Health and Hourly Wages in the UK Labour Market

Evidence from the 2026 Q1 Quarterly Labour Force Survey

Aleksejs Belasovs

2026-06-18

Abstract

Using 7,372 permanent employees aged 21-60 from the UK Quarterly Labour Force Survey (Q1 2026), this paper estimates a semi-log Mincer wage equation to examine whether self-reported health status is associated with hourly wages. An extended model incorporating industry fixed effects, tenure, ethnicity, region, and work-from-home (WFH) status raises the baseline R^2 from 0.221 to 0.305. Under HC1-robust inference, workers with a limiting health condition earn 8.5% less per hour than healthy counterparts; a non-limiting condition carries no significant wage penalty at the 5% level. Structural stability tests reveal intercept differences between health groups but no slope differences, supporting the pooled specification with health dummies. A WFH $\times$ health interaction yields a joint Wald p-value of 0.64, providing no evidence that remote work moderates the health wage penalty. A Heckman two-step selection correction, exploiting the presence of dependent children and marital status as exclusion restrictions, produces a highly significant inverse Mills ratio ($\hat{\lambda} = -0.446$, $p < 0.001$) and reduces the estimated limiting-health penalty to a statistically insignificant coefficient. This suggests the OLS wage gap may partly reflect selective employment into permanent work rather than clear evidence of an independent within-employment wage penalty against health-impaired workers.

1 Introduction

The determinants of wage inequality have long been central to labour economics. Education, experience, and gender feature prominently in the usual Mincer earnings equation [7], but health status has received comparatively less attention in empirical wage equations despite its obvious relevance. Poor health can reduce productivity directly, increase workplace absences, and shift workers toward lower-paying occupations or part-time contracts. At the same time, the composition of health-impaired workers in permanent employment is unlikely to be random as individuals with serious health conditions face a higher hurdle to obtain and retain a permanent job. Those who succeed are therefore unlikely to be a random draw from the health-impaired population: they may differ on unobservable dimensions, including motivation, ability, or willingness to trade pay for security, that also affect wages [5].

This paper asks two questions: first, do permanent employees with self-reported health conditions earn systematically different hourly wages from healthy workers in the 2026 UK labour market? Second, to what extent does the estimated wage gap reflect a genuine within-employment penalty rather than a compositional artefact of who gets hired into permanent work?

We answer these questions using the Q1 2026 wave of the Quarterly Labour Force Survey (QLFS), a nationally representative survey of around 70,000 households and 97,000 individuals per quarter. After restricting to permanent employees aged 21-60, we estimate the baseline Mincer equation augmented with health dummies and then extend it to control for industry, firm size, tenure, ethnicity, region, and WFH status. Heteroscedasticity-robust standard errors are used after Breusch-Pagan and White tests reject homoscedasticity. A Heckman [6] two-step correction addresses the selection of health-impaired workers into permanent employment.

The principal finding is that the OLS wage penalty for limiting health conditions (-8.5%) falls to a statistically insignificant positive coefficient after the Heckman correction. This is more consistent with health-related wage inequality being shaped by barriers to entry into permanent employment than by a clear discriminatory pay differential within employment.

2 Data

2.1 Source and Sample

The data are drawn from the UK Quarterly Labour Force Survey, Q1 (January-March) 2026, a representative random sample of approximately 70,000 households published by the Office for National Statistics. The raw Q1 2026 wave contains around 97,000 individual observations.

The estimation sample is restricted to: (i) employees whose contract type is recorded as permanent (JOBTYP) and excludes temporary, agency and casual workers who have fundamentally different wage-setting. (ii) Individuals aged 21-60 to exclude most full-time students at the lower end and, at the upper end, restrict the sample to prime-age workers and reduce retirement-related selection. (iii) Observations with positive hourly pay and valid education-leaving age (EDAGE). These restrictions leave 7,381 observations with complete wage data; dropping 9 with missing employer tenure gives the estimation sample of **7,372 observations**, used in all models below.

2.2 Variable Construction

The dependent variable is the natural logarithm of gross hourly pay (HOURPAY), which transforms multiplicative wage effects into approximately additive percentage effects. Experience is proxied by $potexp = age - edage$, the conventional Mincer approach. We include a squared term of experience ($potexp^2$) to capture the declining marginal return to experience.

Health status is coded as two binary dummies from the LIMITA variable: `limited` equals one if the respondent reports a health condition that limits the amount of paid work they can do and `nonlimited` equals one if they report a condition that does not limit paid work. The omitted category is healthy workers with no self-reported health condition.

The extended model adds firm size (`small`), four region dummies relative to London and the South East, employer tenure (midpoint-coded from EMPLEN), ethnicity (`nonwhite`), WFH status (`wfh`), and 20 one-letter SIC industry dummies (reference: Section Q, Health and Social Work).

2.3 Summary Statistics

Table 1: Summary statistics: estimation sample (n = 7,372)

	Mean	Std. Dev.	Min	Max
Hourly wage (£)	22.7830	19.6450	0.0700	833.3000
ln(hourly wage)	2.9710	0.5490	-2.6590	6.7250
Potential experience (yrs)	22.3260	11.4260	0.0000	48.0000
Female (= 1)	0.5280	0.4990	0.0000	1.0000
Part-time (= 1)	0.1980	0.3980	0.0000	1.0000
GCSE/O-level (= 1)	0.1100	0.3120	0.0000	1.0000
A-level (= 1)	0.0730	0.2590	0.0000	1.0000
Degree or higher (= 1)	0.4870	0.5000	0.0000	1.0000
Manager/senior official (= 1)	0.3980	0.4900	0.0000	1.0000
Limiting health condition (= 1)	0.0990	0.2980	0.0000	1.0000
Non-limiting health condition (= 1)	0.2390	0.4260	0.0000	1.0000
Employer tenure (yrs)	9.0510	7.9580	0.1250	25.0000
Non-white ethnicity (= 1)	0.1360	0.3430	0.0000	1.0000
Works from home (= 1)	0.3020	0.4590	0.0000	1.0000

The sample is 52.8% female and 19.8% part-time. Nearly half (48.7%) hold a degree or higher qualification, reflecting the restriction to permanent employees in the formal sector. Around 9.9% report a limiting health condition and 23.9% a non-limiting condition; the remainder (66.3%) are healthy. The mean hourly wage is £22.78 (the log-mean implies a geometric mean of £19.51), and 30.2% of workers report working from home at least some of the time. Mean employer tenure is 9.05 years.

3 Econometric Model

3.1 Baseline Specification

The baseline wage equation follows the semi-log Mincer form:

$$\begin{aligned}\ln(hourwage_i) = \beta_0 &+ \beta_1\, potexp_i + \beta_2\, potexp_i^2 + \beta_3\, female_i + \beta_4\, pt_i \\ &+ \beta_5\, gcse_i + \beta_6\, alevel_i + \beta_7\, degree_i + \beta_8\, manager_i \\ &+ \beta_9\, limited_i + \beta_{10}\, nonlimited_i + u_i\end{aligned} \tag{1}$$

The log transformation means coefficients on continuous variables approximate percentage wage changes; for binary dummy variables the exact percentage effect is $100(e^{\hat{\beta}} - 1)\%$. The quadratic in experience captures the concave, hump-shaped experience-earnings profile well-established in the literature [2]. Education dummies are ordered: `gcse` and `alevel` are measured relative to the residual reference group (workers with no formal qualification or with qualifications outside these categories); `degree` captures first and higher degrees. Healthy workers (no health condition) are the omitted reference category, so $\hat{\beta}_9$ and $\hat{\beta}_{10}$ measure the wage differential for limited and non-limited workers relative to healthy workers, conditional on all other covariates.

3.2 Extended Specification

LM tests (reported in Section 5) strongly rejected the baseline specification and motivated the addition of firm size, region, tenure, ethnicity, WFH status, and 20 industry fixed effects:

$$\begin{aligned}\ln(hourwage_i) = \mathbf{x}_i'\beta &+ \gamma_1\, small_i + \gamma_R'\mathbf{region}_i + \gamma_3\, tenure_i \\ &+ \gamma_4\, nonwhite_i + \gamma_5\, wfh_i + \delta'\mathbf{ind}_i + u_i\end{aligned} \tag{2}$$

where $\mathbf{x}_i$ contains the baseline regressors, $\mathbf{region}_i$ the four regional dummies (relative to London/South East), and $\mathbf{ind}_i$ the twenty industry dummies (relative to Section Q, Health and Social Work).

4 Hypotheses

For each regressor we test **individual significance** ($H_0 : \beta_j = 0$). Primary interest lies in β_9 (`limited`) and β_{10} (`nonlimited`): a rejection implies a statistically meaningful wage differential for that health group after controlling for human-capital and demographic characteristics.

Overall significance ($H_0 : \beta_1 = \cdots = \beta_{10} = 0$) is assessed with a joint F-test on all slope coefficients. Rejection confirms that the regressors collectively explain wage variation beyond sampling error [9].

We test **Structural stability** (H_0: the wage equation is the same across health groups) using a Chow [3] F-test and an HC1-robust Wald test on interaction terms. Failure to reject the slopes-only restriction supports the pooled specification with health dummies; rejection would imply that the returns to experience, education, and other covariates differ across health groups.

Finally, we test for **No selection bias** ($H_0 : \lambda = 0$, where λ is the inverse Mills ratio coefficient) using a Heckman [6] two-step correction. Health-impaired workers face a higher barrier to permanent employment, so the observed wage sample may be positively selected. A significant Mills ratio would indicate that OLS estimates of β_9 and β_{10} may be biased and should be interpreted alongside the selection-corrected estimates.

5 Diagnostic Tests

5.1 Functional Form: RESET Test

The Ramsey RESET test [8] adds powers $\hat{y}^2, \hat{y}^3, \hat{y}^4$ to the regression and tests their joint significance. Under correct specification, these terms should add no explanatory power.

Table 2: RESET test (Ramsey 1969), powers 2-4: baseline model

	F-stat	p-value	Verdict
Baseline	4.2316	0.0054	Reject H_0

The baseline model rejects the null of correct functional form ($F = 4.23$, $p = 0.005$), indicating omitted variables or nonlinearities. The MWD test below examines whether the log-linear functional form itself is the source of misspecification; the LM tests subsequently identify specific omitted variable groups.

5.2 MacKinnon-White-Davidson Test

The MacKinnon-White-Davidson [4] test discriminates between a linear level model ($hourwage_i = X_i'\beta + u_i$) and the log-linear form used here ($\ln(hourwage_i) = X_i'\gamma + v_i$). Two auxiliary variables are constructed from the fitted values of each model:

- $z_1 = \ln(\hat{y}_{LIN}) - \hat{y}_{LOG}$: added to the log-linear model to test H_0: the linear specification is correct.
- $z_2 = \exp(\hat{y}_{LOG}) - \hat{y}_{LIN}$: added to the linear model to test H_0: the log-linear specification is correct.

Each null is tested via the t-statistic on the added variable.

Table 3: MacKinnon-White-Davidson test: baseline model

H_0	Added variable	t-stat	p-value	Verdict
Linear is correct	z_1 to log-linear	-2.3053	0.0212	Reject H_0
Log-linear is correct	z_2 to linear	5.3943	< 0.0001	Reject H_0

Both hypotheses are rejected at the 5% level. Rejection of the linear model ($t = -2.31$, $p = 0.021$) confirms that the log-linear form captures wage variation that the linear alternative cannot. Rejection of the log-linear model ($t = 5.39$, $p < 0.001$) echoes the RESET result: the baseline equation omits relevant variables, so neither functional form fits cleanly on its own. As the MWD test does not point unambiguously toward a third alternative, we retain the semi-log specification on theoretical grounds as it is standard in the Mincer earnings literature and produces directly interpretable percentage-effect coefficients. The omitted variable groups identified by the LM tests below are the primary route to reducing the residual misspecification.

5.3 LM Tests for Omitted Variables

The LM (score) test for omitted variables regresses baseline OLS residuals on the baseline regressors plus the candidate variables, and tests the joint significance of the latter using $LM = n \cdot R^2_{aux} \sim \chi^2(q)$, where q is the number of candidate variables.

Table 4: LM tests for omitted variable groups

Variable group	q	LM stat	p-value	Verdict
Regional dummies	4	90.4	< 0.0001	Reject H_0
Firm size	1	125.92	< 0.0001	Reject H_0
Employer tenure	1	34.44	< 0.0001	Reject H_0
Industry (SIC)	20	468.22	< 0.0001	Reject H_0
Ethnicity	1	11.91	0.0006	Reject H_0
Work from home	1	245.54	< 0.0001	Reject H_0

Every group is jointly significant at any conventional level. The largest contributions come from industry ($LM = 468.2$, 20 dummies), WFH status ($LM = 245.5$), and firm size ($LM = 125.9$). These tests motivate the extended specification in Equation 2.

5.4 Heteroscedasticity

Table 5: Heteroscedasticity tests: extended model

Test	LM stat	p-value	Verdict
Breusch-Pagan	75.46	0.0003	Reject H_0
White	686.38	< 0.0001	Reject H_0

Both the Breusch-Pagan ($LM = 75.46$, $p < 0.001$) and White [10] ($LM = 686.38$, $p \approx 0$) tests strongly reject homoscedasticity in the extended model. OLS standard errors are therefore invalid. All inference from this point forward uses HC1 heteroscedasticity-consistent standard errors.

6 Results

6.1 Extended Model Estimates

Table 6: Extended wage equation: HC1 robust standard errors

	Coeff	Std. Err.	p-value
Potential experience	+0.0185***	0.0019	0.0000
Potential experience^2	-0.0003***	0.0000	0.0000
Female	-0.1075***	0.0123	0.0000
Part-time	-0.0468***	0.0152	0.0021

Table 6: Extended wage equation: HC1 robust standard errors

	Coeff	Std. Err.	p-value
GCSE/O-level	-0.0627***	0.0199	0.0016
A-level	+0.0546**	0.0224	0.0146
Degree or higher	+0.2694***	0.0126	0.0000
Manager/senior official	+0.2329***	0.0118	0.0000
Limiting health condition	-0.0888***	0.0201	0.0000
Non-limiting health condition	-0.0213*	0.0121	0.0781
Small firm	-0.1308***	0.0127	0.0000
Rest of England	-0.1675***	0.0261	0.0000
Scotland	-0.1448***	0.0321	0.0000
Wales	-0.2200***	0.0371	0.0000
Northern Ireland	-0.2577***	0.0323	0.0000
Employer tenure (yrs)	+0.0037***	0.0008	0.0000
Non-white ethnicity	-0.0746***	0.0168	0.0000
Works from home	+0.1382***	0.0138	0.0000

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Industry fixed effects (20 SIC section dummies, reference: Section Q - Health and Social Work) included but not reported here; full estimates in Appendix B. $n = 7{,}372$, $R^2 = 0.305$, Adj-$R^2 = 0.301$.

Health coefficients. Workers with a limiting health condition earn 8.5% less per hour than observationally equivalent healthy workers ($\widehat{\beta}_9 = -0.089$, $p < 0.001$). The exact percentage effect is $100(e^{-0.089} - 1) \approx -8.5\%$. Workers with a non-limiting condition earn 2.1% less, but this coefficient is not significant at the 5% level ($p = 0.078$), consistent with the interpretation that a health condition that does not restrict paid work has no material wage effect once other characteristics are held constant.

Human capital. The experience profile is positive and concave as expected: the linear term is positive and the squared term is negative, so the marginal return to experience declines with experience, with the quadratic term implying a peak at $0.0185/(2 \times 0.0003) \approx 31$ years. Degree holders earn 30.9% more than the reference group (no formal or other unclassified qualifications); A-level holders earn 5.6% more. GCSE holders earn 6.1% *less* than the reference group, likely reflecting both the heterogeneous reference category and the concentration of the GCSE group in lower-paying occupations even after controlling for other observables.

Other controls. Female workers earn 10.2% less, part-time workers earn 4.6% less, managers earn

26.2% more, and workers in small firms earn 12.3% less. All regional dummies are negative relative to London and the South East, with the gradient largest for Northern Ireland (-22.7%) and Wales (-19.8%). WFH is associated with a 14.8% wage premium, consistent with WFH being concentrated in higher-skilled, higher-paying occupations [1]. Each additional year of tenure adds 0.37% to wages. Non-white workers earn 7.2% less than white workers, conditional on other characteristics.

6.2 Feature-Screened Addendum

As an additional specification check, I ran a feature-screening exercise over the raw QLFS variables. The screen ranks whole source variables, removes likely wage/pay leakage variables, identifiers, and weights, and then applies backward elimination using BIC. This is appropriate here because many QLFS variables are categorical and would otherwise be split into many dummy columns. The aim is not to let an algorithm choose the causal model, but to see whether the extended specification is missing obvious predictive controls.

The screen points most clearly toward occupational or social class controls, usual hours, and supervisory responsibility. I add these as an addendum to the extended model using `NSECMJ20`, `BUSHR`, and `SUPVIS`. I do not add all screened variables because several are near-duplicates of controls already in the extended model: detailed education duplicates the education dummies, full firm size duplicates the small-firm indicator, and industry and WFH are already included.

Table 7: Feature-screened addendum to the extended model

Model	Limited	p	Non-limited	p	Adj. R2
Extended model	-8.5%	< 0.0001	-2.1%	0.0781	0.301
Feature-screened addendum	-7.6%	< 0.0001	-2.0%	0.0739	0.362

The addendum raises adjusted R^2 from 0.301 to 0.362, which confirms that the screened variables carry meaningful predictive information. The limiting-health coefficient reduces from -8.5% to -7.6% and remains highly significant. The non-limiting health coefficient remains small and statistically insignificant at the 5% level. This strengthens the robustness of the OLS association because the limiting-health gap is not eliminated by adding occupation/class, usual hours, and supervisory responsibility. However, I keep this as an addendum rather than the preferred model because occupational class and hours may be channels through which health affects wages. Including them in the main equation risks controlling away part of the total health-related wage gap. For that reason, the later diagnostics, survey-weighted robustness checks, WFH interaction, and Heckman selection correction continue to use the extended model rather than the feature-screened addendum.

6.3 Multicollinearity

Table 8: Variance Inflation Factors: extended model (main variables)

Variable	VIF
Potential experience	18.43
Potential experience^2	17.96
Female	1.25
Part-time	1.23
GCSE/O-level	1.23
A-level	1.17
Degree or higher	1.53
Manager/senior official	1.13
Limiting health condition	1.07
Non-limiting health condition	1.06
Small firm	1.08
Rest of England	2.85
Scotland	1.97
Wales	1.47
Northern Ireland	1.70
Employer tenure (yrs)	1.27
Non-white ethnicity	1.10
Works from home	1.20

The VIFs for `potexp` and `potexp`^{2} are elevated (~18), but this is a structural consequence of including both a variable and its square in the same model and does not indicate harmful multicollinearity. The coefficients of primary interest (`limited` and `nonlimited`) have VIFs of 1.07 and below which confirms that health status is not collinear with the other regressors. All remaining variables are well below the conventional threshold of 10. Industry dummies (not shown) similarly exhibit VIFs below 2.

6.4 Survey-Weighted Robustness

The QLFS provides survey weights so that sample estimates better represent the UK population. The main models are unweighted because the paper focuses on the conditional wage equation in the observed estimation sample. However, since hourly pay is an income variable, it is important to check whether the health coefficients are sensitive to the income weight `PIWT24`. As a secondary check, we also use the general person weight PWT24. Both weighted models estimate the same extended equation by weighted least squares and retain HC1 robust standard errors.

Table 9: Survey-weighted robustness check

Model	Health group	Coeff	Effect	p-value	n
Unweighted HC1 OLS	Limiting condition	-0.0888***	-8.5%	< 0.0001	7,372
Unweighted HC1 OLS	Non-limiting condition	-0.0213*	-2.1%	0.0781	7,372
Income-weighted WLS (PIWT24)	Limiting condition	-0.0845***	-8.1%	0.0003	7,025
Income-weighted WLS (PIWT24)	Non-limiting condition	-0.0120	-1.2%	0.3555	7,025
Person-weighted WLS (PWT24)	Limiting condition	-0.0965***	-9.2%	< 0.0001	7,372
Person-weighted WLS (PWT24)	Non-limiting condition	-0.0162	-1.6%	0.2326	7,372

The income-weighted estimate for limiting health remains negative and highly significant: the coefficient changes only slightly from -0.0888 in the unweighted model to -0.0845, equivalent to an 8.1% wage penalty. The general person-weighted estimate is also close to the main result at -0.0965, or 9.2%. In both cases, the non-limiting condition remains statistically insignificant at the 5% level. The central OLS finding is therefore not driven by the unweighted composition of the QLFS estimation sample.

7 Structural Stability

A key modelling choice is whether to pool all health groups into a single equation with health dummies, or to estimate separate equations by health status. The Chow [3] test provides a formal assessment.

7.1 Classical Chow F-test

Partitioning the sample into healthy ($n = 4{,}884$) and non-healthy ($n = 2{,}488$) groups and comparing the pooled versus unrestricted RSS using the within-sample formula (same eight-variable formula in all models to ensure comparability):

$$F = \frac{(RSS_R - RSS_{UR})/k}{RSS_{UR}/(n-2k)} = 2.72, \quad p = 0.004$$

This rejects poolability at 5%. However, given that heteroscedasticity is established (Section 5.4), the classical Chow F-statistic is invalid under heteroscedasticity and is reported here for completeness only. The HC1-robust Wald test below is the appropriate inference tool.

7.2 HC1-Robust Interaction Wald Test

Fitting the unrestricted model with full interactions between health group and all slope covariates and testing via an HC1-robust Wald test:

Table 10: HC1-robust Wald test of structural stability

Restriction	df	F-stat	5% critical	p-value	Verdict
Slopes only (16 interactions)	16	1.5712	1.6449	0.0677	Fail to reject H_0
Full (intercepts + slopes)	18	2.4967	1.6053	0.0004	Reject H_0

The slopes-only restriction fails to reject ($F = 1.57$, $p = 0.068$): the returns to experience, education, and other human-capital variables do not differ significantly across health groups. The full restriction, which also tests intercept differences, rejects ($F = 2.50$, $p < 0.001$), but this is driven by the known wage-level gap between healthy and health-impaired workers, which is exactly what the health dummies in the pooled model capture.

Therefore, the pooled model with `limited` and `nonlimited` dummies is the preferred specification. The tests provide no strong evidence that separate equations by health group are needed.

8 Extensions

8.1 Working From Home as a Moderator of the Health Wage Penalty

One motivation for including WFH in the extended model is that remote work may disproportionately benefit health-impaired workers by removing commuting costs, reducing workplace-based fatigue, and allowing flexible scheduling. If so, the health wage penalty should be smaller for WFH workers, and we would expect positive interaction coefficients δ_{lim} and δ_{nl} in:

$$\ln(w_i) = \mathbf{x}_i'\beta + \delta_{lim}(wfh_i \times limited_i) + \delta_{nl}(wfh_i \times nonlimited_i) + u_i$$

Table 11: WFH × health interaction: HC1 robust

Health group	Non-WFH coeff	WFH interaction (δ)	WFH-adjusted coeff
Limiting	-0.0966*** (-9.2%)	+0.0298	-0.0667 (-6.5%)
Non-limiting	-0.0170 (-1.7%)	-0.0134	-0.0304 (-3.0%)

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Joint Wald (H_0: $\delta_{lim} = \delta_{nl} = 0$):

$F = 0.4484$, $p = 0.6387$; fail to reject H_0.

The interaction coefficients are individually and jointly insignificant (joint Wald $F = 0.45$, $p = 0.64$). Although the point estimate suggests a slightly smaller penalty for WFH-enabled limited workers (-6.5% vs -9.2% for non-WFH), this difference is indistinguishable from sampling variation. I find no evidence that the health wage penalty differs by whether the worker is able to work from home. This suggests, but does not prove, that the gap is not primarily driven by physical workplace barriers. This conclusion should nevertheless be treated with caution given the relatively small number of limited WFH workers in the sample.

8.2 Heckman Selection Correction

8.2.1 Motivation

Workers with limiting health conditions are substantially less likely to enter the observed permanent-employee wage sample: the probit coefficient on `lim_s` in the first-stage selection equation is -0.30 ($p < 0.001$), implying a considerable selection probability gap. Those limited workers who are permanently employed are therefore a selected, non-random subset of the limited population. OLS on the employed subsample may conflate the wage effect of health with this selection effect, and the direction of the resulting bias depends on how the unobservables driving employment relate to the unobservables driving wages: if limited workers who obtain permanent jobs are unusually productive or motivated, OLS understates the population penalty; if they instead accept lower pay in exchange for the security of permanent work, OLS overstates the within-employment penalty. The sign of the estimated Mills-ratio coefficient helps distinguish these interpretations within the Heckman framework.

The Heckman [6] two-step estimator is used to adjust for this. In the first step, a probit model of selection into the observed permanent-employee wage sample is estimated on the full working-age population (aged 21-60, $n =$ 39,636). The predicted inverse Mills ratio $\hat{\lambda}_i = \phi(\hat{x}_i'\hat{\gamma})/\Phi(\hat{x}_i'\hat{\gamma})$ is then added to the wage equation as an additional regressor.

8.2.2 Exclusion Restrictions

Credible Heckman estimation requires at least one variable that affects employment probability but not (conditional on the wage regressors) wages. Two exclusion restrictions are used:

- **Dependent children under 16** (`has_dep_child`, from AOHL16): the presence of young children reduces employment probability through childcare costs and time constraints. But after controlling for gender, experience, and part-time status, there is no clear direct effect on the hourly wage of those who do work.
- **Married/cohabiting** (`married`, from MARDY6): partnership status is associated with higher employment probability (possibly reflecting financial motivation or stability). But its direct effect on wages is ambiguous once human-capital controls are included. This restriction is treated as secondary and its validity is assessed by checking whether the main results are materially sensitive to its inclusion.

8.2.3 First Stage

Table 12: First stage: probit of observed wage-sample selection (n = 39,636)

Variable	Coeff	Std. Err.	p-value
Intercept	-2.1344***	0.1189	0.0000
Age	+0.0605***	0.0061	0.0000
Age²	-0.0007***	0.0001	0.0000
Female	+0.0076	0.0148	0.6095
Degree	+0.0995***	0.0168	0.0000
A-level	+0.0921***	0.0303	0.0024
GCSE	+0.0616**	0.0253	0.0149
Limiting condition	-0.2995***	0.0232	0.0000
Non-limiting condition	+0.0854***	0.0182	0.0000
Dependent children†	-0.0806***	0.0172	0.0000
Married/cohabiting†	+0.1386***	0.0176	0.0000

*† Exclusion restriction (excluded from wage equation). Pseudo-R^2 = 0.015. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.*

A limiting health condition reduces the probit index by 0.30 standard deviations ($p < 0.001$), consistent with meaningful health-employment selection. Both exclusion restrictions are significant in the expected directions: dependent children reduce employment probability (-0.081, $p < 0.001$)

and partnership status raises it (+0.139, $p < 0.001$). The modest pseudo-R^2 of 0.015 is typical of cross-sectional employment equations where many unobserved determinants (reservation wages, local labour demand, etc.) are unavailable.

8.2.4 Second Stage and Results

Table 13: Second stage: Heckman-augmented wage equation (HC1 robust)

Variable	OLS coeff	Heckman coeff	Change
Limiting condition	-0.0888***	+0.0222	+0.1109
Non-limiting condition	-0.0213*	-0.0506***	-0.0293
Inverse Mills ratio ($\hat{\lambda}$)	Not estimated	-0.4460***	Not applicable

**** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. HC1 standard errors do not propagate first-stage sampling error; bootstrap standard errors are reported below.*

The inverse Mills ratio is negative and highly significant ($\hat{\lambda} = -0.446$, $p < 0.001$), consistent with substantial selection bias in the OLS estimates. The negative sign indicates that, conditional on observables, workers whose unobservables make permanent employment more likely are estimated to earn less. In this specification, permanent employees do not appear positively selected on unobserved earning power; this is plausible if individuals with high unobserved earning capacity gravitate toward self-employment or contract work offering higher pay or greater flexibility.

Crucially, after adding the Mills ratio, the `limited` coefficient rises from -0.089 to $+0.022$ and becomes statistically insignificant ($p = 0.45$). The mechanics follow directly from the negative Mills coefficient: limited workers have the lowest predicted employment probabilities and hence the largest values of $\hat{\lambda}_i$, so their conditional wage unobservable is the most negative in the sample. Limited workers who do hold permanent jobs are estimated to earn *less* than their observed characteristics predict, consistent with accepting lower pay in exchange for the security, flexibility, or statutory protections of permanent employment. In this specification, OLS does not separate this selection effect from a genuine health penalty and loads it onto the `limited` dummy; after adding $\hat{\lambda}_i$, the remaining within-employment penalty is not statistically distinguishable from zero.

The `nonlimited` coefficient moves from -0.021 to -0.051 and is significant under HC1 inference, although the bootstrap below shows that propagating first-stage sampling error removes this significance. The smaller correction is consistent with the much weaker employment selection for non-limiting conditions (probit coeff: +0.085 vs -0.299 for limiting).

8.2.5 Bootstrap Standard Errors

HC1 standard errors for the Heckman two-step estimator are inconsistent because they treat the inverse Mills ratio as a known regressor, ignoring the sampling variability introduced in the first-stage probit. A pairs bootstrap that re-estimates both stages on each draw provides asymptotically valid standard errors. The procedure resamples $B = 200$ times from the full working-age population ($n = 39{,}636$), re-estimates the probit, recomputes the inverse Mills ratio, and re-estimates the augmented wage equation. The standard deviation of the B coefficient vectors is used as the bootstrap standard error.

Table 14: Heckman estimates: HC1 vs bootstrap standard errors ($B = 200$)

Variable	Coeff	HC1 SE	Bootstrap SE
Limiting condition	+0.0222	0.0297	0.0304
Non-limiting condition	-0.0506	0.0135	0.0581
Inverse Mills ratio ($\hat{\lambda}$)	-0.4460***	0.0892	0.1028

**** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$, based on bootstrap standard errors (normal approximation). Bootstrap and HC1 standard errors agree closely for the `limited` and `mills` coefficients, leaving those conclusions unchanged. For `nonlimited`, however, the bootstrap standard error is roughly four times the HC1 value: the coefficient's significance under HC1 disappears once first-stage sampling error is propagated.*

9 Discussion

The two main empirical results point in different directions regarding the mechanisms behind the health wage gap.

The structural stability tests show no statistically significant difference in the returns to human capital, including experience, education, and managerial status, across health groups. The health wage gap is therefore best represented in this specification as an intercept shift, not as a difference in the wage-experience or wage-education profile. This suggests that health-impaired workers are not penalised at a higher rate as they accumulate human capital, but rather face a level disadvantage.

The Heckman correction, however, reframes what that level disadvantage may represent. The OLS estimate of -8.5% should not be read straightforwardly as a within-employment pay gap: the negative Mills-ratio coefficient implies that limited workers who obtain permanent posts earn less than their observables predict, and OLS may attribute that selection effect to health itself. After controlling for selection, the estimated penalty is no longer statistically significant. This suggests that the OLS estimate is consistent with selection into permanent employment rather than clear evidence of an independent within-employment wage penalty; the disadvantage may lie partly at the entry margin into permanent employment, as suggested by the probit first stage (limiting condition: -0.30, $p < 0.001$).

The null result on the WFH $\times$ health interaction extends this. Even within permanent employment, I find no evidence that WFH status is a meaningful channel through which health affects wages. This may reflect that the WFH wage premium is predominantly an occupational-sorting phenomenon (higher-paying professional roles are more WFH-compatible) rather than a barrier-removal benefit specifically for health-impaired workers.

Several limitations bear on these conclusions. First, the probit pseudo-R^2 is low (0.015), meaning many determinants of employment selection are unobserved; weak selection equations can produce unreliable Mills ratio estimates. Second, two-step HC1 standard errors do not propagate first-stage sampling uncertainty. The $B = 200$ pairs bootstrap reported in Section 8.2 shows the headline conclusion survives this correction. Bootstrap and HC1 standard errors agree closely for the `limited` and `mills` coefficients, but the `nonlimited` bootstrap standard error is roughly four times its HC1 counterpart, so the apparent significance of that coefficient under HC1 should not be over-read. Third, the exclusion restrictions, while plausible, are not bulletproof: marital status in particular may have a direct wage effect through household bargaining or employer preferences.

10 Conclusion

This paper uses a semi-log Mincer wage equation estimated on 7,372 permanent employees from the Q1 2026 QLFS to examine the association between self-reported health status and hourly wages. The extended OLS model finds that workers with a limiting health condition earn 8.5% less per hour than otherwise equivalent healthy workers, while non-limiting conditions carry no statistically significant penalty at the 5% level.

Structural stability testing, using an HC1-robust Wald test in place of the classical Chow test which is invalid under heteroscedasticity, suggests that this gap is mainly an intercept difference: there is no strong evidence that the returns to experience and education differ across health groups, supporting the pooled specification. Augmenting the model with WFH $\times$ health interactions produces a joint Wald p-value of 0.64, providing no evidence that WFH moderates the health wage penalty.

The most substantive finding is that a Heckman selection correction, exploiting variation in dependent children and marital status as exclusion restrictions, produces a highly significant inverse Mills ratio and reduces the estimated limiting-health penalty to a statistically insignificant coefficient (OLS: -8.5%; Heckman: $+2.2\%$, insignificant). A $B = 200$ pairs bootstrap supports this conclusion after allowing for the inconsistency of two-step HC1 standard errors. The OLS wage gap thus appears to reflect selective entry into permanent employment rather than clear evidence of an independent within-employment wage penalty. If this selection interpretation is correct, interventions aimed at reducing the health wage gap may be more effective if targeted at removing barriers to employment entry for health-impaired workers. This can be achieved through disability employment support, flexible hiring practices, or reasonable adjustments rather than focusing only on within-workplace pay equity.

Further Work

Three extensions would strengthen the next version of the analysis. First, the survey-weighted robustness check could be extended to the Heckman selection model, so that both the selection equation and the wage equation incorporate QLFS population weights. Second, the exclusion restrictions could be tested more fully by estimating the Heckman model with dependent children only, marital status only, and both instruments together. This would show whether the reduction of the limiting-health penalty depends on the secondary marital status restriction. Third, the wage equation could be re-estimated after trimming or winsorising extreme hourly wage observations. The current log specification reduces the influence of outliers, but a direct outlier robustness check would make the published version more transparent.

A Variable Definitions

Variable	Definition	QLFS source
`lnhourwage`	ln(gross hourly pay, £)	HOURPAY
`potexp`	Age minus education-leaving age	AGE, EDAGE
`potexp2`	`potexp` squared	Not applicable
`female`	= 1 if female	SEX
`pt`	= 1 if part-time	FTPTWK
`gcse`	= 1 if highest qual. is GCSE/O-level	HIQUAL22
`alevel`	= 1 if highest qual. is A-level	HIQUAL22
`degree`	= 1 if holds first or higher degree	HIQUAL22
`manager`	= 1 if manager or senior official	MANAGER
`limited`	= 1 if health limits paid work	LIMITA
`nonlimited`	= 1 if health condition, no work limit	LIMITA
`small`	= 1 if employer has < 25 employees	MPNR02
`restofeng`	= 1 if region is Rest of England	GOVTOF2
`scotland`	= 1 if region is Scotland	COUNTRY

Variable	Definition	QLFS source
wales	= 1 if region is Wales	COUNTRY
ni	= 1 if region is Northern Ireland	COUNTRY
tenure	Years with current employer (midpoint-coded)	EMPLEN
nonwhite	= 1 if not white ethnicity	ETHUKEUL
wfh	= 1 if works from home or home-based	HOME
sic	One-letter SIC section (20 categories)	INDS07M
has_dep_child	= 1 if has dependent child aged < 16	AOHL16
married	= 1 if married or cohabiting	MARDY6

B Industry Fixed Effects (SIC Section Dummies)

The table below reports the full set of industry coefficients from the extended HC1-robust model. The reference category is Section Q (Health and Social Work), so each coefficient measures the log-wage differential relative to that sector.

Table 15: Industry fixed effects in the extended model, HC1 robust (ref: Section Q - Health and Social Work)

SIC Section	Coeff	Std. Err.	p-value
A - Agriculture, Forestry and Fishing	-0.0001	0.0750	0.9991
B - Mining and Quarrying	+0.3099**	0.1250	0.0131
C - Manufacturing	+0.0929***	0.0226	0.0000
D - Electricity, Gas, Steam and Air Conditioning Supply	+0.2301***	0.0562	0.0000
E - Water Supply, Sewerage and Waste Management	-0.0306	0.0712	0.6677
F - Construction	+0.0855***	0.0279	0.0022
G - Wholesale and Retail Trade	-0.0763***	0.0244	0.0018
H - Transportation and Storage	+0.0828***	0.0279	0.0030
I - Accommodation and Food Service Activities	-0.1891***	0.0344	0.0000
J - Information and Communication	+0.2515***	0.0282	0.0000
K - Financial and Insurance Activities	+0.2203***	0.0289	0.0000
L - Real Estate Activities	+0.0518	0.0557	0.3523
M - Professional, Scientific and Technical Activities	+0.1468***	0.0247	0.0000
N - Administrative and Support Service Activities	+0.0081	0.0342	0.8131
O - Public Administration and Defence	+0.0812***	0.0193	0.0000

P - Education	-0.0048	0.0191	0.8018
R - Arts, Entertainment and Recreation	-0.0729**	0.0354	0.0392
S - Other Service Activities	-0.0262	0.0429	0.5413
T - Activities of Households as Employers	+0.0622	0.1062	0.5579
U - Activities of Extraterritorial Organisations	+0.1168	0.1403	0.4049

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

C Full Diagnostic Output: Baseline Model

Table 16: Baseline model: OLS (for diagnostic reference only)

	Coeff	Std. Err.	p-value
Potential experience	+0.0247***	0.0021	0.0000
Potential experience^2	-0.0004***	0.0000	0.0000
Female	-0.1266***	0.0120	0.0000
Part-time	-0.1159***	0.0154	0.0000
GCSE/O-level	-0.0636***	0.0199	0.0014
A-level	+0.1005***	0.0232	0.0000
Degree or higher	+0.3401***	0.0133	0.0000
Manager/senior official	+0.2383***	0.0120	0.0000
Limiting health condition	-0.1037***	0.0195	0.0000
Non-limiting health condition	-0.0138	0.0136	0.3094
Model summary			
n = 7,372			
$R^2 = 0.2213$			
F = 209.24 ($p < 0.001$)			

References

[1] Luca Bonacini, Giovanni Gallo, and Sergio Scicchitano. “Working from Home and Income Inequality: Risks of a ‘New Normal’ with COVID-19”. In: *Journal of Population Economics* 34 (2021), pp. 303–360.

[2] George J. Borjas. *Labor Economics*. McGraw Hill, 2013.

[3] Gregory C. Chow. “Tests of Equality Between Sets of Coefficients in Two Linear Regressions”. In: *Econometrica* 28.3 (1960), pp. 591–605.

[4] Russell Davidson, James G. MacKinnon, and Halbert White. "Tests for Model Specification in the Presence of Alternative Hypotheses: Some Further Results". In: *Journal of Econometrics* 21.1 (1983), pp. 53–70.

[5] Nick Drydakis. "Health Impairments and Labour Market Outcomes". In: *European Journal of Health Economics* (2009).

[6] James J. Heckman. "Sample Selection Bias as a Specification Error". In: *Econometrica* 47.1 (1979), pp. 153–161.

[7] Jacob A. Mincer. *Schooling, Experience, and Earnings*. Tech. rep. National Bureau of Economic Research, 1974.

[8] J. B. Ramsey. "Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis". In: *Journal of the Royal Statistical Society, Series B* 31.2 (1969), pp. 350–371.

[9] O. Sureiman and C. Mangera. "F-test of Overall Significance in Regression Analysis Simplified". In: *Journal of the Practice of Cardiovascular Sciences* 6 (2020), pp. 116–122.

[10] Halbert White. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity". In: *Econometrica* 48.4 (1980), pp. 817–838.