Haberdashers' School

Occasional Papers Series

Occasional Paper Number Ninety-Nine

Age Pyramids and Job ladders: Climbing US Inequality

Alex Leslie

Emmanuel College, Cambridge. OH.

alexanderaleslie@icloud.com

July 2026

Haberdashers' School Occasional Paper Number Ninety-Nine
July 2026

Age Pyramids and Job Ladders: Climbing US Inequality

Alex Leslie

Abstract

Over the past half century, income inequality in the US has risen markedly. This paper provides empirical evidence that between-occupation differences have increasingly explained income inequality, consistent with a structural shift toward more cognitive work, ultimately disproportionately benefiting the college educated. This paper extends on existing literature by formally testing lifecycle inequality dynamics, demonstrating that within-cohort wage inequality rises rapidly with age before leveling off sharply at 40. Importantly, in line with Guvenen et al. (2022), this paper finds successive cohorts enter the labour market at progressively higher levels of inequality.

1 Literature Review

The table below gives the coauthors and year of publication in the left hand column, followed by an accompanying overview of the paper and its findings, with a focus on how each paper is relevant to mine.

Relevant Literature

Paper	Description
Haltiwanger, Hyatt and Spletzer (2022)	Carry out a variance decomposition separating between-industry and within-industry contributions to wage variation, which is the same technique that I employ. They find that rising earnings inequality from 1996–2018 is increasingly driven by between-industry inequality.

Paper	Description
Acemoglu and Autor (2011)	Document the rise in the college wage premium, replicated in my paper. They discuss explanations for increasing inequality, including labour market polarisation, technological change, and offshoring aligning with my results.
Autor, Levy and Murnane (2003)	Discussion of computerisation and how it complements cognitive tasks, helping to explain trends in the college wage premium and task abstractness highlighted in my paper.
Autor and Dorn (2013)	Analyse labour market polarisation and the growth of low-skill service employment. Importantly for my paper, they classify occupations by routine and abstract task measures.
Goldin and Katz (2008)	Argue that from the late 1970s onwards that educational expansion slowed while skill-biased technological change accelerated providing a useful framework for explaining the change in the college wage premium as well as overall wage inequality.
Deaton and Paxson (1994)	Analyse lifecycle inequality dynamics in income and consumption across multiple countries to test implications of the permanent income hypothesis. They employ cohort-based panel data methods using repeated cross-sections similarly to this paper.
Guvenen, Kaplan, Song and Weidner (2022)	Examine how earnings inequality evolves over the life cycle and compare inequality growth across cohorts. My results are consistent with their finding that younger cohorts enter increasingly unequal labour markets and lifecycle inequality dynamics thereafter remain relatively similar across cohorts. I extend on their work by testing for structural breaks and formally testing the trends in inequality with age that they identify.

2 Data

The main dataset used is the IPUMS Current Population Survey (CPS), which harmonises microdata from the monthly U.S. labor force survey. This is supplemented with Dorn and Autor’s (2013) occupation classification constructed from “task requirements from the fourth edition of the US Department of Labor’s Dictionary of Occupational Titles (US Department of Labor 1977)”. Only data from 1976 onward is used due to incompleteness and/or inconsistency of occupations worked, hours worked and weights inter alia. The table below provides the variables used in my regressions alongside accompanying information, including their use in this project. Moreover, any data cutoffs are discussed.

Variable Definitions

Variable Name	Description
Year	Calendar year corresponding to the CPS survey observation. All results are synthesised to account for the income being reported from the previous calendar year.
lrealwage	Log real wages indexed to 2010 dollars of an individual from the previous calendar year, created from incwage. The bottom cutoff for income was chosen as than the lowest possible full-time working wage over the past 50 years. This was calculated by first taking the lowest federal real minimum wage over the period (which is today's at $7.25). This was then indexed to 2010 dollars ($4.91), and then multiplying by 35 hours per week and 40 weeks per year, giving a value of $6874. Inequality, for the purposes of this paper, will be measured by the variance in real log wages.
Occ90ly	Occupation classification based on the 1990 CPS occupation coding scheme, with data available from 1968 onward. Provides the respondent's reported occupation during the previous calendar year, which is consistent with income data also taken from the previous calendar year. Used to estimate between and within-occupation decompositions.
Occ1990dd	Dorn's occupation classification aggregating U.S. Census occupation codes to a panel of occupations for the 1980, 1990, and 2000 Census, as well as the 2005–2008 ACS. This was used in mapping occupations to task abstractness. Occupations not included in this classification were excluded.
Ind90ly	Industry classification based on the 1990 CPS industry coding scheme. Provides respondents reported industries during the previous calendar year. Used to estimate between-industry and within-industry inequality dynamics over time.
TaskAbstract	A continuous measure of how abstract the work in an occupation is. Higher values indicate more cognitively intensive occupations. Constructed from task requirements from the fourth edition of the US Department of Labor's Dictionary of Occupational Titles.
College	Dummy variable equal to 1 if the value of educ in the CPS dataset is between 110 and 125 inclusive, indicating an individual has completed at least four years of college education.
Fullpart	Dummy variable equal to 1 for individuals working full-time and full-year during the previous calendar year. Only these individuals are considered in this paper.
Sex	Binary variable indicating the respondent's sex.
Race	The respondent's reported race.

Variable Name	Description
Age	Age of the respondent in years at the time of survey collection. Only ages 16-64 are considered in line with Acemolgu and Autor (2011).

3 Methodology

3.1 Occupation and Industry: Variance Decomposition

Following Haltiwanger, Hyatt, and Spletzer (2022), inequality is broken down into the percentage explained by between-industry differences and percentage explained by within-industry differences. To compute this, log real wage is regressed upon all industries, with each industry being a dummy variable.

$$\ln(wage_{it}) = \alpha_t + \sum_{j=1}^{n-1} \beta_{jt} \,\mathrm{Ind}_{ij} + \varepsilon_{it}$$

The R^2 output of this regression for one cross section tells us the percentage of inequality that can be explained by between-industry differences in that year. If R^2 is tracked over time we can see how this percentage evolves over the period. The remaining percentage, 1 - R^2, is necessarily explained by within-industry differences. This same process is then repeated separately for occupations. All regressions are run using the CPS weightings.

The R^2 values were then regressed against time to identify trends in the explanatory power of between industry/occupation differences on inequality. After running this basic OLS model, a quadratic time trend was added, which improved the fit for industry but not for occupation according to the Bayesian Information Criterion (and was hence dropped for occupation). Breusch–Pagan tests do not reject homoskedasticity for both industry and occupation, but Breusch–Godfrey tests strongly indicate serial correlation in the residuals. I therefore report Newey–West errors with three lags to obtain standard errors robust to serial correlation. Finally, the Shapiro-Wilk test was employed on residuals and failed to reject the normality of errors (all diagnostics are reported in the appendix).

3.2 The Cognitive Shift

A central hypothesis in the inequality literature argues there has been a shift from routine manual jobs to more cognitive ones leading to polarisation in salaries (Acemoglu and Autor, 2011). To test this, log real wage is regressed against how abstract the task is (which proxies cognitive requirement) and the R^2 values are plotted over time i.e the proportion of wage variation explained by a job's abstractness over time.

$$\ln(wage_{it}) = \beta_{0t} + \beta_{1t} TaskAbstract_{it} + \varepsilon_{it}$$

To understand the main signature of this potential structural shift, the "college premium" - defined as the additional average pay one receives upon completing a college degree compared to an individual who has not - must be determined. Here, I attempt to estimate the college wage premium while addressing potential problems to causal inference. Log real wage is regressed against a college dummy and controls, with the resulting estimate of the premium being mapped over time.

$$\ln(wage_i) = \alpha + \beta\, College_i + \boldsymbol{X}_i\boldsymbol{\gamma} + u_i \qquad \text{where} \qquad \boldsymbol{X}_i = (Age_i, Age_i^2, Sex_i, Race_i,)$$

The above equation suffers from heteroskedasticity, so robust standard errors are used to correct for this. Variance inflation factors do not indicate severe multicollinearity.

Even with the inclusion of controls, the specification fails the Ramsey RESET test, indicating functional form misspecification. Additionally, β is likely upward biased due to endogeneity stemming from omitted variable bias with data to proxy or measure ability, such as IQ, being unavailable. Suitable instrument variables were searched for including quarter of birth (Angrist and Krueger, 1991), birthplace i.e proximity to a college (Card, 1993) and parental data e.g education, however the CPS lacked the necessary data to make the use of any of these IVs feasible.

However, the Ramsey RESET test can be sensitive and reject for very general reasons (Wooldridge 2012). Thus, despite its limitations, different specifications are reported for the 2019 cross section (the final pre-COVID year). The estimates from equation (2) are then used and the trend of the college premium over time is found.

3.3 Demographic Dynamics

Inequality is shaped by demographic shifts and trends. In line with Guvenen et al. (2022), inequality within each cohort is tracked over time. In this case, the variance in log real wages for the cohort of 24-25 year olds in 1975 is observed and then tracked for the cohort as they age. The cohort was taken over two years to increase sample size. Another cohort of 24-25 year olds is taken in 1985 and the process is repeated. Whilst it is not possible to track individuals over time, the cross-sections of these groups can be treated as a pseudo-panel dataset (Deaton, 1985).

As a result I extend on Guvenen et al. (2022) by regressing variance in income against age for each cohort and testing for a structural break at 40 years old.

$$Inequality_{ct} = \begin{cases} \alpha_c + \beta_1 Age_{ct} + u_{ct}, & \text{if } Age_{ct} < 40 \\ \alpha_c + (\beta_1 + \beta_2) Age_{ct} - 40\beta_2 + u_{ct}, & \text{if } Age_{ct} \geq 40 \end{cases}$$

where c indexes cohorts and t indexes observations every 2 years.

This was accomplished by using a Wald test, testing the hypothesis $\beta_2 = 0$ i.e whether inequality growth within a cohort is constant over the lifecycle. For this analysis, the 2015 cohort is dropped due to insufficient time in the workforce. A joint Wald test then assessed whether the pre-40 slopes were equal across cohorts and separately whether the post-40 slopes were equal across cohorts.

A fixed effects model was then used to remove time invariant, cohort specific factors. No heteroskedasticity was detected when a groupwise modified Wald test was employed. However, autocorrelation was found by the Wooldridge test. To correct for this, clustered standard errors, which assume cross-sectional independence, were used. This assumption could be violated if shocks in a given year, for example a recession, impact multiple cohorts' inequality trajectory; however Peseran's test finds no evidence of such dependence. The results should nonetheless be interpreted cautiously due to the small number of cohorts. Time allowing, I would have increased the number of cohorts tracked. A robust Hausman test confirms FE is preferred to RE. Finally, a first difference equation was run as an additional robustness check.

4 Results

Below are the published results. First, the results from the regressions outlined are printed, followed by an in depth discussion of each regression with accompanying figures.

Main Regression Results

4.1 Between-Occupation/Industry Explanatory Power (R^2 x 100)

Main Regression Results

	(1)	(2)	(3)	(4)	(5)
	Occ OLS	Occ NW	Ind OLS	Ind NW	Ind Quad NW
Time trend	0.102***	0.102***	0.024**	0.024*	-0.119*
	(0.011)	(0.016)	(0.010)	(0.014)	(0.055)
Time trend squared					0.003***
					(0.001)
Constant	31.420***	31.420***	17.502***	18.741***	18.741***
	(0.326)	(0.450)	(0.299)	(0.399)	(0.608)
Observations	50	50	50	50	50
R^2	0.635	0.635	0.101	0.336	0.336

4.2 College Wage Premium for 2019

Main Regression Results

	(1)	(2)	(3)
	Naïve OLS	OLS with controls	Controls Robust
College	0.554***	0.545***	0.543***
	(0.005)	(0.005)	(0.006)
Constant	10.439***	9.093***	9.128***
	(0.003)	(0.028)	(0.032)
Observations	59,986	59,986	59,986
R^2	0.178	0.291	0.292

4.2 Trend in the College Wage Premium

Main Regression Results

	(1)	(2)	(3)
	OLS Linear	OLS Quad	NW Quad
Time trend	0.0052***	0.0124***	0.0124***
	(0.0003)	(0.0008)	(0.0015)
Time trend squared		-0.0001***	-0.0001***
		(0.0000)	(0.0000)
Constant	0.3495***	0.2920***	0.2920***
	(0.0095)	(0.0080)	(0.0186)
Observations	50	50	50
R^2	0.8347	0.9452	0.9452

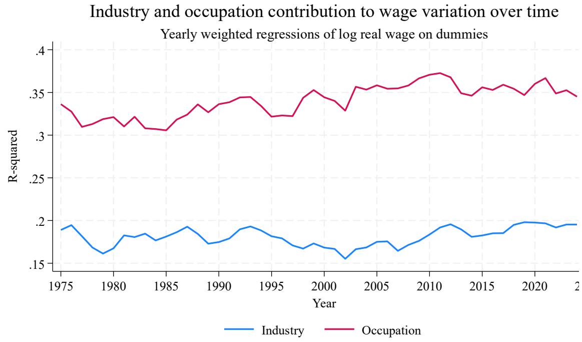
4.3 Lifecycle Inequality Dynamics Across Cohorts

Main Regression Results

Model	Pre-40 slope (β_1)	Post-40 slope ($\beta_1 + \beta_2$)	Structural break (β_2)	Break test p-value	Pre-40 equal p-value	Post-40 equal p-value
Levels piecewise	0.0110***	0.0014	-0.0096***	0.000	0.0296	0.211
Fixed effects	0.0107***	0.0014	-0.0093**	0.002	0.151	0.410
First differences	0.0123***	0.0017	-0.0106***	0.001	0.867	0.903

Notes: Standard errors in parentheses. NW denotes Newey–West standard errors. The final panel reports piecewise lifecycle regressions with a structural break at age 40. Pre-40 and post-40 slopes show biennial inequality growth. The Break test column reports the p-value for the null hypothesis of no structural break. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

4.1



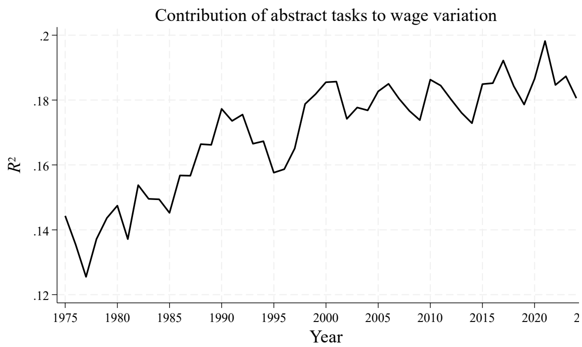
Wage inequality: Variance Decomposition

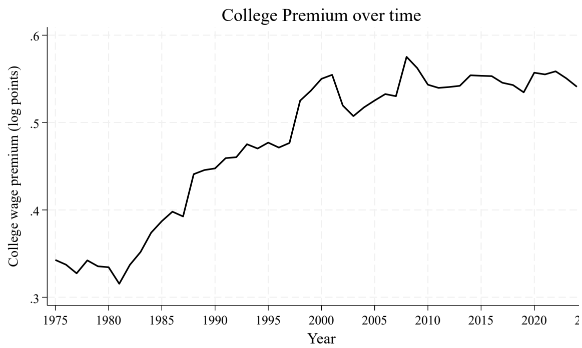
Income inequality as measured by the variance in log real wages, dramatically increased by 53% from 1975 to 2000 before leveling off up until today, marking an overall rise of 45%.

Importantly, between-occupation differences have a statistically significant positive trend explaining an increasing proportion of inequality (5 percentage points more over the period). However, due to the significant absolute rise in wage inequality, within-occupation inequality also rises in absolute terms. Thus within occupation differences accounted for 63.5% of the rise in inequality, with the remaining 36.5% is accounted for by between-occupation differences - highlighting the rising inequality across occupations.

Whilst there is significant evidence at the 5% level of a positive time trend for between-industry differences when running an OLS model, the result is no longer significant at the 5% level when Newey-West errors are used to correct for serial correlation. Regardless, the coefficient on the trend for between-industry differences is negligible at just 0.024% per year over the full period, although over the past 20 years there does appear to be a significant rise in line with the findings of Haltiwanger, Hyatt, and Spletzer's (2022).

4.2

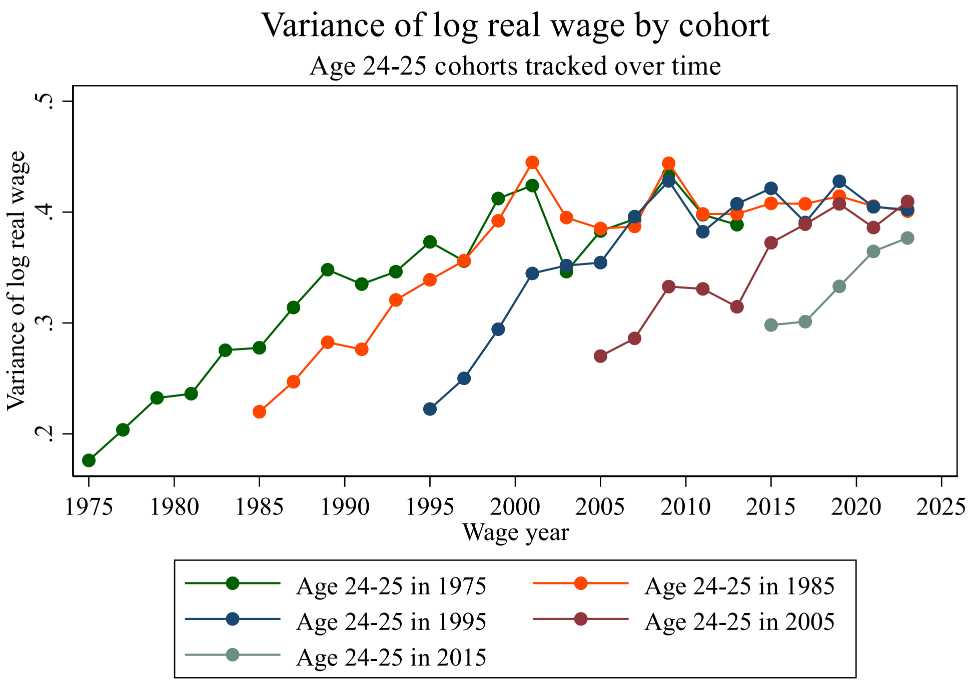




R^2 rises over the period, meaning abstractness of an occupation explains an increasing proportion of inequality. This relationship should be interpreted descriptively, however the rise implies two distinct mechanisms are at work. First, there is a shift toward more cognitive work. This is evidenced by the average abstractness of a job rising on Dorn and Autor's index from 2.70 in 1975 to 3.49 in 2024, marking a 29% rise. Hence, there is a strong case that the reallocation toward more cognitive jobs has structurally changed inequality.

Second, those in more cognitive fields are earning more over time reflected in the rising college premium. As reported, the premium has a statistically significant positive trend, rising from $100 \times (e^{0.338} - 1) = 40.2\%$ in 1975 to $100 \times (e^{0.512} - 1) = 66.9\%$ in 2024. Additionally, the negative quadratic term picks up on the stabilisation of the premium post-2000.

4.3



Each line tracks a cohort initially aged 24–25, observed every two years. Each cohort is taken a decade apart.

The graph shows inequality rising strongly during early working life before flattening later in the lifecycle. If the lines tracking cohorts were parallel, it would mean that inequality would rise to the same extent between cohorts with age. This would mean that the rise in inequality over the past 50 years could instead be seen in increasing inequality between cohorts as they join the labour force.

As reported, the level, first-difference and fixed-effects specifications all provide strong evidence that post-40 lifecycle inequality dynamics are similar across cohorts. Prior to age 40, the levels model rejects equality of cohort slopes at the 5% significance level. However, both the FE and FD specifications fail to reject the null of equal pre-40 slopes across cohorts and are preferred due to their greater robustness to unobserved cohort heterogeneity.

Evidently, inequality for those joining the labour force rises dramatically over the period (seen in the rising starting points for cohorts on the graph), climbing by an astounding 69% from the 1975 to 2015 cohort. Additionally, this paper extends on Guvenen et al. (2022) by finding significant evidence for a structural break at 40.

5 Discussion

Discussion of Results

Topic	Discussion
Between-occupation inequality	Between-occupation differences explain an increasing proportion of income inequality in the US over the past half century (figure 1). This can be explained by the reallocation of labour towards more cognitive occupations which command higher wage premia over time.
Routinization	These findings align with Acemoglu and Autor (2011), who argue rising inequality can be understood through job polarisation, a process where automation hollows out routine middle-skill occupations, coined “routinization” by Autor, Levy and Murnane (2003). Moreover, the process of offshoring contributes to the hollowing out effect. The result is workers being pushed towards opposite ends of the wage spectrum.
College Premium	Polarisation has contributed towards a marked rise in the college wage premium, with college-educated workers increasingly funnelled into higher-paying occupations. Furthermore, as Freeman (1976) forecast, the post-Vietnam War over-supply of graduates corrected itself, while Katz and Goldin (2008) document the rapidly rising demand for skilled labour due to technological change. Together, these forces helped propel the increase in the premium.
Within-occupation inequality	Within-occupation inequality also rose substantially in absolute terms. One possible explanation is demographic ageing. The median age of the workforce increased from 36 in 1978 to 42 in 2024 and, as shown in figure 3, inequality within cohorts rises systematically with age.
The Structural Break	This effect is particularly pronounced before age 40, with the variance of log real wages rising by approximately 0.0055 per year before flattening sharply to around 0.0007 annually thereafter. This pattern is consistent with Ben-Porath’s (1967) model where he predicts declining human capital investment after the mid-career stage, aligning with a structural break in inequality at 40. Huggett et al. (2006) further argue that within-

Topic	Discussion
	cohort inequality expands due to heterogeneity in learning ability, which is most important early in careers.
Inequality patterns across occupations	As documented, inequality across occupations have risen substantially over time. At the extreme end, farmers face inequality six times greater than mail carriers.
Age patterns across occupations and industries	Age distributions differ significantly across occupations and industries. Individuals systematically transition into more abstract occupations with age, particularly during the first decade of working life before this phenomena levels off. Protective services n.e.c. has the youngest mean occupational age at 24, whilst legislators have the oldest at 56. Across industries, dairy product stores exhibit the youngest mean age at 27, compared to 52 for dress-making shops, highlighting substantial heterogeneity in workforce age structures.
Cohort dynamics	These findings support Guvenen et al. (2022), where it is argued that inequality rises similarly within cohorts over the lifecycle, while each successive cohort enters the labour market with a higher initial level of inequality. This reinforces the importance of between-occupation mechanisms, suggesting younger cohorts face an increasingly polarised labour market upon entry.
Limitations	The first key limitation of this analysis was the lack of an available IV or sufficient proxy for education, which could have been used to obtain an unbiased estimate of the college premium. Second, only tracking five cohorts led to less robust results. Future research using richer datasets and with a longer project time horizon may be able to overcome both limitations.

6 Conclusion

This paper affirms the hypothesis that a rising proportion of inequality can be explained by between-occupation differences, with diverging incomes of those within the same occupation still, in absolute terms, driving the majority of inequality growth. This has translated into higher levels of inequality within young adult cohorts, with those holding college degrees benefiting the most. Once working, inequality rises systematically with age and similarly between cohorts. Routinization appears to be the best explanation of the observed phenomena with rapid technological advancement leading to job polarisation and the hollowing-out of routine, middle skilled occupations.

7 Bibliography

Acemoglu, D. and Autor, D. (2011) ‘Skills, Tasks and Technologies: Implications for Employment and Earnings’, in Ashenfelter, O. and Card, D. (eds.) *Handbook of Labor Economics*, Vol. 4B. Amsterdam: Elsevier, pp. 1043–1171.

Angrist, J.D. and Krueger, A.B. (1991) ‘Does Compulsory School Attendance Affect Schooling and Earnings?’, *The Quarterly Journal of Economics*, 106(4), pp. 979–1014.

Autor, D.H. and Dorn, D. (2013) ‘The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market’, *American Economic Review*, 103(5), pp. 1553–1597.

Autor, D.H., Levy, F. and Murnane, R.J. (2003) ‘The Skill Content of Recent Technological Change: An Empirical Exploration’, *The Quarterly Journal of Economics*, 118(4), pp. 1279–1333.

Ben-Porath, Y. (1967) ‘The Production of Human Capital and the Life Cycle of Earnings’, *Journal of Political Economy*, 75(4, Part 1), pp. 352–365.

Card, D. (1993) ‘Using Geographic Variation in College Proximity to Estimate the Return to Schooling’, *NBER Working Paper* No. 4483.

Deaton, A. (1985) ‘Panel Data from Time Series of Cross-Sections’, *Journal of Econometrics*, 30(1–2), pp. 109–126.

Firpo, S., Fortin, N.M. and Lemieux, T. (2011) ‘Occupational Tasks and Changes in the Wage Structure’, *IZA Discussion Paper* No. 5542. Bonn: Institute of Labor Economics (IZA).

Freeman, R.B. (1976) *The Overeducated American*. New York: Academic Press.

Goldin, C. and Katz, L.F. (2008) *The Race between Education and Technology*. Cambridge, MA: Harvard University Press.

Guvenen, F., Kaplan, G., Song, J. and Weidner, J. (2022) ‘Lifetime Earnings in the United States over Six Decades’, *American Economic Journal: Applied Economics*, 14(4), pp. 1–41.

Haltiwanger, J.C., Hyatt, H.R. and Spletzer, J.R. (2022) ‘Industries, Mega Firms, and Increasing Inequality’, *NBER Working Paper* No. 29920. Cambridge, MA: National Bureau of Economic Research.

Huggett, M., Ventura, G. and Yaron, A. (2006) ‘Human Capital and Earnings Distribution Dynamics’, *Journal of Monetary Economics*, 53(2), pp. 265–290.

United States Department of Labor (1977) *Dictionary of Occupational Titles*. 4th edn. Washington, DC: US Government Printing Office.

Wooldridge, J.M. (2012) *Introductory Econometrics: A Modern Approach*. 5th edn. Mason, OH: South-Western Cengage Learning.

US Bureau of Labor Statistics (2026) 'History of Federal Minimum Wage Rates Under the Fair Labor Standards Act, 1938–2009'. Available at: https://www.dol.gov/agencies/whd/minimum-wage/history/chart (Accessed: 10 May 2026).

Federal Reserve Bank of St. Louis (2026) 'Consumer Price Index for All Urban Consumers: All Items in U.S. City Average (CPIAUCSL)'. Available at: https://fred.stlouisfed.org/series/CPIAUCSL (Accessed: 10 May 2026).

OpenAI's ChatGPT was used for technical assistance in debugging Stata syntax and LaTeX formatting.

9 Appendix

Diagnostic Tests

Issue	**Test**	**Results**
Model Specification	Bayesian Information Criterion	Quadratic v Linear Occupation: $-299.6539 > -300.0598$ Industry: $-320.1204 < -308.8488$
Heteroskedasticity	Breusch–Pagan	Occupation: $\chi^2(1) = 0.03, p = 0.8704$ Industry: $\chi^2(1) = 3.05, p = 0.0810$ College: $\chi^2(1) = 354.72, p = 0.0000$
Autocorrelation	Breusch–Godfrey (3 lags)	Occupation: $\chi^2 = 17.197, p = 0.0006$ Industry: $\chi^2 = 27.986, p = 0.0000$
Normality of errors	Shapiro–Wilk	Occupation: $p = 0.2269$ Industry: $p = 0.5811$
Model Specification	Ramsey RESET	College: $p = 0.0000$
Multicollinearity	Variance Inflation Factor	College: Mean VIF = 4.79
Heteroskedasticity	Modified Wald test for groupwise heteroskedasticity	Age inequality: $\chi^2(4) = 1.06, p = 0.9008$
Cross-sectional Dependence	Pesaran	Age inequality: $p = 0.8619$

Issue	Test	Results
FE v RE	Hausman	Age inequality: $\chi^2 = 9.376, p = 0.0092$
Autocorrelation in panel data	Wooldridge	Age inequality: $p = 0.0038$

Notes: In the results column, occupation refers to the regression of an occupation's contribution to explaining inequality on time, and likewise for industry. College refers to the cross-sectional regression used to estimate the college wage premium for the year

2019. Age inequality refers to the diagnostic tests conducted for the lifecycle inequality models in section 4.3.

Table 3: Main Regression Results

5.1 Between-Occupation/Industry Explanatory Power (R^2 x 100)

	(1) Occ OLS	(2) Occ NW	(3) Ind OLS	(4) Ind NW	(5) Ind Quad NW
Time trend	0.102***	0.102***	0.024**	0.024*	-0.119*
	(0.011)	(0.016)	(0.010)	(0.014)	(0.055)
Time trend squared					0.003***
					(0.001)
Constant	31.420***	31.420***	17.502***	18.741***	18.741***
	(0.326)	(0.450)	(0.299)	(0.399)	(0.608)
Observations	50	50	50	50	50
R^2	0.635	0.635	0.101	0.336	0.336

5.2 College Wage Premium for 2019

	(1) Naïve OLS	(2) OLS with controls	(3) Controls Robust
College	0.554***	0.545***	0.543***
	(0.005)	(0.005)	(0.006)
Constant	10.439***	9.093***	9.128***
	(0.003)	(0.028)	(0.032)
Observations	59,986	59,986	59,986
R^2	0.178	0.291	0.292

5.2 Trend in the College Wage Premium

	(1) OLS Linear	(2) OLS Quad	(3) NW Quad
Time trend	0.0052***	0.0124***	0.0124***
	(0.0003)	(0.0008)	(0.0015)
Time trend squared		-0.0001***	-0.0001***
		(0.0000)	(0.0000)
Constant	0.3495***	0.2920***	0.2920***
	(0.0095)	(0.0080)	(0.0186)
Observations	50	50	50
R^2	0.8347	0.9452	0.9452

5.3 Lifecycle Inequality Dynamics Across Cohorts

Model	Pre-40 slope (β_1)	Post-40 slope ($\beta_1 + \beta_2$)	Structural break (β_2)	Break test p-value	Pre-40 equal p-value	Post-40 equal p-value
Levels piecewise	0.0110***	0.0014	-0.0096***	0.000	0.0296	0.211
Fixed effects	0.0107***	0.0014	-0.0093**	0.002	0.151	0.410
First differences	0.0123***	0.0017	-0.0106***	0.001	0.867	0.903

Notes: Standard errors in parentheses. NW denotes Newey–West standard errors. The final panel reports piecewise lifecycle regressions with a structural break at age 40. Pre-40 and post-40 slopes show biennial inequality growth. The Break test column reports the p-value for the null hypothesis of no structural break. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table 5: Diagnostic Tests

Issue	Test	Results
Model Specification	Bayesian Information Criterion	Quadratic v Linear Occupation: $-299.6539 > -300.0598$ Industry: $-320.1204 < -308.8488$
Heteroskedasticity	Breusch–Pagan	Occupation: $\chi^2(1) = 0.03$, $p = 0.8704$ Industry: $\chi^2(1) = 3.05$, $p = 0.0810$ College: $\chi^2(1) = 354.72$, $p = 0.0000$
Autocorrelation	Breusch–Godfrey (3 lags)	Occupation: $\chi^2 = 17.197$, $p = 0.0006$ Industry: $\chi^2 = 27.986$, $p = 0.0000$
Normality of errors	Shapiro–Wilk	Occupation: $p = 0.2269$ Industry: $p = 0.5811$
Model Specification	Ramsey RESET	College: $p = 0.0000$
Multicollinearity	Variance Inflation Factor	College: Mean VIF $= 4.79$
Heteroskedasticity	Modified Wald test for groupwise heteroskedasticity	Age inequality: $\chi^2(4) = 1.06$, $p = 0.9008$
Cross-sectional Dependence	Pesaran	Age inequality: $p = 0.8619$
FE v RE	Hausman	Age inequality: $\chi^2 = 9.376$, $p = 0.0092$
Autocorrelation in panel data	Wooldridge	Age inequality: $p = 0.0038$

Notes: In the results column, occupation refers to the regression of an occupation's contribution to explaining inequality on time, and likewise for industry. College refers to the cross-sectional regression used to estimate the college wage premium for the year 2019. Age inequality refers to the diagnostic tests conducted for the lifecycle inequality models in section 5.3.