use trait language primarily to signal requirements. Conversely, publicly quoted firms impose severe penalties for high-intensity trait mentions, indicating that market pressures may lead these firms to use trait language for compensating differentials or signalling purposes rather than identifying productive characteristics.

Third, the position analysis confirms that linguistic presentation influences wage effects, although most traits carry penalties regardless of placement. The exception is Openness, which shows positive effects when mentioned prominently. The pattern of trait-related wage penalties appears consistently across all measurement approaches: binary indicators, continuous intensity scores, intensity quintiles, and placement analysis all show that personality trait language is associated with lower wages. This consistency across different measurement strategies indicates that the trait wage penalty reflects genuine employer behaviour rather than being driven by particular measurement choices.

These findings indicate that linguistic intensity provides important insights into the trait wage penalty paradox. The relationship between personality traits and wages is not uniform but depends critically on how employers choose to emphasise these characteristics. The evidence suggests that employers use varying intensity levels to signal different messages: moderate emphasis signals job requirements, while high intensity may serve compensating differential functions or target specific demographic groups willing to accept lower wages for perceived better organisational fit.

8 Demographic Targeting

A fourth explanation draws on statistical discrimination models (Aigner and Cain, 1977), suggesting employers use personality-traits to strategically target demographic groups they believe will accept lower wages. For instance, employers might assume women or younger workers value job aspects related to certain personality traits (such as agreeableness or enthusiasm) more highly than monetary compensation. By tailoring advertisements with specific personality trait language, employers may induce self-selection among these groups (Delfgaauw and Dur, 2007; Salop and Salop, 1976).

Recent experimental evidence supports this demographic targeting explanation. Koçak et al. (2023) examined how younger and older job seekers respond to stereotyped traits in job advertisements. Older job seekers' application intentions decreased when job advertisements emphasised traits like "flexibility", which activated negative stereotypes about their age group. Similarly, Koçak and Derous (2025) found that women's attraction to top-executive positions was lower when job advertisements contained men-related stereotyped competencies (such as "leading" or "decision-making") compared to other competencies such as "coaching" or "customer service orientation".

Research on gendered language and wage offers in job advertisements further supports this explanation. Kuhn et al. (2020) found that jobs using feminine-coded language offered wages 7.8 per cent lower than gender-neutral postings. Similarly, Chaturvedi et al. (2021) showed that gendered language in advertisements influenced application patterns, with women more likely to apply for lower-paying positions described with feminine-coded language. This targeting mechanism suggests personality trait language might serve as an implicit demographic filter that allows employers to offer lower wages while still attracting their preferred candidates.

8.1 Empirical Methodology

To test the hypothesis that employers use personality trait language to target specific demographic groups who may accept lower wages, I employ a machine learning approach to identify implicit gender preferences in job advertisements. This builds on recent literature examining gendered language in recruitment. I implement a strategy inspired by Kuhn et al. (2020) and Chaturvedi et al. (2021). While the former implements a Bernoulli naive Bayes classifier to infer the implicit gender preference in job's title, the latter constructs measures of *maleness* and *femaleness* using a multinomial logistic regression classifier on the advertisement explicit preferences. Chaturvedi et al. (2021) demonstrate that implicit hints within job text effectively direct where men and women send their applications, with jobs implicitly directed to women offering lower wages and receiving a higher fraction of female applications.

The UK Equality Act 2010 generally prohibits explicit gender requirements in job advertisements, with exceptions for occupational requirements such as same-gender care roles. Despite these restrictions, employers can use subtler linguistic patterns to signal gender preferences, such as female-related skills, number of qualifications or other gender-stereotyped words like *support, understanding, interpersonal* Abraham et al. (2024); Gaucher et al. (2011). To detect these patterns without restricting the analysis to a dictionary, I use the BERT (Bidirectional Encoder Representations from Transformers) model, a state-of-the-art natural language processing model that understands words in context rather than in isolation.

Unlike traditional methods such as keyword counting or logistic regression, BERT captures complex contextual language patterns. Traditional approaches treat words as independent units, while BERT considers words in relation to their surrounding context. This makes BERT particularly effective at detecting subtle biases in language that simpler methods might miss. For example, phrases like "competitive individual" and "nurturing environment" may carry gendered connotations that are difficult to identify with keyword-based methods but can be captured by contextual models like BERT.

I create a training dataset by identifying advertisements with clearly gendered language. I use predefined gender-associated keywords from Chaturvedi et al. (2021) ("female", "woman", "women", "girl", "lady", "ladies" and "feminine" for female candidates, and "male", "males", "man", "men", "guy", "guys", "boy", "boys", "gent" or "gents" for male candidates) to identify explicitly gendered advertisements. For instance, an explicit women-targeted advertisement stated "*...work in a brand spanking new salon, all decorated to a glam queen standard with friendly girls who love beauty...*", while a men-targeted advertisement contained "*...a big personality who can lead from the front, with a wicked eye for detail, and is a highly organised and effective man manager...*".

I apply several preprocessing steps to the training data before estimating the BERT model. First, to ensure the algorithm learns the underlying patterns of gendered language rather than simply identifying explicit gender terms, once advertisements were categorised as explicit female-, male-, or neutral-oriented, I replaced the gender keywords with neutral tokens, and trained the model on the remaining language.

Then, to avoid false positives from care roles mentioning client gender, I exclude advertisements where gendered terms appear near contextual phrases like "service for", "client is a", or "care for" unless they also contain explicit recruitment language like "female only" or "recruiting for". For instance, advertisements containing phrases like "care for elderly women" are not classified as gender-targeted unless they also contain phrases like "female applicants only", which would indicate a preference for the applicant's gender rather than describing clients.

I remove common stopwords (such as “the”, “and”, “a”) and stem words to their root forms (e.g., converting “caring”, “cares”, and “cared” to “care”). Finally, I stratify the training data by occupation to prevent the model from simply learning occupation-specific language patterns, as some occupations (such as care roles) have a higher prevalence of explicit gender mentions.

From the 11.5 million vacancies analysed, 1.62 per cent were categorised as explicitly targeting women, 1.35 per cent targeting men, and 1 per cent mentioned both genders. This explicitly gendered subset forms the training data for the model, which then classifies the remaining advertisements according to their implicit gender orientation.

The validation metrics show that the model achieves 94.6 per cent accuracy on the test set, with an F1-score of 0.947 averaged across all classes. The model demonstrates consistent performance in classifying gender preferences, with precision rates of 95.4 per cent for female-targeted advertisements, 93.4 per cent for male-targeted, and 95.2 per cent for neutral advertisements. The confusion matrix analysis shows that the model correctly classifies the majority of advertisements, though there are some misclassifications between female and male categories (613 female ads classified as male, and 400 male ads classified as female). The model identifies distinctive linguistic patterns associated with gender targeting, with common terms like “support” appearing across categories but with different weights and contextual usage. A detailed analysis of model performance by occupation is presented in Appendix D.

8.1.1 Empirical models

Using this gender classification, I incorporate implicit gender targeting into the wage model by estimating,for each trait t:

$$\begin{aligned} \ln(wage_i) = \beta_0 + \beta_t Trait_{ti} + \sum_{j \neq t}^{5} \beta_j Trait_{ji} + \sum_{g=1}^{3} \gamma_g Gender\_Target_{gi} \\ + \beta_{gt}(Trait_{ti} \times Gender\_Target_{gi}) + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i \end{aligned} \tag{13}$$

where $Gender\_Target_{gi}$ is a set of categorical indicators for whether advertisement i implicitly targets women, men, or is gender-neutral (based on the BERT classification). The model includes both the main effect of gender targeting ($\gamma_g Gender\_Target_{gi}$) and its interaction with each trait ($\beta_{gt}(Trait_{ti} \times Gender\_Target_{gi})$). The coefficients β_{gt} measure how the wage impact of personality trait t varies depending on the gender targeting category g.

If the demographic targeting hypothesis is correct, the wage penalty observed in the pooled sample could reflect a compositional effect: traits are mentioned more frequently in advertisements targeting demographic groups that receive lower wage offers overall. Under this hypothesis, when estimating wage effects separately by gender targeting category, trait penalties should be substantially reduced or eliminated in both categories. Alternatively, if trait language serves as a strategic targeting mechanism, we would expect systematically larger penalties in female-targeted advertisements where employers seek to attract workers willing to accept lower wages.

For testing both the intensity (Hypothesis 2) and gender targeting (Hypothesis 3) interactions together, I estimate for each trait t:

$$\begin{aligned}\ln(wage_i) = \beta_0 + \beta_t Trait\_Score_{ti} + \sum_{j\neq t}^{5} \beta_j Trait_{ji} + \sum_{g=1}^{3} \gamma_g Gender\_Target_{gi} \\ + \beta_{gt}(Trait\_Score_{ti} \times Gender\_Target_{gi}) + \beta_X X_i + \delta_l + \theta_m + \gamma_k + \phi_j + \epsilon_i\end{aligned} \tag{14}$$

This combined approach allows me to test whether the intensity of trait language interacts with gender targeting to affect wages.

8.2 Results

8.2.1 Gender Targeting and Personality Trait Prevalence

Before analysing wage effects, I first examine whether employers strategically use different personality trait language when targeting specific demographic groups. Table 12 shows that female-targeted advertisements are significantly more likely to mention Openness (3.1 per cent), Extraversion (2 per cent), and Agreeableness (1.1 per cent). No significant differences were found for Conscientiousness or Emotional Stability. This pattern provides initial evidence that employers do use personality trait language differently when targeting female candidates, particularly traits associated with interpersonal skills and openness to experience.

Table 12: Effect of female-targeted ads on personality trait demand

	Agree.	Consc.	Emot.	Extra.	Open.
Effect of Female-targeting	0.011***	-0.002	-0.005	0.020***	0.031***
	(0.003)	(0.004)	(0.006)	(0.006)	(0.006)

Notes: Coefficients show the effect of female gender-targeting on trait demand. Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Regressions control for job requirements, location, time-month, occupation and firm fixed effects. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability.

8.2.2 Wage Effects by Gender Targeting

Using the BERT-based classification model, I analyse how personality trait wage effects differ between female-targeted and male-targeted advertisements. Table 13 presents these results.

Table 13: Total effects of personality traits by gender targeting in job advertisements

	Consc.	Open.	Extra.	Agree.	Emot.
Female-intentioned ads	-0.018**	0.003	-0.031***	-0.023**	-0.046***
	(0.009)	(0.005)	(0.010)	(0.009)	(0.011)
Male-intentioned ads	-0.024***	0.003	-0.041***	-0.024***	-0.039***
	(0.007)	(0.005)	(0.006)	(0.007)	(0.008)

Notes: Total effects (base + interaction). Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. Regressions control for job requirements, location, time-month, occupation and firm fixed effects. Sample size: 9466620 observations.

Personality traits show significant wage penalties across both female-targeted and male-targeted advertisements. This indicates that the pooled wage penalty is not primarily driven by gender targeting. If demographic targeting explained the pooled penalty through compositional effects, penalties should diminish or disappear when estimated separately by gender category. Instead, penalties persist robustly in both categories. Most traits show larger penalties in male-targeted advertisements. Extraversion carries a stronger penalty in male-targeted advertisements (4.1 per cent versus 3.1 per cent), as does Conscientiousness (2.4 per cent versus 1.8 per cent) and Agreeableness (2.4 per cent versus 2.3 per cent). Only Emotional Stability shows a larger penalty in female-targeted advertisements (4.6 per cent versus 3.9 per cent). Openness remains the only trait without significant penalties in either gender category.

These patterns contradict the demographic targeting hypothesis. The persistence of similar penalty structures across gender categories indicates that other mechanisms drive the trait wage penalty. These include compensating differentials, linguistic intensity effects, or organisational signalling rather than strategic demographic targeting.

8.2.3 Trait Intensity and Gender Targeting Interactions

To further explore this relationship, I analyse how trait intensity interacts with gender targeting. Table 14 presents the wage effects across different quintiles of trait intensity for both female-targeted and male-targeted advertisements.

The intensity analysis reveals that gender targeting modifies trait-wage relationships in ways that vary by trait and intensity level. For Conscientiousness, female-targeted advertisements show no significant penalty at low intensity (Q1: 0.1 per cent, non-significant) but increasingly negative effects at higher intensities, reaching 3.9 per cent in the highest quintile. Male-targeted advertisements show penalties beginning at moderate intensities, with effects ranging from 1.2 per cent (Q2, non-significant) to 3.1 per cent (Q5).

Emotional Stability has the most severe penalties across both gender categories, with

Table 14: Effects of personality trait intensity by gender targeting (categorical)

Trait	Gender	Q1	Q2	Q3	Q4	Q5
Agreeableness	Male	-0.001	-0.016	-0.038***	-0.042***	-0.020
		(0.009)	(0.011)	(0.009)	(0.010)	(0.028)
	Female	0.002	-0.008	-0.033***	-0.041***	-0.038***
		(0.011)	(0.013)	(0.012)	(0.012)	(0.014)
Conscientiousness	Male	-0.011	-0.012	-0.038***	-0.024**	-0.031**
		(0.008)	(0.009)	(0.011)	(0.010)	(0.012)
	Female	0.001	-0.014	-0.008	-0.035*	-0.039**
		(0.008)	(0.010)	(0.012)	(0.018)	(0.016)
Emotional Stability	Male	-0.016**	-0.053***	-0.040***	-0.034**	-0.051***
		(0.007)	(0.018)	(0.011)	(0.017)	(0.008)
	Female	-0.025***	-0.075**	-0.035***	-0.039***	-0.054***
		(0.010)	(0.035)	(0.013)	(0.008)	(0.005)
Extraversion	Male	-0.018*	-0.048***	-0.046***	-0.043***	-0.049***
		(0.009)	(0.008)	(0.007)	(0.008)	(0.010)
	Female	-0.001	-0.027**	-0.037***	-0.039***	-0.060***
		(0.012)	(0.011)	(0.013)	(0.011)	(0.011)
Openness	Male	0.012*	0.004	0.001	0.008	-0.009
		(0.006)	(0.007)	(0.006)	(0.007)	(0.010)
	Female	0.001	0.004	0.007	0.005	-0.000
		(0.010)	(0.009)	(0.007)	(0.006)	(0.012)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Q1-Q5 represent quintiles of trait intensity (1=lowest, 5=highest). Regressions control for job requirements, location, time-month, occupation and firm fixed effects. Sample size: 9466620 observations.

particularly pronounced effects in the second quintile for female-targeted advertisements (7.5 per cent) and male-targeted advertisements (5.3 per cent). The penalties generally persist across higher intensity levels, reaching 5.4 per cent for female-targeted and 5.1 per cent for male-targeted advertisements at the highest intensity.

Extraversion shows different patterns across gender targeting. In male-targeted advertisements, penalties are significant across all intensity levels, ranging from 1.8 per cent (Q1) to 4.9 per cent (Q5). Female-targeted advertisements show no significant penalty at the lowest intensity but larger penalties at higher intensities, reaching 6.0 per cent at the highest level.

Agreeableness exhibits similar patterns across gender categories, with both showing non-significant effects at low intensities transitioning to significant penalties of approximately 4.0 per cent at higher intensities (Q3-Q4). This similarity suggests that, regardless of gender targeting, employers use high-intensity Agreeableness language for similar strategic purposes.

Openness maintains its exceptional status, showing no significant wage effects across any intensity level for either gender category, with coefficients remaining small and predominantly positive.

8.2.4 Assessment of the Demographic Targeting Hypothesis

The demographic targeting hypothesis is not supported by the evidence. Three key findings lead to this conclusion.

First, employers do use personality trait language differently when targeting women versus men. Female-targeted advertisements are more likely to mention Openness, Extraversion, and Agreeableness. However, this differential usage does not translate into the wage patterns predicted by demographic targeting theory. If trait language served primarily to attract demographic groups willing to accept lower wages, we would expect larger penalties in female-targeted advertisements. Instead, most traits (Conscientiousness, Extraversion, Agreeableness) have larger penalties in male-targeted advertisements. Only Emotional Stability shows the opposite pattern.

Second, the persistence of significant penalties in both female-targeted and male-targeted advertisements indicates that the pooled wage penalty does not result from compositional effects driven by gender targeting. When estimated separately by gender category, trait penalties remain substantial and significant. This suggests that trait-wage relationships operate within gender categories rather than being explained by the differential targeting of demographic groups that receive different wage offers.

Third, both female-targeted and male-targeted advertisements show similar patterns of increasing penalties with trait intensity. Agreeableness shows identical penalty structures across gender categories. Both show non-significant effects at low intensities that transition to 4.0 per cent penalties at higher intensities. This demonstrates that employers use trait intensity strategically regardless of gender targeting. This further undermines the demographic targeting explanation.

These findings rule out demographic targeting as the primary driver of the wage penalty paradox. The evidence instead supports the mechanisms documented in previous sections. Trait penalties reflect compensating differentials when combined with appropriate amenity bundles (Section 6). They vary systematically with linguistic intensity (Section 7). They correlate with occupation-level characteristics (Section 5). While employers do strategically differentiate trait language by gender targeting, this differentiation does not explain why trait mentions are associated with lower wages.

9 Conclusion

This research has examined the paradoxical relationship between personality trait mentions in job advertisements and posted wages in the UK labour market. While individual-level studies consistently find that traits like Conscientiousness and Emotional Stability increase personal wages, I documented a significant negative relationship between personality trait mentions in job advertisements and posted wages. Using data from 11.7 million online job advertisements posted between 2017 and 2022, I tested four potential explanations for this paradox: correlation with other low-pay characteristics, compensating differentials through amenity provision, linguistic intensity effects, and demographic targeting strategies.

The baseline analysis confirmed the wage penalty paradox. Advertisements mentioning personality traits offered 4.6 per cent lower wages on average, even after controlling for job, occupation, and firm characteristics. Most traits showed significant penalties: Extraversion (3.2 per cent), Emotional Stability (3.1 per cent), Conscientiousness (2.0 per cent), and Agreeableness (1.4 per cent). Only Openness showed a small positive effect (0.7 per cent), though this was only marginally significant.

Correlation with Low-Pay Characteristics

The analysis of correlation with low-pay characteristics provided partial but incomplete explanations for the wage penalty paradox. Occupation fixed effects produced the largest reduction in trait penalties, demonstrating that personality traits appear disproportionately in lower-paying occupations. However, significant penalties persisted even after controlling for occupation, job, and firm characteristics.

The occupation-level analysis revealed that trait penalties are smaller in occupations with higher proportions of women and migrants but larger in occupations with more degree holders and higher mean wages. This pattern—where traits face greater penalties in high-skill, high-wage contexts—contradicts simple productivity-based explanations and suggests strategic use of trait language in different occupational contexts.

Job-level interactions showed that personality traits receive more favourable treatment in entry-level positions and when explicit educational requirements are mentioned. However, traits face amplified penalties when combined with hard skills requirements and temporary contracts. Firm-level analysis provided limited support, with only public listing status showing meaningful effects that amplified rather than explained trait penalties.

Compensating Differentials Through Amenity Provision

The compensating differentials analysis revealed context-dependent support for this mechanism. Simple additive controls for amenities had minimal impact on trait penalties, with the R^2 increasing only marginally from 0.598 to 0.601. This stability suggests that baseline trait penalties do not operate through systematic correlation with amenity provision.

However, the interaction analysis provided evidence for strategic compensating differentials mechanisms. When personality trait language was combined with traditional benefits (health insurance, retirement contributions), workplace culture signals (work environment, human capital development), and quality workplace amenities, trait penalties were reduced or eliminated. These patterns support compensating differentials theory: employers can offer lower wages when trait requirements are bundled with attractive non-monetary compensation and positive organisational culture signals.

Conversely, some amenity combinations amplified trait penalties. Remote work created uniformly larger penalties for most traits, while variable compensation arrangements showed mixed patterns. This heterogeneity suggests that employers strategically deploy compensating differential mechanisms for certain trait-amenity combinations while using personality trait language for different purposes in other contexts.

Linguistic Intensity Effects

The linguistic intensity analysis provided the strongest evidence for strategic employer behaviour and helped explain the wage penalty paradox. The relationship between trait emphasis and wages was highly non-linear: low-intensity mentions often carried minimal or no wage penalties, while high-intensity mentions consistently showed larger penalties.

This intensity gradient suggests that employers use trait mentions strategically depending on the emphasis level. Moderate use likely signals job requirements without triggering wage penalties, while excessive emphasis serves different strategic purposes—potentially compensating differential signals or demographic targeting mechanisms—that justify lower wage offers.

Firm interactions revealed that established, larger firms showed more favourable wage effects for low-intensity trait mentions, suggesting these organisations use trait language primarily to signal requirements. Conversely, publicly quoted firms imposed large penalties for high-intensity trait mentions, indicating that market pressures may lead these firms to use trait language for signalling purposes rather than identifying productive characteristics.

Demographic Targeting

The demographic targeting analysis provided limited support for this explanation. While employers do use personality trait language differently when targeting women versus men—with female-targeted advertisements more likely to mention Openness, Extraversion, and Agreeableness—this differential usage did not translate into the wage patterns predicted by demographic targeting theory.

Most traits showed larger penalties in male-targeted advertisements, with only Emotional Stability showing a larger penalty in female-targeted contexts. The intensity analysis revealed that both female-targeted and male-targeted advertisements showed similar patterns of increasing penalties with trait intensity, suggesting that trait emphasis effects operate through mechanisms beyond demographic targeting.

Synthesis and Implications

These findings collectively show that the personality trait wage penalty paradox reflects multiple overlapping employer strategies rather than a single mechanism. Three key insights emerge about how employers use personality trait language in job advertisements.

First, the relationship between trait mentions and wages is context-dependent. The same trait can signal requirements, serve compensating differential purposes, or function as demographic targeting depending on how it is presented, what amenities are offered, and which type of organisation is hiring. This context-dependency explains why aggregate analyses show wage penalties despite productivity benefits in specific circumstances.

Second, linguistic presentation matters as much as content. The intensity of trait emphasis, its placement within advertisements, and its combination with other requirements influence wage effects. Low-intensity trait mentions often signal requirements without wage penalties, while high-intensity mentions serve strategic signalling purposes that justify lower compensation.

Third, organisational characteristics shape how personality trait language functions in wage determination. Established, larger firms use trait language differently than smaller, publicly traded, or newer organisations. This variation suggests that employer strategy, rather than uniform trait devaluation, drives the observed wage penalties.

Limitations and Future Research

Several limitations affect the interpretation of these results.

The analysis examines posted wages rather than realised compensation or wage growth trajectories. A competing explanation for the wage penalty paradox is that personality traits are non-verifiable characteristics where employers offer lower starting wages but provide steeper wage growth once these attributes become apparent through workplace performance (Altonji and Pierret, 2001; Farber and Gibbons, 1996; Lange, 2007). While cross-sectional vacancy data cannot directly test this mechanism, several patterns in the results are inconsistent with employer learning being the primary driver of the observed penalties.

First, the employer learning literature predicts that non-verifiable characteristics should be particularly important in entry-level positions where workers lack demonstrated track records. Employers facing uncertainty about worker traits should offer lower starting wages in these contexts. However, the job-level interaction analysis (Section 5) shows the opposite pattern. Conscientiousness penalties are significantly smaller in low-seniority positions (interaction term +0.021, reducing the penalty from 3.2 per cent to 1.1 per cent) and disappear entirely when no experience is required (interaction term +0.085). If employer learning drove the wage penalty, we would observe larger penalties in entry-level positions where trait verification is most difficult. The data show smaller or non-existent penalties instead. Second, trait penalties are largest in high-skill occupations where workers typically have longer employment histories. Third, the systematic interaction patterns with amenities (Section 6) and the non-linear intensity effects (Section 7) suggest employers use trait language strategically for multiple purposes beyond addressing verification problems. Low-intensity mentions carry minimal penalties while high-intensity mentions show substantial penalties, indicating deliberate emphasis choices rather than uniform responses to uncertainty.

While employer learning dynamics may contribute to some wage patterns, these inconsistencies indicate it is not the dominant explanation for why advertisements mentioning personality traits are associated with lower posted wages. Future research combining vacancy data with longitudinal employment records could directly examine how initial wage offers relate to subsequent earnings trajectories for workers in trait-emphasising positions. The identification approach treats advertisements not mentioning personality traits as not requiring them. Jobs might implicitly require traits without explicit mention. While extensive controls help address this concern, unmeasured trait requirements could bias results. However, this limitation affects all research examining job advertisement text and does not explain why jobs that explicitly mention traits offer lower wages compared to those that do not.

Data limitations prevented testing whether personality trait requirements correlate with other non-wage job characteristics not captured in advertisement text, or whether employers offering trait-demanding positions provide different benefits not advertised publicly. Future research could address these gaps through employer surveys, applicant-level data, or experimental approaches examining how job seekers respond to different intensities of personality-related language.

Contribution

This research contributes to understanding how personality traits are valued in contemporary labour markets by revealing the strategic complexity underlying employer communication. Rather than simply reflecting productivity differences or uniform trait devaluation, personality trait language in job advertisements serves multiple strategic purposes that vary by context, intensity, and organisational characteristics.

The findings have practical implications for both job seekers and employers. Job seekers should interpret personality trait language in advertisements as potentially signalling organisational culture, compensation structure, and strategic positioning rather than simply describing required characteristics. Employers should consider how trait-related language affects both the quantity and quality of applicants, recognising that high-intensity trait emphasis may signal compensation limitations that could deter qualified candidates.

More broadly, this research shows that job advertisements function as strategic communications rather than straightforward requirement descriptions. As labour markets evolve with technological advancement and changing work arrangements, understanding these strategic dimensions becomes increasingly important for explaining how employers attract workers and how job seekers evaluate opportunities.

By exposing the relationship between personality trait mentions and wages in job advertisements, this research provides new insights into employer communication strategies and their effects on labour market outcomes. The personality trait wage penalty paradox reflects the ways employers use language to signal organisational characteristics, manage compensation expectations, and attract their desired workforce rather than simple productivity-based valuations.

References

Abraham, L., Hallermeier, J., and Stein, A. (2024). Words matter: Experimental evidence from job applications. *Journal of Economic Behavior & Organization*, 225:348–391.

Acemoglu, D., Autor, D., Hazell, J., and Restrepo, P. (2022). Artificial Intelligence and Jobs: Evidence from Online Vacancies. *Journal of Labor Economics*, 40(S1):293–340.

Aigner, D. J. and Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *ILR Review*, 30(2):175–187.

Akerlof, G. A. and Kranton, R. E. (2005). Identity and the economics of organizations. *Journal of Economic Perspectives*, 19(1):9–32.

Alderotti, G., Rapallini, C., and Traverso, S. (2023). The Big Five personality traits and earnings: A meta-analysis. *Journal of Economic Psychology*, 94:102570.

Alekseeva, L., Azar, J., Gine, M., Samila, S., and Taska, B. (2021). The demand for ai skills in the labor market. *Labour Economics*, 71:102002.

Almlund, M., Duckworth, A. L., Heckman, J., and Kautz, T. (2011). Personality psychology and economics. In *Handbook of the Economics of Education*, volume 4, pages 1–181. Elsevier.

Altonji, J. G. and Pierret, C. R. (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, 116(1):313–350.

Arrow, K. (1971). The theory of discrimination. *Labor Economics*.

Banfi, S. and Villena-Roldan, B. (2019). Do high-wage jobs attract more applicants? directed search evidence from the online labor market. *Journal of Labor Economics*, 37(3):715–746.

Barrick, M. R. and Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel Psychology*, 44(1):1–26.

Batra, H., Michaud, A. M., and Mongey, S. (2023). Online job posts contain very little wage information. Working paper 31984, National Bureau of Economic Research.

Borghans, L., Duckworth, A. L., Heckman, J. J., and Ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of Human Resources*, 43(4):972–1059.

Boschetti Adamczyk, W., Delaporte, I., and Escudero, V. (2025). Measuring quality of employment in emerging economies: A methodology for assessing job amenities using big data. Research brief, International Labour Organization.

Bowles, S., Gintis, H., and Osborne, M. (2001). Incentive-enhancing preferences: Personality, behavior, and earnings. *American Economic Review*, 91(2):155–158.

Braakmann, N. (2009). The role of psychological traits for the gender gap in full-time employment and wages: evidence from germany. (162). vailable at SSRN: https://ssrn.com/abstract=3169747 or http://dx.doi.org/10.2139/ssrn.3169747.

Breaugh, J. A. (2013). Employee recruitment. *Annual Review of Psychology*, 64:389–416.

Brenčič, V. (2012). Wage posting: evidence from job ads. *Canadian Journal of Eco-*

nomics/Revue canadienne d'économique, 45(4):1529–1559.

Brenčič, V. and McGee, A. (2023a). Demand for personality traits, tasks, and sorting. Discussion Paper 16576, Institute of Labor Economics (IZA).

Brenčič, V. and McGee, A. (2023b). Employers' demand for personality traits. Discussion Paper 16083, Institute of Labor Economics (IZA).

Burbano, V. C. (2016). Social responsibility messages and worker wage requirements: Field experimental evidence from online labor marketplaces. *Organization Science*, 27(4):1010–1028.

Cable, D. M. and Judge, T. A. (1996). Person–organization fit, job choice decisions, and organizational entry. *Organizational Behavior and Human Decision Processes*, 67(3):294–311.

Cassar, L. and Meier, S. (2018). Nonmonetary incentives and the implications of work as a source of meaning. *Journal of Economic Perspectives*, 32(3):215–238.

Chapman, D. S., Uggerslev, K. L., Carroll, S. A., Piasentin, K. A., and Jones, D. A. (2005). Applicant attraction to organizations and job choice: a meta-analytic review of the correlates of recruiting outcomes. *Journal of Applied Psychology*, 90(5):928.

Chaturvedi, S., Mahajan, K., and Siddique, Z. (2021). Words matter: Gender, jobs and applicant behavior. Discussion Paper 14497, Institute of Labor Economics (IZA).

Cobb-Clark, D. A. and Schurer, S. (2012). The stability of big-five personality traits. *Economics Letters*, 115(1):11–15.

Cobb-Clark, D. A. and Tan, M. (2011). Noncognitive skills, occupational attainment, and relative wages. *Labour Economics*, 18(1):1–13.

Collischon, M. (2020). The returns to personality traits across the wage distribution. *Labour*, 34(1):48–79.

Costa Jr, P. T. and McCrae, R. R. (1992). The five-factor model of personality and its relevance to personality disorders. *Journal of Personality Disorders*, 6(4):343–359.

Cubel, M., Nuevo-Chiquero, A., Sanchez-Pages, S., and Vidal-Fernandez, M. (2016). Do personality traits affect productivity? evidence from the laboratory. *The Economic Journal*, 126(592):654–681.

Delfgaauw, J. and Dur, R. (2007). Signaling and screening of workers' motivation. *Journal of Economic Behavior & Organization*, 62(4):605–624.

Deming (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, 132(4):1593–1640.

Deming and Kahn (2018). Skill requirements across firms and labor markets: Evidence from job postings for professionals. *Journal of Labor Economics*, 36(S1):S337–S369.

Denissen, J. J., Bleidorn, W., Hennecke, M., Luhmann, M., Orth, U., Specht, J., and Zimmermann, J. (2018). Uncovering the power of personality to shape income. *Psychological Science*, 29(1):3–13.

Draca, M., Duchini, E., Rathelot, R., Turrell, A., and Vattuone, G. (2022). Revolution in progress? the rise of remote work in the uk. Working Paper 616, Centre for Competitive Advantage in the Global Economy (CAGE).

Escudero, V., Liepmann, H., and Vergara, D. (2024). Directed Search, Wages, and Non-Wage Amenities: evidence from an Online Job Board. Discussion Papers 17211, Institute of Labor Economics (IZA).

Farber, H. S. and Gibbons, R. (1996). Learning and wage dynamics. *The Quarterly Journal of Economics*, 111(4):1007–1047.

Flinn, C. J., Todd, P. E., and Zhang, W. (2025). Labor market returns to personality: A job search approach to understanding gender gaps. *Journal of Political Economy*, 133(4):1169–1234.

Flory, J. A., Leibbrandt, A., and List, J. A. (2015). Do competitive workplaces deter female workers? a large-scale natural field experiment on job entry decisions. *The Review of Economic Studies*, 82(1):122–155.

Gaucher, D., Friesen, J., and Kay, A. C. (2011). Evidence that gendered wording in job advertisements exists and sustains gender inequality. *Journal of Personality and Social Psychology*, 101(1):109–128.

Gensowski, M. (2018). Personality, IQ, and lifetime earnings. *Labour Economics*, 51:170–183.

Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. *Review of Personality and Social Psychology*, 2(1):141–165.

Goldberg, L. R. (1990). An alternative "description of personality": the big-five factor structure. *Journal of Personality and Social Psychology*, 59(6):1216—1229.

Goldberg, L. R. (1992). The development of markers for the big-five factor structure. *Psychological Assessment*, 4(1):26—-42.

Grugulis, I. and Vincent, S. (2009). Whose skill is it anyway?: "soft" skills and polarization. *Work, Employment and Society*, 23(4):597–615.

Handy, F. and Katz, E. (1998). The wage differential between nonprofit institutions and corporations: Getting more by paying less? *Journal of Comparative Economics*, 26(2):246–261.

Heckman, J. J. and Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, 19(4):451–464.

Heineck, G. (2011). Does it pay to be nice? personality and earnings in the United Kingdom. *ILR Review*, 64(5):1020–1038.

Hershbein, B. and Kahn, L. B. (2018). Do recessions accelerate routine-biased technological change? evidence from vacancy postings. *American Economic Review*, 108(7):1737–1772.

John (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In *Handbook of personality: Theory and research*, pages 66—-100. Guilford Press.

John and Thomsen (2014). Heterogeneous returns to personality: the role of occupational choice. *Empirical Economics*, 47(2):553–592.

Judge, T. A. and Cable, D. M. (1997). Applicant personality, organizational culture, and organization attraction. *Personnel Psychology*, 50(2):359–394.

Khaouja, I., Kassou, I., and Ghogho, M. (2021). A Survey on Skill Identification From Online Job Ads. *IEEE Access*, 9:118134–118153.

Koçak, A. and Derous, E. (2025). Women's intention to apply to top-executive positions: The role of gender meta-stereotypes in job ads. *Sex Roles*, 91(2):4.

Koçak, A., Derous, E., Born, M. P., and Duyck, W. (2023). What (not) to add in your ad: When job ads discourage older or younger job seekers to apply. *International Journal of Selection and Assessment*, 31(1):92–104.

Koch, T. and Zerback, T. (2013). Helpful or harmful? How frequent repetition affects perceived statement credibility. *Journal of Communication*, 63(6):993–1010.

Kristof, A. L. (1996). Person-organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel Psychology*, 49(1):1–49.

Kristof-Brown, A. L., Zimmerman, R. D., and Johnson, E. C. (2005). Consequences of individuals' fit at work: A meta-analysis of person–job, person–organization, person–group, and person–supervisor fit. *Personnel Psychology*, 58(2):281–342.

Krueger, A. B. and Schkade, D. (2008). Sorting in the labor market do gregarious workers flock to interactive jobs? *Journal of Human Resources*, 43(4):859–883.

Kuhn, P. and Shen, K. (2023). What happens when employers can no longer discriminate in job ads? *American Economic Review*, 113(4):1013–1048.

Kuhn, P., Shen, K., and Zhang, S. (2020). Gender-targeted job ads in the recruitment process: Facts from a chinese job board. *Journal of Development Economics*, 147:102531.

Lambrecht, K. (1996). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*, volume 71. Cambridge University Press.

Lange, F. (2007). The speed of employer learning. *Journal of Labor Economics*, 25(1):1–35.

Lyu, W. and Liu, J. (2021). Soft skills, hard skills: What matters most? Evidence from job postings. *Applied Energy*, 300:117307.

Maestas, N., Mullen, K. J., Powell, D., Von Wachter, T., and Wenger, J. B. (2023). The value of working conditions in the United States and implications for the structure of wages. *American Economic Review*, 113(7):2007–2047.

Mahjoub, A. and Kruyen, P. M. (2021). Efficient recruitment with effective job advertisement: an exploratory literature review and research agenda. *International Journal of Organization Theory & Behavior*, 24(2):107–125.

Malhotra, N. K. (1982). Information load and consumer decision making. *Journal of Consumer Research*, 8(4):419–430.

Marinescu, I. and Wolthoff, R. (2020). Opening the black box of the matching function: The power of words. *Journal of Labor Economics*, 38(2):535–568.

Matteson, M. L., Anderson, L., and Boyden, C. (2016). "soft skills": A phrase in search of meaning. *Libraries and the Academy*, 16(1):71–88.

McCrae, R. R. and John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60(2):175–215.

McGee, A. and McGee, P. (2020). Whoever you want me to be: personality and incentives. Discussion Paper 13809, Institute of Labor Economics (IZA).

Mueller, G. and Plug, E. (2006). Estimating the effect of personality on male and female earnings. *ILR Review*, 60(1):3–22.

Nandi, A. and Nicoletti, C. (2014). Explaining personality pay gaps in the uk. *Applied Economics*, 46(26):3131–3150.

Nordman, C. J., Sarr, L. R., and Sharma, S. (2019). Skills, personality traits, and gender wage gaps: evidence from bangladesh. *Oxford Economic Papers*, 71(3):687–708.

Nyhus, E. K. and Pons, E. (2005). The effects of personality on earnings. *Journal of Economic Psychology*, 26(3):363–384.

Nyhus, E. K. and Pons, E. (2012). Personality and the gender wage gap. *Applied Economics*, 44(1):105–118.

Oldford, E. and Fiset, J. (2021). Decoding bias: Gendered language in finance internship job postings. *Journal of Behavioral and Experimental Finance*, 31:100544.

Ployhart, R. E., Weekley, J. A., and Baughman, K. (2006). The structure and function of human capital emergence: A multilevel examination of the attraction-selection-attrition model. *Academy of Management Journal*, 49(4):661–677.

Risse, L., Farrell, L., and Fry, T. R. L. (2018). Personality and pay: do gender gaps in confidence explain gender gaps in wages? *Oxford Economic Papers*, 70(4):919–949.

Roberson, Q. M., Collins, C. J., and Oreg, S. (2005). The effects of recruitment message specificity on applicant attraction to organizations. *Journal of Business and Psychology*, pages 319–339.

Rohrbach-Schmidt, D., Wehner, C., Krueger, S., and Ebner, C. (2023). Wage returns to job tasks and personality traits in germany. *International Journal of Manpower*, 44(9):55–71.

Rosen, S. (1986). The theory of equalizing differences. In *Handbook of Labour Economics*, volume 1, pages 641–692. Elsevier.

Rosenfeld, L., Morville, P., and Arango, J. (2015). *Information architecture: for the web and beyond*. O'Reilly Media, Inc.

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

Salton, G. and Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5):513–523.

Saucier, G. (1994). Mini-markers: A brief version of goldberg's unipolar big-five markers. *Journal of Personality Assessment*, 63(3):506–516.

Sockin, J. (2022). Show me the amenity: Are higher-paying firms better all around? Working Paper 9842, CESifo.

Sparck Jones, K. (1972). A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, 28(1):11–21.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3):355–374.

Stevens, C. D. and Szmerekovsky, J. G. (2010). Attraction to employment advertisements: Advertisement wording and personality characteristics. *Journal of Managerial Issues*, pages 107–126.

Turrell, A., Speigner, B. J., Djumalieva, J., Copple, D., and Thurgood, J. (2019). Transforming naturally occurring text data into economic statistics: The case of online job vacancy postings. Working paper 25837, National Bureau of Economic Research.

Van Hoye, G. and Turban, D. B. (2015). Applicant–employee fit in personality: Testing predictions from similarity-attraction theory and trait activation theory. *International Journal of Selection and Assessment*, 23(3):210–223.

Wei, Y.-C., Chang, C.-C., Lin, L.-Y., and Liang, S.-C. (2016). A fit perspective approach in linking corporate image and intention-to-apply. *Journal of Business Research*, 69(6):2220–2225.

Weinberger, C. J. (2014). The increasing complementarity between cognitive and social skills. *Review of Economics and Statistics*, 96(5):849–861.

Wille, L. and Derous, E. (2018). When job ads turn you down: how requirements in job ads may stop instead of attract highly qualified women. *Sex Roles*, 79:464–475.

Woessmann, L. (2024). Skills and earnings: A multidimensional perspective on human capital. *Annual Review of Economics*, 17.

Woodcock, S. D. (2008). Wage differentials in the presence of unobserved worker, firm, and match heterogeneity. *Labour Economics*, 15(4):771–793.

10 Appendices

Appendix A - Figures and Tables

Table A.1: Personality Trait Coefficients on Log Posted Wage, by 1-digit SOC Occupation

SOC	Consc.	Open.	Extra.	Agree.	Emot.	N	R^2
1	-0.030***	0.017***	-0.034***	-0.010	-0.039***	600,960	0.584
	(0.011)	(0.006)	(0.007)	(0.012)	(0.011)		
2	-0.019***	-0.008	-0.032***	0.020	-0.035***	2,187,332	0.410
	(0.006)	(0.005)	(0.008)	(0.020)	(0.010)		
3	-0.030***	0.014***	-0.049***	-0.033***	-0.056***	1,875,955	0.464
	(0.005)	(0.005)	(0.009)	(0.010)	(0.007)		
4	-0.024***	0.016***	-0.057***	-0.047***	-0.011*	617,840	0.491
	(0.004)	(0.005)	(0.006)	(0.009)	(0.006)		
5	-0.019*	0.021**	-0.010	-0.005	0.008	732,619	0.543
	(0.012)	(0.009)	(0.008)	(0.013)	(0.025)		
6	-0.001	0.018***	-0.013***	-0.014	-0.035***	1,577,798	0.413
	(0.017)	(0.006)	(0.004)	(0.010)	(0.010)		
7	-0.009	0.022**	-0.036***	-0.043***	-0.020**	903,832	0.515
	(0.010)	(0.010)	(0.010)	(0.011)	(0.009)		
8	0.001	-0.031**	-0.039**	0.081	-0.018	485,450	0.573
	(0.012)	(0.015)	(0.018)	(0.076)	(0.030)		
9	-0.015	-0.001	-0.011	0.055*	0.018	855,124	0.537
	(0.013)	(0.008)	(0.014)	(0.033)	(0.030)		

Notes: Each row presents coefficients and standard errors (in parentheses) from separate regressions for each 1-digit SOC occupation. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. The dependent variable is the natural logarithm of posted salary. SOC codes: 1 = Managers, directors and senior officials; 2 = Professional occupations; 3 = Associate professional and technical occupations; 4 = Administrative and secretarial occupations; 5 = Skilled trades occupations; 6 = Caring, leisure and other service occupations; 7 = Sales and customer service occupations; 8 = Process, plant and machine operatives; 9 = Elementary occupations. Standard errors are clustered at the firm level. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. All regressions include controls for soft skills, hard skills, remote work, advertisement length, contract type, work schedule, seniority level, education, experience, amenities availability, and negotiable salary. Fixed effects for location (NUTS2 region), month-year, and firm are included.

Table A.2: **Personality Traits coefficients on Log posted wage, by 2-digit occupation**

SOC	Cons.	Open.	Extra.	Agree.	Emot.	N	R^2
11	-0.037***	0.016***	-0.034***	0.004	-0.019*	390,200	0.614
	(0.013)	(0.005)	(0.008)	(0.015)	(0.010)		
12	-0.005	0.015	-0.025***	-0.027*	-0.055***	208,565	0.553
	(0.009)	(0.014)	(0.009)	(0.016)	(0.019)		
21	-0.023***	-0.011**	-0.049***	-0.002	-0.018**	864,886	0.389
	(0.005)	(0.005)	(0.005)	(0.011)	(0.007)		
22	0.007	-0.019	0.031	0.058*	-0.063***	458,612	0.392
	(0.015)	(0.015)	(0.022)	(0.035)	(0.022)		
23	-0.017**	0.024***	0.001	0.003	-0.041***	170,876	0.440
	(0.007)	(0.008)	(0.019)	(0.012)	(0.015)		
24	-0.027***	-0.005	-0.055***	-0.020***	-0.028***	687,599	0.448
	(0.004)	(0.005)	(0.005)	(0.007)	(0.006)		
31	-0.020***	0.013*	-0.061***	-0.049***	-0.016***	299,461	0.506
	(0.005)	(0.007)	(0.006)	(0.009)	(0.006)		
32	0.016	0.010	-0.051***	-0.037***	-0.029***	150,153	0.477
	(0.011)	(0.013)	(0.011)	(0.006)	(0.008)		
33	-0.042**	0.015	-0.032	0.015	-0.027	9,259	0.699
	(0.021)	(0.020)	(0.023)	(0.049)	(0.029)		
34	-0.022	0.009	-0.023	-0.010	-0.061*	103,482	0.644
	(0.029)	(0.031)	(0.044)	(0.070)	(0.031)		
35	-0.035***	0.010*	-0.051***	-0.019	-0.060***	1,305,530	0.463
	(0.006)	(0.006)	(0.011)	(0.013)	(0.009)		
41	-0.024***	0.015***	-0.056***	-0.043***	-0.013*	539,208	0.475
	(0.004)	(0.005)	(0.007)	(0.011)	(0.007)		
42	-0.018***	0.006	-0.045***	-0.039***	0.005	76,428	0.671
	(0.007)	(0.008)	(0.008)	(0.011)	(0.010)		
51	-0.011	0.048***	-0.043***	-0.016	-0.006	9,189	0.584
	(0.018)	(0.015)	(0.015)	(0.025)	(0.023)		
52	-0.018***	0.011	-0.018**	-0.015	0.011	351,376	0.495
	(0.006)	(0.008)	(0.009)	(0.016)	(0.015)		
53	0.024	0.080*	-0.008	0.013	-0.094	98,435	0.598
	(0.033)	(0.044)	(0.026)	(0.028)	(0.069)		
54	-0.035*	0.033**	0.012	0.001	0.012	271,633	0.378
	(0.019)	(0.015)	(0.013)	(0.019)	(0.036)		
61	0.017**	0.014**	-0.011***	-0.009	-0.028***	1,446,703	0.410
	(0.007)	(0.007)	(0.004)	(0.009)	(0.008)		
62	-0.170*	0.038**	-0.038*	-0.074**	-0.119*	130,177	0.538
	(0.102)	(0.018)	(0.022)	(0.035)	(0.061)		
71	-0.013	0.008	-0.036**	0.023	0.003	259,496	0.619
	(0.018)	(0.015)	(0.014)	(0.020)	(0.023)		
72	-0.009	0.024**	-0.035***	-0.067***	-0.024***	642,649	0.512
	(0.009)	(0.011)	(0.010)	(0.011)	(0.009)		

Continued on next page

Table A.2 – continued from previous page

SOC	Cons.	Open.	Extra.	Agree.	Emot.	N	R^2
81	0.002	-0.005	-0.017*	0.014	-0.005	72,579	0.652
	(0.008)	(0.012)	(0.009)	(0.020)	(0.016)		
82	0.004	-0.035**	-0.043**	0.085	-0.023	412,142	0.577
	(0.013)	(0.016)	(0.022)	(0.077)	(0.037)		
91	-0.021**	-0.012	0.000	0.017	-0.042**	47,292	0.483
	(0.010)	(0.015)	(0.013)	(0.039)	(0.018)		
92	-0.015	-0.000	-0.011	0.056*	0.019	807,273	0.543
	(0.013)	(0.008)	(0.014)	(0.033)	(0.030)		

Notes: Each row presents coefficients and standard errors (in parentheses) from separate regressions for each 2-digit SOC occupation. SOC codes follow the UK Standard Occupational Classification 2010. Full descriptions of occupation categories are available from the Office for National Statistics (www.ons.gov.uk/methodology/classificationsandstandards). "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. The dependent variable is the natural logarithm of posted salary. All regressions include controls for soft skills, hard skills, remote work, advertisement length, contract type, work schedule, seniority level, education, experience, amenities availability, and negotiable salary. Fixed effects for location (NUTS2 region), month-year, and firm are included. Standard errors are clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: Personality Traits coefficients on Log posted wage, by 3-digit occupation

SOC	Cons.	Open.	Extra.	Agree.	Emot.	N	R^2
111	0.012	0.006	-0.040*	-0.017	0.038	6,293	0.837
	(0.025)	(0.026)	(0.021)	(0.039)	(0.045)		
112	-0.024***	-0.014	-0.033***	0.006	-0.037***	96,013	0.668
	(0.007)	(0.010)	(0.009)	(0.015)	(0.013)		
113	-0.019***	0.017***	-0.045***	0.001	-0.023**	117,821	0.472
	(0.006)	(0.006)	(0.006)	(0.011)	(0.011)		
115	0.003	0.049**	-0.021	0.039	-0.076**	6,636	0.697
	(0.018)	(0.023)	(0.018)	(0.039)	(0.031)		
116	-0.058**	0.010	0.018	-0.028	0.016	106,426	0.638
	(0.029)	(0.016)	(0.021)	(0.026)	(0.019)		
117	0.007	0.013	-0.059*	-0.190***	0.009	1,558	0.850
	(0.045)	(0.046)	(0.033)	(0.060)	(0.096)		
118	-0.022	0.031**	-0.027*	0.076*	-0.082***	24,316	0.481
	(0.013)	(0.015)	(0.014)	(0.043)	(0.027)		
119	-0.024*	-0.021	-0.002	0.013	0.028	26,336	0.729
	(0.015)	(0.017)	(0.031)	(0.027)	(0.031)		
121	0.005	-0.241***	0.126	-0.199***	-0.035	688	0.816
	(0.066)	(0.051)	(0.096)	(0.055)	(0.113)		
122	0.012	0.043**	0.002	-0.013	-0.083***	88,550	0.557
	(0.013)	(0.018)	(0.014)	(0.023)	(0.025)		

Continued on next page

Table A.3 – continued from previous page

SOC	Cons.	Open.	Extra.	Agree.	Emot.	N	R^2
124	-0.016	-0.049*	-0.040**	-0.029	-0.012	56,862	0.503
	(0.020)	(0.028)	(0.016)	(0.026)	(0.035)		
125	-0.016	0.020	-0.037***	-0.010	-0.060***	61,343	0.627
	(0.010)	(0.015)	(0.010)	(0.024)	(0.013)		
211	-0.025*	0.017*	-0.050***	0.012	-0.008	24,544	0.570
	(0.013)	(0.010)	(0.013)	(0.022)	(0.022)		
212	-0.006	0.009	-0.041***	0.005	-0.023	155,044	0.416
	(0.008)	(0.008)	(0.009)	(0.020)	(0.014)		
213	-0.027***	-0.018***	-0.049***	-0.006	-0.019**	657,904	0.379
	(0.006)	(0.006)	(0.005)	(0.011)	(0.008)		
214	-0.006	0.021	-0.057***	0.071**	0.115***	11,459	0.640
	(0.019)	(0.018)	(0.019)	(0.030)	(0.035)		
215	-0.046***	0.003	-0.066***	0.018	-0.072*	13,317	0.736
	(0.014)	(0.015)	(0.017)	(0.027)	(0.041)		
221	0.012	-0.016	-0.049	0.010	-0.153***	50,190	0.393
	(0.015)	(0.034)	(0.037)	(0.041)	(0.037)		
222	-0.012	-0.022	0.013	0.010	-0.024	33,197	0.420
	(0.017)	(0.014)	(0.012)	(0.018)	(0.019)		
223	0.013	0.000	0.049*	0.066*	-0.045**	374,840	0.470
	(0.017)	(0.013)	(0.025)	(0.038)	(0.018)		
231	-0.017**	0.024***	0.001	0.003	-0.041***	170,876	0.440
	(0.007)	(0.008)	(0.019)	(0.012)	(0.015)		
241	-0.025***	0.033**	-0.035***	-0.029***	0.007	39,213	0.524
	(0.007)	(0.014)	(0.007)	(0.009)	(0.014)		
242	-0.031***	0.003	-0.054***	-0.014*	-0.027***	422,880	0.430
	(0.005)	(0.006)	(0.006)	(0.007)	(0.008)		
243	-0.022**	-0.007	-0.047***	-0.002	-0.018	74,139	0.479
	(0.009)	(0.011)	(0.009)	(0.019)	(0.015)		
244	0.007	0.017	-0.025**	-0.060***	-0.034*	34,789	0.716
	(0.012)	(0.012)	(0.012)	(0.023)	(0.018)		
245	-0.007	0.052	-0.074*	-0.060	0.076	871	0.830
	(0.037)	(0.038)	(0.039)	(0.058)	(0.054)		
246	-0.027***	-0.002	-0.043***	0.005	-0.031***	82,833	0.555
	(0.006)	(0.008)	(0.009)	(0.023)	(0.011)		
247	-0.015	0.001	-0.040***	0.012	-0.007	27,944	0.563
	(0.010)	(0.010)	(0.009)	(0.018)	(0.011)		
311	-0.003	-0.005	-0.052***	-0.010	0.015	112,425	0.544
	(0.008)	(0.010)	(0.010)	(0.015)	(0.010)		
312	-0.019	-0.002	-0.037**	-0.016	0.013	16,645	0.597
	(0.017)	(0.013)	(0.016)	(0.048)	(0.021)		
313	-0.031***	0.022***	-0.066***	-0.060***	-0.030***	168,785	0.531
	(0.005)	(0.008)	(0.007)	(0.011)	(0.006)		
321	0.011	0.043	-0.025	-0.049	-0.072	9,775	0.518

Continued on next page

Table A.3 – continued from previous page

SOC	Cons.	Open.	Extra.	Agree.	Emot.	N	R^2
	(0.034)	(0.062)	(0.034)	(0.039)	(0.050)		
323	0.018	0.008	-0.050***	-0.035***	-0.023***	140,227	0.470
	(0.011)	(0.013)	(0.011)	(0.006)	(0.008)		
331	-0.042**	0.015	-0.032	0.015	-0.027	9,259	0.699
	(0.021)	(0.020)	(0.023)	(0.049)	(0.029)		
341	-0.007	0.006	-0.020	0.023	-0.029	19,787	0.710
	(0.015)	(0.014)	(0.013)	(0.037)	(0.019)		
342	0.004	-0.089***	-0.055***	0.033	-0.019	30,049	0.673
	(0.012)	(0.015)	(0.013)	(0.024)	(0.017)		
344	-0.030	0.081	-0.008	0.016	-0.113	52,865	0.667
	(0.047)	(0.054)	(0.074)	(0.097)	(0.081)		
351	-0.015	0.161**	-0.188	0.136	0.404	849	0.897
	(0.110)	(0.080)	(0.128)	(0.255)	(0.288)		
352	-0.025	0.027	-0.045***	-0.049**	-0.020	20,679	0.541
	(0.016)	(0.017)	(0.012)	(0.022)	(0.019)		
353	-0.037***	0.026**	-0.058***	-0.054***	-0.054***	315,361	0.474
	(0.006)	(0.013)	(0.008)	(0.016)	(0.011)		
354	-0.033***	0.006	-0.048***	-0.006	-0.052***	642,188	0.502
	(0.008)	(0.007)	(0.012)	(0.014)	(0.013)		
355	0.025	0.044***	-0.097***	-0.019	0.018	2,914	0.718
	(0.015)	(0.016)	(0.035)	(0.015)	(0.031)		
356	-0.046***	-0.001	-0.030**	-0.014	-0.057***	316,191	0.498
	(0.008)	(0.011)	(0.013)	(0.023)	(0.013)		
411	0.045	0.004	-0.076***	-0.053	-0.010	6,719	0.742
	(0.030)	(0.026)	(0.028)	(0.047)	(0.033)		
412	-0.016***	0.025***	-0.058***	-0.040***	-0.039***	246,060	0.508
	(0.005)	(0.008)	(0.006)	(0.009)	(0.007)		
413	-0.032***	0.003	-0.036**	-0.025*	0.015	87,616	0.519
	(0.007)	(0.012)	(0.017)	(0.014)	(0.013)		
415	-0.014***	0.003	-0.028***	-0.026	-0.007	153,110	0.513
	(0.004)	(0.005)	(0.005)	(0.017)	(0.005)		
416	-0.029***	-0.005	-0.065***	-0.032**	-0.008	38,232	0.731
	(0.010)	(0.009)	(0.012)	(0.014)	(0.014)		
421	-0.018***	0.006	-0.045***	-0.039***	0.005	76,428	0.671
	(0.007)	(0.008)	(0.008)	(0.011)	(0.010)		
511	-0.011	0.048***	-0.043***	-0.016	-0.006	9,189	0.584
	(0.018)	(0.015)	(0.015)	(0.025)	(0.023)		
521	-0.008	-0.009	-0.026	-0.095	-0.036	15,616	0.583
	(0.014)	(0.020)	(0.018)	(0.068)	(0.031)		
522	-0.002	-0.007	0.011	-0.023	0.048*	99,264	0.536
	(0.006)	(0.010)	(0.017)	(0.026)	(0.025)		
523	-0.019	0.020	0.004	-0.026	0.010	57,531	0.543
	(0.013)	(0.021)	(0.024)	(0.050)	(0.028)		

Continued on next page

Table A.3 – continued from previous page

SOC	Cons.	Open.	Extra.	Agree.	Emot.	N	R^2
524	-0.018**	0.019**	-0.032***	-0.019	-0.022	173,056	0.474
	(0.008)	(0.008)	(0.007)	(0.015)	(0.022)		
525	0.007	0.008	-0.018	-0.030	-0.069***	3,912	0.680
	(0.018)	(0.030)	(0.027)	(0.030)	(0.025)		
531	0.030	0.106**	-0.006	0.016	-0.095	80,652	0.599
	(0.038)	(0.049)	(0.032)	(0.030)	(0.083)		
532	0.003	-0.072	-0.055**	-0.028	-0.004	8,870	0.664
	(0.020)	(0.054)	(0.024)	(0.046)	(0.057)		
533	-0.013	-0.021	-0.002	-0.066	-0.005	8,331	0.712
	(0.021)	(0.028)	(0.019)	(0.046)	(0.030)		
541	-0.013	0.151**	-0.044	0.029	-0.091	1,309	0.706
	(0.022)	(0.068)	(0.040)	(0.030)	(0.069)		
542	-0.028	0.011	-0.046**	-0.017	0.008	2,761	0.758
	(0.025)	(0.016)	(0.021)	(0.100)	(0.020)		
543	-0.034*	0.031**	0.011	0.002	0.010	262,994	0.373
	(0.020)	(0.015)	(0.013)	(0.019)	(0.037)		
544	-0.071***	0.049*	0.057**	-0.009	0.111	4,325	0.772
	(0.019)	(0.026)	(0.026)	(0.069)	(0.070)		
612	0.002	0.008	-0.019**	-0.010	-0.031***	147,936	0.503
	(0.011)	(0.008)	(0.009)	(0.013)	(0.010)		
613	-0.023	0.002	-0.101**	-0.021	-0.085*	5,886	0.636
	(0.035)	(0.021)	(0.046)	(0.038)	(0.043)		
614	0.018**	0.015**	-0.010**	-0.011	-0.027***	1,292,239	0.401
	(0.008)	(0.008)	(0.004)	(0.010)	(0.008)		
621	-0.029	0.033*	-0.061**	-0.024	-0.009	43,617	0.611
	(0.019)	(0.018)	(0.030)	(0.022)	(0.013)		
622	0.010	-0.010	0.004	-0.102**	0.001	13,135	0.574
	(0.021)	(0.036)	(0.022)	(0.049)	(0.041)		
623	-0.011	0.002	-0.013	0.004	-0.008	19,812	0.674
	(0.010)	(0.016)	(0.010)	(0.020)	(0.013)		
624	-0.312**	0.096**	-0.022	-0.131**	-0.227***	52,889	0.639
	(0.124)	(0.039)	(0.032)	(0.057)	(0.067)		
711	-0.020	-0.015	-0.005	0.015	0.006	194,211	0.579
	(0.017)	(0.012)	(0.014)	(0.022)	(0.025)		
712	-0.064***	-0.008	-0.039*	-0.022	-0.009	43,355	0.796
	(0.019)	(0.025)	(0.022)	(0.027)	(0.026)		
713	0.009	0.054*	-0.077***	-0.012	0.037	21,121	0.778
	(0.026)	(0.029)	(0.027)	(0.027)	(0.047)		
721	-0.012	0.001	-0.006	-0.040***	-0.010	438,368	0.518
	(0.009)	(0.009)	(0.010)	(0.009)	(0.007)		
722	-0.017	0.055***	-0.051***	-0.061***	-0.025	202,452	0.605
	(0.012)	(0.015)	(0.012)	(0.017)	(0.020)		
811	0.010	0.094**	0.065***	0.062	0.043	3,155	0.719

Continued on next page

Table A.3 – continued from previous page

SOC	Cons.	Open.	Extra.	Agree.	Emot.	N	R^2
	(0.022)	(0.041)	(0.024)	(0.138)	(0.042)		
812	0.001	-0.014	-0.017	0.010	0.013	8,122	0.621
	(0.018)	(0.020)	(0.015)	(0.052)	(0.027)		
813	-0.024**	-0.034**	-0.027**	-0.001	-0.013	21,095	0.737
	(0.011)	(0.015)	(0.013)	(0.034)	(0.020)		
814	0.019*	0.011	-0.010	0.005	-0.004	39,571	0.665
	(0.011)	(0.020)	(0.014)	(0.025)	(0.026)		
821	0.009	-0.042**	-0.031	0.123*	-0.053	363,254	0.599
	(0.014)	(0.017)	(0.023)	(0.072)	(0.045)		
822	-0.001	-0.006	-0.016	-0.043	0.022	39,867	0.624
	(0.009)	(0.021)	(0.012)	(0.041)	(0.025)		
823	0.027	-0.023	0.009	-0.020	0.016	8,332	0.663
	(0.017)	(0.020)	(0.014)	(0.034)	(0.020)		
911	-0.024	0.016	-0.049**	0.052	-0.026	3,114	0.710
	(0.022)	(0.054)	(0.023)	(0.063)	(0.042)		
912	-0.034**	-0.007	0.058***	0.053*	-0.047	27,677	0.371
	(0.015)	(0.036)	(0.022)	(0.030)	(0.070)		
913	-0.002	-0.007	0.004	-0.037	-0.036**	16,346	0.610
	(0.012)	(0.012)	(0.015)	(0.032)	(0.017)		
921	0.031	-0.027**	0.085***	0.121***	-0.039	55,434	0.728
	(0.023)	(0.013)	(0.031)	(0.032)	(0.036)		
923	0.006	-0.031***	-0.030**	0.009	-0.001	189,567	0.609
	(0.011)	(0.011)	(0.014)	(0.011)	(0.020)		
924	-0.018**	0.003	-0.019*	-0.006	-0.014	131,362	0.761
	(0.009)	(0.008)	(0.011)	(0.013)	(0.010)		
925	-0.042	0.189***	-0.044	-0.070	-0.151***	1,590	0.819
	(0.060)	(0.071)	(0.065)	(0.074)	(0.058)		
926	-0.016	0.012	-0.016	-0.068**	-0.020	151,947	0.403
	(0.013)	(0.026)	(0.022)	(0.034)	(0.012)		
927	0.028***	0.007	-0.038**	0.027	-0.060**	275,105	0.417
	(0.010)	(0.011)	(0.016)	(0.028)	(0.023)		

Notes: Each row presents coefficients and standard errors (in parentheses) from separate regressions for each 3-digit SOC occupation. SOC codes follow the UK Standard Occupational Classification 2010. Full descriptions of occupation categories are available from the Office for National Statistics (www.ons.gov.uk/methodology/classificationsandstandards). "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. The dependent variable is the natural logarithm of posted salary. All regressions include controls for soft skills, hard skills, remote work, advertisement length, contract type, work schedule, seniority level, education, experience, amenities availability, and negotiable salary. Fixed effects for location (NUTS2 region), month-year, and firm are included. Standard errors are clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Agreeableness intensity interactions with firm characteristics

Firm Characteristic	Effect	Q1	Q2	Q3	Q4	Q5
Firm Age						
	Average firms	-0.005	-0.017	-0.003	-0.016	-0.010
		(0.012)	(0.010)	(0.011)	(0.021)	(0.018)
	High-value firms	0.005	-0.005	0.019	-0.010	-0.028
		(0.013)	(0.017)	(0.018)	(0.021)	(0.034)
Profit						
	Average firms	-0.004	-0.018	-0.006	-0.016	-0.012
		(0.013)	(0.011)	(0.012)	(0.020)	(0.020)
	High-value firms	0.009	-0.020	-0.003	-0.003	0.001
		(0.021)	(0.023)	(0.029)	(0.018)	(0.018)
Firm Size						
	Average firms	-0.002	-0.016	-0.005	-0.017	-0.011
		(0.011)	(0.011)	(0.011)	(0.020)	(0.021)
	High-value firms	0.035***	-0.005	0.002	-0.024	0.002
		(0.013)	(0.013)	(0.015)	(0.024)	(0.022)
Publicly Quoted						
	Non-quoted firms	-0.004	-0.015	-0.003	-0.014	-0.011
		(0.013)	(0.012)	(0.011)	(0.022)	(0.021)
	Quoted firms	-0.011	-0.057	-0.092***	-0.096*	-0.069**
		(0.031)	(0.050)	(0.033)	(0.051)	(0.030)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Reference category is Q0 (no trait mention). All models include full controls and fixed effects. N=1754940. *Average firms* shows coefficients at the mean value of the firm characteristic. *High-value firms* shows effects at 1 SD above mean for continuous characteristics. Firm characteristic means and standard deviations: Firm Age: mean=27 years (SD=21.6); Firm Profit: mean=£185619 (SD=£612396); Firm Size: mean=21845 employees (SD=87181).

Table A.5: Conscientiousness intensity interactions with firm characteristics

Firm Characteristic	Effect	Q1	Q2	Q3	Q4	Q5
Firm Age						
	Average firms	-0.004	-0.007	-0.027**	-0.014	-0.014
		(0.010)	(0.014)	(0.011)	(0.011)	(0.014)
	High-value firms	0.003	-0.017	-0.017	0.006	-0.025*
		(0.012)	(0.017)	(0.015)	(0.014)	(0.013)
Profit						
	Average firms	-0.005	-0.008	-0.027**	-0.013	-0.014
		(0.011)	(0.014)	(0.011)	(0.011)	(0.014)
	High-value firms	0.001	-0.012	-0.029*	-0.011	-0.004
		(0.012)	(0.016)	(0.017)	(0.016)	(0.016)
Firm Size						
	Average firms	0.000	-0.007	-0.027**	-0.011	-0.015
		(0.011)	(0.015)	(0.011)	(0.011)	(0.013)
	High-value firms	0.030*	-0.006	-0.028**	0.001	-0.020
		(0.016)	(0.019)	(0.012)	(0.013)	(0.012)
Publicly Quoted						
	Non-quoted firms	-0.005	-0.004	-0.018*	-0.007	-0.013
		(0.011)	(0.014)	(0.010)	(0.010)	(0.014)
	Quoted firms	-0.021	-0.055*	-0.163***	-0.100**	-0.026
		(0.027)	(0.032)	(0.042)	(0.048)	(0.043)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Reference category is Q0 (no trait mention). All models include full controls and fixed effects. N=1754940. *Average firms* shows coefficients at the mean value of the firm characteristic. *High-value firms* shows effects at 1 SD above mean for continuous characteristics. Firm characteristic means and standard deviations: Firm Age: mean=27 years (SD=21.6); Firm Profit: mean=£185619 (SD=£612396); Firm Size: mean=21845 employees (SD=87181).

Table A.6: Openness intensity interactions with firm characteristics

Firm Characteristic	Effect	Q1	Q2	Q3	Q4	Q5
Firm Age						
	Average firms	0.010	0.001	0.007	0.005	-0.006
		(0.011)	(0.007)	(0.010)	(0.011)	(0.013)
	High-value firms	0.001	0.001	0.010	0.008	0.004
		(0.016)	(0.009)	(0.012)	(0.014)	(0.015)
Profit						
	Average firms	0.008	0.001	0.007	0.005	-0.007
		(0.011)	(0.007)	(0.010)	(0.011)	(0.013)
	High-value firms	0.016	-0.002	0.009	0.006	0.008
		(0.014)	(0.012)	(0.012)	(0.011)	(0.011)
Firm Size						
	Average firms	0.010	0.001	0.007	0.004	-0.000
		(0.011)	(0.007)	(0.010)	(0.010)	(0.013)
	High-value firms	0.027**	-0.005	0.008	-0.004	0.035*
		(0.013)	(0.007)	(0.010)	(0.010)	(0.018)
Publicly Quoted						
	Non-quoted firms	0.011	0.003	0.008	0.007	-0.006
		(0.012)	(0.007)	(0.010)	(0.010)	(0.014)
	Quoted firms	-0.005	-0.027	-0.002	-0.012	-0.022
		(0.039)	(0.022)	(0.038)	(0.037)	(0.043)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Reference category is Q0 (no trait mention). All models include full controls and fixed effects. N=1754940. *Average firms* shows coefficients at the mean value of the firm characteristic. *High-value firms* shows effects at 1 SD above mean for continuous characteristics. Firm characteristic means and standard deviations: Firm Age: mean=27 years (SD=21.6); Firm Profit: mean=£185619 (SD=£612396); Firm Size: mean=21845 employees (SD=87181).

Table A.7: Extraversion intensity interactions with firm characteristics

Firm Characteristic	Effect	Q1	Q2	Q3	Q4	Q5
Firm Age						
	Average firms	-0.011	-0.041***	-0.030**	-0.034**	-0.045***
		(0.011)	(0.012)	(0.013)	(0.015)	(0.013)
	High-value firms	-0.024	-0.068***	-0.034***	-0.038***	-0.023*
		(0.017)	(0.015)	(0.012)	(0.014)	(0.013)
Profit						
	Average firms	-0.011	-0.041***	-0.030**	-0.034**	-0.045***
		(0.011)	(0.013)	(0.013)	(0.015)	(0.013)
	High-value firms	-0.012	-0.062***	-0.027**	-0.033***	-0.045***
		(0.014)	(0.016)	(0.010)	(0.011)	(0.013)
Firm Size						
	Average firms	-0.010	-0.045***	-0.028**	-0.034**	-0.045***
		(0.011)	(0.014)	(0.012)	(0.015)	(0.013)
	High-value firms	-0.008	-0.061***	-0.017	-0.036***	-0.048***
		(0.012)	(0.017)	(0.010)	(0.013)	(0.012)
Publicly Quoted						
	Non-quoted firms	-0.013	-0.039***	-0.029**	-0.034**	-0.045***
		(0.012)	(0.014)	(0.014)	(0.016)	(0.014)
	Quoted firms	0.005	-0.101***	-0.047	-0.025	-0.042
		(0.038)	(0.037)	(0.029)	(0.055)	(0.033)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Reference category is Q0 (no trait mention). All models include full controls and fixed effects. N=1754940. *Average firms* shows coefficients at the mean value of the firm characteristic. *High-value firms* shows effects at 1 SD above mean for continuous characteristics. Firm characteristic means and standard deviations: Firm Age: mean=27 years (SD=21.6); Firm Profit: mean=£185619 (SD=£612396); Firm Size: mean=21845 employees (SD=87181).

Table A.8: Emotional Stability intensity interactions with firm characteristics

Firm Characteristic	Effect	Q1	Q2	Q3	Q4	Q5
Firm Age						
	Average firms	-0.021	-0.024*	-0.032**	0.008	-0.032**
		(0.018)	(0.014)	(0.014)	(0.022)	(0.013)
	High-value firms	-0.025	-0.031**	-0.025	0.016	-0.052***
		(0.018)	(0.012)	(0.021)	(0.029)	(0.016)
Profit						
	Average firms	-0.020	-0.024*	-0.031**	0.008	-0.032**
		(0.018)	(0.014)	(0.014)	(0.022)	(0.013)
	High-value firms	-0.025	-0.020	-0.017	-0.000	-0.056***
		(0.018)	(0.017)	(0.027)	(0.021)	(0.015)
Firm Size						
	Average firms	-0.018	-0.020	-0.029**	0.008	-0.032**
		(0.017)	(0.014)	(0.013)	(0.023)	(0.015)
	High-value firms	0.008	0.010	-0.000	0.011	-0.038**
		(0.017)	(0.014)	(0.019)	(0.030)	(0.018)
Publicly Quoted						
	Non-quoted firms	-0.015	-0.022	-0.033**	0.011	-0.027*
		(0.019)	(0.015)	(0.013)	(0.023)	(0.014)
	Quoted firms	-0.079**	-0.066*	-0.047	-0.047**	-0.101**
		(0.031)	(0.035)	(0.049)	(0.022)	(0.041)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Reference category is Q0 (no trait mention). All models include full controls and fixed effects. N=1754940. *Average firms* shows coefficients at the mean value of the firm characteristic. *High-value firms* shows effects at 1 SD above mean for continuous characteristics. Firm characteristic means and standard deviations: Firm Age: mean=27 years (SD=21.6); Firm Profit: mean=£185619 (SD=£612396); Firm Size: mean=21845 employees (SD=87181).

Table A.9: Effects of personality trait intensity quintiles on wages

Intensity Measure	Any P.T.	Agree.	Consc.	Emot.	Extraversion	Open.
Quintile 1 (Lowest Intensity)	-0.024***	0.022	-0.005	-0.015**	-0.004	0.008
	(0.007)	(0.018)	(0.007)	(0.007)	(0.011)	(0.006)
Quintile 2	-0.039***	-0.005	-0.011	-0.053***	-0.038***	0.009*
	(0.008)	(0.010)	(0.009)	(0.018)	(0.008)	(0.005)
Quintile 3	-0.053***	-0.031***	-0.034***	-0.031***	-0.038***	0.007
	(0.007)	(0.008)	(0.013)	(0.009)	(0.008)	(0.005)
Quintile 4	-0.053***	-0.036***	-0.024**	-0.014	-0.036***	0.011*
	(0.009)	(0.008)	(0.010)	(0.019)	(0.008)	(0.006)
Quintile 5 (Highest Intensity)	-0.071***	-0.024	-0.026**	-0.043***	-0.045***	0.000
	(0.010)	(0.018)	(0.011)	(0.006)	(0.009)	(0.010)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. "Consc." refers to Conscientiousness, "Open." to Openness, "Agree." to Agreeableness, "Extra." to Extraversion, "Emot." to Emotional Stability. Baseline is *no trait mentioned* (trait intensity = 0). Regressions control for job requirements, location, time-month, occupation and firm fixed effects. Sample size: 9888315 observations.

Table A.10: Effects of trait frequency categories on wages

Trait - Number of mentions	1	2	3	4	5+
Agreeableness	-0.014*	-0.018	0.014	-0.031	0.078
	(0.008)	(0.022)	(0.051)	(0.032)	(0.105)
Conscientiousness	-0.017**	-0.033**	-0.029**	-0.045***	-0.058***
	(0.007)	(0.013)	(0.011)	(0.016)	(0.017)
Emotional Stability	-0.032***	-0.020	-0.075***	-0.067**	-0.101***
	(0.006)	(0.025)	(0.017)	(0.026)	(0.020)
Extraversion	-0.028***	-0.037***	-0.059***	-0.072***	-0.092***
	(0.007)	(0.008)	(0.010)	(0.015)	(0.012)
Openness	0.004	0.018**	0.022*	0.027*	-0.041
	(0.004)	(0.007)	(0.013)	(0.015)	(0.073)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Baseline is no trait mention (trait frequency = 0). Regressions control for job requirements, location, time-month, occupation and firm fixed effects. Sample size: 9 888 315 observations.

Appendix B - Amenities' dictionaries

Category	Words searched
A00: General Benefits	
General Benefits	rewardscheme, flexiblebenefit, benefitsincluding, generousbenefits, benefitsplatform, comprehensivebenefits, staffbenefits, benefitsscheme, employeebenefits, competitivebenefits, fantasticbenefits, greatbenefits, excellentbenefits, companybenefits, rangebenefits, standardbenefits, corebenefits, variousbenefits
Benefits Packages	benefitspackage, greatpackage, rewardspackage, salarypackage, remunerationpackage
Recognition Schemes	recognitionschemes, employeemonthaward, monthlyawards, employeerecognition
Referral Schemes	referralscheme, referfriendincentive, referralprogramme, staffreferral, employeereferral, referfriendscheme
Tax Benefits	taxefficientscheme, taxefficientsavings, familytax, taxfreecycle
Loans Facilities	interestfreeloans, mortgagefacility, mileagebank
Relocation Support	relocationpackage, relocationallowance, visarelocation, sponsorshipvisa, visasponsorship, visacosts
Seasonal Benefits	christmas
A01: Bonuses and Commissions	
Basic Compensation	bonus, commission, incentive
Equity Compensation	stockoptions, sharesaving, sharescheme, sharesave, shareincentiveplan
Profit Sharing	profitsharing, profitshare
Performance Pay	performancerelated, payforperformance, performancerecognised, productivitypay
Sales Compensation	salesquota, salesgoal, salestarget
Premium Pay	premiumpay, salarypremium, extrapay, variablepay
Special Payments	paycash, averagepayout, tipsshared, annualpay, anniversaryreward
Savings Schemes	saveearnscheme, savingsscheme
Promotion Based	promotionbased
A02: Overtime Paid	
Overtime	extrahour, overtime, longhour, extrashift, doublepay, unsocialhour
A03: Paid Time Off	
Vacation Holidays	holiday, vacation, freeday
Sick Leave	sickday, sickleave, sicktime, sicknesspay, sickpay, sickdays, sickchild
Medical Leave	healthleave, medicalleave, medicalindemnity
Maternity Benefits	maternitypay, maternityleave, maternityday, maternitypaternity, maternitypolicies, maternitypackage, enhancedmaternity
Paternity Benefits	paternitypay, paternityleave, paternityday
Family Leave	parentalleave, parentaladoption, familypolicies, familydays, adoptionleave
Personal Time	personalday, paidtimeoff, paidleave, paidrest, paidbreak, dayspaid
Special Leave	bereavementleave, trainingleave, educationleave, studyleave, compensationleave, plannedabsence, movinghouseday, birthday, resetdays, earlyfriday

Continued on next page

Table B.1 – continued from previous page

Category	**Words searched**
Volunteer Time	communityvolunteering, annuallyvolunteer, volunteertime, volunteerday, volunteeringday, charitableday, charityprojects
Flexible Time	summerfridays, optionbuy
Wellbeing Time	wellbeingbenefits, wellbeingdays, cashhealthplan, healthcaredentalcare
A04: Health Insurance	
Health Insurance	healthinsurance, offerhealthinsurance, privatehealth, privatehealthinsurance, healthcarescheme, privatemedical, healthplan, medicalplan, healthcareplan, medicalbenefit, medicalcover, medicalinsurance
Dental Vision	dentalbenefit, dentalinsurance, offerdentalinsurance, dentalcoverage, dentalplan, dentalcare, dentaltreatment, dentalcover, includesdental, subsidiseddental, dentalvision, freedental, opticaldental, visioninsurance, visioncoverage
Life Insurance	lifeinsurance, lifecover, lifeaccidentinsurance
Accident Insurance	offeraccidentinsurance, includeaccidentinsurance, personalaccidentinsurance, travelaccidentinsurance
Health Coverage	healthcoverage, healthcontribution, familycoverage, familyhealthcare, healthwellbeing
Free Healthcare	freehealth, freemedical, freehealthcare, freelegaladvice
Health Savings	healthsavings, cashplan, medicash
Wellness Programs	wellnessprogramme, freefluvaccine, freeeyetest, onsiteflushots, annualflu, fluvouchers
Childcare Support	childcarescheme, childcarevouchers
Eye Care	eyecarecover, eyecarediscount, eyecarevoucher
Specialised Coverage	criticalillnesscover, criticalillnessscheme, disabilitycover, illnessinsurance, indemnitycover, mentalhealthfirst
Insurance Schemes	privateinsurance, discountedinsurance, welfarepackage, healthcover
A05: Retirement Contributions	
Pension Schemes	pension
Retirement Plans	retirementcontribution, retirementfund
Employer Contributions	employercontribution, employermatch, companymatch, twicecontribute, doublecontribution
Savings Accounts	savingsaccount, savingscontribution, savingsscheme
Provident Funds	providentfund
Education Support	tuitionreimbursement
Life Benefits	lifeassurance
A06: Food Subsidies and Discounts	
Breakfast Benefits	breakfastincluded, caterbreakfast, freebreakfast, freedailybreakfast, unlimitedbreakfast
Lunch Benefits	freelunch, lunchincluded, caterlunch, lunchallowance, lunchticket
Meal Benefits	freemeal, mealincluded, mealprovided, catermeal, mealsduty, mealsprovided, subsidisedmeal, onshiftmeals, workdaylunches

Continued on next page

Table B.1 – continued from previous page

Category	**Words searched**
Food General	freefood, foodincluded, foodprovided, foodcoupon, foodstamp, foodsubsidy, foodticket
Dinner Benefits	freedinner, dinnerincluded
Drinks Beverages	freedrinks, freecoffee, coffeetea, teacoffee, freebeer, freebeverages, freetea, drinksavailable, drinksfacilities, drinksfridge, freehotdrinks, fridaysdrinks, beerfridays, regulardrinks
Fruit Snacks	freefruit, freshfruit, fruiteveryday, fruitrefreshments, fruitsnacks, dailyfruit, healthysnacks, afternoonsnacks, officesnacks, freeicecream, freecake, freesoft, freeunhealthy, freehealthy
Gym Fitness	gymdiscount, gymmembership, gympass, gymsubsidies, gymnearby, fitnessclasses, freeyoga, swimmembership, yogasessions, weeklyyoga
Transport Benefits	freetransportation, transportationallowance, travelsubsidy, commutesubsidy, transitsubsidy, transportationsubsidy, transportvoucher, travelstipend
Discount Schemes	employeediscount, companydiscount, corporatediscount, staffdiscount, storediscount, retaildiscount, cardiscount, discountcar, discountcard, discountprogramme, discountschemes, discounttravel, discountvarious, discountedcinema, discountedfitness, discountedfood, discountedmembership, discountedstays, discountedsubscription, dealsdiscount, leisurediscount, hoteldiscount, traveldiscount, onlinediscount, instorediscounts, discountlocal
Clothing Uniform	clothesdiscount, companyuniform, comfortableuniform
Accommodation	liveposition, housingincluded, onsiteaccommodation, rentalsubsidy
Vouchers Rewards	cashback, seasontickets, shoppingvouchers, cinematickets, loyaltyrewards, rewardspoints, freetickets, voluntarybenefits
Wellness Perks	wellbeingoffers, wellnessportal, accessheadspace, headspacesubscription, meditationapp
Vehicle Benefits	carleasing, carloanscheme, fuelcard, leasingdeals
Platform Benefits	benefitsplatform, benefitsportal, advantagescheme, leasescheme, huboffers, onlinebenefits, rewardplatform, shoppingportal
Leisure Entertainment	pooltables, personalparcels, freeuse, leisuresavings
Coffee Facilities	excellentcoffeemachine, nespressocoffeemachine
Refreshments	refreshmentsprovided, dailymenu, subsidisedrate
Insurance Extras	bikeinsurance, travelinsurance, travelbenefits
Special Services	deathservice
A07: Office Space and Amenities	
Fitness Facilities	fitnessroom, officegym, onsitefitness, onsitegym, fitnessstudio, officeyoga
Recreational Spaces	gameroom, gamesarea, gamesroom, xboxroom, tabletennis, tennisroom, relaxationarea, breakarea, breakrooms, breakoutarea, chilloutarea
Dining Facilities	companycafeteria, onsitecafeteria, onsitecatering, onsiterestaurant, staffrestaurant, canteen, onsitebistro, stockedkitchen, subsidisedrestaurant

Continued on next page

Table B.1 – continued from previous page

Category	**Words searched**
Office Environment	comfortableoffice, comfortablespace, modernoffice, brandnewoffice, refurbishedoffice, greatoffices, openspace, openplanenvironment, openplanoffice, naturallight, greatviews, openroof
Specialised Rooms	prayerroom, wellnessroom, vapingarea, onsitecinema, onsitenursery
Work Facilities	conferenceroomfacilities, quietworkspaces, quietworkingspaces, coworkingspace
Facility Quality	amenities, facilitiesonsite, onsitefacilities, goodfacilities, greatfacilities, excellentfacilities, fantasticfacilities, modernfacilities, stateartfacilities, rangefacilities, comfortablefacilities, leisurefacilities, sportfacilities
Workplace Description	amazingworkplace, officeincludes, officemodern, developedoffice, stateartbuilding
Outdoor Spaces	roofterrace, rooftopterrace
Service Facilities	showerfacilities, showerfacility, onsiteshower, onsiterestroom, selfservicefacilities, usemicrowave, onsitesubsidised, onsiteswimming
Storage Amenities	stafflockers, staffrestarea
Location Features	locationalsoprovides, locationprovides
A08: Location and Commuting	
Location Quality	centrallocation, excellentlocation, goodlocation, greatlocation, finelocation, ideallocation, perfectlocation, peacefullocation, welllocated
City Center Access	citycenter, towncenter, basedheart, situatedheart, officeheart
Accessibility	easyaccess, easilyaccessible, convenientlylocated, quickaccess, easiestreach, closeaccess
Commuting	shortcommute, longcommute, easycommute, easilycommute, easilycommutable, perfectcommuting, easyreach
Parking Facilities	privateparking, payparking, companygarage, privategarage, carpark, onsitepark, carparking, freeparking, freecarparking, secureparking, siteparking, staffparking, streetparking, parkingavailable, parkingspace, localparking
Public Transport	accesspublictransport, publictransport, reliablepublic, viapublictransport, greattransport, transportlinks, transportationlinks, linkstransport, commuterlinks
Walking Distance	shortwalk, minutewalk, minuteswalk, walkingdistance, walkabledistance, minwalk, withinshortdistance
Cycling Facilities	bike2work, bikesheds, bikestorage, bikeworkscheme, cyclescheme, cycletowork, cyclework, cyclingscheme, taxfreebike, ride2work
Local Transport	buslinks, bustrain, localbus, localtransport, freebus, freelocalbus, discountedbus
Train Stations	nextcentralstation, mainlinestation, metrostationclose, metrostation
Location Context	basedcentre, centrelocated, centrelocation, situatedclose, minutescentral, outsidecitycentre, officenext
Nearby Amenities	localshops, nearlocal, childcareoptions
A09: Work Equipment and Allowances	

Continued on next page

Table B.1 – continued from previous page

Category	Words searched
Remote Work Support	internetcoupon, internetsubsidy, technologysubsidy, homeofficefund, teleworkfund, teleworkstipend, workhomefund, teleworkexpenses, homeofficeexpenses, workhomeexpenses, homeofficeallowance, teleworkallowance, workhomeallowance, teleworksubsidy, workhomesubsidy, homeofficeequipment, teleworkequipment, freebroadband, discountedbroadband, discountsenergybills
Mobile Devices	offermobilephone, includemobilephone, companyphone, freemobile, phoneandlaptop, phonelaptop, phoneprovided, phoneinsurance, phonework, offercompanycellphone, includecompanycellphone
Computing Equipment	offercompanycomputer, includecompanycomputer, offercompanylaptop, includecompanylaptop, companylaptop, twomonitors, twoscreens, dedicateddesk
Company Vehicles	companycar, companyvan, companyvehicle, vehicleallowance, vehiclescheme
Vehicle Expenses	fuelcard, fuelallowance, paidmileage, mileageexpense, carallowance
Company Cards	companycreditcard, companydebitcard
Clothing Equipment	companyuniform, freeuniform, uniformprovided, uniformprovision, provisionuniform, corporateclothing, protectivefootwear, protectiveequipment
Work Equipment	equipmentprovided, necessaryequipment
Perks Boxes	perkbox, perks, perbox, welcomebox
A10: Work Schedule Flexibility	
Flexible Scheduling	flexibleshift, flexibletime, flexiblehour, flexibleschedule, flxshift, flxtime, flxhour, flxschedule, flexibleofficehours, flxoption, flexibleoption
Work Life Balance	worklifebalance, lifebalance, lifework, healthywork, work.life, weeklyrest, familyfriendly
Flexible Work	flexiblework, flxwork, flexibleworking, modernworking, flexibilitywork, flexibilityaroundworking
A11: Workplace Safety and Ergonomics	
Workplace Safety	safeplace, worksafety, safework, safeworking, workplacesafety
Safety Culture	safetyculture, safetyfirst, employeesafety
Security	secureoffices, secureworkplace
Health Safety	healthsafety, awarehealth, focushealth
Safety Compliance	complysafe, complyhealth, compliancehealth, understandhealth, requirementshealth
Safety Regulations	safetylegislation, act1974, fireregulation, safetyact
Safety Policies	safetyrules, safetypolicies
Office Safety	safeoffice, hygienicworking
A12: Job Security	
Job Security	jobsecurity, jobstability
Company Stability	lowturnover, stableenvironment
Severance Benefits	severancepay, severancepackage, layoffcompensation
Unemployment Support	unemploymentinsurance, unemploymentbenefit

Continued on next page

Table B.1 – continued from previous page

Category	Words searched
Income Protection	incomeprotection, indemnityinsurance
A13: Work Environment and Impact on Society	
Workplace Culture	benefitculture, goodculture, companyculture, internalculture, collaborativeculture, positiveculture, workingculture, officeculture
Work Environment	bestworkplace, excellentworkplace, greatworkplace, amazingworkplace, dynamicworkplace, excellentenvironment, goodenvironment, warmenvironment, youngenvironment, multiculuralenvironment, relaxedenvironment, collaborativeenvironment, safeenvironment, socialenvironment, stimulatingenvironment, supportiveenvironment, uniqueenvironment, vibrantworkplace, workenvironment, workingenvironment
Workplace Atmosphere	goodatmosphere, relaxedatmosphere, lovelyatmosphere, casualdress, dresscode, vibrantoffice, sociableoffice, livelyoffice, pleasantoffice
Work Satisfaction	comfortwork, congenialwork, enjoyablework, meaningfulwork, positivework, relaxwork, healthywork, jobrecognition, jobsatisfaction, satisfactionjob, rewarding, fulfilling
Company Values	humanvalues, missionvision, companyvalues, values, promoteequality, valuediversity, responsiblecompany, culturematter
Social Activities	afterworkdrinks, companyevents, companyoutings, companyparties, companyretreats, companysocialevents, companytrips, funnysocialevents, gamesnights, regularsocialevents, regularsocials, socialactivities, socialgatherings, staffevents, teambonding, teambuilding, teamlunch, teamevents, outdooractivities
Company Reputation	bestplaceswork, bestreputation, localreputation, strongcompany
Employee Support	employeeassistance, employeesupport, freecounselling, wellbeingsupport, wellbeingworkshops, prioritisingwellbeing
Work Quality	enjoyableworking, excellentworkplace, fantasticworking, friendlyworking, greatworking, inclusiveworking, supportiveworking
Impact Purpose	hugeimpact, immediateimpact, impactsociety, differencelife, makedifference, makerealdifference, makingimpact, meaningfulproject, peoplelives, positiveimpact, withincommunity, impactworld, makeimpact, realimpact, significantimpact, strongimpact, positivechange, directimpact
Diversity Inclusion	multicultural, equalopportunities, valuemember, feelpart
Team Environment	interdisciplinary, multidisciplinary, dynamicteam, youngteam, greatteam, excitingteam, goodvibes
Work Environment Type	socialworkenvironment, workingethos
A14: Human Capital Development	
Career Growth	buildcareer, movecareerforward, career, progress
Skill Development	buildskills, developingskills, learningdevelopment, personaldevelopment, professionaldevelopment, developmentsupport

Continued on next page

Table B.1 – continued from previous page

Category	Words searched
Learning Opportunities	developmentopportunities, developmentopportunity, growthopportunities, opportunitieslearning, externallearning, continuousopportunities, opportunitiesgrowing, opportunitiesupskill
Company Growth	companygrowth, growingorganisation, growingcompany, growthcompany, growingbusiness, teamexpansion, growingteam
Training Education	training, elearningportal, learningallowance, learningzone, educationsubsidy, payeducation, course
Promotion Advancement	promoteinternally, promotionopportunities, promotionprospects, promotionprocess, promotiondevelopment, meritpromotion, salaryincrease
Merit Based	basemerit, meritincrease, meritbased
Personal Growth	personalgrowth, professionalgrowth, ongoinggrowth, growthopportunity, growthpossible, continuousimprovement, continuingprofessional
A15: Remote Work Opportunities	
Remote Work	conductedremote, fullyremote, homebased, homework, homeworking, jobisremote, jobisvirtual, locationisremote, locationremote, positionisremote, positionremote, remotework, rolewillberemote, telecommuting, telework, virtualjob, workfromhome, workhome, workremote, workingremote, flxwithoffice, remoteposition, roleisremote, remotebasis, hybridwork, hybridremote, workanywhere

Words were scrutinised as not to retrieve false positive words. Each word's context was analysed by sampling the dataset of 100 advertisements per word. When the word's context was not related to personality traits and a straightforward amendment of the word was not possible, this was removed from the dictionary. Words that did not retrieve any match were also not considered.

Before searching for words, modifications were made to some words to avoid false positives and account for different variations of words. The list of changes done to the text is the following:

01 General benefits and packages

"programe" for "programme", "rewards packages" for "rewardscheme", "rewards scheme" for "rewardscheme", "rewards" for "rewardscheme ", "reward" for "rewardscheme ", "flexible benefit" for "flexiblebenefit", "generous benefits" for "generousbenefits", "benefits platform" for "benefitsplatform", "comprehensive benefits" for "comprehensivebenefits", "staff benefits" for "staffbenefits", "recognition schemes" for "recognitionschemes", "benefits scheme" for "benefitsscheme", "employee benefits" for "employeebenefits", "competitive benefits" for "competitivebenefits", "fantastic benefits" for "fantasticbenefits", "great benefits" for "greatbenefits", "excellent benefits" for "excellentbenefits", "company benefits" for "companybenefits", "benefits package" for "benefitspackage", "great package" for "greatpackage", "rewards package" for "rewardspackage", "salary package" for "salarypackage", "remuneration package" for "remunerationpackage", "referral scheme" for "referralscheme", "range benefits" for "rangebenefits",

”standard benefits” for ”standardbenefits”, ”core benefits” for ”corebenefits”, ”benefits including” for ”benefitsincluding”, ”various benefits” for ”variousbenefits”, ”christmas” for ”christmas”, ”refer friend incentive” for ”referfriendincentive”, ”mileage bank” for ”mileagebank”, ”interest free loans” for ”interestfreeloans”, ”tax efficient scheme” for ”taxefficientscheme”, ”tax efficient savings” for ”taxefficientsavings”, ”family tax” for ”familytax”, ”mortgage facility” for ”mortgagefacility”, ”referral programme” for ”referralprogramme”, ”staff referral” for ”staffreferral”, ”tax free cycle” for ”taxfreecycle”, ”employee referral” for ”employeereferral”, ”refer friend scheme” for ”referfriendscheme”, ”employee month award” for ”employeemonthaward”, ”monthly awards” for ”monthlyawards”, ”visa costs” for ”visacosts”, ”relocation package” for ”relocationpackage”, ”relocation allowance” for ”relocationallowance”, ”visa relocation” for ”visarelocation”, ”sponsorship visa” for ”sponsorshipvisa”, ”visa sponsorship” for ”visasponsorship”, ”employee recognition” for ”employeerecognition”,

02 Payment and compensation variations

”bonuses” for ”bonus ”, ”commissions” for ”commission ”, ”incentives” for ”incentive ”, ”profit share” for ”profitsharing”, ”stock option” for ”stockoptions”, ”option stock” for ”stockoptions”, ”share profit” for ”profitsharing”, ”profit sharing” for ”profitsharing”, ”sharing profit” for ”profitsharing”, ”share profits” for ”profitsharing”, ”performance base” for ”performancerelated”, ”performance based” for ”performancerelated”, ”base performance” for ”performancerelated”, ”promotion base” for ”promotionbased”, ”base promotion” for ”promotionbased”, ”promotion based” for ”promotionbased”, ”extra pay” for ”extrapay”, ”pay extra” for ”extrapay”, ”sales goal” for ”salesgoal”, ”goal sales” for ”salesgoal”, ”sales target” for ”salestarget”, ”target sales” for ”salestarget”, ”premium pay” for ”premiumpay”, ”pay premium” for ”premiumpay”, ”salary premium” for ”salarypremium”, ”premium salary” for ”salarypremium”, ”productivity pay” for ”productivitypay”, ”pay productivity” for ”productivitypay”, ”sales quota” for ”salesquota”, ”quota sales” for ”salesquota”, ”variable pay” for ”variablepay”, ”pay for performance” for ”payforperformance”, ”performance pay” for ”payforperformance”, ”decommission” for ”decomm”, ”commissioners” for ”missioners”, ”share saving” for ”sharesaving ”, ”share schemes” for ”sharescheme ”, ”share scheme” for ”sharescheme ”, ”shares scheme” for ”sharescheme ”, ”tips shared” for ”tipsshared ”, ”share tips” for ”tipsshared ”, ”average payout” for ”averagepayout ”, ”annual pay” for ”annualpay”, ”share save” for ”sharesave”, ”anniversary reward” for ”anniversaryreward”, ”save earn scheme” for ”saveearnscheme”, ”savings scheme” for ”savingsscheme”, ”share incentive plan” for ”shareincentiveplan”, ”profit share” for ”profitshare”, ”performance recognised” for ”performancerecognised”

03 Work hours and overtime

”extra hour” for ”extrahour”, ”hour extra” for ”extrahour”, ”long hour” for ”longhour”, ”hour long” for ”longhour”, ”extra shift” for ”extrashift”, ”shift extra” for ”extrashift”, ”double pay” for ”doublepay”, ”pay double” for ”doublepay”, ”unsocial hour” for ”unsocialhour”

04 Time off

”breaks” for ”break”, ”sick day” for ”sickday”, ”day sick” for ”sickday”, ”sick leave” for ”sickleave”, ”leave sick” for ”sickleave”, ”sick time” for ”sicktime”, ”time sick” for ”sicktime”, ”health leave” for ”healthleave”, ”leave health” for ”healthleave”, ”medical leave” for ”medicalleave”, ”leave medical” for ”medicalleave”, ”maternity pay” for ”maternitypay”, ”pay maternity” for ”maternitypay”, ”maternity leave” for ”maternityleave”, ”leave maternity” for ”maternityleave”, ”maternity day” for ”maternityday”, ”day maternity” for ”maternityday”, ”paternity pay” for ”paternitypay”, ”pay paternity” for ”paternitypay”, ”paternity leave” for ”paternityleave”, ”leave paternity” for ”paternityleave”, ”paternity day” for ”paternityday”, ”day paternity” for ”paternityday”, ”bereavement leave” for ”bereavementleave”, ”leave bereavement” for ”bereavementleave”, ”personal day” for ”personalday”, ”day personal” for ”personalday”, ”paid time off” for ”paidtimeoff”, ”time off paid” for ”paidtimeoff”, ”paid leave” for ”paidleave”, ”leave paid” for ”paidleave”, ”free day” for ”freeday”, ”day free” for ”freeday”, ”training leave” for ”trainingleave”, ”leave training” for ”trainingleave”, ”education leave” for ”educationleave”, ”leave education” for ”educationleave”, ”sickness pay” for ”sicknesspay”,”compensation leave” for ”compensationleave”,”parental leave” for ”parentalleave”, ”study leave” for ”studyleave”,”medical indemnity” for ”medicalindemnity”,”paid rest” for ”paidrest”, ”parental adoption” for ”parentaladoption”,”planned absence” for ”plannedabsence”,”maternity paternity” for ”maternitypaternity”, ”maternity . paternity” for ”maternitypaternity”,”maternity , paternity” for ”maternitypaternity”,”wellbeing benefits” for ”wellbeingbenefits”, ”maternity policies” for ”maternitypolicies”,”sick days” for ”sickdays”,”moving house day” for ”movinghouseday”, ”charity projects” for ”charityprojects”,”sick child” for ”sickchild”,”community volunteering” for ”communityvolunteering”, ”annually volunteer” for ”annuallyvolunteer”,”option buy” for ”optionbuy”,”paid break” for ”paidbreak”, ”birthday” for ”birthday”,”days paid” for ”dayspaid”,”sick pay” for ”sickpay”, ”vacation” for ”vacation”,”maternity package” for ”maternitypackage”,”volunteer time” for ”volunteertime”, ”summer fridays” for ”summerfridays”,”enhanced maternity” for ”enhancedmaternity”,”maternity pay” for ”maternitypay”, ”reset days” for ”resetdays”,”volunteer day” for ”volunteerday”,”adoption leave” for ”adoptionleave”, ”family policies” for ”familypolicies”,”family days” for ”familydays”,”volunteering day” for ”volunteeringday”, ”cash health plan” for ”cashhealthplan”,”healthcare dental care” for ”healthcaredentalcare”,”early friday” for ”earlyfriday”, ”charitable day” for ”charitableday”,”wellbeing days” for ”wellbeingdays”

05 Health insurance and benefits

”join family healthcare” for ”join family”, ”family healthcare team” for ”family team”, ”medical benefit” for ”medicalbenefit”, ”benefit medical” for ”medicalbenefit”, ”free health” for ”freehealth”, ”health free” for ”freehealth”, ”free medical” for ”freemedical”, ”medical free” for ”freemedical”, ”health contribution” for ”healthcontribution”, ”contribution health” for ”healthcontribution”, ”health coverage” for ”healthcoverage”, ”coverage health” for ”healthcoverage”, ”offer health insurance” for ”offerhealthinsurance”, ”health insurance offer” for ”offerhealthinsurance”, ”include health insurance” for ”includehealthinsurance”, ”health insurance include” for ”includehealthinsurance”, ”dental benefit” for ”dentalbenefit”, ”benefit dental” for ”dentalbenefit”, ”free dental” for ”freedental”, ”dental free” for ”freedental”, ”dental coverage” for ”dentalcoverage”, ”coverage dental” for ”dentalcoverage”, ”offer dental insurance” for ”offerdentalinsurance”, ”dental insurance offer” for ”offerdentalinsurance”, ”include dental insurance” for ”includedentalinsurance”, ”dental insurance include” for ”includedentalinsurance”, ”vision coverage” for ”visioncoverage”, ”coverage vision” for ”visioncoverage”, ”offer vision insurance” for ”offervisioninsurance”, ”vision insurance offer” for ”offervisioninsurance”, ”include vision insurance” for ”includevisioninsurance”, ”vision insurance include” for ”includevisioninsurance”, ”offer accident insurance” for ”offeraccidentinsurance”, ”accident insurance offer” for ”offeraccidentinsurance”, ”include accident insurance” for ”includeaccidentinsurance”, ”accident insurance include” for ”includeaccidentinsurance”, ”offer life insurance” for ”offerlifeinsurance”, ”life insurance offer” for ”offerlifeinsurance”, ”include life insurance” for ”includelifeinsurance”, ”life insurance include” for ”includelifeinsurance”, ”health savings” for ”healthsavings”, ”savings health” for ”healthsavings”, ”private insurance” for ”privateinsurance”, ”insurance private” for ”privateinsurance”, ”family coverage” for ”familycoverage”, ”coverage family” for ”familycoverage”, ”onsite flu shots” for ”onsiteflushots”, ”child care scheme” for ”childcarescheme”, ”childcare scheme” for ”childcarescheme”, ”health care scheme” for ”healthcarescheme”, ”childcare vouchers” for ”childcarevouchers”, ”child care vouchers” for ”childcarevouchers”, ”private medical” for ”privatemedical”, ”dental vision insurance” for ”dentalvisioninsurance”, ”medical plan” for ”medicalplan”, ”private health” for ”privatehealth”, ”health wellbeing plans” for ”healthwellbeingplans”, ”private health insurance” for ”privatehealthinsurance”,

”life assurance” for ”lifeassurance”, ”private healthcare scheme” for ”healthcarescheme”, ”health plan” for ”healthplan”, ”cash plan” for ”cashplan”, ”optical dental” for ”opticaldental”, ”life insurance” for ”lifeinsurance”, ”eye care cover” for ”eyecarecover”, ”vision insurance” for ”visioninsurance”, ”flu vouchers” for ”fluvouchers”, ”dental insurance” for ”dentalinsurance”, ”health insurance” for ”healthinsurance”, ”free healthcare” for ”freehealthcare”, ”dental plan” for ”dentalplan”, ”discounted insurance” for ”discountedinsurance”, ”health care plan” for ”healthcareplan”, ”dental vision” for ”dentalvision”, ”dental cover” for ”dentalcover”, ”includes dental” for ”includesdental”, ”eyecare voucher” for ”eyecarevoucher ”, ”medicash” for ”medicash”, ”life cover” for ”lifecover”, ”health wellbeing” for ”healthwellbeing”, ”indemnity cover” for ”indemnitycover”, ”annual flu” for ”annualflu”, ”critical illness cover” for ”criticalillnesscover”, ”critical illness scheme” for ”criticalillnessscheme”, ”dental care” for ”dentalcare”, ”dental treatment” for ”dentaltreatment”, ”disability cover” for ”disabilitycover”, ”eye care discount” for ”eyecarediscount”, ”eye care voucher” for ”eyecarevoucher ”, ”welfare package” for ”welfarepackage”, ”health cover” for ”healthcover”, ”travel accident insurance” for ”travelaccidentinsurance”, ”free health” for ”freehealth”, ”free legal advice” for ”freelegaladvice”, ”life accident insurance” for ”lifeaccidentinsurance”, ”medical cover” for ”medicalcover”, ”mental health first” for ”mentalhealthfirst”, ”personal accident insurance” for ”personalaccidentinsurance”, ”subsidised dental” for ”subsidiseddental”, ”family healthcare” for ”familyhealthcare”, ”free medical” for ”freemedical”, ”illness insurance” for ”illnessinsurance”, ”free flu vaccine” for ”freefluvaccine”, ”free eye test” for ”freeeyetest”, ”wellness programme” for ”wellnessprogramme”, ”medical insurance” for ”medicalinsurance”

06 Retirement and savings

”company match values” for ”company values”, ”company match demand” for ”company demand”, ”team company match ” for ”team company ”, ”retirement contribution” for ”retirementcontribution”, ”contribution retirement” for ”retirementcontribution”, ”retirement fund” for ”retirementfund”, ”fund retirement” for ”retirementfund”, ”savings account” for ”savingsaccount”, ”account savings” for ”savingsaccount”, ”employer contribution” for ”employercontribution”, ”contribution employer” for ”employercontribution”, ”employer match” for ”employermatch”, ”match employer” for ”employermatch”, ”company match” for ”companymatch”, ”match company” for ”companymatch”, ”tuition reimbursement” for ”tuitionreimbursement”, ”reimbursement tuition” for ”tuitionreimbursement”, ”savings contribution” for ”savingscontribution”, ”contribution savings” for ”savingscontribution”, ”savings scheme” for ”savingsscheme”, ”scheme savings” for ”savingsscheme”, ”provident contribution” for ”providentcontribution”, ”contribution provident” for ”providentcontribution”, ”provident fund” for ”providentfund”, ”fund provident” for ”providentfund”, ”twice contribute” for ”twicecontribute”, ”double contribution” for ”doublecontribution”, ”life assurance” for ”lifeassurance”

07 Food, subsidies and discounts

”tea coffee” for ”teacoffee”, ”cinemas” for ”cinema”, ”discounts” for ”discount”, ”memberships” for ”membership”, ”subsidized” for ”subsidised”, ”breakfast included” for ”breakfastincluded”,”cater breakfast” for ”caterbreakfast”,”clothes discount” for ”clothesdiscount”, ”company uniform” for ”companyuniform”,”dinner included” for ”dinnerincluded”,”employee discount” for ”employeediscount”, ”food included” for ”foodincluded”,”free breakfast” for ”freebreakfast”,”free dinner” for ”freedinner”, ”free drinks” for ”freedrinks”,”free food” for ”freefood”,”free lunch” for ”freelunch”,”free meal” for ”freemeal”, ”free transportation” for ”freetransportation”,”gym discount” for ”gymdiscount”,”live position” for ”liveposition”, ”lunch included” for ”lunchincluded”,”meal included” for ”mealincluded”,”transportation allowance” for ”transportationallowance”, ”travel subsidy” for ”travelsubsidy”,”cater lunch” for ”caterlunch”,”cater meal” for ”catermeal”, ”commute subsidy” for ”commutesubsidy”,”discount coupon” for ”discountcoupon”,”food coupon" for "foodcoupon", "food stamp" for "foodstamp","food subsidy" for "foodsubsidy","food ticket" for "foodticket","housing included" for "housingincluded", "lunch ticket" for "lunchticket","on site accommodation" for "onsiteaccommodation","rental subsidy" for "rentalsubsidy", "transit subsidy" for "transitsubsidy","transportation subsidy" for "transportationsubsidy","transport voucher" for "transportvoucher", "travel stipend" for "travelstipend","voluntary benefits" for "voluntarybenefits","cash back" for "cashback", "wellbeing offers" for "wellbeingoffers","lease scheme" for "leasescheme","advantage scheme" for "advantagescheme", "afternoon snacks" for "afternoonsnacks","benefits platform" for "benefitsplatform","benefits portal" for "benefitsportal", "bike insurance" for "bikeinsurance","car discount" for "cardiscount","car leasing" for "carleasing", "car loan scheme" for "carloanscheme","cinema tickets" for "cinematickets","comfortable uniform" for "comfortableuniform", "company discount" for "companydiscount","corporate discount" for "corporatediscount","daily fruit" for "dailyfruit", "daily menu" for "dailymenu","deals discount" for "dealsdiscount","death service" for "deathservice","discount car" for "discountcar","discount card" for "discountcard", "discount programme" for "discountprogramme","discount schemes" for "discountschemes","discount travel" for "discounttravel", "discount various" for "discountvarious","discounted cinema" for "discountedcinema","discounted fitness" for "discountedfitness", "discounted food" for "discountedfood","discounted membership" for "discountedmembership","discounted stays" for "discountedstays", "discounted subscription" for "discountedsubscription","drinks fridge" for "drinksfridge","food provided" for "foodprovided","unlimited breakfast" for "unlimitedbreakfast", "free coffee" for "freecoffee","free daily breakfast" for "freedailybreakfast","free fruit" for "freefruit", "free healthy" for "freehealthy","free hot drinks" for "freehotdrinks","free ice cream" for "freeicecream","free membership" for "freemembership", "free refreshments" for "freerefreshments","free soft" for "freesoft","free tea coffee" for "freeteacoffee", "free unhealthy" for "freeunhealthy","free yoga" for "freeyoga","fresh fruit" for "freshfruit","fruit every day" for "fruiteveryday", "fruit refreshments" for "fruitrefreshments","fruit snacks" for "fruitsnacks","fuel card" for "fuelcard","gym discount" for "gymdiscount", "gym membership" for "gymmembership","gym pass" for "gympass","gym subsidies" for "gymsubsidies","healthy snacks" for "healthysnacks", "hotel discount" for "hoteldiscount","hub offers" for "huboffers","leisure discount" for "leisurediscount","leisure savings" for "leisuresavings", "loyalty rewards" for "loyaltyrewards","lunch allowance" for "lunchallowance","meal provided" for "mealprovided", "meals duty" for "mealsduty","meals provided" for "mealsprovided","office snacks" for "officesnacks","online benefits" for "onlinebenefits","online discount" for "onlinediscount", "onshift meals" for "onshiftmeals","refreshments provided" for "refreshmentsprovided","retail discount" for "retaildiscount", "reward platform" for "rewardplatform","rewards points" for "rewardspoints","season tickets" for "seasontickets","shopping vouchers" for "shoppingvouchers","staff discount" for "staffdiscount", "store discount" for "storediscount","subsidised meal" for "subsidisedmeal","subsidised rate" for "subsidisedrate", "swim membership" for "swimmembership","tea coffee" for "teacoffee","travel benefits" for "travelbenefits", "travel discount" for "traveldiscount","travel insurance" for "travelinsurance","wellness portal" for "wellnessportal", "access headspace" for "accessheadspace","beer fridays" for "beerfridays","coffee tea" for "coffeetea", "drinks available" for "drinksavailable","drinks facilities" for "drinksfacilities","fitness classes" for "fitnessclasses", "free beer" for "freebeer","free beverages" for "freebeverages","free tea" for "freetea","fridays drinks" for "fridaysdrinks", "gym nearby" for "gymnearby","headspace subscription" for "headspacesubscription","instore discounts" for "instorediscounts", "meditation app" for "meditationapp","personal parcels" for "personalparcels","pool tables" for "pooltables", "excellent coffee machine" for "excellentcoffeemachine","nespresso coffee machine" for "nespressocoffeemachine", "free coffee machine" for "freecoffeemachine","shopping portal" for "shoppingportal","workday lunches" for "workdaylunches", "yoga sessions" for "yogasessions","discount local" for "discountlocal","free cake" for "freecake", "free tickets" for "freetickets","free use" for "freeuse","regular drinks" for "regulardrinks","weekly yoga" for "weeklyyoga"

08 Office space and amenities

”areas” for ”area”,”comfortable office environment” for ”good office environment”, ”fitness room” for ”fitnessroom”,”game room” for ”gameroom”,”office gym” for ”officegym”,”office yoga” for ”officeyoga”, ”comfortable office” for ”comfortableoffice”,”comfortable space” for ”comfortablespace”,”open space” for ”openspace”, ”natural light” for ”naturallight”,”company cafeteria” for ”companycafeteria”,”conference room facilities” for ”conferenceroomfacilities”, ”amenities” for ”amenities”,”amazing workplace” for ”amazingworkplace”,”brand new office” for ”brandnewoffice”, ”break area” for ”breakarea”,”break rooms” for ”breakrooms”,”breakout area” for ”breakoutarea”,”canteen” for ”canteen”, ”chillout area” for ”chilloutarea”,”comfortable facilities” for ”comfortablefacilities”,”coworking space” for ”coworkingspace”, ”developed office” for ”developedoffice”,”facilities onsite” for ”facilitiesonsite”,”fantastic facilities” for ”fantasticfacilities”, ”fitness studio” for ”fitnessstudio”,”games area” for ”gamesarea”,”games room” for ”gamesroom”,”good facilities” for ”goodfacilities”, ”great facilities” for ”greatfacilities”,”great facilities” for ”greatfacilities”,”great views” for ”greatviews”, ”leisure facilities” for ”leisurefacilities”,”modern facilities” for ”modernfacilities”,”modern office” for ”modernoffice”, ”office includes” for ”officeincludes”,”office modern” for ”officemodern”,”onsite cafeteria” for ”onsitecafeteria”, ”onsite catering” for ”onsitecatering”,”onsite cinema” for ”onsitecinema”,”onsite facilities” for ”onsitefacilities”, ”onsite fitness” for ”onsitefitness”,”onsite gym” for ”onsitegym”,”onsite kitchen” for ”onsitekitchen”, ”onsite nursery” for ”onsitenursery”,”onsite recreation” for ”onsiterecreation”,”onsite restaurant” for ”onsiterestaurant”, ”onsite restroom” for ”onsiterestroom”,”onsite shower” for ”onsiteshower”,”onsite subsidised” for ”onsitesubsidised”, ”onsite swimming” for ”onsiteswimming”,”open plan environment” for ”openplanenvironment”,”open plan office” for ”openplanoffice”, ”prayer room” for ”prayerroom”,”quiet work spaces” for ”quietworkspaces”,”range facilities” for ”rangefacilities”, ”refurbished office” for ”refurbishedoffice”,”relaxation area” for ”relaxationarea”,”roof terrace” for ”roofterrace”, ”rooftop terrace” for ”rooftopterrace”,”self service facilities” for ”selfservicefacilities”,”shower facilities” for ”showerfacilities”, ”shower facility” for ”showerfacility”,”sport facilities” for ”sportfacilities”,”staff rest area” for ”staffrestarea”, ”staff restaurant” for ”staffrestaurant”,”stocked kitchen” for ”stockedkitchen”,”subsidised restaurant” for ”subsidisedrestaurant”, ”table tennis” for ”tabletennis”,”tennis room” for ”tennisroom”,”use microwave” for ”usemicrowave”, ”vaping area” for ”vapingarea”,”wellness room” for ”wellnessroom”,”xbox room” for ”xboxroom”,”staff lockers” for ”stafflockers”, ”excellent facilities” for ”excellentfacilities”,”great offices” for ”greatoffices”,”location also provides” for ”locationalsoprovides”, ”location provides” for ”locationprovides”,”onsite bistro” for ”onsitebistro”,”open roof” for ”openroof”,”quiet working spaces” for ”quietworkingspaces”, ”state art facilities” for ”stateartfacilities”

09 Location and commuting

”spaces” for ”space”, ”city center” for ”citycenter”,”central location” for ”centrallocation”,”excellent location” for ”excellentlocation”, ”good location” for ”goodlocation”,”easy access” for ”easyaccess”,”private parking” for ”privateparking”,”pay parking” for ”payparking”, ”company garage” for ”companygarage”,”private garage” for ”privategarage”,”short commute” for ”shortcommute”,”long commute” for ”longcommute”, ”local shops” for ”localshops”,”access public transport” for ”accesspublictransport”,”based centre” for ”basedcentre”, ”based heart” for ”basedheart”,”bike 2 work” for ”bike2work”,”bike sheds” for ”bikesheds”,”bike storage” for ”bikestorage”, ”bike work scheme” for ”bikeworkscheme”,”bike2 work” for ”bike2work”,”bus links” for ”buslinks”,”bus train” for ”bustrain”, ”car park” for ”carpark”,”car parking” for ”carparking”,”central location” for ”centrallocation”,”centre located” for ”centrelocated”, ”centre location” for ”centrelocation”,”childcare options” for ”childcareoptions”,”commuter links” for ”commuterlinks”, ”conveniently located” for ”convenientlylocated”,”coworking space” for ”coworkingspace”,”cycle scheme” for ”cyclescheme”,”cycle to work” for ”cycletowork”, ”cycle work” for ”cyclework”,”cycling scheme” for ”cyclingscheme”,”discounted bus” for ”discountedbus”,”easily accessible” for ”easilyaccessible”, ”easily commutable" for "easilycommutable","easy commute" for "easycommute","easy reach" for "easyreach", "excellent location" for "excellentlocation","fine location" for "finelocation","free bus" for "freebus", "free carparking" for "freecarparking","free local bus" for "freelocalbus","free parking" for "freeparking", "good location" for "goodlocation","great location" for "greatlocation","great transport" for "greattransport", "leasing deals" for "leasingdeals","links transport" for "linkstransport","local bus" for "localbus", "local transport" for "localtransport","min walk" for "minwalk","minute walk" for "minutewalk", "minutes central" for "minutescentral","near local" for "nearlocal","next central station" for "nextcentralstation", "office heart" for "officeheart","office next" for "officenext","on site park" for "onsitepark","onsite park" for "onsitepark", "outside city centre" for "outsidecitycentre","public transport" for "publictransport","reliable public" for "reliablepublic", "ride2 work" for "ride2work","secure parking" for "secureparking","short walk" for "shortwalk", "site parking" for "siteparking","situated close" for "situatedclose","situated heart" for "situatedheart", "staff parking" for "staffparking","street parking" for "streetparking","tax free bike" for "taxfreebike","town center" for "towncenter", "transport links" for "transportlinks","transportation links" for "transportationlinks","via public transport" for "viapublictransport", "walking distance" for "walkingdistance","well located" for "welllocated","within short distance" for "withinshortdistance", "easiest reach" for "easiestreach","easily commute" for "easilycommute","ideal location" for "ideallocation", "local parking" for "localparking","mainline station" for "mainlinestation","metro station close" for "metrostationclose", "minutes walk" for "minuteswalk","parking available" for "parkingavailable","parking space" for "parkingspace", "peaceful location" for "peacefullocation","perfect commuting" for "perfectcommuting","perfect location" for "perfectlocation", "quick access" for "quickaccess","walkable distance" for "walkabledistance","state art building" for "stateartbuilding","metro station" for "metrostation"

10 Work equipment and technology

"maintain company" for "maintain comp ","maintain equipment" for "maintain equ ","uniforms" for "uniform", "internet coupon" for "internetcoupon", "coupon internet" for "internetcoupon", "internet subsidy" for "internetsubsidy", "subsidy internet" for "internetsubsidy", "technology subsidy" for "technologysubsidy", "subsidy technology" for "technologysubsidy", "home office fund" for "homeofficefund", "fund home office" for "homeofficefund", "telework fund" for "teleworkfund", "fund telework" for "teleworkfund", "telework stipend" for "teleworkstipend", "stipend telework" for "teleworkstipend", "work home fund" for "workhomefund", "fund work home" for "workhomefund", "telework expenses" for "teleworkexpenses", "expenses telework" for "teleworkexpenses", "home office expenses" for "homeofficeexpenses", "expenses home office" for "homeofficeexpenses", "work home expenses" for "workhomeexpenses", "expenses work home" for "workhomeexpenses", "home office allowance" for "homeofficeallowance", "allowance home office" for "homeofficeallowance", "telework allowance" for "teleworkallowance", "allowance telework" for "teleworkallowance", "work home allowance" for "workhomeallowance", "allowance work home" for "workhomeallowance", "telework subsidy" for "teleworksubsidy", "subsidy telework" for "teleworksubsidy", "work home subsidy" for "workhomesubsidy", "subsidy work home" for "workhomesubsidy", "home office equipment" for "homeofficeequipment", "equipment home office" for "homeofficeequipment", "telework equipment" for "teleworkequipment", "equipment telework" for "teleworkequipment", "offer mobile phone" for "offermobilephone", "mobile phone offer" for "offermobilephone", "include mobile phone" for "includemobilephone", "mobile phone include" for "includemobilephone", "offer company cellphone" for "offercompanycellphone", "company cellphone offer" for "offercompanycellphone", "include company cellphone" for "includecompanycellphone", "company cellphone include" for "includecompanycellphone", "offer company computer" for "offercompanycomputer", "company computer offer" for "offercompanycomputer", "include company computer" for "includecompanycomputer", "company computer include" for "includecompanycomputer", "offer company laptop" for "offercompanylaptop", "company laptop offer" for "offercompanylaptop", "include company laptop" for "includecompanylaptop", "company laptop include" for "includecompanylaptop", "close access" for "closeaccess","company car" for "companycar","company credit card" for "companycreditcard", "company debit card" for "companydebitcard","company laptop" for "companylaptop","company phone" for "companyphone", "company uniform" for "companyuniform","company van" for "companyvan","company vehicle" for "companyvehicle", "corporate clothing" for "corporateclothing","dedicated desk" for "dedicateddesk","discounts energy bills" for "discountsenergybills", "equipment provided" for "equipmentprovided","free broadband" for "freebroadband","free mobile" for "freemobile", "free uniform" for "freeuniform","free uniform" for "freeuniform","fuel card" for "fuelcard", "necessary equipment" for "necessaryequipment","paid mileage" for "paidmileage","perkbox" for "perkbox", "perks" for "perks","phone and laptop" for "phoneandlaptop","phone insurance" for "phoneinsurance", "phone laptop" for "phonelaptop","phone provided" for "phoneprovided","protective footwear" for "protectivefootwear", "provision uniform" for "provisionuniform","two monitors" for "twomonitors","two screens" for "twoscreens", "uniform provided" for "uniformprovided","uniform provision" for "uniformprovision","vehicle allowance" for "vehicleallowance", "vehicle scheme" for "vehiclescheme","welcome box" for "welcomebox","protective equipment" for "protectiveequipment", "phone work" for "phonework","car allowance" for "carallowance","discounted broadband" for "discountedbroadband", "fuel allowance" for "fuelallowance","mileage expense" for "mileageexpense","perbox" for "perbox"

11 Work flexibility

"flexible shift" for "flexibleshift", "shift flexible" for "flexibleshift", "flexible time" for "flexibletime", "time flexible" for "flexibletime", "flexible work" for "flexiblework", "work flexible" for "flexiblework", "flexible hour" for "flexiblehour", "hour flexible" for "flexiblehour", "flexible schedule" for "flexibleschedule", "schedule flexible" for "flexibleschedule", "work life balance" for "worklifebalance", "life work balance" for "worklifebalance", "family friendly" for "familyfriendly", "friendly family" for "familyfriendly", "weekly rest" for "weeklyrest", "rest weekly" for "weeklyrest", "flexibility around working" for "flexibilityaroundworking", "flexibility work" for "flexibilitywork", "flexible office hours" for "flexibleofficehours", "flexible working" for "flexibleworking", "flx hour" for "flxhour", "flx schedule" for "flxschedule","flx shift" for "flxshift", "flx time" for "flxtime","flx work" for "flxwork","life balance" for "lifebalance","life work" for "lifework", "flx schedule" for "flxschedule","flx work" for "flxwork","healthy work" for "healthywork", "modern working" for "modernworking","family friendly" for "familyfriendly","work . life" for "work.life", "worklife balance" for "worklifebalance","work life balance" for "worklifebalance","flx option" for "flxoption", "flexible option" for "flexibleoption", "flx shift" for "flexiblework", "flx with" for "flexiblework", "flx basis" for "flexiblework", "flx about" for "flexiblework", "hours are flx" for "flexiblework", "flx annual" for "flexiblework", "is flx" for "flexiblework", "flx schedule" for "flexiblework", "fully flx" for "flexiblework", "flx pay" for "flexiblework", "flx opportunities" for "flexiblework", "flx subcontract" for "flexiblework", "flx between" for "flexiblework", "flx start" for "flexiblework", "flx dayst" for "flexiblework", "flx within" for "flexiblework", "choose flx" for "flexiblework", "times flx" for "flexiblework", "flx schedules" for "flexiblework", "flx vacancies" for "flexiblework", "scheme flx" for "flexiblework", "time flx" for "flexiblework", "flx and optional" for "flexiblework", "flx self employed" for "flexiblework", "flx in relation to" for "flexiblework", "flx job" for "flexiblework", "options flx" for "flexiblework", "flx rotas" for "flexiblework", "flx employment" for "flexiblework", "days are flx" for "flexiblework", "freelance and flx" for "flexiblework"

12 Workplace safety and ergonomics

"safe place" for "safeplace", "place safe" for "safeplace", "safety culture" for "safetyculture", "culture safety" for "safetyculture","employee safety" for "employeesafety", "safety employee" for "employeesafety", "work safety" for "worksafety", "safety work" for "worksafety","safety first" for "safetyfirst", "first safety" for "safetyfirst", "secure offices" for "secureoffices", "safe work" for "safework","health safety" for

”healthsafety”, ”safety rules” for ”safetyrules”, ”secure workplace” for ”secureworkplace”, ”focus health” for ”focushealth”,”safe working” for ”safeworking”, ”workplace safety” for ”workplacesafety”, ”aware health” for ”awarehealth”, ”comply safe” for ”complysafe”, ”safety legislation” for ”safetylegislation”, ”act 1974” for ”act1974”,”fire regulation” for ”fireregulation”, ”safety act” for ”safetyact”,”comply health” for ”complyhealth”, ”compliance health” for ”compliancehealth”, ”understand health” for ”understandhealth”,”requirements health” for ”requirementshealth”, ”safe working” for ”safeworking”,”safety policies” for ”safetypolicies”,”safe office” for ”safeoffice”,”hygienic working” for ”hygienicworking”

13 Job security

”severance pay” for ”severancepay”, ”pay severance” for ”severancepay”, ”severance package” for ”severancepackage”, ”package severance” for ”severancepackage”, ”layoff compensation” for ”layoffcompensation”, ”compensation layoff” for ”layoffcompensation”, ”unemployment insurance” for ”unemploymentinsurance”, ”insurance unemployment” for ”unemploymentinsurance”, ”unemployment benefit” for ”unemploymentbenefit”, ”benefit unemployment” for ”unemploymentbenefit”, ”low turnover” for ”lowturnover”, ”turnover low” for ”lowturnover”, ”job security” for ”jobsecurity”, ”security job” for ”jobsecurity”, ”job stability” for ”jobstability”, ”stability job” for ”jobstability”, ”low turnover” for ”lowturnover”,”job security” for ”jobsecurity”,”job stability” for ”jobstability”, ”indemnity insurance” for ”indemnityinsurance”,”income protection” for ”incomeprotection”,”stable environment” for ”stableenvironment”, ”severance pay” for ”severancepay”,”severance package” for ”severancepackage”,”layoff compensation” for ”layoffcompensation”, ”unemployment insurance” for ”unemploymentinsurance”,”unemployment benefit” for ”unemploymentbenefit”

14 Work environment and culture

”working conditions” for ”work conditions”,”working hour” for ”work hour”,”benefit culture” for ”benefitculture”, ”best workplace” for ”bestworkplace”,”comfort work” for ”comfortwork”,”congenial work” for ”congenialwork”, ”dynamic workplace” for ”dynamicworkplace”,”enjoyable work” for ”enjoyablework”,”excellent environment” for ”excellentenvironment”, ”good atmosphere” for ”goodatmosphere”,”good culture” for ”goodculture”,”good environment” for ”goodenvironment”, ”great place work” for ”greatplacework”,”healthy work” for ”healthywork”,”human values” for ”humanvalues”, ”job recognition” for ”jobrecognition”,”meaningful work” for ”meaningfulwork”,”mission vision” for ”missionvision”, ”positive work” for ”positivework”,”promote equality” for ”promoteequality”,”relaxed environment” for ”relaxedenvironment”, ”relax work” for ”relaxwork”,”strong company” for ”strongcompany”,”warm environment” for ”warmenvironment”, ”young environment” for ”youngenvironment”,”multiculural environment” for ”multiculuralenvironment”,”afterwork drinks” for ”afterworkdrinks”, ”best places work” for ”bestplaceswork”,”best reputation” for ”bestreputation”,”casual dress” for ”casualdress”, ”collaborative culture” for ”collaborativeculture”,”collaborative environment” for ”collaborativeenvironment”, ”company culture” for ”companyculture”,”company events” for ”companyevents”,”company outings” for ”companyoutings”, ”company parties” for ”companyparties”,”company retreats” for ”companyretreats”,”company social events” for ”companysocialevents”, ”company trips” for ”companytrips”,”company values” for ”companyvalues”,”culture matter” for ”culturematter”,”difference life” for ”differencelife”, ”dress code” for ”dresscode”,”employee assistance” for ”employeeassistance”,”employee support” for ”employeesupport”, ”enjoyable working” for ”enjoyableworking”,”excellent workplace” for ”excellentworkplace”,”fantastic working” for ”fantasticworking”, ”free counselling” for ”freecounselling”,”friendly working” for ”friendlyworking”,”funny social events” for ”funnysocialevents”, ”games nights” for ”gamesnights”,”good vibes” for ”goodvibes”,”great working” for ”greatworking”,”great workplace” for ”greatworkplace”, ”huge impact” for ”hugeimpact”,”immediate impact” for ”immediateimpact”,”impact society” for ”impactsociety”, ”inclusive working” for ”inclusiveworking”,”inclusive working” for ”inclusiveworking”,”internal culture” for ”internalculture”, ”job satisfaction” for ”jobsatisfaction”,”lively office” for ”livelyoffice”,”local reputation” for ”localreputation”, ”lovely atmosphere” for ”lovelyatmosphere”,”make difference” for ”makedifference”,”make real difference” for ”makerealdifference”, ”making impact” for ”makingimpact”,”meaningful project” for ”meaningfulproject”,”multi cultural” for ”multicultural”, ”outdoor activities” for ”outdooractivities”,”people lives” for ”peoplelives”,”pleasant office” for ”pleasantoffice”, ”positive culture” for ”positiveculture”,”positive impact” for ”positiveimpact”,”prioritising wellbeing” for ”prioritisingwellbeing”, ”regular social events” for ”regularsocialevents”,”regular socials” for ”regularsocials”,”relaxed atmosphere” for ”relaxedatmosphere”, ”responsible company” for ”responsiblecompany”,”rewarding” for ”rewarding”,”safe environment” for ”safeenvironment”, ”satisfaction job” for ”satisfactionjob”,”sociable office” for ”sociableoffice”,”social activities” for ”socialactivities”, ”social environment” for ”socialenvironment”,”social gatherings” for ”socialgatherings”,”staff events” for ”staffevents”, ”stimulating environment” for ”stimulatingenvironment”,”supportive environment” for ”supportiveenvironment”, ”supportive working” for ”supportiveworking”,”team bonding” for ”teambonding”,”team building” for ”teambuilding”, ”unique environment” for ”uniqueenvironment”,”value diversity” for ”valuediversity”,”values” for ”values”, ”vibrant office” for ”vibrantoffice”,”vibrant workplace” for ”vibrantworkplace”,”wellbeing support” for ”wellbeingsupport”, ”wellbeing workshops” for ”wellbeingworkshops”,”within community” for ”withincommunity”,”work environment” for ”workenvironment”, ”working culture” for ”workingculture”,”working environment” for ”workingenvironment”,”direct impact” for ”directimpact”, ”equal opportunities” for ”equalopportunities”,”fulfilling” for ”fulfilling”,”impact world” for ”impactworld”, ”make impact” for ”makeimpact”,”working culture” for ”workingculture”,”working ethos” for ”workingethos”, ”office culture” for ”officeculture”,”real impact” for ”realimpact”,”significant impact” for ”significantimpact”, ”social work environment” for ”socialworkenvironment”,”strong impact” for ”strongimpact”,”positive change” for ”positivechange”, ”valuemember” for ”valuemember”, ”feelpart” for ”feelpart”,”interdisciplinary” for ”interdisciplinary”,”multidisciplinary” for ”multidisciplinary”, ”dynamicteam” for ”dynamicteam”,”youngteam” for ”youngteam”,”great team” for ”greatteam”,”team lunch” for ”teamlunch”, ”team events” for ”teamevents”,”exciting team” for ”excitingteam”

15 Learning and development

”learning disabilities” for ”learnd disabilities”, ”base merit” for ”basemerit”, ”merit base” for ”basemerit”,”build career” for ”buildcareer”, ”career build” for ”buildcareer”, ”company growth” for ”companygrowth”,”growth company” for ”companygrowth”,”education subsidy” for ”educationsubsidy”,”growth opportunity” for ”growthopportunity”, ”growth possible” for ”growthpossible”,”merit increase” for ”meritincrease”,”merit promotion” for ”meritpromotion”, ”pay education” for ”payeducation”,”personal growth” for ”personalgrowth”,”professional development” for ”professionaldevelopment”, ”professional growth” for ”professionalgrowth”,”promotion process” for ”promotionprocess”,”salary increase” for ”salaryincrease”, ”build skills” for ”buildskills”,”career” for ”career”,”development opportunities” for ”developmentopportunities”, ”development opportunity” for ”developmentopportunity”,”elearning portal” for ”elearningportal”, ”external learning” for ”externallearning”,”growth opportunities” for ”growthopportunities”, ”learning allowance” for ”learningallowance”,”learning development” for ”learningdevelopment”, ”move career forward” for ”movecareerforward”,”personal development” for ”personaldevelopment”, ”professional development” for ”professionaldevelopment”,”promote internally” for ”promoteinternally”, ”promotion opportunities” for ”promotionopportunities”,”promotion prospects” for ”promotionprospects”, ”training” for ”training”,”developing skills” for ”developingskills”,”development support” for ”developmentsupport”, ”ongoing growth” for ”ongoinggrowth”,”continuous opportunities” for ”continuousopportunities”, ”promotion prospects” for ”promotionprospects”,”growing organisation” for ”growingorganisation”, ”growing company” for ”growingcompany”,”opportunities learning” for ”opportunitieslearning”,

”learning development” for ”learningdevelopment”,”learning zone” for ”learningzone”, ”growth company” for ”growthcompany”,”continuous improvement” for ”continuousimprovement”, ”growing business” for ”growingbusiness”,”continuing professional” for ”continuingprofessional”, ”merit based” for ”meritbased”,”opportunities growing” for ”opportunitiesgrowing”, ”opportunities upskill” for ”opportunitiesupskill”,”development opportunities” for ”developmentopportunities”, ”promotion opportunities” for ”promotionopportunities”,”promotion development” for ”promotiondevelopment”, ”teamexpansion” for ”teamexpansion”,”growingteam” for ”growingteam

16 Remote Work availability

”conducted remote” for ”conductedremote”, ”fully remote” for ”fullyremote”, ”home based” for ”homebased”, ”home work” for ”homework”, ”job is virtual” for ”jobisvirtual”, ”location is remote” for ”locationisremote”, ”location remote” for ”locationremote”, ”position is remote” for ”positionisremote”, ”position remote” for ”positionremote”, ”remote work” for ”remotework”, ”role will be remote” for ”rolewillberemote”, ”virtual job” for ”virtualjob”, ”work from home” for ”workfromhome”, ”work home” for ”workhome”, ”work remote” for ”workremote”, ”working remote” for ”workingremote”, ”flx with office” for ”flxwithoffice”, ”remote position” for ”remoteposition”, ”role is remote” for ”roleisremote”, ”remote basis” for ”remotebasis”, ”hybrid work” for ”hybridwork”, ”hybrid remote” for ”hybridremote”, ”tele commuting” for ”telecommuting”, ”tele work” for ”telework ”, ”no telecommuting” for ”no”, ”does not offer teleworker” for ”no”, ”no taking work home” for ”no”, ”take work home” for ”no”

Appendix C - Term Frequency - Inverse Document Frequency (TF-IDF) and Cosine Similarity

TF-IDF

The Term Frequency - Inverse Document Frequency (TF-IDF) is a statistical measure used to evaluate how important a word is. It assigns a weight to each term (word) in a document (vacancy text) in a corpus (set of all vacancies). It comprises two components: Term Frequency (TF) and Inverse Document Frequency (IDF).

Term frequency $tf(t,d)$ is the simplest approach where each term t is assigned a weight equal to the number of times it appears in the document d. Formally, it is defined as:

$$tf(t,d) = \text{frequency of term } t \text{ in document } d \tag{15}$$

Inverse Document Frequency $idf(t)$, on the other hand, is a measure of the amount of information the term provides, i.e., if it's common or rare across all documents. It is the logarithmically scaled inverse fraction of the documents that contain the term. Formally, it is defined as:

$$idf(t) = \log\left(\frac{N}{df(t)}\right) \tag{16}$$

where N is the total number of documents in the corpus and $df(t)$ is the number of documents in the corpus that contain the term t.

The TF-IDF weight of a term is the product of its TF and IDF:

$$tfidf(t,d) = tf(t,d) \times idf(t) \tag{17}$$

Cosine Similarity

In the context of Information Retrieval, each document is represented as a vector, and the similarity between the two documents is quantified using the cosine of the angle between their vectors. Text vectorisation involves transforming textual data into a numerical format so that it can be analyzed mathematically. One common approach is to use a bag-of-words model. In this model, I create a vocabulary of unique terms (words) present in the corpus and represent each document as a vector in this high-dimensional space. Each dimension corresponds to a unique term from the vocabulary.

For instance, consider a corpus with two documents: Document 1: "Cat sits on the mat." and Document 2: "Dog sits on the log.". The vocabulary for this corpus would be cat, dog, sits, on, the, mat, log. Each document is then represented as a 7-dimensional vector based on the frequency of these words. For example, the vector for Document 1 would look like [1, 0, 1, 1, 1, 1, 0], representing the presence of 'cat', 'sits', 'on', 'the', and 'mat', and the absence of 'dog' and 'log'. Once the documents are represented as vectors, the cosine similarity can be calculated. The cosine similarity measures the cosine of the angle between these two vectors. Mathematically, it is the dot product of the vectors divided by the product of their magnitudes (Euclidean lengths). This measure ranges from -1 to 1, where 1 indicates identical vectors, 0 indicates orthogonality (no similarity), and -1 indicates completely opposite vectors.

Given two documents d_1 and d_2, the cosine similarity is defined as follows:

$$sim(d_1, d_2) = \frac{V(d_1) \cdot V(d_2)}{||V(d_1)|| \times ||V(d_2)||} \tag{18}$$

where $V(d_1)$ and $V(d_2)$ are the vector representations of documents d_1 and d_2 respectively, $V(d_1) \cdot V(d_2)$ is the dot product of the vectors, and $||V(d_1)||$ and $||V(d_2)||$ are their Euclidean lengths.

Appendix D - BERT Model for Gender Classification

This appendix provides technical details on the BERT model used to identify implicit gender targeting in job advertisements.

Model Architecture and Implementation

I implemented a fine-tuned BERT model using the pre-trained bert-base-uncased model from the Hugging Face Transformers library. The model architecture consists of the BERT encoder with 12 transformer layers and a classification head for the three categories: female-targeted, male-targeted, and neutral.

During implementation, I added special tokens `[GENDER]` and `[CONTEXT]` to the tokenizer vocabulary to handle gender-specific terms and contextual phrases during preprocessing.

Data Preprocessing and Feature Engineering

The data preprocessing pipeline involved several steps to prepare the text data for model training:

Identifying Explicitly Gendered Advertisements

To create the training dataset, I identified advertisements with explicit gender preferences using two sets of criteria:

Gender Keywords: I used the following sets of keywords to identify potential gender preferences:

- Female indicators: "female", "woman", "women", "girl", "girls", "lady", "ladies", "feminine"
- Male indicators: "male", "males", "man", "men", "guy", "guys", "boy", "boys", "gent", "gents"

Contextual Filtering: To distinguish between references to worker gender versus client gender, I applied contextual rules based on two comprehensive lists:

- Contextual clues that indicate client references (for exclusion):

  "service for", "patients", "assist a", "men client", "support to", "supports women", "supporting", "care for", "care to", "caring for", "client is a", "therapy for", "treatment for", "work for", "work with", "services for", "support for", "assistance to", "assistance for", "programme for", "counselling for", "advice for", "guidance for", "empower", "service is for", "dedicated to", "abuse", "in need", "working with", "pregnant", "isle of man"

- Gender-specific phrases that indicate worker requirements (for inclusion):

  "applicants only", "applicant only", "female only", "men only", "male only", "male support", "recruit", "recruiting for", "be female", "be male", "available for a female", "encourage"

If a gender keyword appeared within the same sentence as a contextual clue (within 10 words before or after), it was assumed to be referring to a client or service user rather than a job applicant, unless a gender-specific phrase was also present indicating an explicit worker requirement.

Text Cleaning and Normalisation

For model training, I applied the following text processing steps:

1. Basic Text Cleaning:
   - Convert to lowercase
   - Remove punctuation and special characters
   - Remove numbers and digit-only tokens
2. Linguistic Processing:
   - Remove common stopwords (e.g., "the", "and", "a")
   - Apply stemming to reduce words to their root forms
   - Remove very short words (fewer than 4 characters)
3. Gender Word Handling:
   - Replace explicit gender keywords with the `[GENDER]` token
   - Replace contextual phrases with the `[CONTEXT]` token

This preprocessing ensures the model learns underlying linguistic patterns rather than relying on explicit gender terms. For example, an original text like:

> "We are looking for female care assistants to provide care for elderly women clients"

Would be transformed to:

"looking [GENDER] care assist provide [CONTEXT] elderly [GENDER] client"

This preprocessing removes stopwords, applies stemming (e.g., "assistants" becomes "assist"), and replaces gender terms and contextual phrases with special tokens.

Training Data Stratification

To prevent occupation-specific biases, I stratified the training data by occupation codes. For each occupation-gender combination, I sampled a minimum of 10 and a maximum of 5,000 advertisements. This balanced approach ensures that occupations with high concentrations of explicit gender preferences (such as care roles) do not dominate the training data.

Model Training

The model was trained using the following procedure:

1. Split the explicitly gendered advertisements into training (70 per cent), validation (15 per cent), and test (15 per cent) sets with stratification
2. Apply class weighting to handle imbalanced classes
3. Train for 3 epochs with early stopping based on validation loss

The training process took a total of 87,614 seconds (approximately 24.3 hours) across three epochs. As shown in Table D.1, the model demonstrated significant improvement during training, with training loss decreasing from 0.23 in the first epoch to 0.08 by the third epoch. Validation accuracy improved from 90.3 to 94.7 per cent, while validation loss decreased from 0.26 to 0.16.

Table D.1: Model Training Metrics by Epoch

Epoch	Training Loss	Validation Loss	Validation Accuracy	Training Time (s)
1	0.230	0.260	0.903	8,440
2	0.164	0.180	0.939	48,905
3	0.084	0.165	0.947	30,269

Model Performance and Validation

The final model achieved an overall accuracy of 94.7 per cent on the test set, with comparable performance across all three classification categories. As shown in Table D.2, the model demonstrated high precision and recall for all target classes, with the neutral class showing the highest F1-score (0.950).

Table D.2: Final Classification Metrics

Class	Precision	Recall	F1-score	Support
Female	0.954	0.938	0.946	13,307
Male	0.934	0.953	0.943	11,787
Neutral	0.952	0.948	0.950	7,481
Macro average	0.947	0.947	0.947	32,575
Weighted average	0.946	0.946	0.946	32,575

The confusion matrix in Table D.3 shows that the model correctly classified the majority of advertisements in each category. The most common misclassifications were between female and male categories, with 613 female-targeted advertisements incorrectly classified as male-targeted, and 400 male-targeted advertisements incorrectly classified as female-targeted.

Table D.3: Confusion Matrix

	Predicted Female	Predicted Male	Predicted Neutral
True Female	12,488	613	206
True Male	400	11,235	152
True Neutral	199	187	7,095

Application to the Analysis Dataset

While the model was trained on explicitly gendered advertisements, for the wage analysis in the main paper, I excluded all advertisements with explicit gender preferences (2.8 per cent of the total dataset). This methodological choice focuses the analysis on implicit rather than explicit gender targeting, which better captures subtle linguistic patterns that might influence wage offers without overt discrimination. After applying the trained model to the remaining advertisements without explicit gender mentions, 28 per cent were classified as female-targeted, 70 per cent as male-targeted, and 2 per cent as neutral. This substantial gender imbalance in the predicted classifications provides important context for interpreting the wage effects across gender categories and reflects implicit gender associations in occupational language.

Performance Across Occupations

To validate the model's consistency across different occupational contexts, I evaluated its performance on subsets of advertisements grouped by Standard Occupational Classification (SOC) codes. Table D.4 presents a selection of occupations with their corresponding accuracy and gender distribution ratios.

The model maintained high accuracy across most occupational categories, with accuracies ranging from 81.8 to 98.6 per cent. Notably, the model performed well in occupations with both balanced and imbalanced gender distributions. For instance, in

Table D.4: Model Performance Across Selected Occupations

SOC	Occupation Type	N	Acc.	F. Ratio	M. Ratio
213	IT Prof.	1,324	0.968	0.554	0.415
223	Healthcare Prof.	2,129	0.964	0.309	0.356
354	Sales Prof.	1,580	0.965	0.478	0.291
356	Public Services	1,630	0.971	0.483	0.270
614	Caring Services	2,324	0.818	0.331	0.338
721	Customer Service	930	0.984	0.859	0.114
821	Transport Drivers	917	0.977	0.110	0.796
927	Elementary Services	1,483	0.986	0.365	0.183

Notes: SOC = Standard Occupational Classification; N = Sample size; Acc. = Accuracy; F. Ratio = Female Ratio; M. Ratio = Male Ratio

Transport Drivers (SOC 821), which has a strong male bias (79.6 per cent), the model achieved 97.7 per cent accuracy. Similarly, in Customer Service Occupations (SOC 721), which has a strong female bias (85.9 per cent), the model achieved 98.4 per cent accuracy.

The lowest accuracy (81.8 per cent) was observed in Caring Personal Services (SOC 614), which has a more balanced gender distribution. This likely reflects the challenging nature of gender classification in care-related contexts, where advertisements frequently mention both male and female clients, potentially creating ambiguous patterns for the model to learn.

Word Importance Analysis

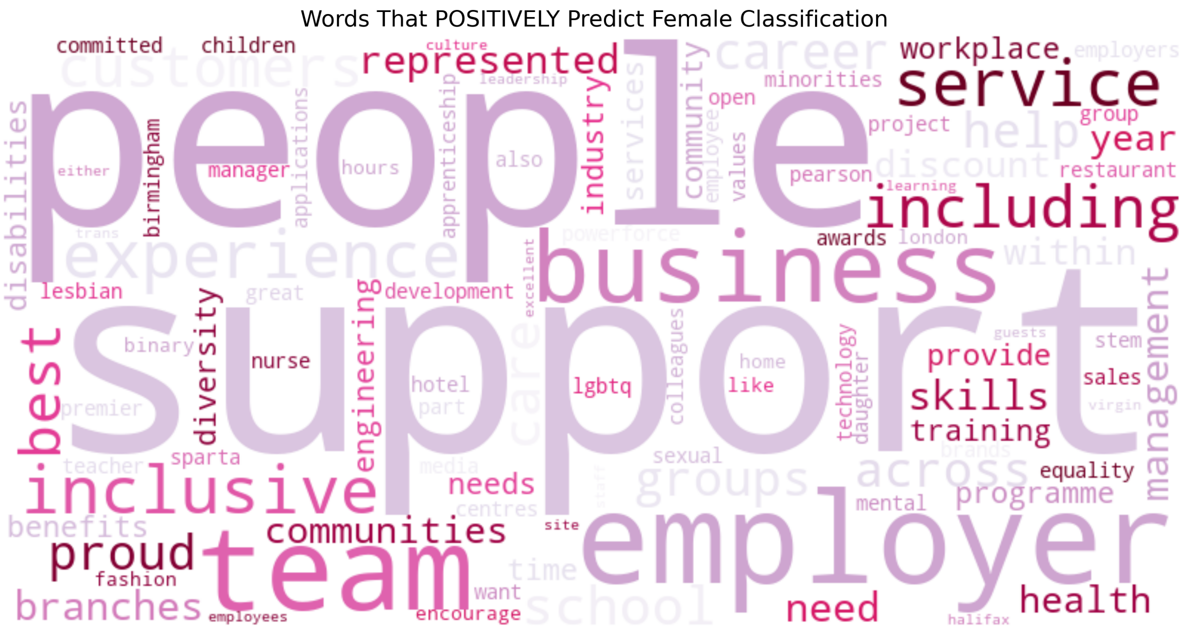
To understand the linguistic features associated with gender targeting, I extracted words with the highest predictive power from the model. Table D.5 presents the top 10 words most predictive of female-targeted, male-targeted, and neutral advertisements.

Table D.5: Top 10 Most Predictive Words by Gender Category

Female Words	Score	Male Words	Score	Neutral Words	Score
support	0.0054	support	0.0057	support	0.0085
people	0.0044	manager	0.0050	requirement	0.0078
employer	0.0034	service	0.0047	people	0.0072
team	0.0029	mail	0.0042	roles	0.0059
business	0.0027	royal	0.0023	employees	0.0049
experience	0.0025	team	0.0021	candidates	0.0045
inclusive	0.0025	delivery	0.0020	avon	0.0041
service	0.0024	people	0.0018	school	0.0036
including	0.0023	skills	0.0017	adult	0.0035
customers	0.0021	driving	0.0016	hotel	0.0035

Notably, some terms appear as predictive across multiple gender categories, though with different weights. “Support” is the top predictor across all three categories, but with a higher weight for neutral advertisements (0.0085) compared to male (0.0057) or female (0.0054) advertisements. This does not indicate a lack of differentiation between

Figure D.1: Word cloud of words predictive of female targeting



categories, but rather illustrates that gender targeting is often expressed through subtle variations in how common terms are used and combined, rather than through entirely separate vocabularies.

The model is likely capturing context-dependent usage patterns—the same word may appear in different linguistic environments across gender categories. For instance, "support" might co-occur with care-related terms in female-targeted advertisements, with technical terms in male-targeted advertisements, and with more formal recruitment language in neutral advertisements. The variation in predictive weights across categories (particularly the higher weight in neutral advertisements) suggests that these contextual differences are meaningful for classification.

"People" also appears in all three categories, while "team" appears in both female and male categories, again indicating that the model relies on contextual patterns rather than isolated keywords. This aligns with BERT's design as a contextual language model that considers words in relation to their surrounding text.

Despite these overlaps, distinctive patterns emerge. Female-targeted advertisements are more associated with terms like "inclusive", "employer", and "business", while male-targeted advertisements show stronger associations with "manager", "mail", "royal", and "driving". Neutral advertisements are distinctively associated with more formal recruitment terminology such as "requirement", "roles", "employees", and "candidates".

Word clouds visualizing these predictive words are shown in Figures D.1, D.2 and D.3.

Figure D.2: Word cloud of words predictive of male targeting

Limitations

The approach has several limitations:

1. Training Data Limitations: The model is trained on explicitly gendered advertisements, which may not be representative of the full advertisement population.

2. Occupation-Specific Language: Despite stratification, occupation-specific language patterns may still influence gender classification. The variability in performance across occupations, as shown in Table D.4, indicates that context-specific factors affect model accuracy.

3. Reinforcement of Stereotypes: The model may reinforce existing gender stereotypes in occupational language, as it learns patterns from human-written advertisements that may contain implicit biases.

4. Binary Framework: The approach implements a primarily binary gender framework that does not account for non-binary gender identities.

5. Temporal Stability: Language patterns associated with gender may evolve over time, affecting the model's performance on advertisements from different time periods.

Despite these limitations, the model provides a valuable tool for examining the relationship between gender targeting, personality trait language, and wage offers in job advertisements.

Figure D.3: Word cloud of words predictive of neutral targeting