Empirical Economics
https://doi.org/10.1007/s00181-023-02394-x

REPLICATION STUDY

The puzzle of changes in employment and wages in routine task-intensive occupations

Pallab Ghosh[1] · **Zexuan Liu**[2]

Received: 23 December 2020 / Accepted: 16 February 2023

Abstract

Autor and Dorn (Am Econ Rev 103(5):1553–1597, 2013) provide an explanation of the polarization of US employment and wages for the period 1980–2005. Using the 1980 Census and 2005 American Community Survey data, this study replicates the estimation results of Autor and Dorn (2013) for employment polarization in all major occupation groups and qualitatively matches the wage polarization results. Also, we investigate the puzzle of why employment and wages changed in opposite directions only in clerical and administrative support occupations in 1980–2005.

Keywords Routine task-intensive occupations · Routine employment share · Wage polarization

Mathematics Subject Classification J24 · J31 · R23

1 Introduction

Although Autor and Dorn (2013) test a hypothesis to explain the concurrent polarization of employment and wages, a question remains unanswered: Why did employment and wages change in the opposite direction only in clerical and administrative support occupations in 1980–2005?[1] The purpose of this replication study is to investigate

[1] Autor and Dorn (2013) show that only clerical employment and wages moved strongly in opposite directions, exhibiting a pattern opposite to those of the other five broadly classified occupations that the authors consider.

✉ Pallab Ghosh
pallab.ghosh@ou.edu

Zexuan Liu
zxliu@nau.edu.cn

1 Department of Economics. 308 Cate Center Drive, Room 158 Cate 1 (CCD1), University of Oklahoma, Norman, OK 73072, USA

2 Institute of Politics and Economics, Nanjing Audit University, Nanjing 211815, Jiangsu Province, People's Republic of China

Published online: 28 February 2023

the robustness of two major findings of Autor and Dorn (2013) related to changes in occupational employment shares and wage levels during 1980–2005. Additionally, using the researchers' framework and data, we attempt to understand why in routine task-intensive occupations such as clerical, administrative support, and retail sales, employment and wages changed in opposite directions.

During the period 1980–2005, the employment share of clerical, administrative support, and retail sales occupations declined about 8%, whereas hourly wages increased approximately 7.4%.[2] To explain these results Autor and Dorn (2013) argue that the increase in wages of clerical workers in that period was probably not due to the rise in demand for routine task-based jobs. Hence, the wage growth in clerical, administrative support, and retail sales occupations is most likely due to the selection of displaced workers from routine task-intensive occupations from 1980 to 2005. One of the primary motivations of this study is to test this hypothesis empirically and also provide supportive empirical evidence.

We contribute to analyzing this puzzle of inconclusive evidence of falling demand for highly routine task-based occupations by studying the relative displacement of skilled and unskilled workers within such occupations. Also, we estimate the impact of the 1980's routine occupation share of a local labor market on the wage growth from 1980–2005 in various quantiles of the wage distribution function. First, we show that only the least productive non-college-educated workers were displaced in the clerical occupations. In contrast, the employment share of relatively more skilled workers increased during the same period. Hence, the workers' average productivity in clerical and administrative service occupations increased, explaining why the correlation between changes in aggregate employment and wages is positive only in clerical occupations for the period 1980–2005.

Second, we empirically test the hypothesis that the least productive non-college-educated workers were displaced in the clerical occupations because the rapid increase in personal computer adoption rate in work during 1980–2005 replaced many routine task-intensive clerical jobs which are in general performed by those low-skill workers. Third, using the Chetverikov et al. (2016)'s instrumental variable quantile regression (IVQR) method, we provide additional information on why the average wage in clerical and administrative support occupations increased during 1980–2005. The IVQR estimates suggest that the wage gains of the relatively more productive non-college-educated workers who belong to the upper-tail of the wage distribution are approximately 57% more compare to the median workers. As a result, the average wage of the least productive low-skilled clerical workers increased during the period 1980–2005.

Fourth, consistent with the finding of non-random selection of displacement of noncollege workers in clerical occupations, the IVQR results also show that the wages of the least productive non-college-educated workers marginally decreased during the period 1980–2005. In contrast, the decline in wages of least efficient noncollege workers in two other high routine content occupation groups are sufficiently large to lead to a negative wage effect, which Autor and Dorn (2013) find in their Table 7

[2] As shown in Appendix Table 1, the employment share of clerical, administrative support, and retail sales occupations fell from 22.2 to 20.4, and hourly wages increased from 2.42 to 2.60.

Panel B. These results also help to explain why the wage effect of the 1980's routine occupation share is positive only in clerical and administrative support occupations in 1980–2005 and negative in two other high routine content occupation groups, precision production and craft, and machine operators and assemblers.

To summarize, this study replicates the Autor and Dorn (2013) estimation results of wage and employment polarization. Also, we provide empirical evidence of non-random selection of workers displaced in highly routine task-intensive occupations during the period 1980–2005, which can entangle the puzzle of why wages increased despite routine employment shares declined only in administrative support and retail occupations. Furthermore, using the IVQR method, we find that the average wage growth from 1980–2005 in clerical occupations is positive because workers who belong in the wage distribution's upper-tail experienced significant wage growth.

The outline of the paper is as follows. Before presenting our empirical results in Sect. 3, we first describe the Autor and Dorn (2013)'s framework and Chetverikov et al. (2016)'s instrumental variable quantile regression model in Sect. 2 since we use these two approaches to obtain the estimation results. We conclude the paper in Sect. 4.

2 Econometric model

To investigate the question, why employment and wages change in the opposite direction only in clerical and administrative support occupations, first, we use the empirical model Autor and Dorn (2013) have used for explaining the job polarization in the U.S. from 1980 to 2005. Next, we move beyond the linear regression model and implement Chetverikov et al. (2016)'s IVQR method to provide supportive evidence of our empirical findings from the Autor and Dorn (2013) approach.

2.1 Autor and Dorn (2013) framework: employment and wage changes in major occupations

Autor and Dorn (2013) use census microdata to calculate the mean wage within each commuting zone (CZ) and find that employment share declines in production and operation occupations among low-skill males and clerical and sales staff among non-college-educated female workers. However, to explain the differential wage changes among unskilled workers in low-skill service and routine task-intensive occupations, Autor and Dorn (2013) use individual-level data from 1980 Census and 2005 American Community Survey (ACS) data and estimate the following wage equation using 2SLS:

$$\ln w_{ijskt} = \gamma_{jk} + \lambda_k \left\{ RSH_{j,1980} \times 1\left[t = 2005\right] \right\} + X_i' \beta_t + \delta_{kt} + \phi_{st} + e_{ijkt} \quad (1)$$

where subscript i denotes workers, j represents commuting zones, s denotes states, k is the occupation, and the time period is $t \in \{1980, 2005\}$. $RSH_{j,1980}$ denotes the 1980's employment share in routine task-intensive occupations in a commuting zone j. The notation γ_{jk} denotes commuting zone-occupation group fixed effects and δ_{kt} and ϕ_{st} are the time trends for occupation groups and states.

The set of control variables X includes an intercept, an interaction between the time dummy and the share of workers in an occupation group whose 1980 wage was below the federal or state minimum wage of 2005, nine dummies for education levels, a quadratic term of potential experience, dummies for married, nonwhite, and foreign-born, and interactions of all individual-level controls with the time dummy. The pooled model includes a female dummy and its interaction term with the time dummy. Hourly wages are defined as yearly wage or salary income divided by the product of weeks worked and usual weekly hours. The coefficient of interest λ_k measures the impact of a CZ's routine employment share in 1980 on wage growth during 1980–2005, and the model is estimated for each occupation group k separately.[3] Following Autor and Dorn (2013), we have deflated the hourly wage variable by the national Personal Consumption Expenditure index, which does not account for local or regional price levels. We report the CZ-level clustered standard errors because RSH does not vary within commuting zones.

To address the endogeneity concern regarding what determines the differences in production structure across CZs, Autor and Dorn (2013) consider the industry structure ($E_{l,j,1950}$) in the 1950 s to estimate the long-run quasi-fixed component of routine occupation share ($\widetilde{RSH}$) for each CZ by using the following expression:

$$\widetilde{RSH}_j = \sum_{l=1}^{L} E_{l,j,1950} \times R_{l,-j,1950} \tag{2}$$

where $R_{l,-j,1950}$ denotes the national routine occupation share in industry l in the 1950 s.[4] This measure is most likely a valid instrument because it was determined three decades prior to 1980, and hence, we expect it to be uncorrelated with unobserved local labor market factors reflected in e_{ijkt} in Eq. (1). Autor and Dorn (2013) show that this instrument has strong predictive power for explaining regional variation in routine occupation shares in 1980.

To validate our empirical approach for investing the puzzle Autor and Dorn (2013) point out in their Table 7, first we use the regression model in Eq. (1) and instrument, $\widetilde{RSH}_j$ to replicate their empirical results. Autor and Dorn (2013) provide both data and Stata codes for all the tables in their article except Table 7, probably because of the massive size of the data file. That is why we show the replication results for Tables 1–6 of Autor and Dorn (2013) in the Appendix and focus only on Table 7 of the original article, which shows the employment and wage polarization results. Specifically, in Table 7 of the original article, Autor and Dorn (2013) show the impact of the 1980's routine occupation share in a CZ on the changes in employment and wages of non-college-educated workers in six major occupation groups.

To replicate the Autor and Dorn (2013)'s Table 7 results, we draw observations from the 1980 Census and 2005 American Community Survey (ACS) and exclude self-

[3] The six occupation groups are (i) service, (ii) transportation, construction, mechanics, mining, and farming, (iii) managers, professional, finance, and public safety, (iv) administrative support and retail sales, (v) precision and production craft workers, and (vi) machine operators and assemblers.

[4] In Eq. (2), the subscript $-j$ implies that to calculate $\widetilde{RSH}_j$, Autor and Dorn (2013) do not include the state s to which the CZ j belongs.

employed workers and farm-workers. We also follow all the data cleaning procedures that are described in the Appendix of Autor and Dorn (2013).[5] In Table 1, we report Autor and Dorn (2013)'s Table 7 replication results. The two panels of Table 1 show the impact of CZ's routine-intensity on employment and wage growth during 1980–2005 estimated using Eq. (1). As shown in the upper panel of Table 1, we have matched all the coefficients and standard errors of the original article's employment polarization results.[6] Thus, similar to Autor and Dorn (2013), we also find that between 1980 and 2005, noncollege employment share within a commuting zone declines in all three high routine intensive occupations: administrative support, production, and machine operators. The decline in employment share was more pronounced in production and operator occupations among men and in clerical and sales occupations among women.

The lower panel of Table 1 shows the replication results of wage polarization, which is shown in the lower panel of Table 7 in Autor and Dorn (2013). We obtain those estimates using the 2SLS regression models for six occupation groups separately. The 2SLS point estimate of 0.35 for clerical, administrative support and retail sales occupations implies that a CZ's 10 percentage point higher routine employment share in 1980 induced a 3.5 percentage point of wage growth of clerical workers during the period 1980–2005. Similar to Autor and Dorn (2013), we have also found that the wage effects are uniformly positive for all three occupation groups with low routine task content in columns 1–3. However, the point estimates for transport, construct, mechanics, mining, and farm occupations reported in column 2 are not significant. In sum, the three occupation categories with low routine content experienced heterogeneous growth in noncollege employment and wages during 1980–2005.

The wage estimates for occupations with high routine content experienced differential growth than the three low routine content occupation groups. As shown in column 6 of Panel B in Table 1, the wage effects for operative and assembler occupations are negative and similar in magnitude (but opposite in sign) to the estimates for administrative support occupations. The column 5 wage estimates for precision production and craft occupations are uniformly negative but are not statistically significant. Therefore, similar to Autor and Dorn (2013) Table 7 results, we also find that the clerical employment and wages changed in opposite directions in routine task-intensive labor markets, and this pattern of change is incompatible with the other five occupations groups.

2.2 Instrumental variable quantile regression models

The motivation of using the IVQR approach is to estimate the heterogeneous impacts of the initial routine employment share in a local labor market on the wage growth across the wage distribution function of clerical workers. We use these distributional wage effect results to infer the average wage effect estimated by Autor and Dorn (2013) in Table 7. Furthermore, estimating the distributional wage effects for two other high routine content occupation groups, precision production and craft, and machine operators and assemblers, we investigate why the wage effect in routine occupation

[5] We briefly describe the data cleaning procedures in the "Appendix".

[6] Autor and Dorn (2013) show employment polarization results in Table 7 panel A.

Table 1 Routine employment share and change in occupational employment shares and wage levels within commuting zones, 1980–2005: 2SLS and reduced Form OLS estimates (replication of Autor and Dorn (2013) Table 7) (Dependent variable: 10 × annual change in share of noncollege employment by occupation; log real hourly wage

		I. Occupations with low routine content			II. Occupations with high routine content		
		Service occs (1)	Transport, construct, mechanics, mining, farm (2)	Managers, prof, tech, finance, public safety (3)	Administrative support, retail sales (4)	Precision production, craft workers (5)	Machine operators, assemblers (6)
Panel A: Change in share of noncollege employment							
(i) All	Share of routine occs	0.192***	0.248***	0.028	− 0.277***	− 0.085***	− 0.107**
		(0.035)	(0.037)	(0.029)	(0.038)	(0.017)	(0.044)
(ii) Males	Share of routine occs	0.210***	0.246***	−0.043	− 0.055*	− 0.145***	− 0.213***
		(0.027)	(0.046)	(0.036)	(0.030)	(0.026)	(0.046)
(iii) Females	Share of routine occs	0.253***	0.002	0.117***	− 0.431***	− 0.028**	0.087
		(0.073)	(0.045)	(0.030)	(0.062)	(0.012)	(0.055)

Table 1 continued

		I. Occupations with low routine content			II. Occupations with high routine content		
		Service occs (1)	Transport, construct, mechanics, mining, farm (2)	Managers, prof, tech, finance, public safety (3)	Administrative support, retail sales (4)	Precision production, craft workers (5)	Machine operators, assemblers (6)
Panel B: log hourly wage of noncollege employment							
(i) All	Share of routine occs × 2005	0.379***	0.036	0.464***	0.350***	− 0.066	− 0.386***
		(0.093)	(0.099)	(0.112)	(0.082)	(0.109)	(0.085)
(ii) Males	Share of routine occs × 2005	0.303**	0.039	0.339**	0.201**	− 0.056	− 0.361***
		(0.127)	(0.097)	(0.148)	(0.098)	(0.139)	(0.105)
(iii) Females	Share of routine occs × 2005	0.321***	0.306*	0.635***	0.473***	− 0.219	− 0.413***
		(0.098)	(0.181)	(0.114)	(0.092)	(0.139)	(0.107)

Panel A: Each coefficient is based on a separate 2SLS regression with $N = 2166$ (3 time periods × 722 commuting zones). Models include an intercept, state dummies, and time dummies, and are weighted by start of period commuting zone share of national population. The routine occupation share is instrumented by interactions between the 1950 industry mix measure interacted with time dummies. Robust standard errors in parentheses are clustered on state. Panel B: Each row presents coefficients from one pooled OLS reduced form regression with $N = 5{,}363{,}963/2{,}844{,}441/2{,}519{,}522$ in rows i/ii/iii. Observations are drawn from the 1980 Census and 2005 American Community Survey (ACS), and exclude self-employed and farm workers. The instrument (share of routine occupations predicted by industry structure in 1950) is interacted with a dummy for the observations of year 2005. All models include an intercept, commuting zone-occupation group fixed effects, time trends for occupation groups and states, an interaction between the time dummy and the share of workers in an occupation group whose 1980 wage was below the federal or state minimum wage of 2005, nine dummies for education levels, a quartic in potential experience, dummies for married, nonwhite, and foreign-born, and interactions of all individual level controls with the time dummy. Pooled sex models also include a female dummy and its interaction with the time dummy. Hourly wages are defined as yearly wage and salary income divided by the product of weeks worked times usual weekly hours. Robust standard errors in parentheses are clustered on commuting zones. Observations are weighted by each worker's share in total labor supply in a given year

***Significant at the 1 percent level
**Significant at the 5 percent level
*Significant at the 10 percent level

share is positive only in clerical occupations among the three high routine content occupation groups.

Using Chetverikov et al. (2016)'s IVQR method, we estimate the impact of 1980's routine intensity in a CZ on the wage growth of clerical workers in various quantiles of the wage distribution function. To apply the IVQR method, we slightly modify the Autor and Dorn (2013)'s linear regression framework and estimate the following equation:

$$\Delta ln\bar{w}_{jt} = \tilde{\gamma}_{jt} + \tilde{\lambda}_{\tau} RSH_{j_{1980}} + X'_{j1980}\tilde{\beta}_{\tau} + \tilde{\delta}_{t=2005} + \tilde{e}_{jt} \tag{3}$$

where the subscript j denotes commuting zones, t represents the decade 1980–2005, and s denotes states.[7] Hence, in this setup, a group is the combination of j and t which represents a given CZ in the given decade 1980–2005.[8]

The variable of interest $RSH_{j_{1980}}$ represents 1980's employment share of routine task-intensive occupations in a commuting zone j, and X_{j1980} are characteristics of the CZs used in Autor and Dorn (2013) and $\tilde{\delta}_{t=2005}$ is a time dummy for the end period 2005. The dependent variable $\Delta ln\bar{w}_{jt}$ denotes the change in log hourly wage from 1980 to 2005 within a commuting zone j. However, to implement the grouped IV quantile regression method, we follow the Chetverikov et al. (2016) approach and replace $\Delta ln\bar{w}_{jt}$ in Eq. (3) with Δlnw^{u}_{jt}, the change in the u-quantile of log wages in the CZ for the decade 1980–2005.

The coefficient of interest $\tilde{\lambda}_{\tau}$ represents the impact of the 1980's RSH on the wage growth during the period 1980–2005 at the τth quantile of the wage distribution. A concern in the above linear quantile regression model is that the 1980's growth of low-skill service occupations in a local labor market is probably correlated with the unobserved labor demand shocks in the subsequent periods. To address this endogeneity concern, we use the Autor and Dorn (2013)'s instrument $\widetilde{RSH}_j$ because, as we discussed, it is most likely uncorrelated with a local labor market's unobserved heterogeneity and strongly correlated with the 1980's RSH in a commuting zone j.[9]

3 Results

To provide a plausible explanation of the countervailing directions of change in clerical employment and wage growth in routine task-intensive labor markets, first, we present some empirical evidence of non-random displacement of least productive noncollege workers from administrative occupation. Furthermore, using the IVQR models, we estimate the distributional impacts of 1980's RSH on wage growth of both college and noncollege workers separately to infer about the average wage effects obtained by Autor and Dorn (2013).

[7] Unlike Eq. (1), we do not have the notation k in Eq. (3) because we estimate this model only for the administrative support occupations.

[8] We have 722 CZ-level observations in both periods 1980 and 2005. Hence, $\Delta ln\bar{w}_{jt}$ also have 722 commuting zone-level observations for the decade 1980–2005.

[9] Note that the instrument $\widetilde{RSH}_g$ is defined in Eq. (2) and also described in Sect. 2.1.

Table 2 College and non-college workers employment ratio in the administrative support, retail sales occupations during the period 1980–2005

	1980	1990	2000	2005	1980–2005
Panel A: Labor supply ratio in All CZs					
All	0.254	0.230	0.222	0.209	−0.044
College graduate	0.117	0.136	0.140	0.133	0.016
HS graduate and HS dropout	0.137	0.095	0.082	0.077	−0.060
College grad/ HS grad and HS dropout	0.851	1.436	1.720	1.735	0.883
Panel B: Employment ratio in CZs above median RSH in 1980					
All	0.254	0.230	0.222	0.209	– 0.045
College graduate	0.117	0.136	0.140	0.133	0.016
HS graduate and HS dropout	0.137	0.095	0.082	0.076	−0.061
College grad/ HS grad and HS dropout	0.852	1.437	1.722	1.737	0.885
Panel C: Employment ratio in CZs above median RSH in 1980					
All	0.202	0.200	0.205	0.204	0.002
College graduate	0.082	0.106	0.116	0.116	0.034
HS graduate and HS dropout	0.120	0.094	0.089	0.088	−0.032
College grad/ HS grad and HS dropout	0.680	1.121	1.309	1.318	0.638

To aggregate the data at CZ-level, we weight observations by the product of labor supply weight and a CZ's geographical weight based on national share of population

3.1 Evidence of non-random displacement of low skill workers

This study's primary motivation is to explain why wages increased and routine employment shares fell only in administrative support and retail occupations during 1980–2005. To address this question, we examine the change in the employment share of different types of workers based on education and wage within clerical occupations during the period 1980–2005. Following Autor et al. (2008) we classify college and post-college graduates as skilled workers and all noncollege graduates as unskilled workers. As an alternative measure, following Autor and Dorn (2013) we use wage as a proxy of workers' skill levels. The results from both these measures are shown in Tables 2 and 3.

In the top panel of Table 2, we report the changes in the employment share of college and noncollege workers from 1980 to 2005. As shown, the employment share of college graduate workers increased approximately 14%. In contrast, the employment share of noncollege graduate workers decreased approximately 44%.[10] That is why the high skill to low skill employment ratio increased from 0.85 in 1980 to 1.74 in 2005. Therefore, these results suggest that the overall employment declined in the administrative support and retail sales occupations over the period 1980–2005 mainly because of the displacement of relatively less-skilled workers in those clerical occupations.

[10] As shown in Table 2, during the period 1980–2005, the employment share of noncollege workers decreased by −0.60, which is approximately 44% (= −0.060/0.137), and college workers increased by 0.016, which is around 14% (= 0.016/0.117).

Table 3 Fraction of workers in the administrative support, retail sales occupations during the period 1980–2005

	1980	1990	2000	2005	1980–2005
Panel A: All CZs					
Minimum wage to 10	0.381	0.328	0.270	0.251	−0.130
Hourly wage 10–15	0.322	0.341	0.329	0.308	−0.013
Hourly wage 15–20	0.157	0.167	0.194	0.203	0.046
Hourly wage 20 and above	0.140	0.165	0.208	0.238	0.098
Panel B: CZs above Median RSH in 1980					
Minimum wage to 10	0.366	0.310	0.254	0.236	−0.131
Hourly wage 10–15	0.327	0.345	0.329	0.306	−0.021
Hourly wage 15–20	0.161	0.173	0.200	0.209	0.048
Hourly wage 20 and above	0.146	0.172	0.217	0.250	0.104
Panel C: CZs below Median RSH in 1980					
Minimum wage to 10	0.510	0.505	0.433	0.399	−0.111
Hourly wage 10–15	0.275	0.297	0.323	0.333	0.058
Hourly wage 15–20	0.120	0.106	0.132	0.143	0.023
Hourly wage 20 and above	0.095	0.092	0.111	0.125	0.030

To obtain the real hourly wage rate, we deflated the nominal hourly wage variable by the national Personal Consumption Expenditure index. Based on the real hourly wage, we classify the administrative support workers into four groups, as shown in the table. Each group represents the share of all administrative workers within a CZ

To provide further evidence of non-random displacement of noncollege workers in administrative support occupations, we classify the CZs in two groups based on above and below national median RSH in 1980, and the summary statistics results are reported in panels B and C, respectively. Consistent with our intuition, we find that the change in high-to-low skill employment ratio is approximately 28% higher in community zones with an initially larger routine task-intensive employment shares.[11] These results support our initial claim that relatively less-skilled workers were most likely displaced from the administrative support occupations and also address the concern that an increase in the high-to-low-skill ratio could be due to only an expansion of high-skilled workers.

In Table 3, we use the measure hourly wage rate as a proxy of workers' skill level to classify them into four groups and report the trends of employment share of those groups during 1980–2005.[12] Consistent with the findings in Table 2, the Table 3 top panel results show that the employment share of relatively less efficient workers whose hourly wages are between the state minimum wage and 15 dollars was decreased by −0.14. In contrast, the employment share of relatively more productive workers whose hourly wages are above 15 dollars moved in strongly countervailing

[11] As shown in Panel B and Panel C of Table 2, the high-to-low-skill labor supply ratio in above-median CZs is 0.247 (= 0.885-−0.638) more than the CZs, which had RSH below the national median in 1980.

[12] These four groups are: (i) minimum wage to 10 dollars, (ii) 10–15 dollars, (iii) 15–20 dollars, and (iv) equal and more than 20 dollars.

directions in clerical and administrative support occupations and increased by 0.14. Furthermore, comparing the panel B and C results, we find that this pattern of changes in the employment share of more productive and least efficient workers is stronger in community zones with initially larger routine task-intensive employment shares.

Since Autor and Dorn (2013) show that the pattern of opposite employment and wage effects is unique to administrative support and retail sales occupations, one concern is whether the non-random displacement of noncollege workers also occurred in two other high routine content occupation groups (i) precision production and craft, and (ii) machine operators and assemblers. To address this concern, we repeat the Tables 2 and 3 results for two other high routine content occupation groups and report the results in Appendix Tables 7–10. These descriptive statistics results show that a similar pattern of non-random displacement of low-skill workers that we find in Tables 2 and 3 did not occur in two other high routine content occupation groups. Hence, these findings strengthen our argument.[13] To summarize, Tables 2 and 3 descriptive statistics results from two separate skill measures unanimously suggest that the average productivity of the workers in clerical and administrative support occupations increased during 1980–2005 because the relatively less efficient non-college-educated clerical workers were displaced.

3.1.1 Mechanism

To examine a channel why the least productive clerical workers were displaced in clerical and administrative support occupations, we explore how rapid changes in the personal computer (PC) adoption rate in work during 1980–2005 affected the high-to-low-skill employment ratio and report the results in Table 4. The intuition of this mechanism is that the rapid technological innovations related to the cost of computers drove PC adoption at work in 1980–2005 and that eventually replaced many routine task-intensive jobs. Therefore, the workers who performed highly routine tasks lost their jobs in clerical occupations.

We follow the Basso et al. (2020) approach to construct a measure of computer adoption at the commuting zone level, which represents the fraction of workers who uses computers at work in a given industry within a CZ. We use a simple linear regression model to estimate the impact of PC adoption rate on the employment ratio of high skill over low skill workers.[14] A threat to the OLS identification approach is that the changes in computerization rate are most likely correlated with the unobserved demand shocks related to the employment ratio of skilled and unskilled workers in a local labor market. Therefore, we employ an instrumental variable strategy that accounts for the potential endogeneity of computerization.

[13] As shown in Appendix Tables 7 and 8, the employment share of both college and noncollege workers decreased in two other high routine content occupation groups. We do not observe this pattern in administrative support and retail sales occupations in Table 3. Furthermore, Comparing the results of Table 3 with Appendix Tables 9 and 10, we find that the pattern of changes in the employment share of more productive and least efficient workers in clerical occupations is different than the pattern we observe in two other high routine content occupation groups.

[14] The data for personal computer adoption rate in work in the US and the measure of computerization are discussed in the Appendix.

Table 4 Computer adoption and change in employment of college, noncollege workers and the ratio of college over noncollege workers within commuting zones, 1980–2005

	Δ fraction of college Workers in 1980–2005		Δ fraction of non-college Workers in 1980–2005		Δ college/non-college Workers in 1980–2005	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
Change in PC adoption	− 0.137***	− 0.391***	− 0.332***	− 0.557***	4.607***	5.560***
Rate in 1980–2005	(0.027)	(0.069)	(0.028)	(0.045)	(0.546)	(0.768)
All other controls	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.568	0.453	0.588	0.490	0.507	0.502
No of observations	722	722	722	722	722	722

Each coefficient is based on a separate regression where the column heading denotes the dependent variable. Models include an intercept, state dummies, and time dummies, and are weighted by start of period commuting zone share of national population. Following Basso et al. (2020) we use the following measure of computer adoption at commuting zone level: $\mathrm{PC}_{j,t}^{US} = \sum_m \omega_{c,m,t} \times \frac{\mathrm{PC}_{US,m,t}}{L_{US,m,t}}$ where $\omega_{j,m,t} = \frac{L_{j,m,t}}{\sum_m L_{j,m,t}}$ and $L_{j,m,t}$ is the employment in industry m at commuting zone j in year t. Thus, $\mathrm{PC}_{US,m,t}$ denotes the total number of computers at work in the US in industry m at year t. Following Autor et al. (2013) and Acemoglu and Restrepo (2017) we construct a non-US exposure to computer adoption variable by using an analogous measure from the PC adoption in high-income European countries and use it as the instrument for the change in the PC adoption rate variable. The discussions of the measure of PC adoption rate and the construction of the instrument are shown in the "Appendix". Robust standard errors in parentheses are clustered on commuting zones. Observations are weighted by the 1980's commuting zone share of national population

To construct an instrument for changes in PC adoption rate ($\text{PC}^{US}_{j,t}$) in the US, we follow a similar approach of Autor et al. (2013) and Acemoglu and Restrepo (2017). Specifically, we construct an instrument variable ($\text{PC}^{EU}_{j,t}$) for the US exposure to computerization using an analogous measure constructed from the PC adoption in high-income European countries and 1950's industry structure in the US. The construction of this Bartik-type instrumental variable is shown in "Appendix". We present the first-stage estimates for the instrumental variable in Appendix Table 15 and show that the predictive relationship between $\text{PC}^{US}_{j,t}$ and $\text{PC}^{EU}_{j,t}$ is sizable and highly significant even at 1 percent level. The first-stage F-stat and the Kleibergen-Paap Wald weak IV test result suggest that our IV is not weak. We provide a discussion on the validity of the instrument in the Appendix.

The 2SLS estimates in columns 1 and 2 confirm that the changes in PC adoption rate are positively associated with the loss of employment of high-skill and low-skill workers in administrative support and retail sales occupations. However, the impact of the ΔPC adoption rate is much higher for low-skill workers because automation replaced many routine task jobs. As a result, in column 3, we find that the impact of ΔPC rate in 1980–2005 is positive and statistically significant, suggesting that the rapid changes in computerization during 1980–2005 displaced more low-skilled workers than high-skilled workers in clerical occupations.

Based on our findings in Tables 2 and 3, we infer that clerical workers' wages increased in 1980–2005, most likely not due to the rise in demand for routine task-based jobs. Instead, because of the non-random selection of displaced noncollege workers from the routine task-intensive occupations. Autor and Dorn (2013) argue that the selection of displaced low-skill workers is probably because the traditional clerical tasks are replaced because of automation. Our Table 4 results provide empirical evidence of this hypothesis. Therefore, the remaining clerical and administrative jobs have become more skill demanding and less routine task intensive. This phenomenon can explain why the correlation between changes in aggregate employment and wages is positive in clerical occupations from 1980 to 2005.

3.2 Distributional impacts of RHS on change in wages from 1980–2005

To provide supportive evidence of non-random displacement of noncollege workers in clerical occupations, we implement the IVQR models introduced by Chetverikov et al. (2016). In Table 5, we report the IVQR point estimates of RSH_{j1980} on the wage growth from 1980 to 2005 for five quantiles of the wage distribution. The robust standard errors are clustered by state, and observations are weighted by a CZ's share of the national population in 1980. The upper panel results for skilled workers show that the 1980's routine employment share in a CZ positively affected the wage gains from 1980 to 2005 in the 10th quantile and above of the wage distribution of college workers in routine task-intensive occupations. However, there is a considerable variation in wage gains across the wage distribution. As shown, the IVQR point estimate at the median is 0.77, which is approximately four times the marginal impact at $\tau = 0.1$ and approximately half of the point estimate at $\tau = 0.9$.

Table 5 IVQR estimates of the 1980's routine occupation share on the change in conditional wage distribution from 1980 to 2005 for college and noncollege workers

	10th quantile	25th quantile	50th quantile	75th quantile	90th quantile
College workers					
Share of routine occs	0.205	0.846	0.772	0.936	1.577
	(0.304)	(0.245)	(0.229)	(0.256)	(0.234)
All other controls	Yes	Yes	Yes	Yes	Yes
No of observations	1444	1444	1444	1444	1444
Noncollege workers					
Share of routine occs	− 0.103	0.284	0.562	0.464	0.884
	(0.229)	(0.194)	(0.155)	(0.198)	(0.251)
All other controls	Yes	Yes	Yes	Yes	Yes
No of observations	1444	1444	1444	1444	1444

The IV quantile regression point estimates are obtain from the Chetverikov et al. (2016) method. The dependent variable is change in log hourly wages of noncollege workers from 1980 to 2005. The independent variable is the routine occupation share in 1980, which is instrumented by the 1950's industry mix measure. Observations are weighted by a commuting zone's share on national population at the initial period, 1980. Robust standard errors are clustered by state. The control variables are the same as Table 7 panel A in the Autor and Dorn (2013)

The IVQR estimates of noncollege workers in Table 5 show that the impact of the 1980's RSH on wage growth at the 10th quantile is -0.10, suggesting that the real wage of the least productive non-college-educated clerical workers declined about 10% probably because of automation. The intuition is that the least efficient low-skilled workers probably did not have the skills to adjust with automation and hence their relative productivity fell compare to others during the period 1980–2005. Furthermore, to investigate whether similar pattern wage growth in clerical occupations also exists in the other two occupational groups with a high routine content, we report the IVQR results for two other routine task-intensive occupation groups, (i) machine operators and (ii) precision production and craft in Appendix Table 11. We find that the decline in wages in the lower half of the skill distribution in these two high routine content occupation groups are sufficiently large to lead to a negative wage effect on average, which Autor and Dorn (2013) find in Table 7.

In addition to Table 5 results, we report the IVQR estimates for the entire wage distribution in Fig. 1. Specifically, in the two panels of Fig. 1, we separately show the IVQR point estimates for skilled and unskilled workers at $\tau \in \{0.05, 0.1, \ldots, 0.95\}$ along with pointwise 95% confidence intervals. Figure 1 shows that the IVQR estimates monotonically increase along with the wage distribution of skilled workers and unskilled laborers within the clerical and administrative support occupations. As shown in Panel A, the IVQR estimates of high-skilled workers are approximately constant at 0.75 from 25th to 80th quantile, and lie in between 1.45 to 1.82 in the upper-tail of the wage distribution. In contrast, in Panel B, we see that the impact of 1980's RSH on the wage growth of low-skilled workers is negative in the lower tail of the distribution, lies in between 0.45 and 0.60 in the middle, and is greater than 0.75 in the 85th quantile and above.

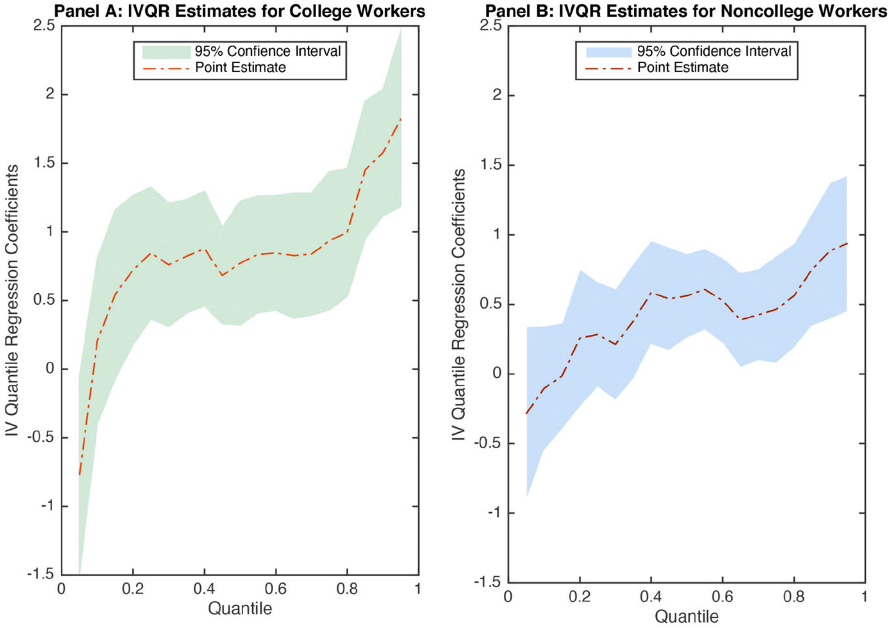

Fig. 1 Effect of routine task-intensive employment share on conditional wage distribution, 1980–2005. The IV quantile regression point estimates are obtain from the Chetverikov et al. (2016) method. The dependent variable is change in log hourly wages of noncollege workers from 1980 to 2005. The independent variable is the routine occupation share in 1980, which is instrumented by the 1950's industry mix measure. Observations are weighted by a commuting zone's share on national population at the initial period, 1980. Robust standard errors are clustered by state

Autor and Dorn (2013) only consider non-college-educated workers to explain the wage polarization and employment polarization in the US. Our IVQR results suggest that Autor and Dorn (2013)'s estimated average positive effect of the 1980's routine occupation share is primarily driven by large positive marginal effects for non-college-educated clerical workers in the upper tercile of the distribution. As shown in the right panel of Fig. 1, the marginal effects in the upper-tail of the wage distribution are approximately one and a half times as large as the average effect. Furthermore, the IVQR point estimates in the left panel of Fig. 1 suggest that similar results also hold for college-educated clerical workers. As shown, most productive skilled workers within the routine-task intensive occupations obtained the maximum wage growth because of the automation.

4 Conclusion

In Table 1, we show that our replication results of employment polarization across six major occupation groups are identically matched with Autor and Dorn (2013). The wage polarization replication results are also almost the same as the original article's Table 7 panel B results. In addition, the Tables 2 and 3 summary statistics results unanimously show that relatively less productive low-skill workers in clerical, administrative support, and retail sales occupations were displaced in 1980–2005, suggesting that the average productivity of the clerical worker went up during that period. As a result, the average hourly wages increased by approximately 7.4% despite routine employment shares fell about 8% in administrative support and retail occupations during the period 1980–2005. Furthermore, we show a plausible mechanism why the least productive noncollege workers were displaced in clerical occupations. Specifically, we show that the replacement of many routine task-intensive clerical jobs is associated with the rapid expansion of computer adoption in work during 1980–2005.

Using the Chetverikov et al. (2016)'s instrumental variable quantile regression approach, we provide further evidence of the non-random displacement of workers in clerical and administrative support occupations during 1980–2005. Specifically, the IVQR results show that the wage gains of non-college-educated workers are positive throughout the wage distribution except for its lower tail. Furthermore, the effect of the 1980's routine occupation share on average wage growth from 1980–2005 in clerical, administrative support, and retail sales occupations is positive because of the large wage gains of non-college-educated workers in the upper tercile of the wage distribution. Therefore, our IVQR estimation results provide additional supportive evidence of why clerical workers' average wage increased during 1980–2005.

To conclude, the displacement of less-productive labor and high returns of relatively more efficient non-college-educated clerical workers can explain the puzzle of why the average wage increased despite routine employment shares fell in administrative support and retail occupations during 1980–2005.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00181-023-02394-x.

References

Acemoglu D, Restrepo P (2017) Secular stagnation? The effect of aging on economic growth in the age of automation. Am Econ Rev 107(5):174–179

Autor D, Dorn D (2013) The growth of low-skilled service jobs and polarization of the US labor market. Am Econ Rev 103(5):1553–1597

Autor D, Katz L, Kearney M (2008) Trends in U.S. wage inequality: revising the revisionists. Rev Econ Stat 90(2):300–23

Autor D, Dorn D, Hanson G (2013) The China syndrome: local labor market effects of import competition in the United States. Am Econ Rev 103(6):2121–2168

Basso G, Peri G, Rahman AS (2020) Computerization and immigration: theory and evidence from the United States. Can J Econ 53(4):1457–1494

Chetverikov D, Larsen B, Palmer C (2016) IV quantile regression for group-level treatments, with an application to the distributional effects of trade. Econometrica 84(2):809–833