Innovation, Wages, and Polarization in China

Belton M. Fleisher
fleisher.1@osu.edu
Ohio State University, & Hunan University Center for Economics, Finance, and Management Studies

William H. McGuire
wmcguire@uw.edu
University of Washington Tacoma

Yaqin Su
yaqinsu2013@gmail.com
Hunan University Center for Economics, Finance, and Management Studies

Min Qiang Zhao
kent_zhao@xmu.edu.cn
WISE, Xiamen University

Abstract

Using data from CHIPS 1995-2013, we find polarization of employment from middle-income Skilled jobs to work in the Unskilled and Self-Employment job categories. This redistribution of employment is consistent with the automation of routine noncognitive tasks in the skilled sector as analyzed in numerous works cited above. While the Unskilled and Self-Employment jobs remain below median income, the redistribution of employment has not been associated with a commensurate polarization of labor income. We find no evidence of polarization of either employment or income at the upper end of the job-skill spectrum

Key Words polarization; innovation; wage growth; China
JEL Codes JEL Codes J24; J31; O30; D33

1. Introduction

The substitution of machines and of more- for less-skilled workers in advanced economies has received intensive analysis in recent papers by Autor & Dorn (2013) and Tan (2013) and many previous publications that they cite, including those of Acemoglu (2010) and Acemoglu & Autor (2011). The development of capital- and skilled-worker intensive technology in advanced economies in the later decades of the 20th century has led to the automation of routine tasks and the redistribution of workers to jobs at both ends of the skill spectrum. The resulting reallocation of workers to lower- and higher-skill jobs resulted in wage- and worker polarization that contributed to a slowing of wage increases among workers in middle-skill jobs.

We search for evidence of a similar phenomenon in China over the 18-year period 1995-2013. China's integration into the world economy has allowed the country to benefit from technology at the world frontier, enabling the substitution of machines for routine tasks. The possible impact of technology on job polarization in a broad range of developing economies is examined by Maloney and Molina (2016). This new technology affects polarization through two channels: (i) its adoption reduces employment of workers in routine manual and cognitive tasks, moving displaced workers to less-routine employment at the lower end of the wage spectrum; (ii) its creation increases demand for workers in higher-skilled occupations, leading to polarization at the upper end of the wage distribution.

Redistribution of employment across job types associated with unequal wage growth may have contributed to the rising income inequality and polarization in China that has been explored in a considerable body of research. Khan, Schettino, & Gabriele (2017) report that polarization of incomes has reduced the share of those who receive middle-level incomes in the benefits of China's economic growth and that those classified as "middle class" have moved toward the lower end of the income distribution. Molero-Simarro (2017) relates the growing share of China's top incomes to the increasing importance of non-labor income's impact on inequality.

We address changes in the distribution of incomes associated with employment in jobs ranked by their average skills as reflected in job-related income by examining the proportion of workers employed in eight job categories and the corresponding changes in the jobs' mean incomes and worker characteristics. We find evidence of employment polarization from 1995 to 2013 for jobs paying below median income in 1995, resulting in a redistribution of workers to the bottom of the wage-skill distribution. At the same time, mean income in the lowest-paying job category rose from approximately 60% to about 65% of income in the category ranked highest in 1995. We find no evidence of polarization for jobs at the upper end of the 1995 income distribution.

The next section contains our methodology; section 3 discusses our data and summary statistics; section 4 reports the counterfactual simulations; and section 5 concludes.

2. Methodology: Simulating Job Choices.

We analyze changes in job choices over job categories ranked according to their skills as reflected in reported job-related incomes by modeling changes in employment shares (proportions of a worker cohort found in each job category) with a standard multinomial logit approach to job choice. We then use the estimation results to construct counterfactual simulations of how employment shares would have evolved between 1995 and 2013 under two scenarios: (i) the probabilities of choosing a job in 2013 were applied to the population of 1995; and (ii) if the probabilities of choosing a job in 1995 were applied to selected population characteristics of 2013.

Multinomial Logit Estimation.

We assume workers maximize utility by choosing a job offering the highest return to their cognitive and noncognitive skills, given the job's working conditions (e.g., safety, comfort, working hours flexibility, etc.). Employers minimize costs by choosing workers whose marginal product (a function of their cognitive and noncognitive skills) does not exceed their marginal contribution to the value of output. We take workers' skills to be predetermined, but preferences for job characteristics associated with working conditions (safety, job amenities, flexibility in work hours, and so on) may reasonably be assumed to be wage- or income elastic. In this framework, the impact of a technology shock will be observed in changes in the probability that a worker of given characteristics will be observed in a given job[1]. We simulate these probabilities using estimation results from a multinomial logit equation. The indirect utility of choosing a choice j is specified as follows:

---

[1] Cortes, Jaimovich, & Siu (2017) refer to the *propensity* to work in a given job. Fleisher, McGuire, Wang, & Zhao (2018) analyze the impact of wage-induced technology change on aggregate employment and wages in China over the period 1996-2007.

$$U_{ipjt} = \beta^{w}_{jt} W_{pt} + \beta^{E}_{jt} Edu_{it} + \beta^{A}_{jt} Age_{it} + \beta^{F}_{jt} Female_{it} + \delta_{pjt} + \varepsilon_{ipjt} \qquad (1)$$

where

- the dependent variable = 1 if individual *i* who lives in province *p* is observed in job *j* = 1…9 representing the eight job categories plus those not working in survey years *t* = 1995, 2002, or 2013;
- *Edu* denotes years of schooling;
- *Age* is the individual's age;
- *Female* is a dummy indicator for female individuals;
- δ_{pjt} captures province and year fixed effects that are choice-specific;
- W_{pt} is the province-year specific wage taken from the Mincer equation (2), net of workers' schooling, age, and gender.

$$LnWage_{ipt} = \alpha^{S}_{t} Edu_{it} + \alpha^{X}_{t} Age_{it} + \alpha^{F}_{t} Female_{it} + \eta_{pt} + \varepsilon_{ipt} \qquad (2)$$

where

- *Wage* is reported work income for worker *i* in survey year *t;*
- *Edu* is reported years of schooling;
- *Age* is worker's age
- *Female* is a dummy variable for female; and
- η is a set of dummy variables for province, year and their interaction terms.

Counterfactual Simulations.

Estimation results of equation (1) provide the basis for simulating counterfactual changes in worker employment shares over job categories between 1995 and 2013. We conduct counterfactual simulations (CFS) of the changes in job-category shares that would occur: (i) if the estimated coefficients of equation (1) for 2013 were applied to the 1995 values of the right-hand variables of equation (1); and (ii) if the estimated coefficients of equation (1) for 1995 were applied to the 2013 values of the right-hand variables of equation (1). These simulated channels are indicated as $\Pr(\beta_T, W_0, X_0)$ and $\Pr(\beta_0, W_1, X_1)$, respectively [2]. We then calculate the differences between $\Pr(\beta_T, W_0, X_0)$ and $\Pr(\beta_0, W_1, X_1)$ and $\Pr(\beta_0, W_0, X_0)$. ($\Pr(\beta_0, W_0, X_0)$ equals the observed employment shares for 1995 except for minor deviations due to deleting

[2] Recall that we use W_{pt} in our MNL model.

problematic observations.)

The expression $\beta_T - \beta_0$ captures changes not only changes in labor demand (e.g., technology change) but also in individuals' job tastes. Details on this procedure are reported in the Appendix. We turn on the channels (β, *Wage*, *Edu*, *Age*, and *Female*) individually and the non-β channels altogether, and examine how each scenario (holding other channels fixed to the 1995 level) is able to account for the total changes in employment shares.

3. Data and Summary Statistics.

Our primary source of data is the CHIP urban waves conducted in 1995, 2002, and 2013. The CHIP surveys do not sample every province in China, and the coverage of provinces changes slightly year after year. We use data from provinces represented in all three waves: Beijing, Shanxi, Liaoning, Jiangsu, Anhui, Henan, Hubei, Guangdong, Chongqing, Sichuan, Yunnan, and Gansu. Observations with missing data and for provinces not included in all provinces are deleted. Collectively, the data we use account for approximately 70% of the total observations in the CHIP data for 1995 and 2002 and approximately 50% in 2013

The CHIP questionnaires request information on respondents' occupations and in some cases industrial sectors, which vary in number and detail among the CHIP waves. We aggregate the more detailed job categories available in the 2013 wave into the eight categories we can match in the 1995 and 2002 waves. This aggregation has the further advantage of avoiding jobs so finely defined that they have too few observations for meaningful statistical analysis. The availability of valid answers to questionnaire items on sources of income requires us to measure wages using reported annual income, which we deflate using provincial indexes of consumer prices.

Sample statistics for the CHIPS job categories are reported in table 1. Job-type definitions are reported in column (1). Mean employment shares, schooling, age, school years, real income, and proportion of female respondents in each job categories are reported in columns (2a) to (6c), respectively. Mean schooling affords a measure of worker cognitive skills. Age reflects important components of human-capital accumulation, and gender mix reflects a mixture of employer discrimination, physical human capital, and work attitudes and tastes for tasks and working environment.

Overview of the CHIP Sample Statistics: Employment and Income Polarization. Figures 1-3 illustrates changes in employment shares and mean labor income across 7 job categories for the periods 1995-2013, 1995-2002, 2002-2013 respectively. Observations along the horizontal axes are ranked in increasing order of their mean income similarly to the ranking of jobs by Autor and Dorn (2013) [3]. The most striking observation is that there has been a massive shift of workers from the Skilled category, where mean income was approximately 95% of that in Self-Employed and Clerical & Office (the median job categories) in 1995 to the Unskilled category, on the left tail. The unskilled employment share more than doubled between 1995 and 2013 while the share of workers in the skilled category declined to approximately one-third of its 1995 level. The Self-Employed is the only other category to increase its share of workers between 1995 and 2013. We find no evidence of employment-share polarization to the right tail of job categories ranked by their 1995 incomes.

These observations reported in the preceding paragraph are consistent with those of Khan, Schettino, & Gabriele (2017) who report polarization toward the lower end of the job-income spectrum. The subperiods illustrated in figures 2 and 3 both display a substantial growth in the employment share of unskilled workers and a downward trend left-to-right in share growth across skill categories. The Skilled Worker jobs include most workers classified as Operatives in the Construction, Mining, and Manufacturing sectors and are thus likely to contain a substantial portion of routine manual tasks while Clerical/Office jobs are likely to contain routine cognitive tasks. Both types of tasks are subject to automation. Over the same periods, the Clerical/Office job category exhibits no change in employment share 1995-2002, but exhibits modest decline over the period 2002-2013 that dominates the entire 1995-2013 period. The Skilled job category exhibits higher wage growth than its lower-wage adjacent job categories (Unskilled and Self-Employed) over the period 1995-2002, while from 2002 to 2013, Skilled jobs experienced lower wage growth than the Unskilled/Service job category and somewhat higher growth than the Self-Employed category. These patterns are consistent with the conjecture that increasing demand for

[3] Autor and Dorn's access to a far larger data base and catalog of occupational descriptions allow them to rank nearly 400 job types into decile bins in ascending order of mean wage, whereas our far more limited sample requires us simply to rank individual job categories and available data preclude accurate measurement of hourly wage. Autor and Dorn's categorization of jobs into those with worker tasks that can be more readily automated follows the work of Autor, Levy, & Murnane (2003).

services provided by workers in the Unskilled category has buffered the downward-directed wage pressure coming from the availability automation-displaced workers who had performed routine tasks in the Skilled and perhaps in the Clerical/Office categories.

Mean real income growth for Owner/Managers stands out as highest among all job categories in both 1995-2002 and 2002-2013. Our data do not permit us to separate labor and capital income for this group, and while the relatively high-income growth among Owner-Managers contributed to total-income polarization as analyzed in Khan, *et al.* (2018), we do not take this growth as evidence supporting growing returns to the ability of Owner-Managers to perform complex tasks.

It is critical not to interpret the soaring share of workers in the unskilled-job category as representing a decline in the abilities of workers in China to carry out specific tasks. As indicated in table 1, mean schooling increased by 1.7 years and by approximately 2 years respectively for workers in the Unskilled- and Skilled-job categories between 1995 and 2013, and annual income growth was similar to that in other job categories, with the exception of the owner/manager group.

Overview of the CHIP Sample Statistics: Schooling and Experience. Although Chinese workers were better-educated on average in 2013 than in1995, we see in figure 4 reveals a striking divergence between the schooling composition of Skilled and Unskilled. We attribute this divergence to changes in the worker attributes required for performing the increasingly automated tasks of Skilled jobs, where workers' mean schooling increased by almost 2 years. Workers with less schooling evidently found their ways into Unskilled/Service jobs, where mean schooling years increased by 1.6 years, and into Self-Employment, where mean schooling increased by 1.3 years[4].

---

[4] Estimation of the Mincer equation (2), reported in table 2, shows that the rate of return to an additional year of schooling within Unskilled Worker jobs was 4% in 1995 and 2002 and 5% in 2013, whereas in Skilled Worker jobs, the comparable returns were 3%, 4%, and 7%, indicating a substantial gain in returns to schooling for workers in Skilled jobs relative to those in Unskilled jobs after 2002We noted above that employment growth in the Clerical/Office category exhibited a mixed pattern of employment and income growth that would be consistent

Figure 5 reveals a significant divergence between Skilled and Unskilled jobs in workforce experience. While average workforce experience during the 1995-2002 period rose by about one-half year per annum for skilled workers, it declined sharply in the subsequent period. These patterns are consistent with the hypothesis that the jobs in the Skilled category increasingly require workers who have received training in the techniques required to operate the higher-technology equipment used in more-automated production.

In the Clerical/Office category average workforce experience rose by about 4 months each year over the 1995-2002 period but fell by nearly two months annually between 2002 and 2013. Perhaps human capital acquired on the job depreciated less rapidly among Clerical/Office workers than among Skilled job holders after 2002. This conjecture invites further investigation.

4. Counterfactual Simulations.

Marginal probabilities at the mean values of the independent variables from the estimation results for equation (1) provide information for conducting counterfactual simulations of the changes in the distribution of worker shares over the job categories. Although the principal use of these estimates is to construct the counterfactual simulations illustrated table 4, reported in table 5 and graphically presented in figures 6 and 7a-7c, it is instructive to evaluate some of the coefficients that are reported in table 3.

We focus on the two job categories exhibiting the most dramatic employment-share shifts 1995-2013, Skilled and Unskilled jobs in columns (7) and (8). The intercepts shift in opposite directions, negatively for the Skilled and positively for the Unskilled. The coefficients of the Provincial Wage variable (taken from the Mincer equation (2) net of individual human-capital variables) indicate that high-wage provinces were more likely to employ skilled workers and fewer unskilled workers in 1995, and both coefficients trend toward insignificantly different from zero in 2013. The marginal probabilities of the Education variable increase from a significant -0.24 to a significant -0.12 for Skilled workers; for the Unskilled, both coefficients are

---

with automation of routine tasks. Annual growth in mean education years for Clerical/Office jobs between 2002 and 2013 was second only to that for skilled workers over the same period. The rate of return to an additional year of schooling for Clerical/Office workers rose from 3% to 9% between 1995-2002 and 2002-2013, one percentage point more than the increase for workers in the Professional job category.

highly significant, negative, and greater in absolute value than for the Skilled, and they increase algebraically from 1995 to 2013. The Age coefficients indicate the trend toward younger workers in Skilled jobs noted above, while for the Unskilled the Age coefficients trend positively. The overall change 1995-2013 for the coefficient of the Female variable is negligible for the Skilled category and sharply positive for the Unskilled, indicating an increased probability for females to employed as Unskilled workers relative to Skilled.

The procedure for mapping the equation (1) estimation results into the Counterfactual Simulations (CFS) is described in the Appendix. For further clarity we illustrate one of the counterfactuals in table 4. Our example presents the counterfactual of applying the 2013 betas to the 1995 values of X and W so that we can calculate the counterfactual change in employment shares, $Pr(\beta_T, W_0, X_0)$- $Pr(\beta_0, W_0, X_0)$. The top subset represents $Pr(\beta_0, W_0, X_0)$ for individuals who graduated from senior junior middle school, and the bottom represents the counterfactual $Pr(\beta_T, W_0, X_0)$ for the same group. For example, comparing the two simulations, the probability that an individual age 23-27 who had graduated from senior middle school would have been a Skilled worker in 1995 under the assumption that 2013 coefficients (β_T) applied (bottom set) is 0.05, compared to 0.26 in the top set, where β_0 applies, a difference of -21 percentage points (-21%). A similar comparison for the Unskilled category indicates that the probability under the β_0 assumption is 0.19, while the β_T counterfactual is 0.51, a difference of 32%.

Table 5 reports similarly calculated percent-point differences for various CFS across the 7 job categories discussed above plus Other and Not Working. We compare $Pr(\beta_T, W_0, X_0)$ - $Pr(\beta_0, W_0, X_0)$, and $Pr(\beta_0, W_1, X_1)$ - $Pr(\beta_0, W_0, X_0)$, and $Pr(\beta_0, W_0, X_0)$ - $Pr(\beta_0, W_0, X_0)$ over all job categories in figure 6. We see that the β CFS closely accounts for the Skilled, Unskilled, and Self-Employed 1995-2013 differences in employment-share growth, $Pr(\beta_0, W_0, X_0)$ - $Pr(\beta_0, W_0, X_0)$, suggesting the importance of induced technology change. In contrast the Education, Age, and Wage CFS comes closer to $Pr(\beta_0, W_0, X_0)$ - $Pr(\beta_0, W_0, X_0)$ for the Clerical/Office and Professional job categories than does the β CFS.

Owner/Managers exhibits virtually no change in employment shares 1995-2013; consequently, there is little to account for. We conjecture that policies aimed at economizing State-Owned Enterprises was an important factor underlying the congruence of $Pr(\beta_T, W_0, X_0)$ - $Pr(\beta_0, W_0, X_0)$, and $Pr(\beta_0, W_1, X_1)$ - $Pr(\beta_0, W_0, X_0)$ for Directors and Managers of Government Agencies.

Counterfactual Simulation: Interactions among Demographics, Wages, and Probabilities. Our CFS procedures do not incorporate interactions between the "beta" counterfactuals (βCFS) and those that apply the 2013 values of Wage, Age, and Female to the 1995 CHIP observations. Nevertheless, we believe that it is meaningful to report how much the beta channel alone can account for the total changes in employment shares without incorporating interactions among the channels. We gain a little more insight into possible interactions among the channels results by comparing the beta CFS and combined wage, age, and female (X) CFS within job categories across schooling and age groups, which we illustrate in figures 7a-7f.

Figure 7a illustrates the actual and counterfactual trends across age- and education groups for Skilled Workers. The βCFS generally tracks closely with the observed changes in employment shares, and both exhibit an upward trend across age groups (very pronounced) and across the schooling levels. In sharp contrast, the W&X CFS substantially over accounts for observed growth of employment shares and exhibits a pronounced downward pattern across age groups at all schooling levels, trending lower among college graduates. Although both the β- and W&X CFS lie mostly in the negative ranges, they are slightly positive among older workers at schooling levels Junior Secondary and higher. These positive flows presumably reflect the balance of positive growth in mean work experience for the Skilled category 1995-2002 and the reverse (albeit smaller in annual value) 2002-2013 exhibited in figure 5.

For the Unskilled, in figure 7b, it should not surprise that we see contrasting relationships and trends among the CFS revealed in figure 7a. While the βCFS and observed series are roughly parallel, the gaps are larger gaps than for the Skilled jobs. The age trends are mainly positive for the two lower schooling levels and negative for graduates of senior secondary school and college/university. The W&X CFS are nearly the reverse of those for the Skilled workers, all negative (under accounting for the observed changes) and positively sloped across age (pronouncedly) and schooling groups.

Figure 7c exhibits the observed series and CFS for the Self Employed. As for the Skilled and Unskilled, the βCFS and Actual series track fairly closely, but they change relative positions with age. Both rise with age at the two lower schooling. Both the βCFS and Actual series trend downward with schooling, approaching zero for college graduates, while the W&X CFS substantially under accounts , trending negatively across the schooling levels.

Figures 7d and 7e display the observed and CFS series for Clerical/Office and Professional workers, the two job categories exhibiting no evidence that the βCFS matches actual employment-share growth in table 5 or figure 6. In both job categories, the actual changes in employment shares are rather closely matched by both the β- and W&X CFS. Both CFS tend to increasingly under account at higher schooling levels in both figures 7d and 7e, but within the Professional category, the β CFS and observed changes in employment-share closely match for middle- and older ages while the W&X CFS diverges toward substantial under-accounting.

Finally, figure 7f presents the relationship between the observed share changes for those not working. There is no relationship between $Pr(\beta_T, W_0, X_0) - Pr(\beta_0, W_0, X_0)$ and observed share changes in table 5 and figure 6 for those not working, and the actual changes in employment shares are rather closely matched by both the β- and W&X CFS in figure 7f. We see a substantial movement of workers above middle age into work that we conjecture, based on our preceding discussions, is facilitated by the opening up of opportunities for the self-employed and the less-skilled in the services sector.

5. Conclusion. Using data for 12 major provinces covered in CHIPS 1995, 2002, and 2013, we observe polarization of employment from middle-income Skilled jobs to work in the Unskilled category and into Self-Employment. This redistribution of employment is consistent with the automation of routine noncognitive tasks in the skilled sector as analyzed in numerous works cited above. While the Unskilled and Self-Employment jobs remain below median income, our finding that the redistribution of employment has not been associated with a commensurate polarization of labor income is consistent with results reported by Cheng and Wan (2015) that business income and, somewhat less so, labor income have reduced polarization of incomes over the period 1978-2010. In fact, mean income growth among for the Unskilled has exceeded that for workers in Skilled jobs. We attribute growth in demand for output of the Service sector as a major factor cushioning the impact of the flow of workers from the Skilled to the Unskilled and Self-Employment job categories as a major factor supporting increased employment. We do find evidence of significant automation of routine cognitive tasks that might have occurred in the Clerical/Office job category.

In contrast to Ge, Sun, and Zhao (2013) we observe a substantial decline in share of our sample population not working (Out of Labor Force). Differences in analysis by age, gender, schooling, and job category between their study and ours make it difficult to sort out reasons for this discrepancy, but a contributing factor may be that Ge, Sun, & Zhao is based on data from the China Census of population covering all provinces. Our evidence suggests that the major increase in non-workers has arisen among individuals at or less than middle-age. Clerical/Office, Professional, and to a lesser extent, Self-Employment and work in the Unskilled jobs, most likely in the Service Sector, has contributed to the older population choosing employment over not working and retirement.

In contrast to employment trends in economies that are closer to the world technology frontier, we find no evidence of employment polarization in the right-tail of the job-category wage distribution. Professional and Owner-Manager job categories, which should encompass a relatively high proportion of non-automatable tasks associated with technology development as suggested by Acemoglu and Autor (2011) among others exhibit almost stable employment shares so that there is no tendency for the right-tail of the distribution of shares to to mirror what has occurred at the left extreme.

References.

Acemoglu, Daron, 2010. When Does Labor Scarcity Encourage Innovation? *Journal of Political Economy* 118, 1037-1078.

Acemoglu, Daron, & Autor, David, 2011. Skills, Tasks, and Technologies: Implications for Employment and Earnings. ;*Handbook of Labor Economics* 2011, 1044-1171.

Autor, David H. & Dorn, David, 2013. The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review* 103, 1553-1597.

Autor, David H., Levy, Frank, & Murnane, Richard, 2003. The Skill Content of Recent Technological Change: An Empirical Exploration. *Quarterly Journal of Economics* 118, 1279-1333.

Cheng, Wang & Wan, Guanghua, 2015. Income Polarization in China: Trends and Changes. *China Economic Review* 36, 58-72.

Cortes, Guido Matias, Jaimovich, Nir, & Siu, Henry E., 2017. Disappearing Routine Jobs: Who, How, and Why? *Journal of Monetary Economics*. 91, 69-87.

Fleisher, Belton M., McGuire, William H., Smith, Adam Nicholas, & Zhou, Mi, 2015. Knowledge Capital, innovation, and Growth in China. *Journal of Asian Economics*39, 31-42.

Ge, Peng, Sun, Wenkai, & Zhao, Zhong, 2018. Automation Technology and Employment Structures in China:1990 to 2015. Unpublished Manuscript. Renmin University of China.

Khan, Haider Ali, Schettino, Francesco, & Gabriele, Alberto, 2018. Polarization and the Middle Class in China: a Non-Paremetric Evaluation Using CHNS and CHIP Data. Unpublished manuscript. https://mpra.ub.uni-muenchen.de/86133

Molero-Simarro, Ricardo, 2017. Inequality in China Revisited. The Effect of Functional Distribution of Income on Urban Top Incomes, The Urban-Rural Gap and the Gini Index, 1978-2015. *China Economic Review* 42, 101-117.

Maloney, William F. & Molina, Carlos, 2016. Are Automation and Trade Polarizing Developing Country Labor Markets, Too? *Policy research Working Paper 7922. Policy Research Working Paper 7922.* http://documents.worldbank.org/curated/en/869281482170996446/pdf/WPS7922.pdf

Figure 1
Annual Growth in Mean Income and Employment Share (1995-2013)

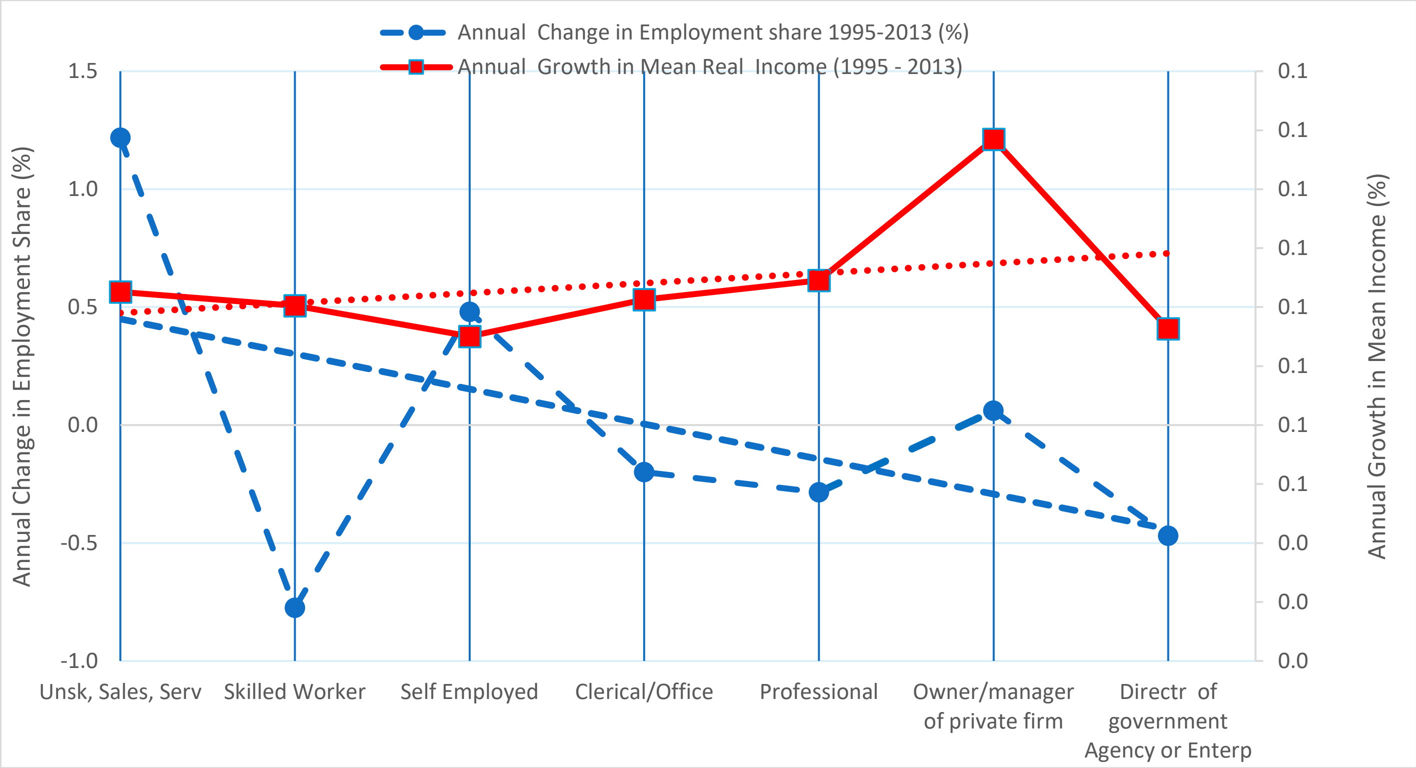

Figure 2 Annual Growth in Mean Income and Employment Share (1995-2002)

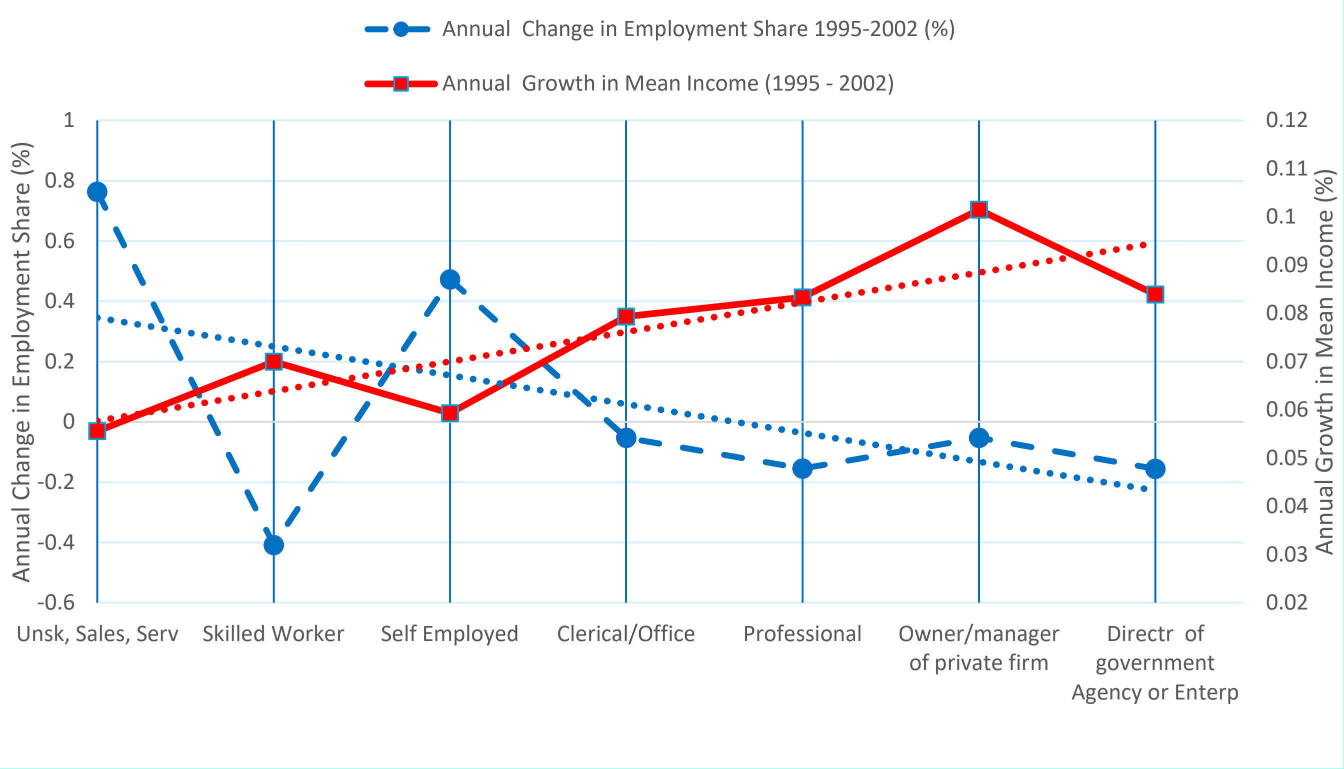

Figure 3 Annual Growth in Mean Income and Employment Shares (2002-2013)

Figure 4
Annual Change in Mean Years of Education

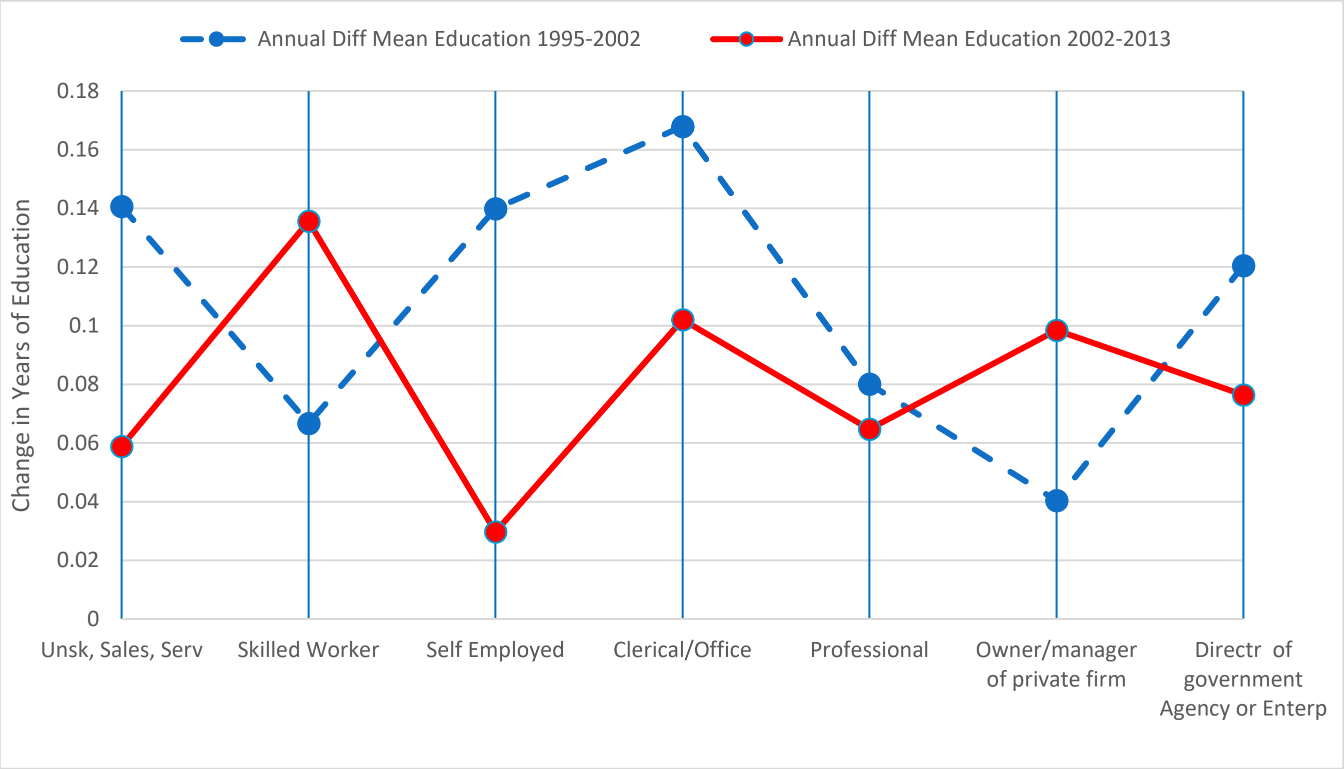

Figure 5
Annual Change in Mean Years of Work Experience

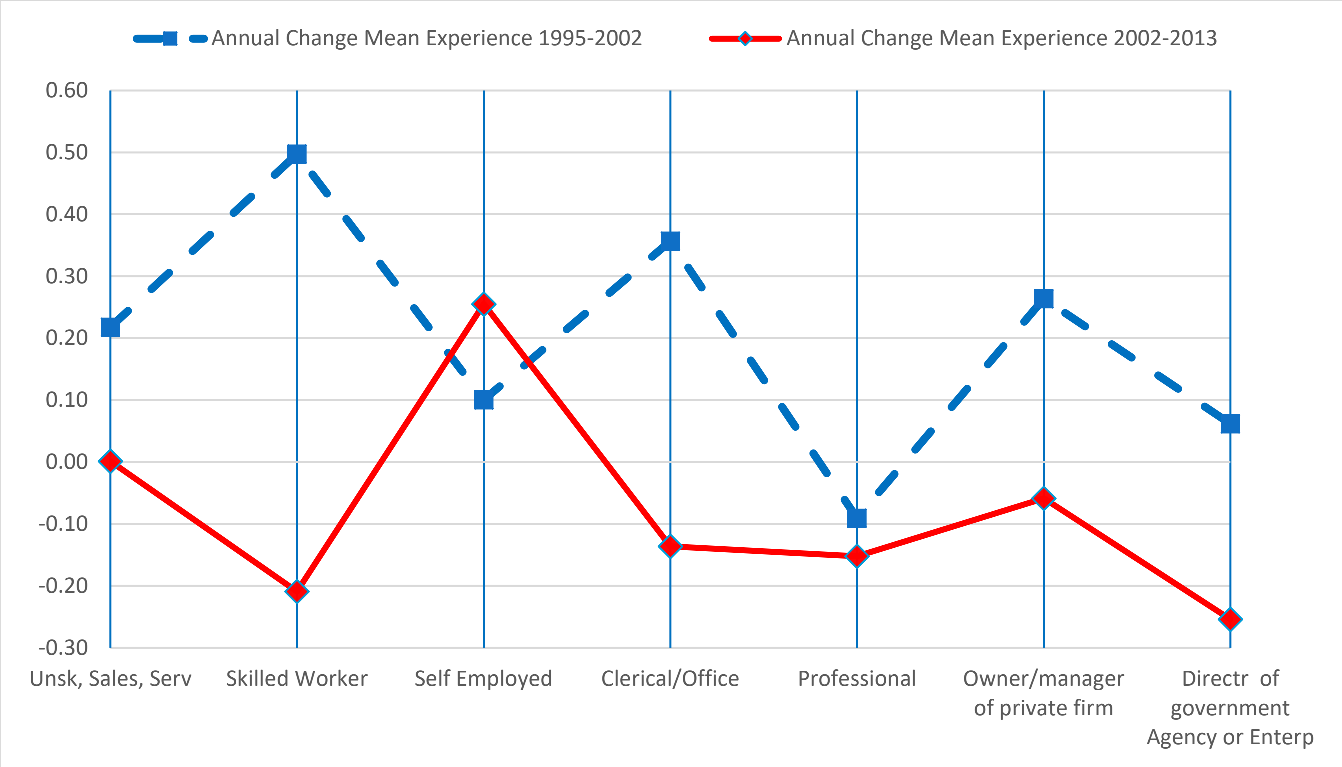

Figure 6 Counterfactual Simulation: All Job Categories

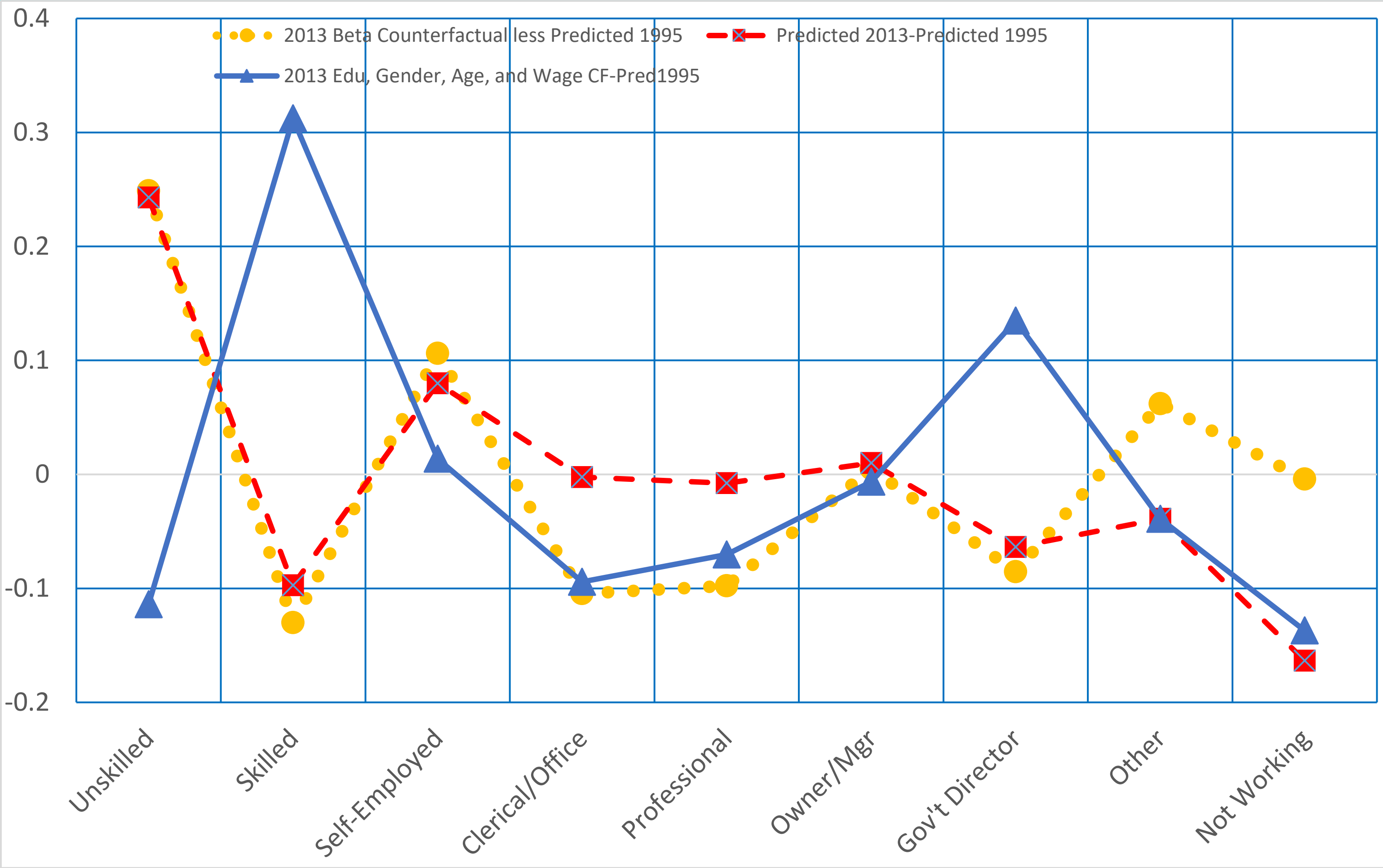

Figure 7a Actual and Counterfactual Change in Employment for Skilled Workers (1995-2013)

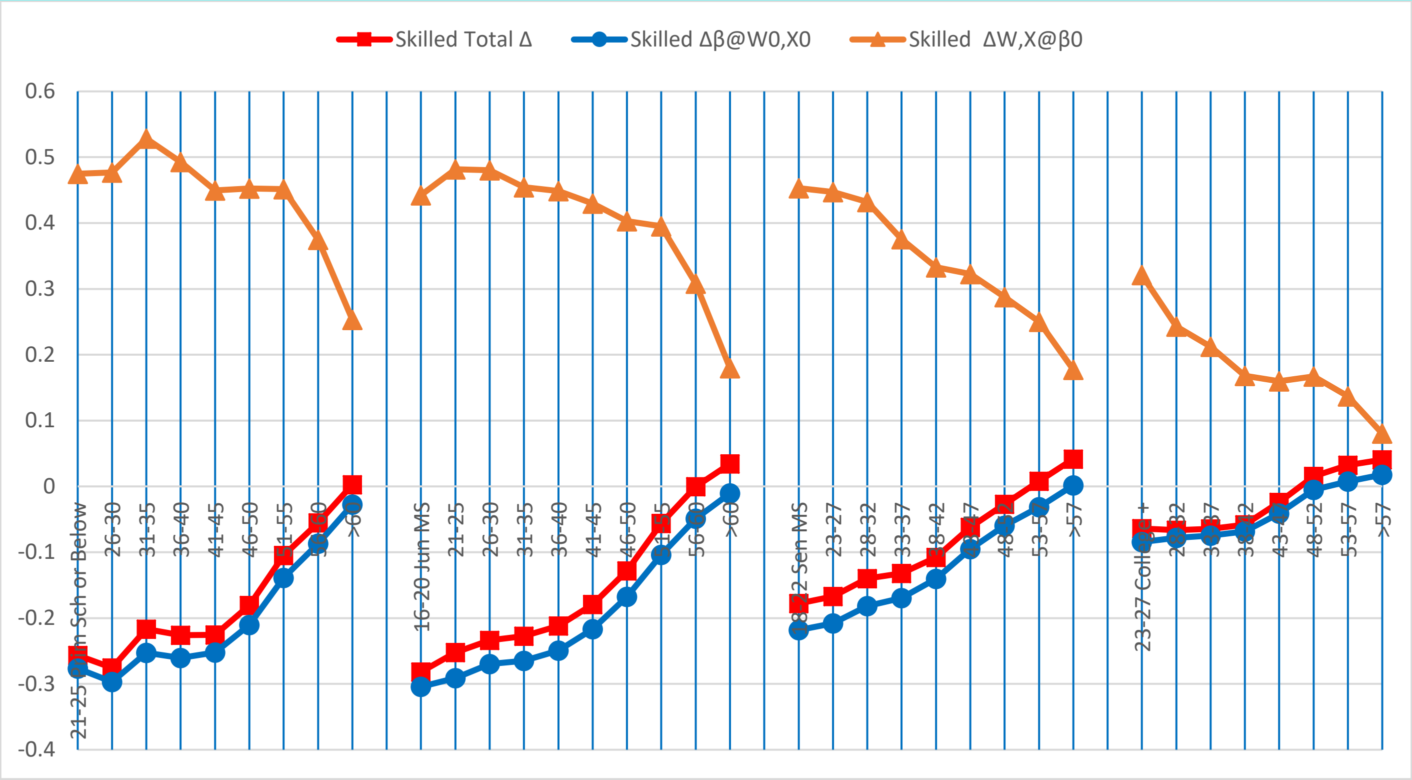

Figure 7b Actual and Counterfactual Change in Employment for Unskilled Workers (1995-2013)

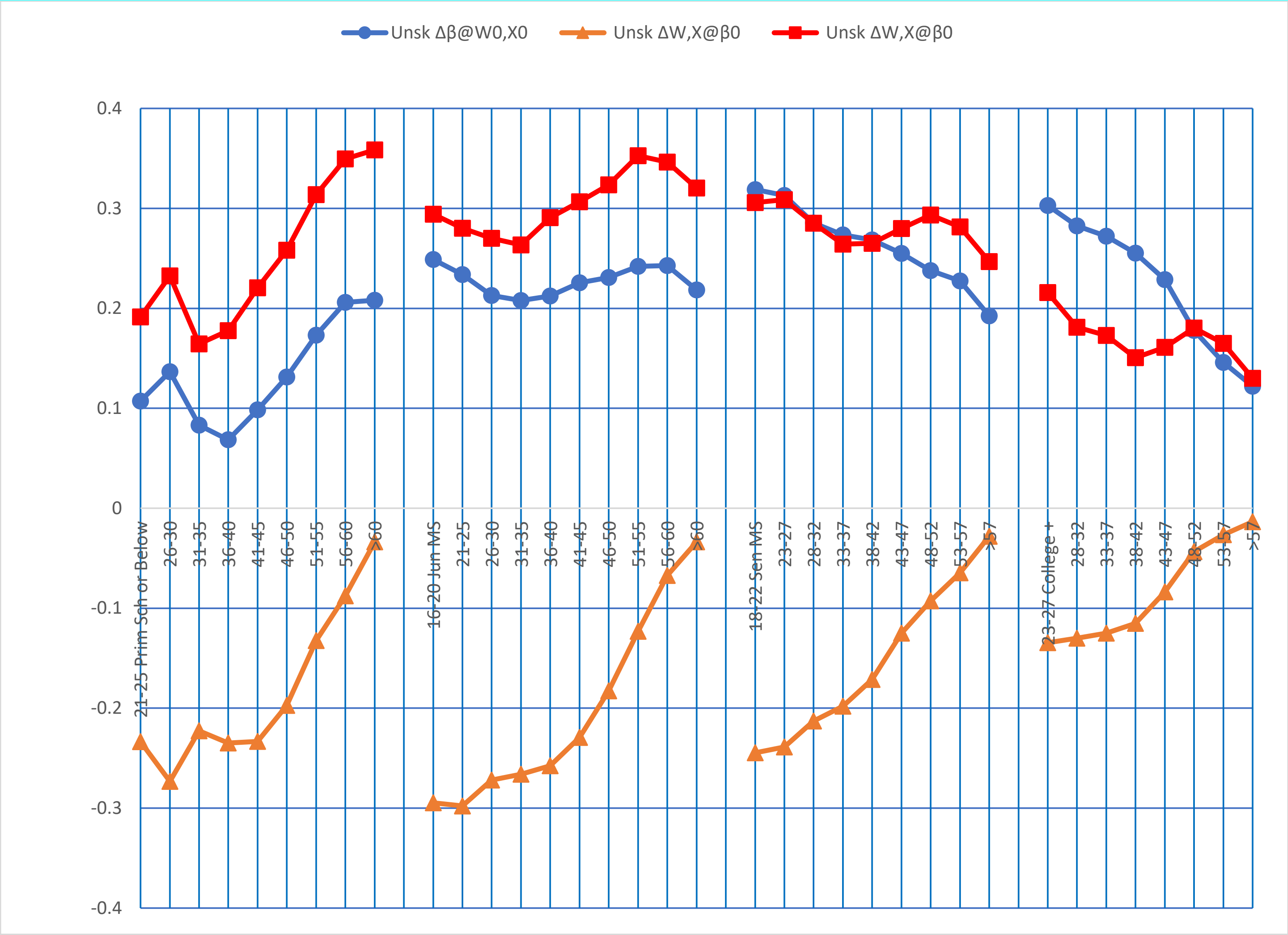

Figure 7c Actual and Counterfactual Change in Employment for Self-Employed Workers (1995-2013)

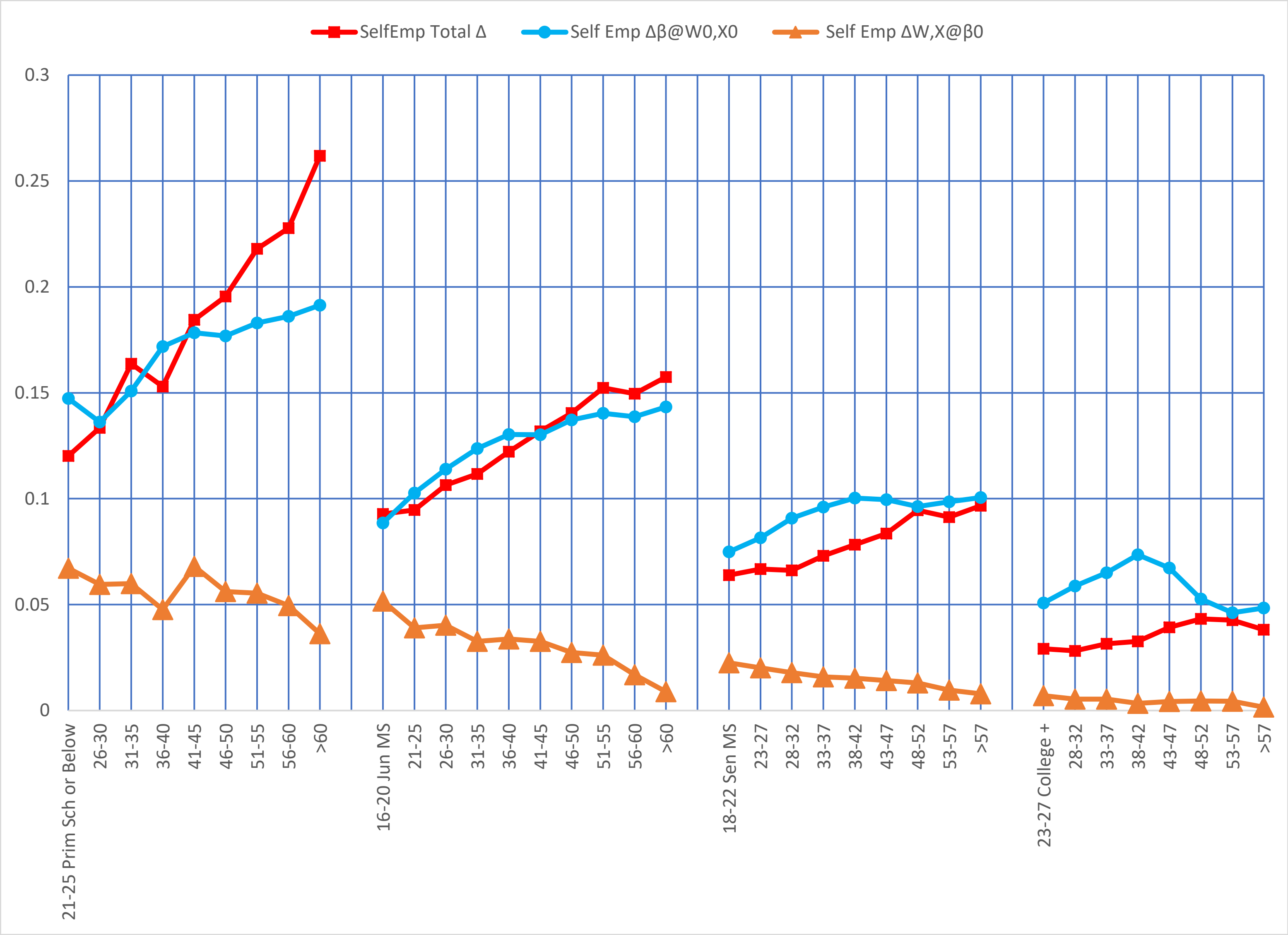

Figure 7d Actual and Counterfactual Change in Employment for Clerical and Office Workers (1995-2013)

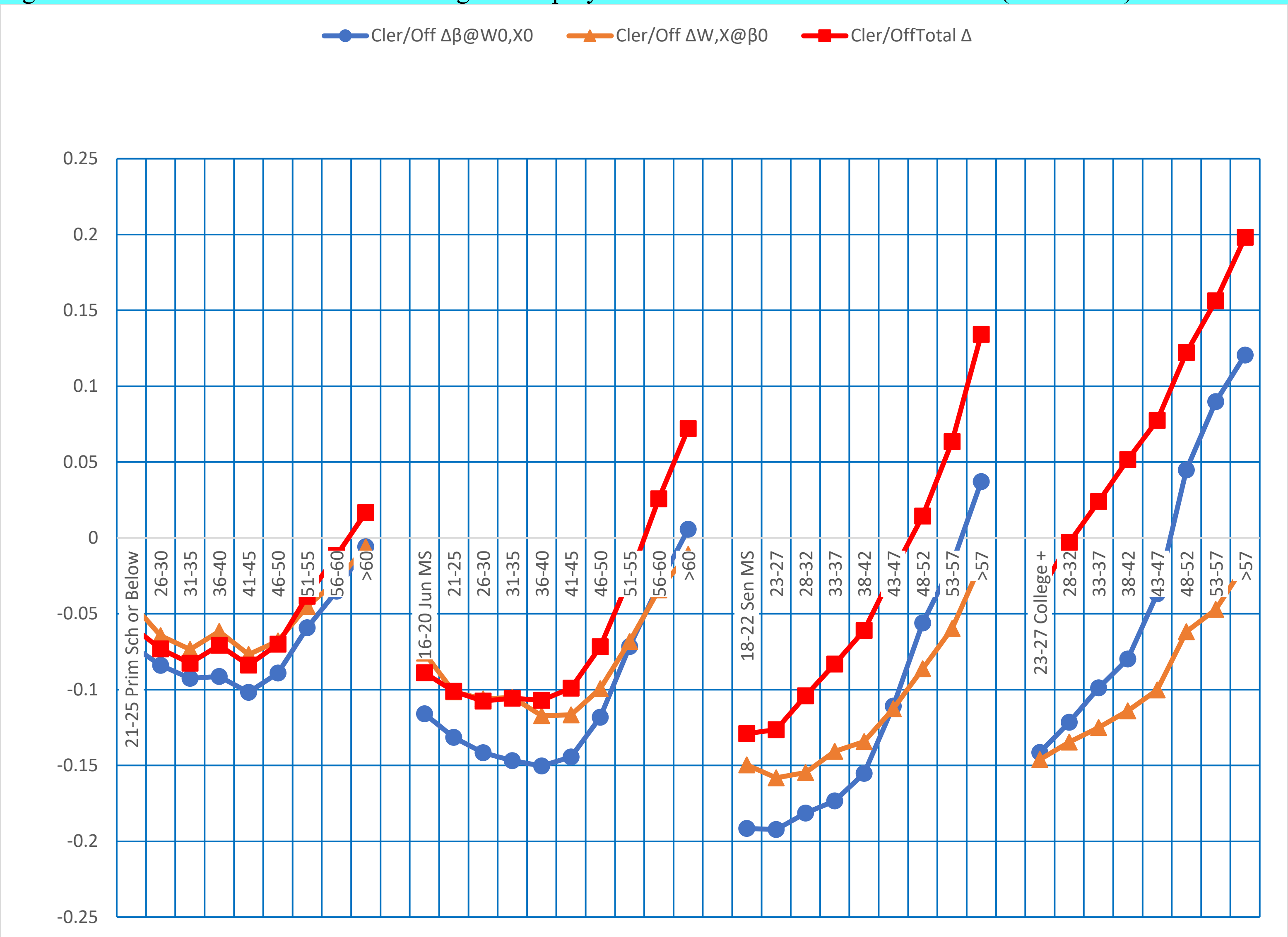

Figure7e Actual and Counterfactual Change in Employment for Professional Workers (1995-2013)

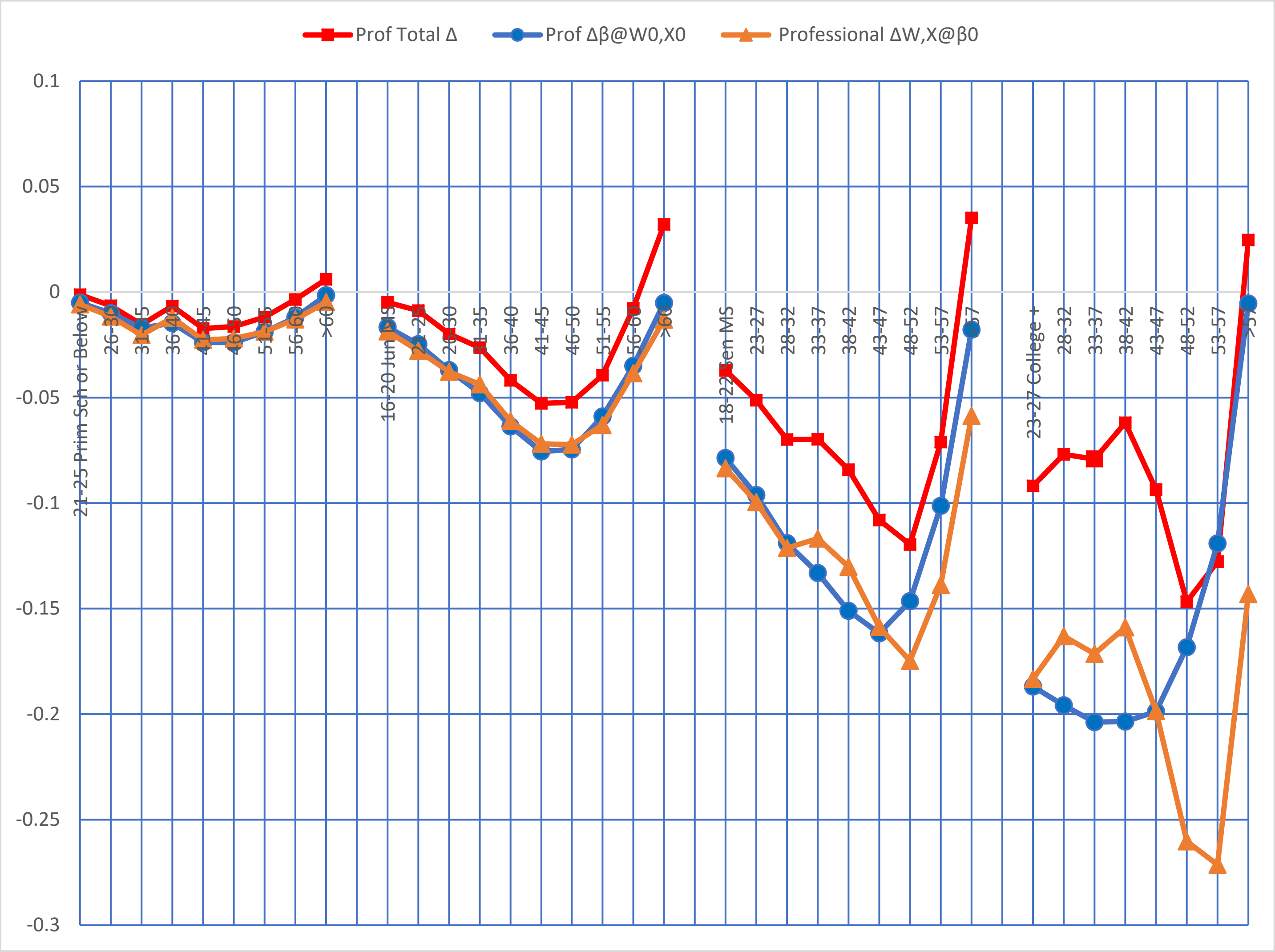

Figure 7f Actual and Counterfactual Change in Proportion Not Working (1995-2013)

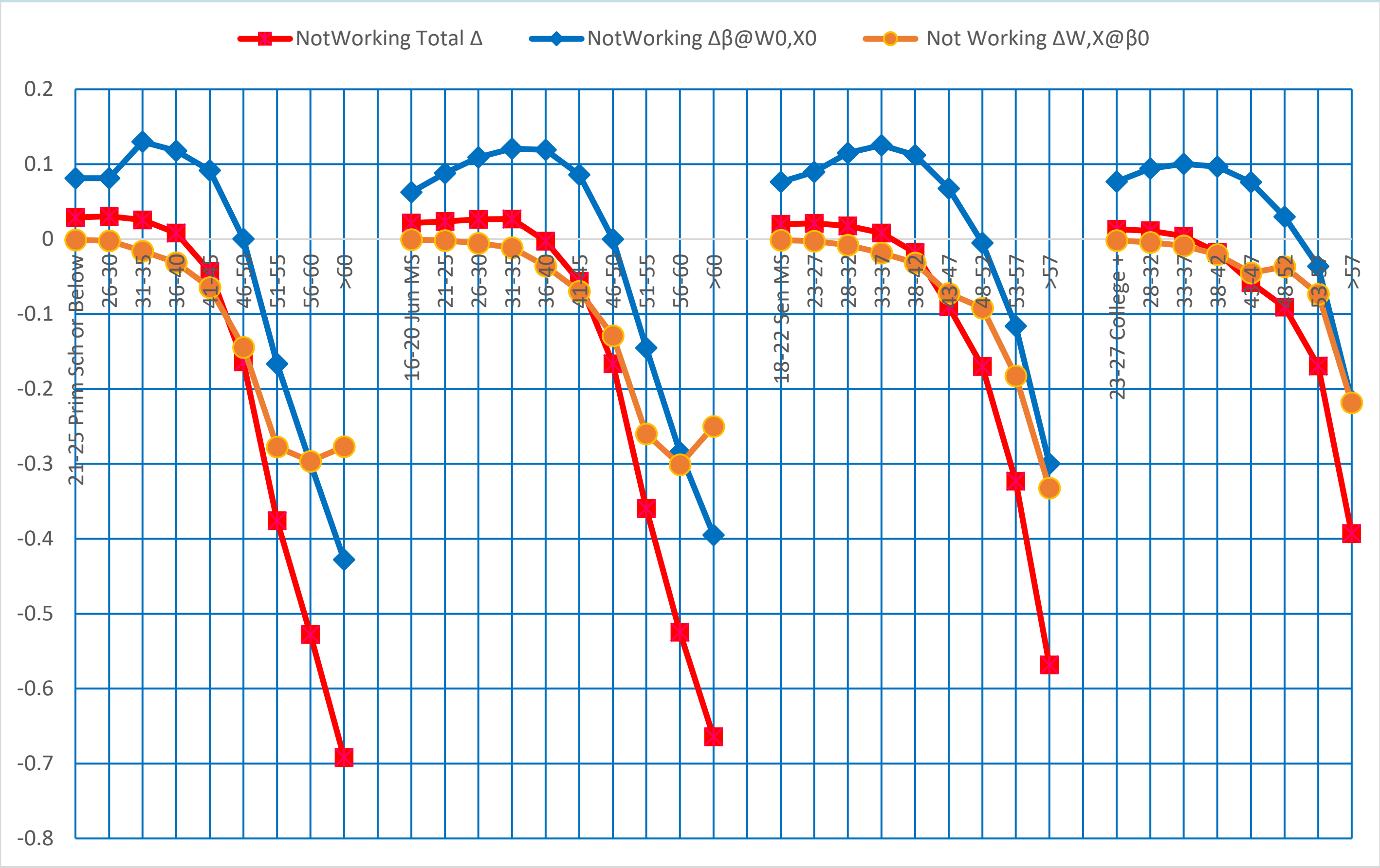

Table 1 Sample Statistics

(1)	Employment Share			Mean Age			Mean School Years			Mean Real Income (RMB)		
	(2a)	(2b)	(2c)	(3a)	(3b)	(3c)	(4a)	(4b)	(4c)	(5a)	(5b)	(5c)
	1995	2002	2013	1995	2002	2013	1995	2002	2013	1995	2002	2013
Unskilled, Sales, Service	0.167	0.220	0.386	37	39	40	8.93	9.92	10.56	1213	1771	5059
Skilled Worker	0.217	0.188	0.077	37	41	40	9.60	10.08	11.57	1445	2321	5802
Self Employed	0.010	0.043	0.096	39	40	44	8.22	9.20	9.52	1538	2302	5660
Clerical/Office	0.208	0.204	0.172	38	40	40	10.94	12.11	13.23	1554	2652	6344
Professional	0.223	0.212	0.172	41	40	39	12.69	13.25	13.96	1794	3143	7736
Owner/Manager Private Firm	0.008	0.004	0.019	41	43	43	10.69	10.98	12.06	1846	3633	11784
Director Govt. Agency/Enterprise	0.116	0.105	0.032	45	45	43	12.19	13.03	13.87	2027	3564	7618
Other	0.052	0.023	0.047	36	42	41	9.44	10.30	10.63	1280	1672	4591
Not Working	0.2840	0.4540	0.0586	58	56	46	8.72	8.9	10.08	-	-	-

Job Category	Proportion Female		
	(6a)	(6b)	(6c)
Unskilled, Sales, Service	0.62	0.59	0.45
Skilled Worker	0.39	0.29	0.29
Self Employed	0.47	0.44	0.43
Clerical/Office	0.51	0.51	0.46
Professional	0.49	0.47	0.48
Owner/Manager Private Firm	0.43	0.30	0.34
Director Govt Agency/Enterprise	0.23	0.21	0.34
Other	0.67	0.52	0.37
Not Working	0.55	0.59	0.52

Source: CHIPS 1995, 2002, & 2013 and authors' calculations. Sample sizes are 14,853, 14,646, and 9,751 in 1995, 2002, and 2013, respectively.

Table 2 Estimation Results Mincer Equation (Equation 2)

Variables	Coefficients Standard Errors
Edu Yrs x Year 1995	0.04*** (0.00)
Edu Yrs x Year 2002	0.07*** (0.00)
Edu Yrs x Year 2013	0.08*** (0.00)
Age x Year 1995	0.02*** (0.00)
Age x Year 2002	0.02*** (0.00)
Age x Year 2013	0.01*** (0.00)
Female x Year 1995	-0.14*** (0.00)
Female x Year 2002	-0.19*** (0.00)
Female x Year 2013	-0.28*** (0.00)
Year 2002	0.24*** (0.00)
Year 2013	1.38*** (0.00)
Constant	6.08*** (0.00)
Observations	31,584
R-squared	0.50

Robust pval in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3 Multinomial Logit Estimation Results (Equation 1)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	Not Working	Owner/Manager Of Private Firm	Self-Employed	Professional	Government Director	Clerical/ Office Staff	Skilled Worker	Unskilled Worker	Other
Year 2002	0.535***	-0.0243*	0.0594***	0.03	-0.0529**	-0.346***	-0.01	-0.113**	-0.0839***
	0.00	-0.07	0.00	-0.44	-0.02	0.00	-0.85	-0.04	0.00
Year 2013	1.213***	-0.0219*	0.0426**	0.03	-0.229***	-0.631***	-0.349***	-0.07	0.01
	0.00	-0.05	-0.04	-0.52	0.00	0.00	0.00	-0.22	-0.76
Prov. Wage. x Year 1995	0.01	-0.0543**	0.04	-0.04	0.08	-0.03	0.272**	-0.199**	-0.0733*
	-0.95	-0.01	-0.13	-0.64	-0.20	-0.76	-0.02	-0.04	-0.05
Prov. Wage. x Year 2002	-0.289***	0.01	-0.02	0.177**	0.109*	0.14	0.16	-0.201**	-0.0825**
	-0.01	-0.38	-0.32	-0.04	-0.07	-0.17	-0.15	-0.03	-0.04
Prov. Wage. x Year 2013	-0.176***	0.00	-0.01	0.04	0.0614***	0.0876***	0.05	-0.03	-0.0324**
	0.00	-0.75	-0.40	-0.13	0.00	-0.01	-0.12	-0.30	-0.01
Edu Yrs x Year 1995	-0.00211**	0.00	-0.0074***	0.058***	0.0169***	0.0178***	-0.0237***	-0.0532***	-0.00639***
	-0.05	-0.18	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Edu Yrs x Year 2002	-0.0174***	0.00	-0.0064***	0.0515***	0.0184***	0.0283***	-0.0272***	-0.0422***	-0.00463***
	0.00	-0.69	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Edu Yrs x Year 2013	-0.0228***	0.00	-0.0062***	0.0414***	0.0181***	0.0305***	-0.0124***	-0.0418***	-0.00695***
	0.00	-0.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Age x Year 1995	0.0219***	0.00	-0.0010***	0.00258***	0.00263***	0.00462***	0.00711***	-0.0123***	-0.00206***
	0.00	-0.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Age x Year 2002	0.0179***	0.00	0.00109***	0.000798**	0.00208***	0.00263***	0.00533***	0.00955***	-0.000550**
	0.00	-0.74	0.00	-0.01	0.00	0.00	0.00	0.00	-0.02
Age x Year 2013	0.00430***	0.000225***	0.00	0.00166***	0.00319***	0.00154***	0.00345***	0.00647***	0.000868***
	0.00	0.00	-0.16	0.00	0.00	0.00	0.00	0.00	0.00
Female x Year 1995	0.0949***	0.00	-0.00950**	0.0189***	-0.0630***	0.00	-0.101***	0.0493***	0.0176***
	0.00	-0.16	-0.03	0.00	0.00	-0.56	0.00	0.00	0.00
Female x Year 2002	0.128***	-0.00671***	0.00734***	0.00	-0.0701***	0.0252***	-0.144***	0.0671***	0.00
	0.00	-0.01	0.00	-0.58	0.00	0.00	0.00	0.00	-0.42
Female x Year 2013	0.0616***	-0.00248*	0.00	0.0370***	-0.0129*	0.0259***	-0.102***	0.00	-0.0104***
	0.00	-0.06	-0.65	0.00	-0.10	-0.01	0.00	-0.57	0.00
Observations	39,264	39,264	39,264	39,264	39,264	39,264	39,264	39,264	39,264

pval in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4 Counterfactual Simulation Examples

Pr(·|β0, w0, x0)									
Schooling and Age Groups	Not Working	Owner/Manager of Private Firm	Self-Employed	Professional	Government Director	Clerical/Office Staff	Skilled Worker	Unskilled Worker	Other
SMS 18-22	0.00	0.01	0.01	0.13	0.04	0.25	0.27	0.22	0.08
23-27	0.01	0.01	0.01	0.16	0.05	0.25	0.26	0.19	0.07
28-32	0.01	0.01	0.01	0.19	0.07	0.25	0.23	0.18	0.06
33-37	0.04	0.01	0.01	0.20	0.08	0.24	0.22	0.16	0.05
38-42	0.07	0.01	0.01	0.23	0.11	0.23	0.19	0.12	0.04
43-47	0.15	0.01	0.01	0.25	0.14	0.19	0.14	0.08	0.03
48-52	0.25	0.01	0.00	0.26	0.16	0.15	0.10	0.05	0.02
53-57	0.41	0.01	0.00	0.20	0.16	0.11	0.07	0.03	0.01
>57	0.66	0.00	0.00	0.11	0.12	0.05	0.03	0.01	0.00
Pr(·|βT, w0, x0)									
Schooling and Age Groups	Not Working	Owner/Manager of Private Firm	Self-Employed	Professional	Government Director	Clerical/Office Staff	Skilled Worker	Unskilled Worker	Other
SMS 18-22	0.08	0.01	0.08	0.05	0.00	0.05	0.05	0.54	0.13
23-27	0.09	0.01	0.09	0.06	0.00	0.06	0.05	0.51	0.13
28-32	0.13	0.01	0.10	0.07	0.00	0.07	0.05	0.46	0.12
33-37	0.16	0.01	0.10	0.07	0.00	0.07	0.05	0.43	0.11
38-42	0.18	0.01	0.11	0.08	0.01	0.07	0.04	0.39	0.11
43-47	0.22	0.01	0.10	0.09	0.01	0.08	0.04	0.34	0.10
48-52	0.25	0.02	0.10	0.11	0.01	0.10	0.04	0.29	0.09
53-57	0.29	0.02	0.10	0.10	0.01	0.09	0.04	0.26	0.08
>57	0.36	0.02	0.10	0.10	0.02	0.09	0.03	0.20	0.07
23-27	0.08	0.01	0.05	0.18	0.01	0.13	0.06	0.38	0.10
28-32	0.10	0.01	0.06	0.18	0.01	0.13	0.06	0.35	0.10
33-37	0.12	0.01	0.07	0.17	0.01	0.13	0.06	0.33	0.09
38-42	0.15	0.02	0.08	0.16	0.01	0.12	0.05	0.32	0.10
43-47	0.17	0.02	0.07	0.18	0.02	0.14	0.05	0.27	0.09
48-52	0.17	0.02	0.05	0.25	0.03	0.17	0.04	0.19	0.07
53-57	0.18	0.03	0.05	0.28	0.04	0.18	0.04	0.15	0.06
>57	0.23	0.03	0.05	0.25	0.06	0.17	0.03	0.13	0.05

Table 5 Summary of Counterfactuals. Shares of All Observations in Each Category

	Unskilled	Skilled	Self-Employed	Clerical/ Office	Professional	Owner /Mgr	Gov't Director	Other	Not Working
Predicted 2013-Predicted 1995	24.29%	-9.73%	7.99%	-0.24%	-0.78%	1.01%	-6.36%	-3.89%	-16.34%
2013 Education Counterfactual Less Predicted 1995	-2.20%	-1.91%	-0.11%	0.41%	3.46%	0.004%	0.64%	-0.56%	0.27%
2013 Gender Counterfactual Less Predicted 1995	-0.25%	0.73%	0.020%	0.044%	0.039%	0.014%	0.42%	-0.13%	-0.89%
2013 Age Counterfactual less Predicted 1995	0.51%	0.80%	0.053%	1.14%	1.36%	0.057%	0.57%	0.13%	-4.62%
2013 Wage Counterfactual Less Predicted 1995	-11.11%	35.67%	1.82%	-11.03%	-11.22%	-0.65%	8.21%	-3.87%	-7.82%
2013 Beta Counterfactual less Predicted 1995	24.88%	-13.00%	10.62%	-10.46%	-9.79%	0.50%	-8.53%	6.21%	-0.41%
2013 Edu, Gender, Age, and Wage CF-Pred1995	-11.40%	31.23%	1.38%	-9.44%	-7.03%	-0.65%	13.50%	-3.89%	-13.70%

Notes:

(i) Predicted 1995 Levels = Pr(·|β0, w0, x0); Predicted 2013 Levels = Pr(·|βT, wT, xT); 2013 Betas counterfactuals are obtained by substituting the MNL βT for β0 in each observation;

(ii) 2013 Education, Gender, and Age counterfactuals are averages over all age and schooling groups obtained by grouping 1995 observations into bins of appropriate X-value ranges so that all 2013 observations are matched with a bin

Appendix

Counterfactual simulation

We use the estimates from the multinomial logit occupational choice model to conduct counterfactual simulations. We are interested in how the distribution of workers across occupations would change between 1995 and 2013 if we only allowed one of the "channels" affecting the occupational distribution to operate at once. These channels include: 1) The distribution of education levels, 2) the distribution of gender, 3) the distribution of age, 4) the provincial wage level, and 5) the year-specific coefficient estimates ("betas") from the multinomial logit estimation.

Education, gender, and age

In order to simulate the effect of a change in the distribution of one demographic characteristic, we need to close off all other "channels" and allow only the distribution of that one characteristic to change between 1995 and 2013. We do this by creating demographic groups of individuals in each year of the data who share unique combinations of age range, gender, and education. For example, one demographic group might be composed of males, aged 36-40, who have completed senior middle school.[5] For simplicity, we assign the group-specific mean values of age and years of education to all individuals within the same demographic group. We then assign weights to these groups based on their share of the sample in a given year

Rather than change the years of education, age, or gender of particular individuals, we perform the counterfactual simulation by changing the weights assigned to the demographic groups. For example, we perform the education counterfactual as follows: Call $\omega_{a,g,e,t}$ the share of a demographic group sharing a common age range (a), gender (g), and level of education (e), in a given year (t).

[5] We ensure the demographic groups match between 1995 and 2013 by eliminating any groups exist in only one of the sample years. This only eliminates four individuals from the total sample.

Call $\omega_{a,g,t}$ the share of a demographic group sharing a common age range and gender in a given year. Note that $\omega_{a,g,t} = \sum_{e=1}^{4} \omega_{a,g,e,t}$. To simulate a change in the distribution of education, we simply replace $\omega_{a,g,e,t}$ as follows:

$$\omega_{a,g,e,t} = \omega_{a,g,e,t+1} * \left({}^{\omega_{a,g,t}}/_{\omega_{a,g,t+1}}\right)$$

We then calculate the predicted probabilities of appearing in each occupation for each demographic group. Finally, we calculate the weighted sum of probabilities across demographic groups within an occupation, using $\omega_{a,g,e,t}$ as the weight.

The procedure is essentially the same for modeling changes in the age and gender distribution. For a change in the gender distribution, we would replace $\omega_{a,g,e,t}$ with:

$$\omega_{a,g,e,t} = \omega_{a,g,e,t+1} * \left({}^{\omega_{a,e,t}}/_{\omega_{a,e,t+1}}\right)$$

For a change in the age distribution, we would replace $\omega_{a,g,e,t}$ with:

$$\omega_{a,g,e,t} = \omega_{a,g,e,t+1} * \left({}^{\omega_{g,e,t}}/_{\omega_{g,e,t+1}}\right)$$

Provincial Wage Level

The counterfactual for the provincial wage level is somewhat simpler. Since the provincial wage level is province-year specific, rather than individual specific, we can simply assign the 2013 provincial wage level to all individuals in the same province in 1995. The rest of the procedure is the same as described above; we calculate the predicted probabilities of appearing in each occupation for each demographic group and use $\omega_{a,g,e,t}$ as the weight to calculate the weighted sum of probabilities within occupations, across demographic groups.

Year-Specific "Betas"

We specify our multinomial logit model to allow the coefficients to vary across years of the sample. We believe that variation in these "betas" captures technological change, though we acknowledge it may also capture other structural changes in the labor market. Isolating the effect of the changes in "betas" does not require changing any of the observed characteristics of the individuals in the sample, or their probability weights. Instead, we simply tell the estimator to treat the 1995 values of the variables as if they were the 2013 values when we predict the probability of each demographic group of appearing in each occupation. As above, we then use $\omega_{a,g,e,t}$ to calculate the weighted sum of probabilities within occupations, across demographic groups.