Supplement to:

Cheng, Siwei, Bhumika Chauhan, and Swati Chintala. 2019. "The Rise of Programming and the Stalled Gender Revolution." Sociological Science 6: 321-351.

## Online Appendix A Historical Trends in Computerization and Gender Inequality: A Descriptive Overview

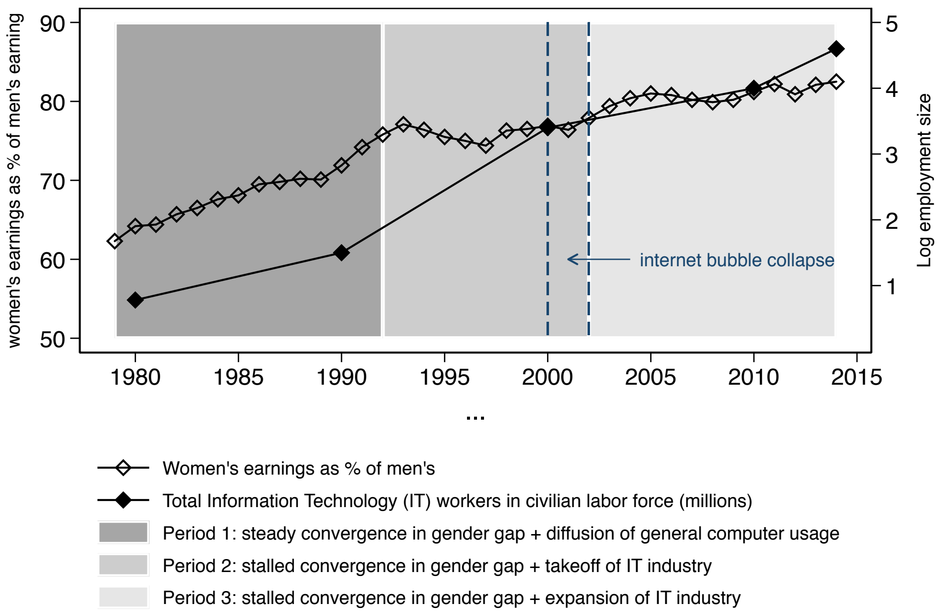
To present a descriptive overview of the relevant historical context, Figure A1 pairs up two historical trends in the United States from 1980 to 2015. The first is the trend in aggregate gender inequality, indicated by women's median weekly earnings as a percentage of men's earnings among full-time workers, and the second is the trend in total employment size in Information Technology (IT) workers - a proxy measure for advances in computer technology. Both aggregate measures come from data published by the Bureau of Labor Statistics.[1] The trends cover three periods, as illustrated by areas of different shades in the figure. The focus of our empirical analysis in this study is on post-mid-1990s years because that is the period when the rise of programming took place, but we start our macro-level trends from the 1980s to provide a fuller background.

The first period runs from 1980 to the early 1990s. General computer usage was diffusing across the workplace, coinciding with a steady convergence of the gender earnings gap. Tasks related to general computer usage, such as typing, bookkeeping, and scheduling, compliment the productivity of office jobs, which have traditionally been performed by women (Blau and Kahn, 2000; Krueger, 1993; Weinberg, 2000; Yamaguchi, 2013). Manual tasks have largely remained the purview of men and falling returns to these are seen to explain the narrowing wage gap between men and women (Black and Spitz-Oener, 2010).

The second period starts from the mid-1990s, where the convergence of the gender earnings gap slowed down and even reverted moderately. Meanwhile, the diffusion of computers had plateaued following its proliferation in the 1980s (Blau and Kahn, 2006; Friedberg, 2003). During this period, the IT industry started to take off, resulting in not just a growing share of workers taking up IT-related occupations, but also in an increase in the complexity and level of specialization in these occupations (Beckhusen, 2016). But the takeoff is found to

[1]Data on trends in employment in IT workers come from a 2013 report (Csorny, 2013) and data on trends in gender inequality come from a 2016 report (BLS, 2016).

Figure A1: Aggregate Trends in the Gender Earnings Gap and Total Employment in Information Technology




**NOTE:** This figure provides a background overview of historical trends in aggregate gender inequality and IT employment. Men's and women's median weekly earnings among full-time, year-round workers come from calculations by Bureau of Labor Statistics based on CPS data. The earnings comparisons are on an aggregate level and do not control for many important determinants that can help explain earnings differences. The employment size of IT workers was calculated by the Bureau of Labor Statistics, based on data from U.S. Census Bureau, Equal Employment Opportunity Supplementary Reports from the 1980, 1990, 2000 censuses and 2010 and 2014 American Community Surveys. The trends cover three periods.

be unequal by gender, with women being less likely than men to major in computer science (Abbate, 2012) and to participate in the IT workforce (Zarrett et al., 2006).

The third period starts around 2002. The IT industry resurged in the early 2000s after a setback due to the collapse of the internet bubble, and continued to expand. The job tasks for these IT workers are increasingly specialized and often involve intensive programming, data processing, or web development. Meanwhile, from 2012 to 2017, real average hourly wage in the information industry increased at a faster rate than the average of all private-sector employees (BLS, 2017). Moreover, this rise of programming-intensive occupations has extended *beyond* the IT industry that directly provides services involving computer infrastructures and computer software. Consider the occupation of *operations research analysts* as an example. Workers in this occupation do not usually work in an IT industry. However, as we have shown, this occupation ranks among the top 20 in over 300 detailed occupations in terms of programming intensity. According to the Bureau of Labor Statistics' online Occupational Outlook Handbook[2], the main tasks of operations research analysts are to "use advanced mathematical and analytical methods to help organizations investigate complex issues, identify and solve problems, and make better decisions." Given this job description, a large part of the tasks performed by these practitioners are likely to involve coding, database processing, and computer-based quantitative analysis. Similar examples of non-IT, programming-intensive occupations include actuaries, statisticians, and mechanical engineers. Another case for the rise of programming extending *beyond* the IT industry is the emergence of the financial technology (also known as FinTech) industry, in which new technology, such as operating systems, software and machine learning algorithms, is being applied to the finance industry to improve financial activities and services (Dapp et al., 2014). Both examples illustrate that the rise of programming intensity is pervasive and is making its way across many industries.

And yet, aggregate statistics also suggest that, the rise of programming-intensive jobs

[2]Link: https://www.bls.gov/ooh/math/operations-research-analysts.htm (accessed on 10/27/2017)

is uneven by gender. The proportion of women in IT occupations has decreased continuously from 31.0% in 1990 to 25.0% in 2014 (Beckhusen, 2016). Both male and female IT workers earn higher income than the population average. However, in contrast to the dramatic increase in IT-related earnings gains over time among men, female IT workers' earnings premium (relative to earnings in all occupations) increased at a much slower speed (Beckhusen, 2016). These gender differentials in the IT industry — in terms of both employment size and economic returns — also coincide with a continued stall in the progress toward overall gender equality, as presented in Figure A1. Women's median earnings as a percentage of men's earnings increased only minimally by 1.7 percentage points from 79.4 % in 2003 to 81.1 % in 2015, exhibiting a much flatter growth pattern than in the 1980s.

## Online Appendix B Procedure for Constructing Occupation-level Skill-intensity Measures in the O*NET Data

*Setting up O*NET Data*

We use the Occupational Information Network (O*NET), to construct the occupation-level skill measures. The O*NET database, developed by the U.S. Department of Labor, provides information about key attributes of occupations in the US. The 21.3 version of the O*NET database, on which our analysis is based, includes data collected over 2003 and 2016. In this database, incumbents, occupational experts, and analysts provide ratings or frequencies, as appropriate, for each element of O*NET according to a scale provided in the O*NET questionnaires.

The O*NET database provides us with not only the key occupation-level measures of programming skill intensity and general computer skill intensity, but also a variety of other occupation-level skill intensity measures, such as importance of verbal or quantitative skills, which are used as control variables in our wage regression models in Step 3.

Of the various aspects of occupational characteristics provided in datasets from O*NET, we focus on five main datasets for constructing occupation-level skill-intensity measures: The *Skills* dataset contains elements that are seen to be developed capacities that facilitate learning of or performance of activities that occur across jobs. The elements in the *Abilities* dataset constitute "enduring attributes of the individual that influence performance." *Work Styles* refer to personal characteristics that can affect how well someone performs in an occupation. The *Knowledge* dataset includes elements that represent organized sets of principles and facts applying in general domains. Lastly, the *Generalized Work Activities* dataset has elements that are common acts performed in jobs. Further, each element or observed variable is associated with two scores – level and importance - representing, respectively, the required level of competence of that occupational characteristic and the importance of that occupational characteristic for each occupation. Since only the importance scores are available for every characteristic that we include in our analysis, we use only the importance scores for our calculations. Supplementary analysis shows that where both level and importance scores are available, they tend to be highly associated. Importance scores are measured on a 1-5 scale. For ease of interpretation, we follow the O*NET recommended formula to convert scores to a scale of 100: S = ( (O - L) / (H - L) ) * 100, where S is the standardized score, O is the original rating score on one of the three scales, L is the lowest possible score on the scale, and H is the highest possible score on the scale.

*Converting O*NET SOC coding scheme to Census 1990 scheme*

O*NET-SOC is a Standard Occupational Classification System (SOC) based on the 2010 SOC scheme. To be able to merge the skill-intensity measures from O*NET with CPS data we converted O*NET's SOC coding scheme to the Census 1990 3-digit coding scheme used in the CPS data (OCC1990). This involved a multi-stage process because no direct crosswalks exist between O*NET-SOC and OCC1990 codes. We used available crosswalks to convert O*NET-SOC codes to SOC 2010 codes[3] then to OCC2000 codes[4] and finally to OCC1990 codes[5]. The O*NET-SOC coding scheme contains more detailed occupational categories than both the Census coding schemes. Therefore, when converting between the two, we aggregated the scores of the O*NET-SOC occupations that were merged into a single Census 2000 occupation. For the conversion from Census 2000 to Census 1990 coding scheme, we used the conversion factors provided in the OCC2000 to OCC1990 crosswalks to calculate weighted means, and these are the final scores used in the analysis.

*Constructing the Occupation-level Skill-intensity Measures*

We largely follow the multi-dimensional skill factors proposed in Liu and Grusky (2013)'s recent study to construct our skill-intensity measures. We make two major modifications to their original coding scheme. First, as a key focus of our study is to isolate

[3]BLS. "2010 Standard Occupational Classification." BLS. Retrieved June 26, 2017 (https://www.bls.gov/soc/soc_structure_2010.pdf).

[4]BLS."2000 Standard Occupational Classification" Retrieved June 26, 2017 (https://www.bls.gov/soc/2000/soc-structure-2000.pdf).

[5]IPUMS. "IPUMS USA | Occupation and Industry Codes and Documentation." Retrieved June 26, 2017 (https://usa.ipums.org/usa/volii/occ_ind.shtml).

programming skills from computer-related skills in general, we separate their original "computer" skill factor into two sub categories: (1) a programming skill factor; and (2) a general computer skills factor. Second, we added a 'Technical Miscellaneous' factor to capture elements such as equipment selection and installation. This decision was based on the results of preliminary exploratory factor analysis that suggested that this latent factor explains a large portion of the variance. Some variables included by Liu and Grusky do not appear in the later O*NET dataset. For example, the "Basic Skills-Entry Requirement" elements that Liu and Grusky deploy (namely, "Getting Information' and "Information Organization') have not been available since O*NET 3.0 at least. We substituted these with other elements from O*NET whose definitions most closely resembled that of the missing variables.

After assigning elements to skill categories, we conducted confirmatory factor analysis to obtain occupation-level factor scores for each skill measure.[6] We ran several iterations of the factor analysis, eliminating elements that had factor loadings of less than 0.50. Table C2 in Online Appendix C presents our final list of latent skill factors and their factor loadings on related O*NET elements. The final skill factors are: cognitive (verbal, quantitative, analytic), creative, technical (programming, general computer, science and engineering, technical miscellaneous), and social (managerial, carework).

[6]In STATA, we used the –confa– command, while also using the structural equation method of –sem– command. Both yielded approximately the same factor loadings for the latent factors.

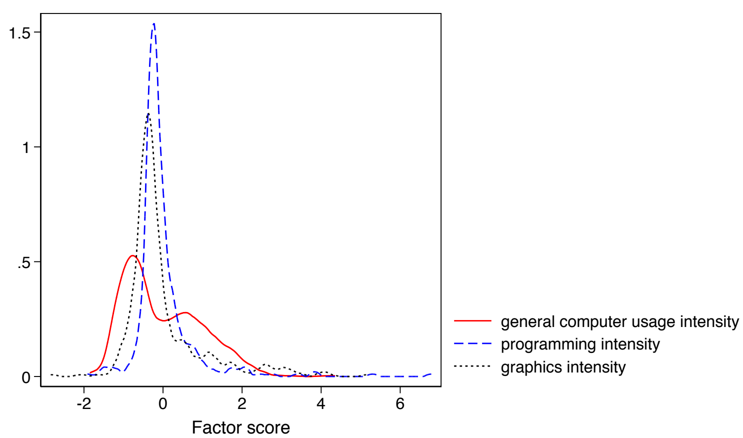
# Online Appendix C Supplementary Tables and Figures for Occupation-level Skill- and Task-intensity Measures

Table C1: Factor Loadings of Task-intensity Measures

| Variable | Factor1 | Factor2 | Factor3 |
|---|---|---|---|
| Programming | 0.703 | 0.360 | 0.358 |
| Spreadsheets or databases | 0.353 | 0.756 | 0.449 |
| Word processing or desktop publishing | 0.190 | 0.873 | 0.387 |
| Scheduling | 0.327 | 0.827 | 0.317 |
| Emails | 0.268 | 0.870 | 0.388 |
| Graphics | 0.262 | 0.441 | 0.858 |

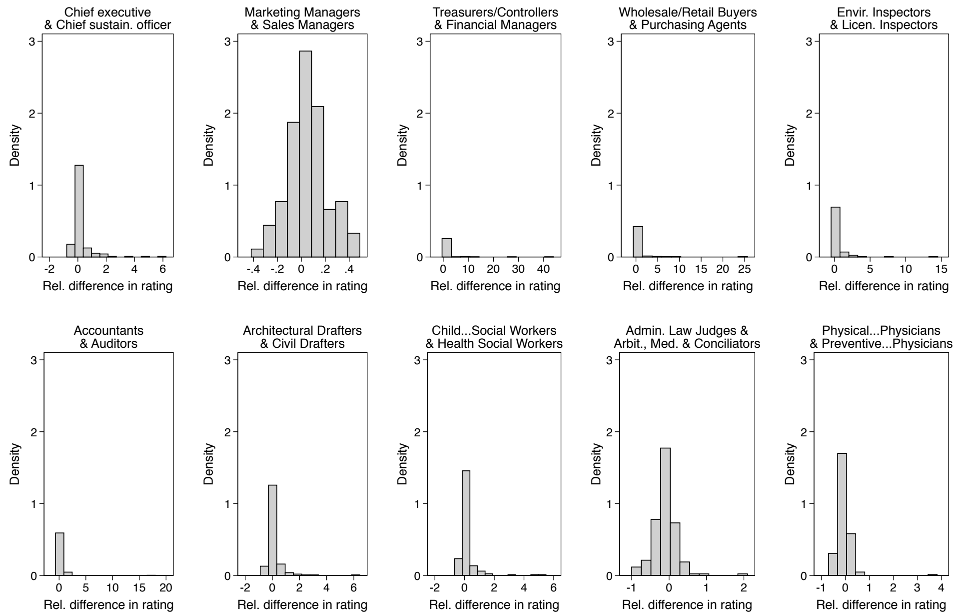
**NOTE:** Occupation-level task intensity measures are based on data from the 1997, 2001, and 2003 waves of the Computer and Internet Use supplement of the CPS. Factors are rotated to represent orthogonal dimensions.

Figure C1: Density Distribution of Factor Scores for Specific Computer-related Tasks




**NOTE:** Occupation-level task intensity measures are based on data from the Computer and Internet Use supplement of the CPS. Factors are rotated to represent orthogonal dimensions.

Figure C2: Distributions of Relative Differences between Incumbent Ratings and Occupational Expert Ratings




**NOTE:** Occupation-level skill intensity measures are based on data from the Occupational Information Network. Pairs of occupations are chosen so that they are similar in skill requirement.

Table C2: Occupation-level Skill Factors, O*NET Items, and Factor Loadings

| Skill Factor | Factor Loading |
|---|---|
| **COGNITIVE** | |
| **Verbal** | |
| Oral Comprehension | 1 |
| Oral Expression | 1.17 |
| Written Expression | 1.56 |
| Reading Comprehension | 1.36 |
| Written Comprehension | 1.38 |
| Active Listening | 1.05 |
| Writing | 1.56 |
| Speaking | 1.17 |
| Interpreting the Meaning of Information for Others | 1.38 |
| Communicating with Persons Outside the Organization | 1.4 |
| **Quantitative** | |
| Mathematical Reasoning | 1 |
| Number Facility | 0.82 |
| Mathematics | 0.92 |
| **Analytic** | |
| Problem Sensitivity | 1 |
| Deductive Reasoning | 1.35 |
| Inductive Reasoning | 1.44 |
| Information Ordering | 0.75 |
| Category Flexibility | 0.71 |
| Speed of Closure | 0.7 |
| Flexibility of Closure | 0.59 |
| Fluency of Ideas | 1.45 |
| Critical Thinking | 1.45 |
| Active Learning | 1.62 |
| Complex Problem Solving | 1.44 |
| Judgement and Decision Making | 1.28 |
| Systems Analysis | 1.75 |
| Systems Evaluation | 1.71 |
| Identifying Objects, Actions, and Events | 0.81 |
| Processing Information | 1.39 |
| Analyzing Data or Information | 2.01 |
| Analytical Thinking | 1.72 |
| Monitoring | 0.83 |
| Getting Information | 0.96 |
| **CREATIVE** | |
| **Creative** | |
| Originality | 1 |
| Fine Arts | 0.65 |
| Thinking Creatively | 1.53 |
| Innovation | 0.92 |
| **TECHNICAL** | |
| **Programming** | |
| Programming | (stand-alone item) |

**Table C2 – continued from previous page**

| Skill Factor | Factor Loading |
|---|---|
| **General computer** | |
| Computers and Electronics | 1 |
| Interacting with Computers | 1.62 |
| Documenting Recording Information | 0.75 |
| **Science and Engineering** | |
| Science | 1 |
| Operations Analysis | 0.71 |
| Technology Design | 0.69 |
| Production and Processing | 0.99 |
| Engineering and Technology | 1.97 |
| Design | 1.75 |
| Building and Construction | 1.32 |
| Mechanical | 1.82 |
| Mathematics | 1.02 |
| Physics | 1.51 |
| Chemistry | 1 |
| Monitor Processes, Materials, or Surroundings | 0.5 |
| Estimating the Quantifiable Characteristics of Products, Events, or Information | 0.8 |
| **Technical Miscellaneous** | |
| Equipment Selection | 1 |
| Installation | 0.55 |
| Operation Monitoring | 1.02 |
| Operation and Control | 1.21 |
| Equipment Maintenance | 1.24 |
| Troubleshooting | 1.11 |
| Repairing | 1.24 |
| Quality Control Analysis | 0.82 |
| Inspecting Equipment, Structures, or Material | 1.04 |
| Controlling Machines and Processes | 1.25 |
| Operating Vehicles, Mechanized Devices, or Equipment | 1.11 |
| Repairing and Maintaining Mechanical Equipment | 1.31 |
| Repairing and Maintaining Electronic Equipment | 0.81 |
| **SOCIAL** | |
| **Managerial** | |
| Time Management | 1 |
| Management of Financial Resources | 1.31 |
| Management of Material Resources | 1.19 |
| Management of Personnel Resources | 1.79 |
| Administration and Management | 1.72 |
| Economics and Accounting | 1.35 |
| Personnel and Human Resource | 1.57 |
| Developing Objectives and Strategies | 1.82 |
| Scheduling Work and Activities | 1.69 |
| Organizing, Planning, and Prioritizing | 1.37 |
| Coordinating the Work and Activities of Others | 1.67 |
| Developing and Building Teams | 1.65 |
| Guiding, Directing, and Motivating Subordinates | 2 |
| Provide Consultation and Advice to Others | 1.74 |
| Performing Administrative Activities | 1.38 |

**Table C2 – continued from previous page**

| Skill Factor | Factor Loading |
|---|---|
| Staffing Organizational Units | 1.84 |
| Monitoring and Controlling Resource | 1.66 |
| Leadership | 1.63 |
| **Carework** | |
| Social Perceptiveness | 1 |
| Instructing | 1.05 |
| Service Orientation | 1.21 |
| Therapy and Counseling | 1.39 |
| Assisting and Caring for Others | 1.5 |
| Cooperation | 0.68 |
| Concern for Others | 1.21 |
| Social Orientation | 1.28 |

**NOTE:** Occupation-level skill intensity measures are based on data from the Occupational Information Network.

Table C3: Correlation Matrix for Occupational Skill Measures

| | verbal | quantitative | analytic | creative | programming | computer | sci & engineering | tech misc | managerial | carework |
|---|---|---|---|---|---|---|---|---|---|---|
| verbal | 1 | | | | | | | | | |
| quantitative | 0.473*** | 1 | | | | | | | | |
| analytic | 0.860*** | 0.664*** | 1 | | | | | | | |
| creative | 0.620*** | 0.305*** | 0.657*** | 1 | | | | | | |
| programming | 0.288*** | 0.676*** | 0.455*** | 0.309*** | 1 | | | | | |
| computer | 0.625*** | 0.604*** | 0.501*** | 0.267*** | 0.591*** | 1 | | | | |
| sci & engineering | 0.312*** | 0.600*** | 0.575*** | 0.441*** | 0.538*** | 0.316*** | 1 | | | |
| tech misc | -0.275*** | 0.0953 | -0.000971 | -0.0983 | 0.189** | -0.124 | 0.539*** | 1 | | |
| managerial | 0.606*** | 0.435*** | 0.730*** | 0.566*** | 0.184** | 0.276*** | 0.367*** | 0.00294 | 1 | |
| carework | 0.203** | -0.309*** | 0.112 | 0.0786 | -0.387*** | -0.198** | -0.351*** | -0.178** | 0.349*** | 1 |

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

**NOTE:** Occupation-level skill intensity measures are based on data from the Occupational Information Network.

Table C4: List of Ten Top-ranked Occupations by Skill-specific Factor Score

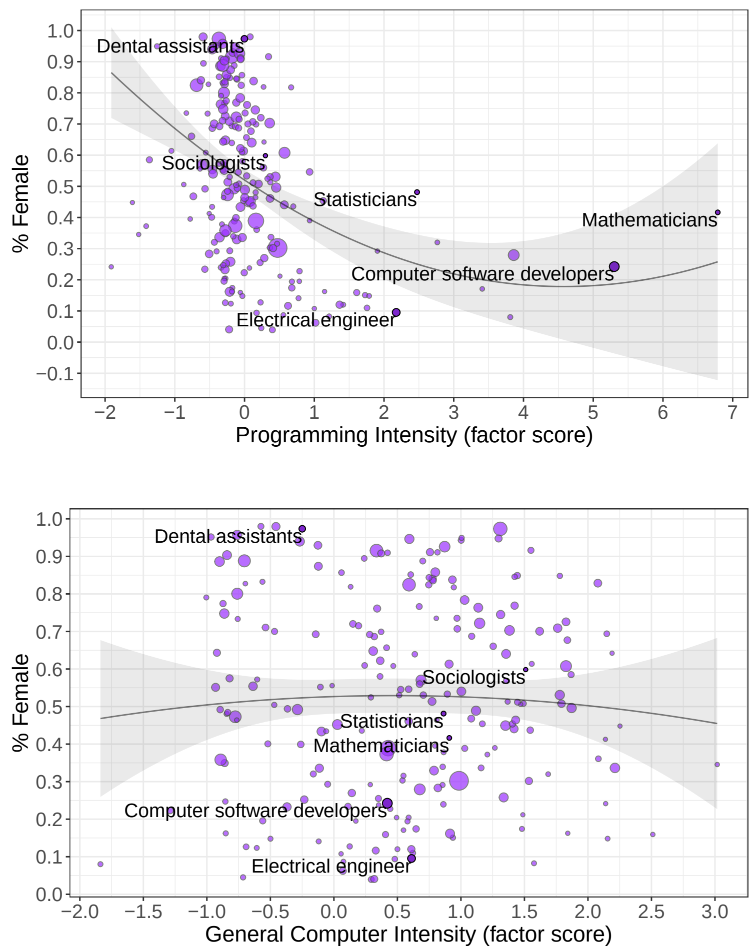
| Skill Type | Ten Top-Ranked Occupations |
|---|---|
| **Cognitive** | |
| Verbal | Psychologists; Lawyers; Clergy; Atmospheric and space scientists; Technical writers; Authors; Editors and reporters; Judges; Social Scientists, n.e.c.; Sociologists |
| Quantitative | Mathematical scientists n.e.c.; Actuaries; Physists and astronomers; Statisticians; Accountants and auditors; Chemical engineers; Economists; Aerospace engineers; Operations and system researchers and analysts; Civil Engineers |
| Analytic | Medical scientists; Actuaries; Chief executives & general administrators, public administration; Physicians; Physicists and astronomers; Agricultural engineers; Aerospace engineers; Psychologists; Judges; Mining engineers |
| **Creative** | |
| Creative | Authors; Painters, sculptors, craft-artists, and artist printmakers; Architects; Dancers; Physicists and astronomers; Designers; Actors and directors; Photographers; Sales support occupations n.e.c.; Editors and reporters |
| **Technical** | |
| Programming | Computer programmers; Tool programmers, numerical control; Statisticians; Computer systems analysts and scientists; Mining engineers; Operations and systems researchers and analysts; Physicists and astronomers; Computer operators; Actuaries; Peripheral equipment operators |
| Computer | Sales workers, parts; Computer programmers; Computer operators; Atmospheric and space scientists; Peripheral equipment operators; Aerospace engineers; Correspondence clerks; Data-entry keyers; Computer systems analysts and scientists; Actuaries |
| Science and Engineering | Agricultural engineers; Mechanical engineers; Nuclear engineers; Chemical engineers; Aerospace engineers; Metallurgical and materials engineers; Engineers n.e.c.; Civil engineers; Mining engineers; Marine engineers and naval architects |
| Technical Miscellaneous | Tool programmers, numerical control; Firefighting occupations; Mechanical engineering technicians; Industrial engineering technicians; Electrical and electronic technicians; Chemical technicians; Engineering technicians, n.e.c.; Office machine operators, n.e.c.; Duplicating machine operators; Science technicians, n.e.c. |

**Table C4 – continued from previous page**

| Skill Type | Top-Ranked Occupations |
|---|---|
| **Social** | |
| Managerial | Managers, medicine and health; Postmasters and mail superintendents; Administrators, education and related fields; Chief executives & general administrators, public administration; Supervisors, police and detectives; Personnel and labor relations managers; Purchasing managers; Managers, food serving and lodging establishments; Urban planners; Managers, service organizations, n.e.c. |
| Carework | Counselors, educational and vocational; Physical therapists; Clergy; Social workers; Licensed practical nurses; Religious workers, n.e.c.; Psychologists; Occupational therapists; Therapists, n.e.c.; Respiratory therapists |

**NOTE:** Occupation-level skill intensity measures are based on data from the Occupational Information Network.

Figure C3: Occupation-level Task-Intensity and Percentage of Female Workers




**NOTE:** Occupation-level task intensity measures are based on data from the 1997, 2001, and 2003 waves of the Computer and Internet Use supplement of the CPS; occupation-level skill intensity measures are based on data from the Occupational Information Network. Dot sizes in this figure correspond to employment sizes of the corresponding occupation. On each plot, six occupations are labeled (with occupation titles marked next to darker-shaded dots) as examples. The solid line and shaded area represent the quadratic fit and 95% confidence interval respectively.

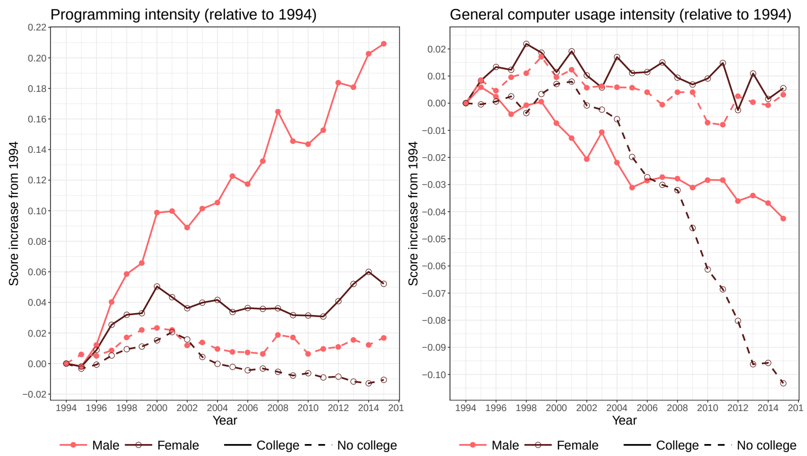
## Online Appendix D Descriptive Statistics for Demographic, Educational, and Employment & Job Characteristics

| | Mean/% (S.D.) |
|---|---|
| **Demographic variables** | |
| Age | 39.4393 (8.5) |
| Female | 50.77% |
| Black | 12.74% |
| Asian | 5.12% |
| Other race/ethnicity | 1.68% |
| | |
| **Educational attainment** | |
| Less than high school | 11.46% |
| Some college | 27.60% |
| BA degree | 20.70% |
| MA degree | 7.06% |
| PhD/Professional degree | 2.74% |
| | |
| **Employment & job characteristics** | |
| No union coverage | 84.38% |
| Member of labor union | 14.16% |
| Union but not a member | 1.46% |
| | |
| Full-time worker | 80.34% |
| | |
| Industry: agricultural | 2.01% |
| Industry: mining | 0.47% |
| Industry: construction | 6.57% |
| Industry: manufacturing | 11.72% |
| Industry: transportation | 6.37% |
| Industry: wholesale | 2.87% |
| Industry: retail | 12.09% |
| Industry: finance | 5.74% |
| Industry: business service | 6.28% |
| Industry: personal service | 2.66% |
| Industry: entertainment | 1.48% |
| Industry: professional | 22.35% |
| Industry: public administration | 4.17% |
| Industry: military | 0.09% |
| N | 3,136,881 |

**NOTE:** All statistics are weighted to represent the population at each year.

## Online Appendix E

Figure E1: Gender-specific Trends in Average Factor Scores of General Computer Usage Intensity and Programming Intensity in Task-based Measures (Relative to 1994 Level)




**NOTE:** Data source: 1994-2015 waves of the merged outgoing rotation groups files of the CPS. All statistics are weighted to represent the population in each year.

# Online Appendix F Comparing $R^2$ Statistics for the Skill/Task Model against the Fixed-effect Model

In this appendix, we provide a formal assessment of the extent to which our wage regression models explain the between-occupation differences in wages. The main model employed in our study relies on a set of occupation-level task- and skill-intensity measures to capture the occupational differences in wages (hereafter, "skill/task model"). Another strategy, one that has also been adopted by previous literature on occupational sex segregation, is to include occupation-specific fixed effects to absorb the maximum amount of between-occupation variations (hereafter, "fixed-effect model"). The fixed-effect model is typically used when the aim is to assess the extent to which wage variations can be attributed to between-occupation variations, but it does not allow for assessment of the effects of particular occupation-level characteristics because all occupational variations are already absorbed in the dummies. However, the fixed-effect model can be used as a benchmark for assessing how well our skill/task model explains the total and between-occupation variations.

To what extent does our skill/task model explain the total between-occupation variations as shown in the fixed-effect model? We examine this by presenting the $R^2$ and incremental $R^2$ statistics for three model specifications (no occupation controls, skill/task model, fixed-effect model), as shown in Table F1. We take the ratios of $R^2$ and incremental $R^2$ between the skill/task model and fixed effect model in Column D of Table F1. The ratio of $R^2$ measures the percentage of *total* wage variation (i.e. between- and within-occupation) explained by the skill/task model relative to the fixed-effect model, and the ratio of incremental $R^2$ measures the percentage of *between-occupation* wage variation explained by the skill/task model relative to the fixed-effect model. The table suggests that the skill/task model does a very good job explaining the total wage variations (about 94% among men and 89% among women), and a reasonably good job explaining the between-occupation portion of wage variations (about 72% among men and 58% among women). In sum, these analyses assure us that our skill/task model provides a reasonably good fit for the data.

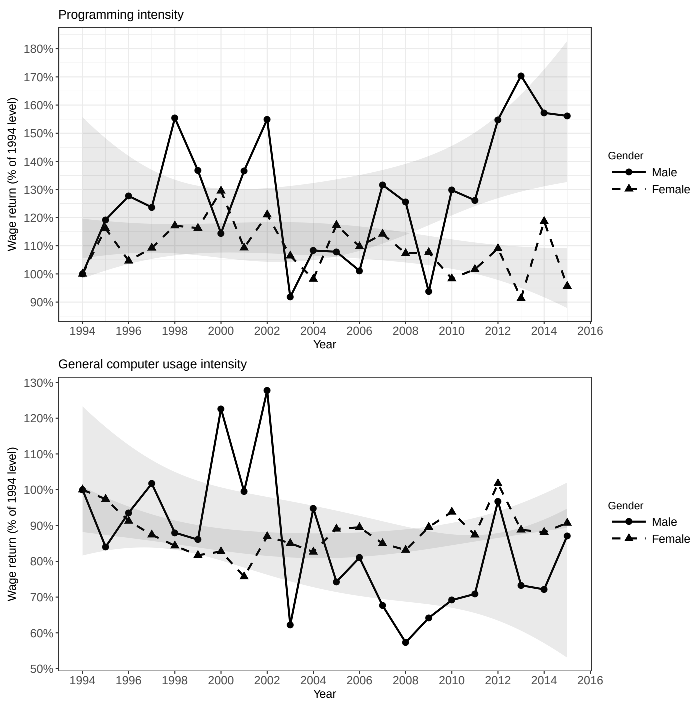
Table F1: $R^2$ and Incremental $R^2$ in Skill/Task Models versus Fixed-effect Models with Occupation Dummies by Gender and Period

| | A | B | C | D |
|---|---|---|---|---|
| | No occupation variables | Skill/task model | Fixed-effect model | Ratio B/C |
| **Men, 1994-2004** | | | | |
| $R^2$ | 0.275 | 0.334 | 0.356 | **93.82%** |
| Incremental $R^2$ | — | 0.059 | 0.081 | **72.84%** |
| **Men, 2005-2015** | | | | |
| $R^2$ | 0.279 | 0.331 | 0.351 | **94.30%** |
| Incremental $R^2$ | — | 0.052 | 0.072 | **72.22%** |
| **Women 1994-2004** | | | | |
| $R^2$ | 0.269 | 0.324 | 0.363 | **89.26%** |
| Incremental $R^2$ | — | 0.055 | 0.094 | **58.51%** |
| **Women, 2005-2015** | | | | |
| $R^2$ | 0.258 | 0.307 | 0.343 | **89.50%** |
| Incremental $R^2$ | — | 0.049 | 0.085 | **57.65%** |

**NOTE:** Data source: 1994-2015 waves of the merged outgoing rotation groups files of the Current Population Survey. All statistics are weighted to represent the population in each year. Column A includes only demographic, educational, and employment characteristics but no occupation variables. Skill/task models (Column B) are identical to Model M4 used in the main analysis of this study. Fixed-effect models (Column C) replace the occupation-level skill and task measures with 386 dummy variables for each detailed 3-digit Census occupation (with one dummy held as the reference group). Incremental $R^2$ refers to the change in $R^2$ from Column A to Column B or C.

## Online Appendix G

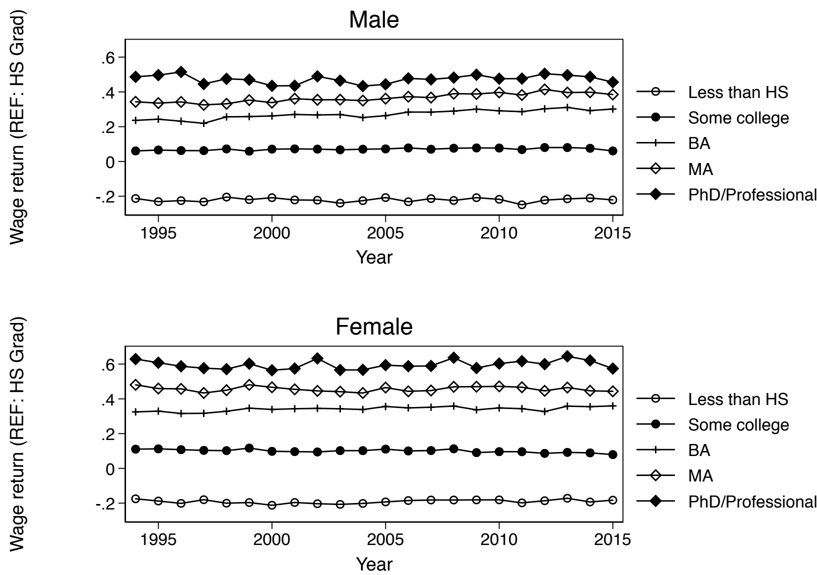
Figure G1: Gender-specific Trends in Wage Returns (as % of 1994 Level) to Programming Intensity and General Computer Usage Intensity, Unsmoothed




**NOTE:** Data source: 1994-2015 waves of the merged outgoing rotation groups files of the CPS. Occupation-level task intensity measures are based on data from the Computer and Internet Use supplement of the CPS. Hourly wages are adjusted to 1999 dollars. All statistics are weighted to represent the population in each year.

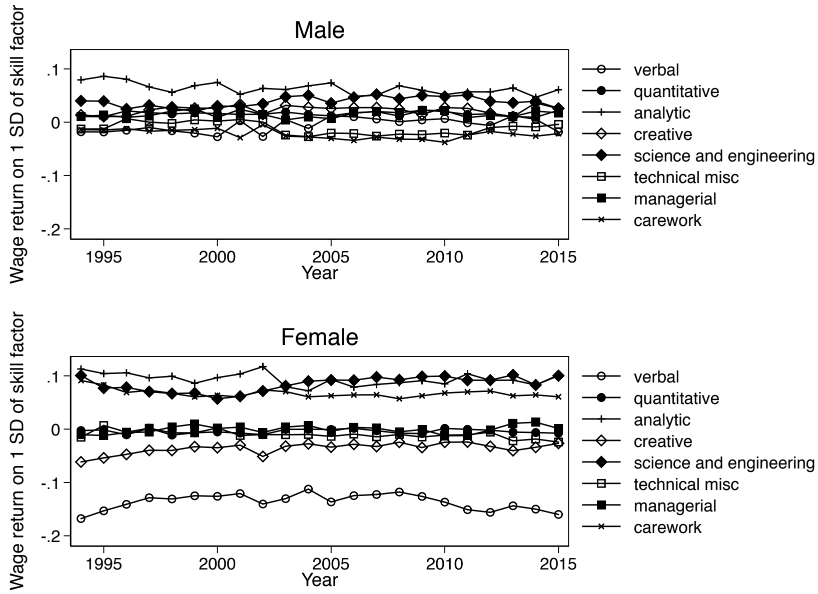
# Online Appendix H Trends in Wage Returns to Education, Occupational Skills, and Race

Figure H1: Gender-specific Trends in Wage Returns to Educational Attainment (Reference Category = High School Graduates)



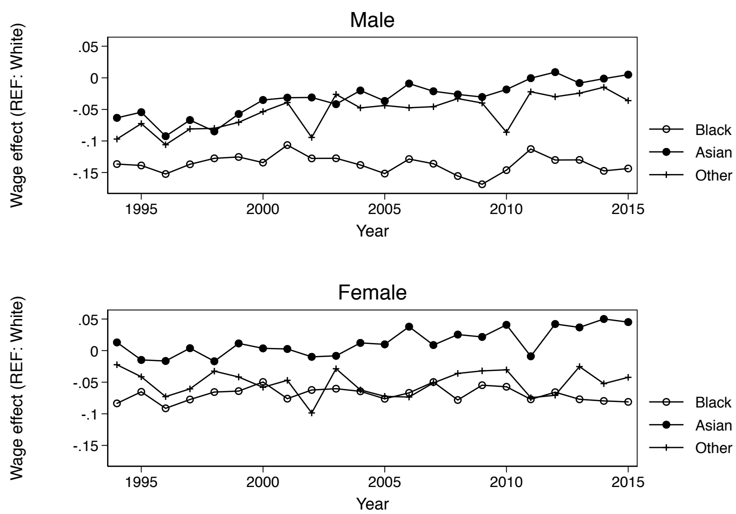

**NOTE:** Data source: 1994-2015 waves of the merged outgoing rotation groups files of the CPS. Hourly wages are adjusted to 1999 dollars. All statistics are weighted to represent the population in each year.

Figure H2: Gender-specific Trends in Wage Returns to Occupational Skill Importance Scores




**NOTE:** Data source: 1994-2015 waves of the merged outgoing rotation groups files of the CPS. Occupation-level task intensity measures are based on data from the Computer and Internet Use supplement of the CPS. Occupation-level skill intensity measures are based on data from the Occupational Information Network. Hourly wages are adjusted to 1999 dollars. All statistics are weighted to represent the population in each year.

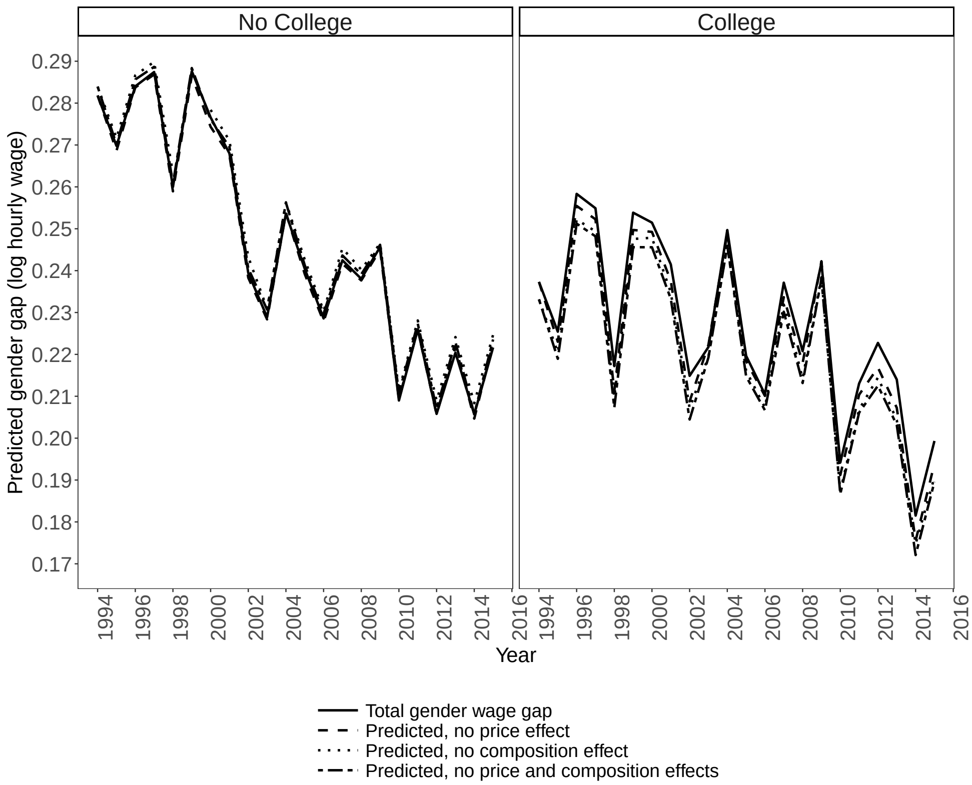
Figure H3: Gender-specific Trends in Wage Returns to Race (Reference Category = White)




**NOTE:** Data source: 1994-2015 waves of the merged outgoing rotation groups files of the CPS. Hourly wages are adjusted to 1999 dollars. All statistics are weighted to represent the population in each year.

## Online Appendix I

Figure I1: Trajectories of the Composition and Price Effects of Programming Intensity on the Gender Gap in Log Hourly Wage, Unsmoothed




**NOTE:** Data source: 1994-2015 waves of the merged outgoing rotation groups files of the CPS. Occupation-level task intensity measures are based on data from the Computer and Internet Use supplement of the CPS. Hourly wages are adjusted to 1999 dollars. All statistics are weighted to represent the population in each year. Technical procedure for the counterfactual decompositions are described in Online Appendix J.

## Online Appendix J Technical Notes on Counterfactual Simulations for Composition and Price Effects

In Step 4 of our empirical analysis, we predicted two counterfactual trajectories for the gender wage gap. This appendix section presents some technical details on constructing these counterfactual simulations.

*Fixing Composition at Base Year*

In the first counterfactual trajectory, we "turn off" the composition effect — that is, we fix the levels of occupational programming intensity for men and women at their respective levels of the beginning year of this period (year 1994). To do so, we predict the mean log hourly wage for men and women as follows:

$$E(\hat{Y}^{men}|\text{no composition}) = \beta_0^m + \beta^{prog,m} \cdot \delta^{prog,m1994} + \text{other covariates,} \tag{1}$$

$$E(\hat{Y}^{women}|\text{no composition}) = \beta_0^w + \beta^{prog,w} \cdot \delta^{prog,w1994} + \text{other covariates,} \tag{2}$$

where $\delta^{prog,m1994}$ and $\delta^{prog,w1994}$ represent the average programming intensity in 1994 for men and women respectively. We then take the difference between the two as the predicted counterfactual trajectory of gender wage gap under the condition that the composition effect is removed:

$$E(\Delta Y|\text{no composition}) = E(\hat{Y}^{men}|\text{no composition}) - E(\hat{Y}^{women}|\text{no composition}) \tag{3}$$

*Fixing Price (Wage Returns) at Base Year*

Similarly, in the second counterfactual trajectory, we "turn off" the price effect by fixing the wage returns to occupational programming intensity for men and women at their 1994 levels:

$$E(\hat{Y}^{men}|\text{no price}) = \beta_0^m + \beta^{prog,m1994} \cdot \delta^{prog,m} + \text{other covariates,} \tag{4}$$

$$E(\hat{Y}^{women}|\text{no price}) = \beta_0^w + \beta^{prog,w1994} \cdot \delta^{prog,w} + \text{other covariates,} \tag{5}$$

where $\beta^{prog,m1994}$ and $\beta^{prog,w1994}$ represent the wage returns to programming intensity in 1994 for men and women respectively. Then, the counterfactual trajectory of gender wage gap under the condition that the price effect is removed can be expressed as:

$$E(\Delta Y|\text{no price}) = E(\hat{Y}^{men}|\text{no price}) - E(\hat{Y}^{women}|\text{no price}) \tag{6}$$

The trajectories $E(\Delta Y|\text{no composition})$ and $E(\Delta Y|\text{no price})$ are thus plotted in Step 4 of the main analysis.

*Fixing Composition and Price at Base Year*

Lastly, we assess the extent to which the combination of price and composition effects can explain the slow convergence of the gender wage gap. To do so, we predict the following counterfactuals by fixing both composition and price at base year (1994):

$$E(\hat{Y}^{men}|\text{no composition, no price}) = \beta_0^m + \beta^{prog,m1994} \cdot \delta^{prog,m1994} + \text{other covariates,} \tag{7}$$

$$E(\hat{Y}^{women}|\text{no composition, no price}) = \beta_0^w + \beta^{prog,w1994} \cdot \delta^{prog,w1994} + \text{other covariates,} \tag{8}$$

We then predict the counterfactual trajectory for gender wage gap under the assumption of no price and composition effects:

$$\begin{aligned} E(\Delta Y|\text{no composition, no price}) = E(\hat{Y}^{men}&|\text{no composition, no price}) \\ -E(\hat{Y}^{women}&|\text{no composition, no price}) \end{aligned} \tag{9}$$

*Why the Total Effect Does not Equal the Sum of Separate Effects*

We note that the combined effects of price and composition do not equal the sum of their separate effects. This is because the price and composition effects affect the predicted wages *multiplicatively* rather than additively. To see this, consider an example in which the predicted wage ($Y$) is determined by the product of composition ($X$) and price ($\beta$), so that the predicted wages at Time 0 and Time 1 are:

$$E(Y_0) = \beta_0 X_0 \tag{10}$$

$$E(Y_1) = \beta_1 X_1. \tag{11}$$

When we "fix" composition, $E(Y_1)$ becomes:

$$E(Y_1|\text{no composition}) = \beta_1 X_0. \tag{12}$$

When we "fix" price, $E(Y_1)$ becomes:

$$E(Y_1|\text{no price}) = \beta_0 X_1. \tag{13}$$

After re-arranging the above equations together, it can be shown that the total effect equals:

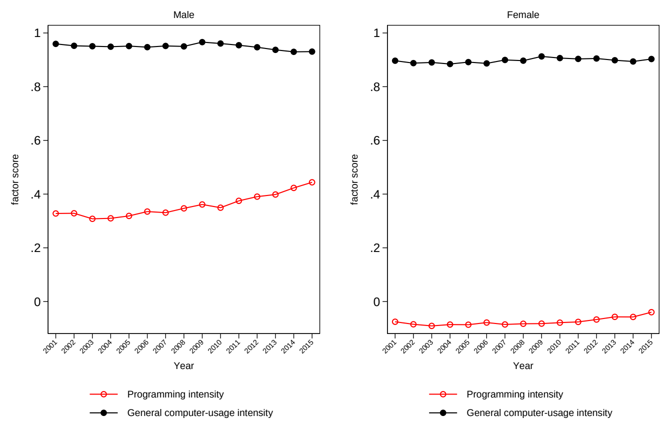
$$\begin{aligned} E(Y_1|\text{ no composition, no price}) &= \beta_1 X_1 - \beta_0 X_0 \\ &= \underbrace{\beta_0 X_1 - \beta_0 X_0}_{\text{composition effect}} + \underbrace{\beta_1 X_0 - \beta_0 X_0}_{\text{price effect}} + \underbrace{(\beta_1 - \beta_0)(X_1 - X_0)}_{\text{multiplicative effect}} \end{aligned} \tag{14}$$

Equation 14 illustrates that the total effect is the sum of two separate effects as well as an overlapping portion. This is the reason why the combination of price and composition effects does not exactly equal the sum of the two separate effects.

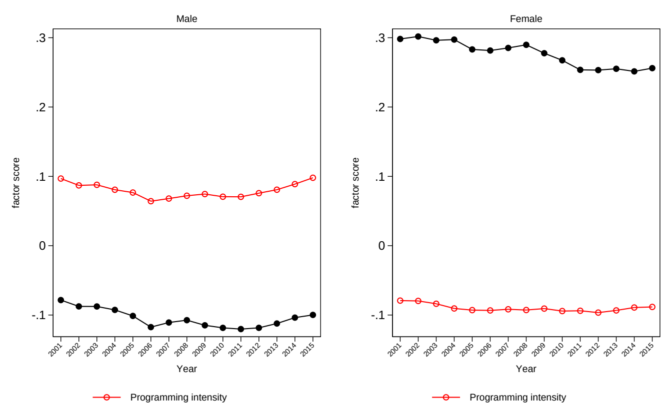
# Online Appendix K Gender-specific Trends in Programming and General Computer Usage Intensity in the American Community Survey Data

Figure K1: Gender-specific Trends in Average Factor Scores of General Computer Usage Intensity and Programming Intensity in Task-based Measures in ACS Data

**Panel A: College**



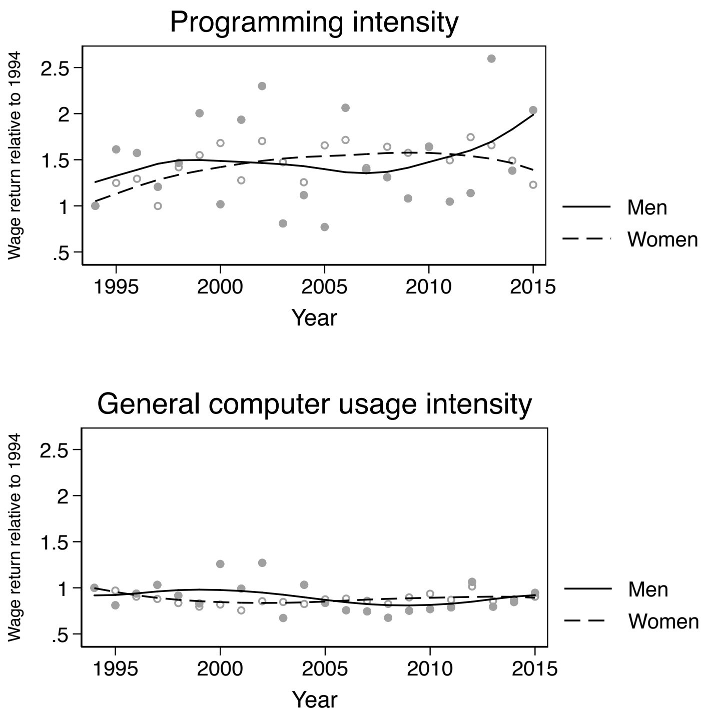

**Panel B: No College**




**NOTE:** Data source: 2001-2015 waves of the American Community Survey. All statistics are weighted to represent the population in each year.

## Online Appendix L Excluding Top 10 Programming-intensive Occupations

Figure L1: Trends in Wage Returns to Programming Intensity, with Top 10 Programming-intensive Occupations Excluded

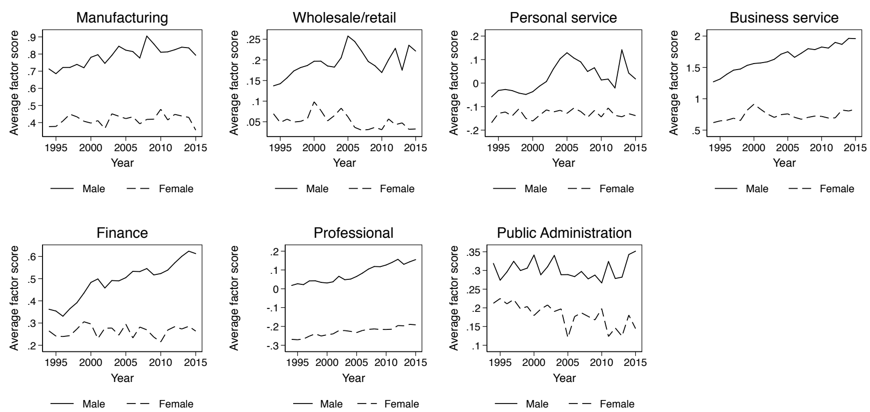



**NOTE:** Data source: 1994-2015 waves of the merged outgoing rotation groups files of the CPS. All statistics are weighted to represent the population in each year. The dots in the figure are point estimates of wage returns.

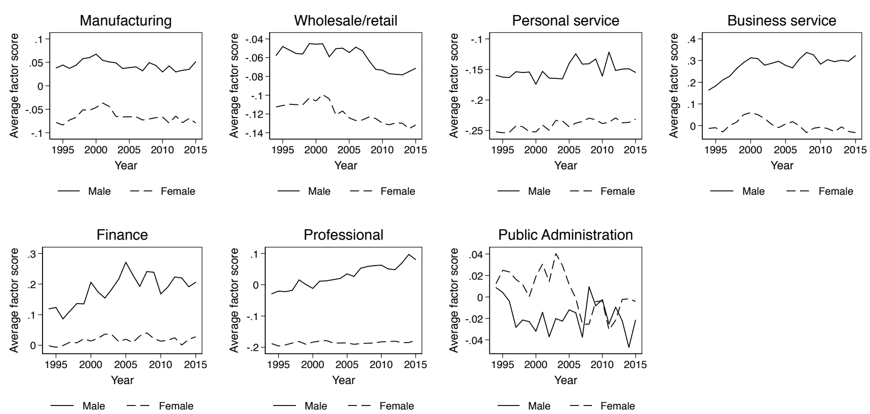
# Online Appendix M Supplementary Analysis: Gender-specific Trends in Average Factor Scores by Major Industrial Categories

Figure M1: Gender-specific Trends in Average Factor Scores of Programming Intensity by Major Industrial Categories

**Panel A: College**




**Panel B: No college**



## Online Appendix N Visualizing Results for Composition and Price Changes within Census Occupational Classification Schemes in the CPS Data

Figure N1: Visualizing Composition Changes within Census Occupational Classification Schemes in the CPS Data

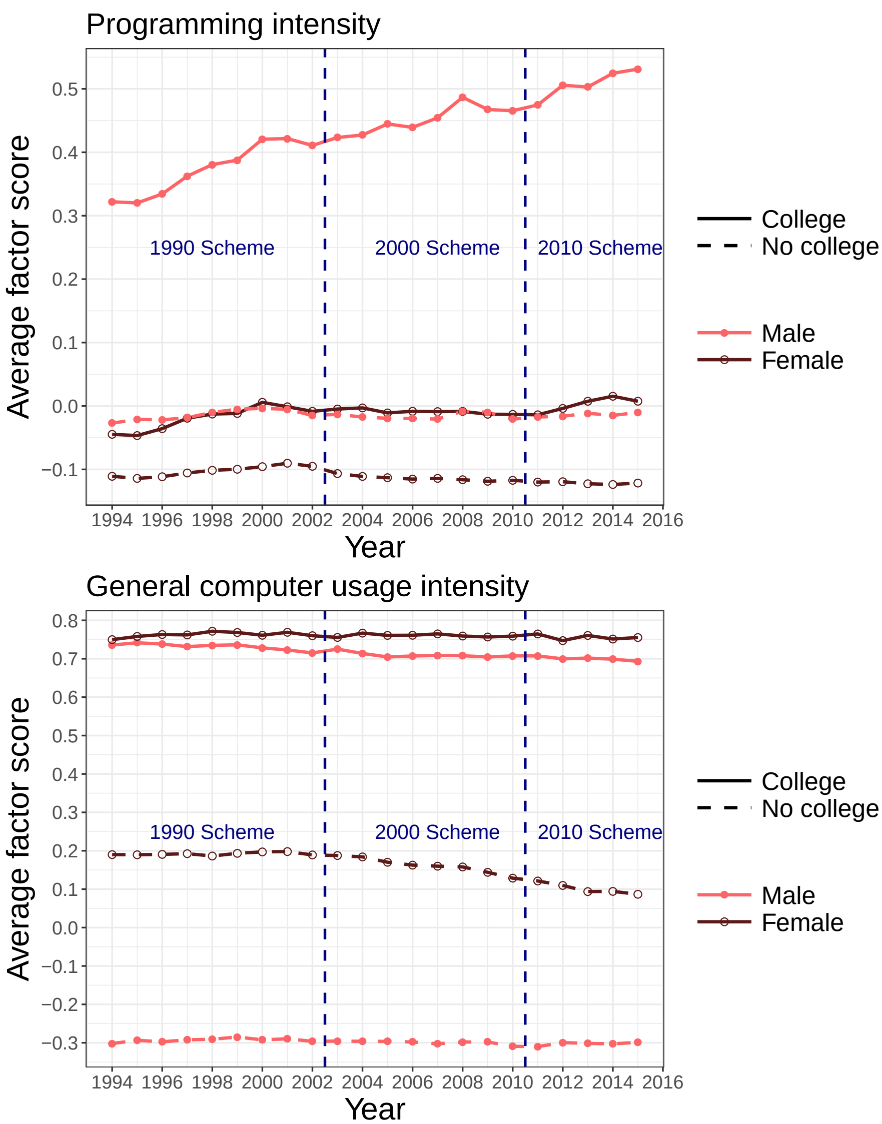
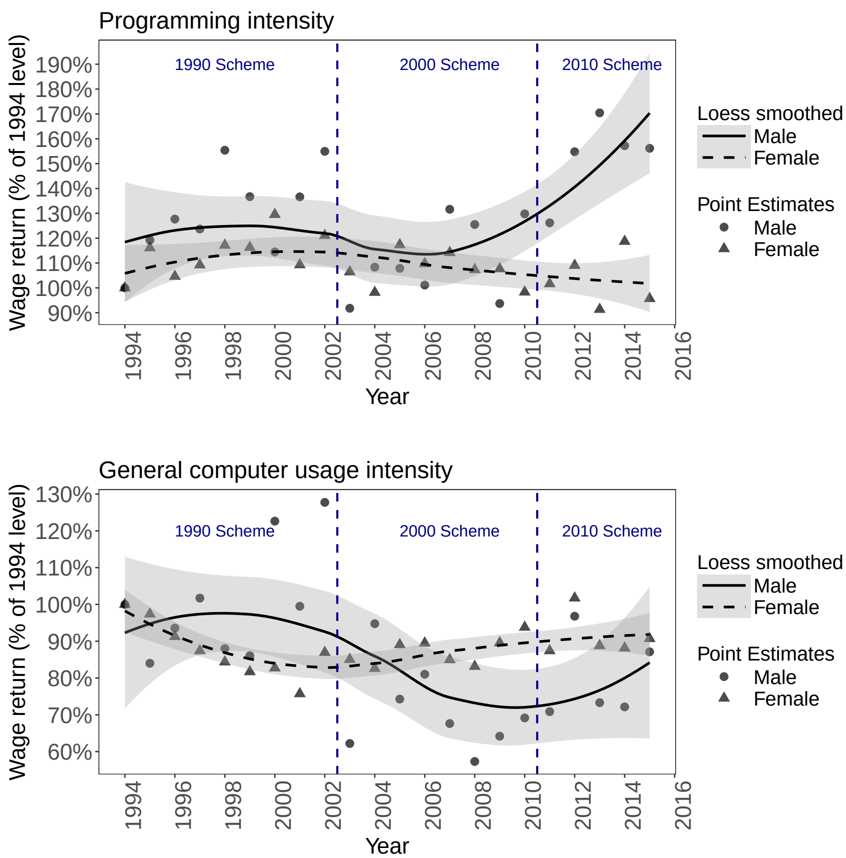


Figure N2: Visualizing Price Changes within Census Occupational Classification Schemes in the CPS Data



## References

Abbate, J. (2012). *Recoding gender: Women's changing participation in computing*. MIT Press.

Beckhusen, J. (2016). *Occupations in Information Technology*. US Department of Commerce, Economics and Statistics Administration, US Census Bureau.

Black, S. E. and A. Spitz-Oener (2010). Explaining women's success: technological change and the skill content of women's work. *The Review of Economics and Statistics 92*(1), 187–194.

Blau, F. D. and L. M. Kahn (2000). Gender differences in pay. Technical report, National Bureau of Economic Research.

Blau, F. D. and L. M. Kahn (2006). The us gender pay gap in the 1990s: Slowing convergence. *Industrial and Labor Relations Review 60*(1), 45–66.

BLS (2016). Highlights of women's earnings in 2015, U.S. Bureau of Labor Statistics.

BLS (2017). The economics daily, real hourly earnings up 2.2 percent over the year in information industry on the internet.

Csorny, L. (2013). Careers in the growing field of information technology services: Beyond the numbers. *US Bureau of Labor Statistics*.

Dapp, T., L. Slomka, D. B. AG, and R. Hoffmann (2014). Fintech–The digital (r)evolution in the financial sector. *Deutsche Bank Research, Frankfurt am Main*.

Friedberg, L. (2003). The impact of technological change on older workers: Evidence from data on computer use. *ILR Review 56*(3), 511–529.

Krueger, A. B. (1993). How computers have changed the wage structure: Evidence from microdata, 1984–1989. *The Quarterly Journal of Economics 108*(1), 33–60.

Liu, Y. and D. B. Grusky (2013). The payoff to skill in the third industrial revolution. *American Journal of Sociology 118*(5), 1330–1374.

Weinberg, B. A. (2000). Computer use and the demand for female workers. *ILR Review 53*(2), 290–308.

Yamaguchi, S. (2013). Changes in Returns to Task-Specific Skills and Gender Wage Gap.

Zarrett, N., O. Malanchuk, P. E. Davis-Kean, and J. Eccles (2006). Examining the Gender Gap in IT by Race: Young Adults' Decision to Pursue an IT Career. In J. M. Cohoon and W. Aspray (Eds.), *Women and Information Technology: Research on Underrepresentation.* MIT Press.