Northwestern | INSTITUTE FOR POLICY RESEARCH

RESEARCH EXCELLENCE • POLICY IMPACT

IPR Working Paper Series

WP-24-16

One Cohort at a Time: A Perspective on the Declining Gender Pay Gap

Jaime Arellano-Bover
Yale University

Nicola Bianchi
Northwestern University and IPR

Salvatore Lattanzio
Bank of Italy

Matteo Paradisi
Einaudi Institute for Economics and Finance

Version: July 12, 2024

DRAFT

Please do not quote or distribute without permission.

Abstract

This paper studies the interaction between the decrease in the gender pay gap and the stagnation in the careers of younger workers, analyzing data from the United States, Italy, Canada, and the United Kingdom. The researchers propose a model of the labor market in which a larger supply of older workers can crowd out younger workers from top-paying positions. These negative career spillovers disproportionately affect the career trajectories of younger men because they are more likely than younger women to hold higher-paying jobs at baseline. The data strongly support this cohort-driven interpretation of the shrinking gender pay gap. The whole decline in the gap originates from (i) newer worker cohorts who enter the labor market with smaller-than-average gender pay gaps and (ii) older worker cohorts who exit with higher-than-average gender pay gaps. As predicted by the model, the gender pay convergence at labor-market entry stems from younger men's larger positional losses in the wage distribution. Younger men experience the largest positional losses within higher-paying firms, in which they become less represented over time at a faster rate than younger women. Finally, they document that labor-market exit is the sole contributor to the decline in the gender pay gap after the mid-1990s, which implies no full gender pay convergence for the foreseeable future. Consistent with the framework, the researchers find evidence that most of the remaining gender pay gap at entry depends on predetermined educational choices.

1 Introduction

The gender pay gap has been decreasing in many high-income economies since the mid-1970s. However, this decrease slowed down around the turn of the twenty-first century for reasons that are not yet well understood.[1] Throughout the same period in which the earnings differentials between men and women have been shrinking, younger workers have been faring progressively worse compared to older workers, experiencing a widening age pay gap (Rosolia and Torrini, 2007; Bianchi and Paradisi, 2023), lower likelihood of receiving promotions and reaching higher-paying jobs (Bianchi et al., 2023), and decreasing lifetime income (limited to male workers; Guvenen et al., 2022). In this paper, we document why the worsening in the careers of younger workers is related to the narrowing of the gender pay gap and its subsequent slowdown. In doing so, we provide a series of novel facts that contribute to the existing knowledge on trends in the gender pay gap.

We first extend a model of the labor market with cross-cohort spillovers between younger and older workers, such as the one discussed in Bianchi and Paradisi (2023), to include gender groups. The overarching goal of this conceptual framework is to illustrate how an increased supply of older workers can narrow the gender pay gap by damaging the labor-market outcomes of younger men more than those of younger women. We then use a combination of survey and administrative data for the United States, Italy, Canada, and the United Kingdom to validate key assumptions of the model and test its main predictions, which all find empirical support.[2] Possibly contrary to popular belief, we establish that gains in women's labor-market outcomes were not the primary drivers of the quick decrease in the gender pay gap that many high-income economies experienced between the 1970s and the 1990s. Relatedly, the more recent slowdown in the closing of the gender pay gap has not stemmed from a stagnation of the labor-market gains experienced by women.

In the model, firms employ younger and older workers in more productive top-level jobs and less productive bottom-level ones. To match the data, we introduce a gender pay gap in favor of men by assuming that firms pay a higher cost for hiring women in top jobs, which leads to their underrepresentation at the top of firms' hierarchies. This differential cost could result from statistical or taste-based discrimination, among other factors.

Furthermore, we make the careers of older and younger workers interconnected so that an improvement in the labor-market conditions of the former can come at a cost for the latter. Specifically, we assume that firms cannot change incumbent older workers' wages

[1] See, for example, England, Levine, and Mishel (2020) for the United States.

[2] Current Population Survey data for the United States (1976-2019) and Social Security data for Italy (1976-2019) are the main focus of our analysis, while smaller-scale survey data for Canada (1973-2011) and the UK (1976-2019) mainly serve as robustness checks.

and job allocation. Moreover, firms pay both a wage and an *organizational cost* to create and maintain a new top job. The latter cost captures the notion that new top jobs often come with managerial responsibilities and decision-making power that either stem from a growing business or are carved out from other existing units. Consequently, firms cannot always expand the top part of their hierarchies to accommodate all younger workers who are qualified to receive a promotion (Lazear, Shaw, and Stanton, 2018). In this context, Bianchi and Paradisi (2023) shows that an increase in the number of older workers crowds out younger employees from top positions and slows their career progression, especially in firms with higher organizational costs. This paper develops several additional predictions about the consequences of these cross-cohort effects on the gender pay gap.

We first show that the assumption that changes in the gender pay gap happen mainly across cohorts, which is one of the main premises of the model, finds full empirical support.[3] We show that the gender pay gap has narrowed in all countries in our sample for at least the last four decades. Crucially, this decline has stemmed from reductions happening across subsequent cohorts, rather than over the life cycle of any given cohort. Moreover, the recent deceleration in the closing of the total gender pay gap in the economy, which started in the mid-1990s, coincided with a slowdown in the closing of the gap between cohorts.

Beyond this initial descriptive evidence, we formally quantify how much of the shrinkage in the gender pay gap stems from cross-cohort convergence between men and women. Specifically, we compute the *between-cohort* change in the gender pay gap, which shuts down any variation in the gap taking place over the cohorts' life cycle. In practice, we assign to all individuals in our sample the mean earnings that workers of the same gender and age earned at 25 years old. Therefore, changes in the between-cohort gender pay gap reflect variations in mean earnings close to labor-market entry or changes in the relative size of birth cohorts. In all four countries, more than the entire decline in the aggregate gender pay gap in weekly earnings can be accounted for by a progressive decrease in the between-cohort component. For example, the between-cohort change in the US equals 127 percent of the total decline in the gender pay gap between 1976 and 2019. Our analysis thus reveals that the trend in the aggregate gender pay gap comes entirely from the entry of younger cohorts with smaller earnings differentials and the exit of older cohorts with larger differentials.

Next, we test the model's prediction that workforce aging narrows the gender pay gap across cohorts by worsening the outcomes of younger men more than those of younger women.

[3] A model inspired by the one in Card and Lemieux (2001) in which older workers are perfect substitutes for younger men and imperfect substitutes for younger women can produce this initial result without additional firm-level constraints. However, this model is not compatible with further findings, such as no relevant improvements in the outcomes of younger women at labor-market entry or the subsequent slowdown in the closure of the gender pay gap.

This gender differential arises because, at baseline, younger men are more likely to be employed in top jobs, which become progressively more crowded due to the increased supply of older workers. Empirically, we examine where men and women at age 25 ranked over time in the overall pay distribution. The narrowing of the gender pay gap at labor-market entry, which lasted until the mid-1990s, was indeed driven by younger men falling closer to younger women in the pay distribution rather than by younger women experiencing disproportionate improvements in labor-market outcomes. In the US, the average rank of younger men at age 25 fell from the 50th percentile of the wage distribution in 1976 to the 39th percentile in 1995, while the mean position of women at age 25 remained fairly stable around the 30th percentile during the same period. After the mid-1990s, the positions of younger men and younger women followed the same flat trajectory, a trend that we further explore toward the end of the paper. The same findings hold if we analyze directly the probability of holding higher-ranked positions within firms' hierarchies (managerial jobs with high wages), rather than the percentiles in the pay distribution.

We then discuss changes in the distribution of younger men and younger women across different types of firms. The model indicates that having more older workers in top jobs can affect worker sorting because the negative effects on the careers of younger men are larger within higher-paying firms.[4] Consistent with this prediction, we document that, compared with younger women, younger men experienced larger positional losses within the pay distributions of all types of firms. However, these losses were the largest within higher-paying firms, which younger men left at a higher rate. The result is that younger men joined younger women in being less represented among higher-paying firms by the end of the sample period.

Finally, we quantify that the convergence of the gender pay gap at labor-market entry (for brevity, *convergence at entry*) can explain one third of the total decline in the gender pay gap during the last forty years, while the rest is accounted for by the exit of older cohorts from the labor market (*convergence through exit*). However, if we focus only on the last twenty years of data, we find that the gender pay gap has been shrinking due almost exclusively to the convergence through exit, given that the convergence at entry stopped in the mid-1990s. Therefore, in contrast to forecasts based on trends in the *aggregate* gender pay gap (for example, see World Economic Forum (2023)), which usually predict that the mean earnings of men and women will match in a few decades, we project that the gender pay gap is not slated to disappear in the high-income countries in our sample. At best, and in the absence of structural breaks in the labor market, the gender pay gap will converge to

[4] In the model, more productive, higher-paying firms have higher retention rates. Therefore, an increase in the number of older workers in the economy affects these firms disproportionately. In alignment with the theory, Bianchi and Paradisi (2023) has documented that older workers have become increasingly concentrated in higher-paying firms.

the level observed among recent labor-market entrants.

To better understand why the convergence at entry has stalled and why the closure of the gender pay gap has slowed down, we show that a substantial part of the remaining entry gap results from different choices of college majors among young men and young women. For the US, we use American Community Survey data to predict the average weekly earnings for college graduates based on their college major. We find that the major-predicted entry pay gap (i) has remained steady over the last three decades, and (ii) constitutes approximately 63 percent of the total gender pay gap at entry for college graduates. These findings are consistent with the predictions of our stylized framework, which indicates that a larger number of older workers cannot further shrink the gender pay gap when most of the remaining earnings differential between younger men and younger women depend on education choices that predate entry into the labor market.

In conclusion, our findings imply that pay convergence has happened one cohort at a time. A decline in the gender gap at labor-market entry initially spurred it. However, during the past two decades, it has been driven solely by the retirement of older cohorts with larger earnings differentials. Even more disappointingly, the male-female convergence in entry outcomes that lasted until the mid-1990s did not stem from improved prospects for younger women but rather from disproportionately worse outcomes for younger men.

Many prior papers have documented the existence of different types of gender gaps in the labor market (Altonji and Blank, 1999; Azmat and Petrongolo, 2014; Bertrand, 2020), as well as the recent convergence in the gender pay gap in most high-income economies (Olivetti and Petrongolo, 2016; Blau and Kahn, 2017). Two papers related to our analysis are Goldin (2014) and Blundell, Lopez, and Ziliak (2024), which show that the gender gap in earnings has become progressively smaller for younger cohorts in the United States while increasing within each cohort.

We build upon this evidence by focusing further on the cross-cohort convergence, using observations over a long period and for different countries. In this respect, our main contribution is to show that cohort-driven effects can fully account for the dynamics of the gender pay gap over the past forty years (i) partially through the *inflows* of newer cohorts with lower gaps for the first two decades, and (ii) exclusively through the *outflows* of older cohorts with higher gaps for the last two decades. This latter result leads to the novel conclusion that we do not expect the gender pay gap to close in the future, since the outflow of older cohorts will only reduce the gap to the level observed among newer cohorts. Moreover, we show with both theory and data that the trend in the gender pay gap is linked to the seemingly unrelated phenomenon of career spillovers between older and younger workers. Lastly, our evidence on the characteristics of the remaining gender pay gap among new entrants corroborates prior findings on the importance of policies, attitudes, and allocations that impact younger women's experiences before or at labor-market entry.[5]

Prior work has studied the importance of economic conditions at the time of labor-market entry for the career progression of new entrants. For example, Kahn (2010) and Oreopoulos, von Wachter, and Heisz (2012) have documented that macroeconomic conditions at the time of college graduation affect the career of new entrants for several years after entry. More closely related to gender differentials, Bovini, De Philippis, and Rizzica (2023) and Foliano et al. (2023) have examined the existence of the gender pay gap at labor-market entry and its evolution over time. We contribute to this body of work first by quantifying the importance of conditions at labor-market entry in shaping the trajectory followed by the *aggregate* gender pay gap. One of our main findings is that a greater gender balance in the earnings of new entrants not only was relevant for the gender pay gap of the directly affected cohorts, but was also a major driver of the closure of the economy-wide gender pay gap.

Moreover, building on recent work on the worsening outcomes of younger workers (Dabla-Norris, Pizzinelli, and Rappaport, 2023; Bianchi and Paradisi, 2023), we show that the shrinking gap in entry wages is predominantly the result of a race to the bottom of the wage distribution, in which the labor-market outcomes of younger men have worsened substantially more than those of younger women. In addition, we document that the remaining gap at entry is largely related to pre-entry factors, such as the field of study among college graduates. In doing so, we expand on the evidence in Bertrand (2018) and Sloane, Hurst, and Black (2021) by analyzing data from more countries and explicitly linking the gender gap in pay at labor-market entry with that in major-predicted earnings.

Finally, our findings contribute to prior papers on the interconnectedness of the careers of older and younger workers (Bertoni and Brunello, 2021; Boeri, Garibaldi, and Moen, 2022; Ferrari, Kabátek, and Morris, 2023; Guaitoli and Pancrazi, 2023; Mohnen, forthcoming). For example, Bianchi et al. (2023) and Bianchi and Paradisi (2023) have shown that having more older workers at the top of firms' hierarchies has restricted access to higher-paying positions for younger workers, who have fallen toward the bottom of the wage distribution. These career spillovers are more negative in firms that face more binding constraints on adding top slots and in firms in which older workers are more likely to extend their careers (higher-paying ones). We contribute to these findings by developing a stylized model of the labor

[5] Examples of factors relevant in this context include gender stereotypes and norms in education (Carlana, 2019) and in the labor market (Fernández, Fogli, and Olivetti, 2004; Kleven, 2023; Charles, Guryan, and Pan, forthcoming), gender differences in the choice of the field of study (Porter and Serra, 2020), family-leave policies (Blair and Posmanick, 2023; De Quinto and González, 2024; Karademir, Laliberté, and Staubli, 2024; Kleven et al., 2024), or first-job sorting across heterogeneous employers and industries (Card, Cardoso, and Kline, 2016; Olivetti and Petrongolo, 2016; Morchio and Moser, 2023; Arellano-Bover, 2024; Casarico and Lattanzio, 2024).

market with both cross-cohort spillovers between younger and older workers and a positive gender pay gap. We then empirically test key predictions from this framework to show why more restricted access to top jobs disproportionately worsens younger men's outcomes.

The remainder of this paper is structured as follows. Section 2 discusses the conceptual framework. Section 3 describes our data sources. Section 4 quantifies the contribution of cross-cohort effects in shrinking the aggregate gender gap. Section 5 zooms in on entry wages and compares the positions of younger men and younger women in the wage distribution both within and between firms. Section 6 provides further insights into the sources of cross-cohort convergence and analyzes the drivers of the remaining gender pay gap at entry. Section 7 concludes.

2 A Model of the Gender Gap with Cross-Cohort Spillovers

This section discusses the consequences of introducing gender into a stylized framework of the labor market with cross-cohort (or career) spillovers. Cross-cohort spillovers occur when demographic shifts in the workforce affect the career progression of workers in different age groups. Bianchi et al. (2023) and Bianchi and Paradisi (2023) have already established both theoretically and empirically that an increase in the supply of older workers can limit access to higher-paying jobs for younger workers, thus widening the pay gap between the two age groups. In contrast to older workers, younger workers have experienced increasing challenges in reaching the top segments of the wage distribution and higher-ranked job levels in many high-income countries. Moreover, the pay inequality between older and younger workers has widened more within firms that face more binding constraints on adding higher-ranked positions and in firms in which older workers have become more entrenched.

By introducing gender considerations in a model similar to the one used in Bianchi and Paradisi (2023), the stylized framework in this section shows that negative cross-cohort spillovers on younger workers can contribute to shrinking the pay gap between men and women.

Production. We start from an economy with a price-taking representative firm. There is a fixed supply of l_y younger workers and l_o older workers. Workers differ with respect to their gender so that the labor supply of each age group can be written as the sum of men and women ($l_a = m_a + f_a$). The firm employs these labor inputs to perform a top job t and a bottom job b. Production occurs through the production function $AY(L_y, L_o)$, where A is a productivity shifter, $Y_{L_a} > 0$, and $Y_{L_a,L_a} < 0$ $\forall a \in \{y, o\}$. Moreover, younger and older workers are complements in production such that $Y_{L_y,L_o} > 0$. The inputs L_y and L_o are efficient units of younger and older labor, respectively: $L_a = \theta_{a,t}(m_{a,t} + f_{a,t}) + \theta_{a,b}(m_{a,b} + f_{a,b})$, where $\theta_{a,j}$ is the marginal productivity in job $j \in \{t, b\}$ of workers in the age group $a \in \{y, o\}$. We assume that $\theta_{a,t} > \theta_{a,b}\ \forall\, a$ to make all workers more productive in the top job.

Cross-cohort spillovers and gender pay gap. We introduce two key features to generate negative career spillovers and a positive gender pay gap in this model.

First, we introduce the possibility of cross-cohort spillovers on employment levels and wages. Following Bianchi and Paradisi (2023), we assume that the firm inherits a stock of older workers from period -1 and cannot change their job assignment and wages. Specifically, these *legacy workers* in job j are equal to $\rho_j l_{o,j}^{-1}$, where ρ_j is the retention rate in job j and $l_{o,j}^{-1}$ is the number of older workers in job j in period -1. Moreover, the firm pays a quadratic organizational cost proportional to the parameter $\kappa > 0$ to create and maintain K slots at the top ($K = l_{o,t} + m_{y,t} + f_{y,t}$). This cost captures the idea that the firm cannot always create higher-ranked positions, regardless of whether it can afford to pay the associated wages (similar to the models discussed in Lazear, Shaw, and Stanton (2018), Bianchi et al. (2023), and Bianchi and Paradisi (2023)). In fact, top jobs often entail managerial responsibilities, decision-making power, and complex tasks. Therefore, creating a new top job requires some slack both in the firm's organizational capacity (in the form of available high-level responsibilities) and in its payroll budget.

Second, we introduce a gender pay gap in favor of men, a pattern that all data sources available to us support. Specifically, although men and women are perfect substitutes in production, (i) the firm pays a quadratic cost proportional to c_g for employing younger workers of gender g in the top job, and (ii) this cost is higher for younger women ($c_f > c_m$).[6] These parameters, which make younger women less concentrated in top jobs, can be microfounded as either taste-based or statistical discrimination.[7]

Wage formation and timing. Following Acemoglu and Restrepo (2023), we assume that wages in the top job pay an exogenous rent over wages in the bottom job: $w_{a,t}^g = \mu_a w_{a,b}^g$, where $\mu_a > 1$ is the exogenous rent for age group a and $w_{a,j}^g$ is the wage in job j for workers in age group a and gender group $g \in \{m, f\}$.[8]

The timing is as follows. First, the firm is endowed with older workers from period -1. Then, given a set of wages for younger workers, the firm decides how many younger men and younger women to slot in the top and bottom jobs by equating their marginal revenue products in the two positions to their marginal costs. Based on these decisions, the

[6] In principle, we could apply these costs also to older workers. However, they would be redundant because the firm does not choose older workers' employment in period 0.

[7] Alternatively, we could rewrite the production function assuming that (i) men and women are imperfect substitutes in the top job, and (ii) women are less productive than men in the top job.

[8] Acemoglu and Restrepo (2023) show that these wedges can be microfounded using either efficiency wages or bilateral wage bargaining.

firm allocates younger workers randomly between the top and bottom jobs until its labor demands in the two positions are satisfied. Finally, the production is realized, and the firm pays all workers.

The firm problem. The firm problem is to choose the number of younger men and younger women (hereafter, shortened to *younger men and women* for ease of exposition) to employ in top and bottom jobs that maximizes its profits, as follows:

$$\max_{m_{y,t},f_{y,t},m_{y,b},f_{y,b}} AY\left(L_y, L_o\right) - \sum_{g\in\{m,f\}}\sum_{a\in\{y,o\}}\sum_{j\in\{t,b\}}\left(w^g_{a,j}g_{a,j}\right) - \frac{\kappa}{2}K^2 - \sum_{g\in\{m,f\}}\left(\frac{c_g}{2}g^2_{y,t}\right).$$

Appendix B discusses the full solution of the firm problem and provides all the proofs of the following results. Moreover, it includes several extensions: an alternative source of the positive gender pay gap, a different parametrization of the organizational cost of top jobs, an endogenous labor supply without full employment, and no exogenous rents in wages.

In equilibrium, there is a gender gap in employment in top jobs. Specifically, the marginal revenue product of labor of younger men and women in bottom jobs, and hence their wages in both bottom and top jobs, are the same. However, the number of younger women in top jobs is lower than the number of younger men in those jobs because the cost of employing the former is higher than the cost of employing the latter: $c_f > c_m$. In fact, it is optimal for the firm to keep a constant ratio between younger men and women in top jobs below 1: ${}^{f_{y,t}}/_{m_{y,t}} = {}^{c_m}/_{c_f} = \delta_f < 1$.

In this labor market, cross-cohort spillovers are crucial drivers of the trend in the gender pay gap. By construction, the gender pay gap can change only if the pay gap between younger men and women differs from that found among older workers. To this end, we consider the effects of an increase in the number of legacy older workers in top jobs inherited by the firm in period 0 ($l^{-1}_{o,t}$) on the mean wages of younger men and women.[9] The mean wage $\bar{w}_{y,g}$ of younger workers of gender g changes as follows:

$$\frac{\partial \bar{w}_{y,g}}{\partial l^{-1}_{t,o}} = \underbrace{\frac{1}{g_y}\left(\mu_y - 1\right) w_{y,b}\frac{\partial g_{y,t}}{\partial l^{-1}_{o,t}}}_{\text{Career spillovers}} + \underbrace{\left(\frac{1}{g_y}\left(\mu_y g_{t,y} + g_{b,y}\right)\right)\frac{\partial w_{y,b}}{\partial l^{-1}_{o,t}}}_{\text{Wage level}}, \tag{1}$$

for each $g \in \{m, f\}$.

An increase in the number of older workers induces the firm to add slots at the top of its hierarchy due to the complementarity in production between younger and older workers.

[9] The implicit assumption is that firms made past hiring decisions without considering future changes (i) in the relative size of cohorts, (ii) in the length of workers' careers, and (iii) in the economic growth rate. We capture this dynamic consideration within our static framework by studying an unforeseen increase in the number of older workers after period −1.

However, there is a threshold $\bar{\kappa}$ for the cost parameter κ above which this endogenous increase in top jobs K does not compensate for the increased supply of older workers.

Hence, when $\kappa > \bar{\kappa}$, the first component of Equation (1) describes negative career spillovers. An increased supply of older workers at the top restricts the career opportunities of younger workers by decreasing their chances of reaching higher-paying jobs ($\frac{\partial g_{y,t}}{\partial l_{o,t}^{-1}} < 0$). In contrast, the second component of Equation (1) is positive and refers to changes in the level of wages paid to younger workers in both bottom and top jobs. Having more older workers increases the wages of younger workers (i) due to the complementarity of younger and older workers in the production function (Freeman, 1979; Welch, 1979; Berger, 1985) and (ii) due to the fact that younger workers become more likely to be in bottom jobs and, therefore, their marginal revenue product of labor increases.

Next, we focus further on the crowding out of younger workers from top positions. The model predicts that these negative career spillovers are larger in magnitude among younger men as long as the latter are more concentrated in top jobs than younger women ($\frac{m_{y,t}}{m_y} > \frac{f_{y,t}}{f_y}$). Therefore, by blocking younger workers from top positions, a larger supply of older workers can narrow the preexisting gender pay gap by compressing the earnings of younger men and women toward the bottom of the distribution.

The rationale behind this comparative statics is that many high-income economies have experienced an increase in the supply of older workers in recent decades. In the United States, the mean age of the population increased from 32.5 years in 1976 to 38.7 years in 2019 (Figure A1, Panel A), while the mean age of the private-sector workforce increased from 37.5 in 1976 to 41.6 in 2019 (Figure A1, Panel B).[10]

In this framework, an increase in the retention rate of older workers and a decrease in the rate of economic growth each produce the same negative spillovers on the employment outcomes of younger workers. The data confirm that these two phenomena have coexisted with population aging: workers have progressively experienced longer careers (Figure A1, Panel C), a slower GDP growth rate (Figure A1, Panel D), and a lower degree of firm dynamism (Decker et al., 2014; Bianchi and Paradisi, 2023).

Introducing skills. In this extension, we discuss how our framework can accommodate the slower decline in the gender pay gap that started in the mid-1990s. A key prediction from the baseline model is that younger men are more concentrated in top jobs than younger women and, therefore, suffer more harshly from these positions being progressively occupied by older workers. When younger men and women become equally concentrated in top jobs

[10] In Italy, the median population age increased from 32.7 in 1976 to 46.0 in 2019 (World Population Prospects 2022, United Nations). In addition to the US and Italy, Bianchi and Paradisi (2023) has documented that the workforce has been aging in at least eleven other high-income countries.

($\frac{m_{y,t}}{m_y} \approx \frac{f_{y,t}}{f_y}$), further increases in the supply of older men can still harm the prospects of younger workers but cannot do so differentially across genders. After this point, factors other than the initial job assignment, such as gender imbalances that predate entry into the labor market, become the primary drivers of changes in the gender pay gap.

Here, we assume that each worker enters the labor market with either high (h) or low (l) skills, which represent cross-worker differences in college major choices among other pre-labor-market factors. Moreover, each job is divided into two different tasks, and there is a one-to-one correspondence between skills and tasks. The younger workers' vector of efficient units of labor is $\mathbf{L_y} = (L_{y,h}, L_{y,l})$, and workers with different skills are complements in the production function. The rest of the firm problem is unchanged.

In this static framework, we model a scenario in which younger men and women are concentrated equally in top jobs by assuming that the number of older workers in top jobs is large enough that no higher-ranked positions remain available for younger workers in period 0. We then study what happens to the mean wages of younger workers when the number of older workers in top jobs increases even further:

$$\frac{\partial \bar{w}_{y,g}}{\partial l_{o,t}^{-1}} = \sum_{s \in \{h,l\}} \left(\frac{g_{y,s}}{g_y} A \theta_{y,b,s} Y_{L_{y,s} L_o} \theta_{o,t,s} \rho_t \right) > 0,$$

for each $g \in \{m, f\}$.

Whether this wage increase is larger for men or women crucially depends on their distribution across different skills/tasks. We can further show this point by assuming that the complementarity between younger and older workers is proportional to a task's marginal product: $Y_{L_{y,s} L_o} = Y_{L_{y,s}} \cdot C(\mathbf{L_y}, L_o)$, where $C(\cdot, \cdot)$ is the same across tasks.[11] In this case, the change in mean wages of younger women is larger than that of younger men if

$$\left(\frac{f_{y,h}}{f_y} - \frac{m_{y,h}}{m_y} \right) \left(Y_{L_{y,h}} \theta_{y,b,h} \theta_{o,t,h} - Y_{L_{y,l}} \theta_{y,b,l} \theta_{o,t,l} \right) > 0.$$

Assuming that high-skill tasks are more productive than low-skill tasks, the inequality holds if women are overrepresented in high-skill tasks: $\frac{f_{y,h}}{f_y} > \frac{m_{y,h}}{m_y}$. Given that prior research has established that men choose higher-return college majors (skill h in the context of our framework) more than women do (for example, see Black et al. (2008), Bertrand (2020), Huneeus et al. (2021), Sloane, Hurst, and Black (2021), and Bovini, De Philippis, and Rizzica (2023)), an increased supply of older workers is unlikely to further shrink the gender pay gap.

[11] This is true if the production function is a Cobb-Douglas or a nested CES (see Appendix B).

Heterogeneous firms. Finally, we replace the representative firm with N heterogeneous firms to study how the gender pay gap varies within and between different firms. Each firm n chooses four wages ($w^m_{y,b,n}$, $w^f_{y,b,n}$, $w^m_{y,t,n}$, and $w^f_{y,t,n}$) such that all marginal revenue products of labor equate to the marginal costs.[12] We drop the assumption that wages in the top job pay a rent over wages in the bottom job so that the presence of more older workers can have more complex effects on the wage schedule of different types of firms. In line with prior research (Antwi and Phillips, 2013; Ruffini, 2022), higher-paying firms (higher A_n) have a higher retention rate of older workers (higher $\rho_{t,n}$).

Moreover, we make two assumptions to make the computations more tractable. First, the cost of hiring younger workers for top jobs is linear, instead of quadratic. Second, the organizational cost of top jobs has a sharp discontinuity so that all firms face a binding constraint on the number of top slots: $\overline{K}_n = l_{o,t} + m_{y,t} + f_{y,t}$, where $\overline{K}_n$ is fixed. This assumption allows us to focus on the empirically relevant scenario in which firms cannot expand their organizations at the top. Appendix B outlines the firm problem and the optimal personnel policies.

On the labor-supply side, worker i of age group a and gender g derives the following utility when working in job j and firm n: $U_{i,a,j,n} = \log\left(w^g_{a,j,n}\right) + \xi_{i,a,j,n}$, where $\xi_{i,a,j,n}$ represents the idiosyncratic preference of worker i over job j of firm n. We assume that $\xi_{i,a,j,n}$ is unobserved by firms and follows a type-1 extreme distribution with a parameter $\sigma > 0$.

This extension produces two additional insights. First, an increase in the supply of older workers decreases the probability of younger men holding top jobs more than that of younger women within all firms, regardless of their productivity level. However, these negative career spillovers are larger in magnitude for both younger men and women within higher-productivity and higher-paying firms because the number of older workers increases more in these firms.

Second, given that younger men are more likely to be displaced from the top jobs of higher-paying firms, they are also more likely than younger women to migrate toward the bottom jobs of other firms. Appendix B outlines under what conditions younger men move from higher-paying to lower-paying firms, a result that finds empirical support.

Summary. In the rest of the paper, we test several predictions of this stylized framework:

1. Cross-cohort differences in the gender pay gap are the main source of decline in the aggregate gender pay gap (Section 4).

2. The gender pay gap narrows across cohorts because younger men lose more positions

[12] Following Card et al. (2018) and Lamadon, Mogstad, and Setzler (2019), we assume that firms disregard the indirect effects of their choices on other firms.

in the overall pay distribution and in firms' hierarchies, relative to younger women (Section 5.1).

3. These cross-cohort spillovers are more negative among younger men than among younger women within all firms. For all younger workers, they are larger in magnitude among higher-paying firms (Section 5.3).

4. Compared to younger women, younger men leave higher-paying firms at a higher rate (Section 5.3).

5. When the concentration of younger men and women in top jobs becomes more balanced and predetermined education choices account for a large share of the gender pay gap, the gender pay gap stops shrinking (Section 6).

3 The Data

This paper uses a combination of administrative and survey data with 376,814,659 observations from four high-income countries: the United States, Italy, Canada, and the United Kingdom (Table A1). Due to their much larger sample size and informational depth, our discussion focuses on the US and Italy, and uses data from Canada and the UK primarily to show the robustness of the main findings.

3.1 US: Current Population Survey and American Community Survey

Most of our US analyses rely on forty-four years (1976-2019) of repeated cross-sections from the Annual Social and Economic March Supplement of the Current Population Survey (CPS), which we accessed through IPUMS (Flood et al., 2022). We impose similar sample restrictions across all the datasets available to us: we limit our sample to individuals who were between 25 and 64 years old, had worked in the private sector for at least 24 weeks during the past year, and had earned strictly positive earnings.

The CPS data allow us to construct three compensation measures: (i) the annual wage and salary income, (ii) weekly earnings, obtained by dividing the annual wage and salary income by the number of weeks worked during the previous year, and (iii) hourly earnings, obtained by dividing weekly earnings by the usual number of hours worked per week. All compensation measures are expressed in 2015 USD, using the CPI provided by the Bureau of Labor Statistics. Moreover, they are winsorized yearly at the 99.9$^{\text{th}}$ percentile from above.

Since the CPS does not include information on college graduates' field of study, we complement some of our analyses with data from the 2009–2019 waves of the American Community Survey (ACS), also accessed through IPUMS (Ruggles et al., 2023).

3.2 Italian Social Security Data

Our analysis of the Italian data mainly relies on confidential administrative data provided by the Italian Social Security Institute (INPS). This dataset comprises forty-four years (1976-2019) of matched employer–employee records for the whole population of private-sector, nonagricultural firms with at least one salaried employee. In each year, we focus on workers who were between 25 and 64 years old, had worked at least 24 weeks, had earned strictly positive earnings, and had not retired by December 31. Unlike the CPS data, the INPS dataset matches workers to firms, a feature that allows us to study the trends in the gender pay gap both within and between firms.

The Italian administrative data include two main compensation measures: (i) total annual earnings, which comprise all forms of gross labor compensation, and (ii) full-time-equivalent (FTE) weekly earnings, computed as the ratio between total annual earnings and FTE working weeks. Both variables are expressed in 2015 euros using the CPI provided by the OECD. Moreover, they are winsorized at the 99.9$^{\text{th}}$ percentile from above.[13] We complement the administrative data with the Labor Force Survey for the years 2009-2019 to shed light on the changes in college graduates' fields of study.

3.3 Luxembourg Income Study

We leverage survey data from the Luxembourg Income Study (LIS) database for two additional high-income countries that have long time series and sufficiently large sample sizes: Canada (1973-2019) and the United Kingdom (1976-2019). We can compute weekly earnings only for Canada, while total yearly labor earnings are available in both countries. Both earnings variables are expressed as 2011 purchasing-power-parity US dollars using the conversion tables directly provided by LIS. Moreover, they are winsorized at the 99.9$^{\text{th}}$ percentile from above. Whenever possible, we apply the same sample restrictions that we used for the US and Italian data. Specifically, we keep workers who were between 25 and 64 years old, had worked at least 24 weeks (available only in Canada), and had strictly positive labor earnings.

4 The Cross-Cohort Decline in the Gender Pay Gap

4.1 Aggregate and Cohort-Specific Trends in the Gender Pay Gap

Aggregate gap. For at least the last four decades, the United States and Italy have been experiencing a decrease in the gender gap in weekly earnings (Figure 1). Between 1976 and 2019, the gender gap shrank by 0.47 log points or 59 percent (relative to the baseline level of

[13] Annual earnings are also winsorized at €3,000 from below to address a few instances in which yearly compensation is implausibly low.

0.8 log points) in the US and by 0.19 log points or 57 percent (relative to the baseline level of 0.33 log points) in Italy.[14] If we replace yearly earnings for weekly earnings, the decrease in the gender gap remains large, at -59 percent in the US and -26 percent in Italy (Figure A2).

The shrinkage in the gender gap extends to the other two countries in our dataset, appearing to be a generalized trend within high-income economies (Figure A3). Specifically, we find that the gender gap decreased by 0.29 log points or 43 percent in Canada between 1973 and 2019 (weekly earnings) and by 0.63 log points or 60 percent in the United Kingdom between 1976 and 2019 (yearly earnings).

Gap between and within cohorts. We now start to highlight the importance of cross-cohort effects in driving the overall negative aggregate trend in the gender pay gap (Figure 2). Specifically, for various worker cohorts (defined by their birth year), we plot the trend in the gender gap from the year in which all workers in the cohort turned 25 years old to the year in which they turned 50 years old.[15] This graph shows three main results that will inform most of the further analysis in the rest of the paper.

First, consistent with Prediction 1, the aggregate downward trend in the gender gap stems entirely from the fact that younger worker cohorts entered the labor market with smaller initial differences in the log weekly earnings of men and women. In the US, the gender gap at age 25 decreased from 0.55 log points for the cohorts born between 1947 and 1951 to 0.19 log points for the cohorts born between 1967 and 1971. Similarly, in Italy the gender gap at age 25 declined from 0.21 log points for the cohort born in 1951 to 0.09 log points for the cohort born in 1971.[16]

Second, the gender gap almost always increases over the life-cycle of each cohort.[17] Moreover, this within-cohort increase becomes steeper across subsequent worker cohorts, a trend that works against closing the aggregate gender gap. If we consider the 1951 cohort, the gender gap increased between age 25 and age 30 by 0.07 log points in the US and 0.01 log points in Italy. The same increase for the 1971 cohort was equal to 0.17 log points in the US and 0.03 log points in Italy. While our static framework is not fully equipped to capture these within-cohort dynamics, these results are consistent with many prior findings in the literature. For example, the child penalty incurred by mothers in the labor market (Kleven,

[14]In the United States, the aggregate gender pay gap was largely stable between 1955 and 1975 (Bailey, Helgerman, and Stuart, forthcoming).

[15]We stop at 50 years old to limit the influence coming from selection into retirement.

[16]In the CPS data, we create groups of five cohorts to increase precision. In this case, we start reporting the cohort-level gaps when all cohorts within a group turn 25 years old. In the remainder, we refer to these groups using their youngest cohort.

[17]The flattening toward the end of each cohort's career is consistent with more negative selection into early retirement among women (Goldin and Mitchell, 2017).

Landais, and Søgaard, 2019) and women's propensity to move towards more flexible but lower-paying jobs (Goldin, 2014) are two factors that contributed to increasing the gender pay gap after labor-market entry, especially among more recent worker cohorts.[18]

Third, convergence at labor-market entry across subsequent worker cohorts significantly slowed after 1995 (see also Figure A4). As already discussed, the gender pay gap at age 25 dropped by 0.36 log points in the US and by 0.12 log points in Italy between the 1951 cohort and the 1971 cohort, but then declined by only 0.01 log points in the US and by 0.02 log points in Italy by the 1991 cohort. The end of this form of convergence coincided with a slowdown in the closure of the aggregate gender pay gap.

All these findings hold across alternative definitions of earnings, as well as in Canada and the United Kingdom (Figure A5).

4.2 Decomposing the Gender Pay Gap Within and Between Cohorts

In this section, we introduce a more formal test of Prediction 1, which states that most of the decline in the gender pay gap has been taking place across worker cohorts. To this end, we propose a decomposition of the change in the gender pay gap within and between worker cohorts.

First, we write the average log earnings of birth-year cohort c of gender g at time t ($w_{c,g,t}$), as the sum of two terms: $w_{c,g,t} = w^e_{c,g} + \Delta w_{c,g,t}$, where $w^e_{c,g}$ represents cohort c's mean log earnings at the time of entry into the labor market (or alternatively, at an early career stage), and $\Delta w_{c,g,t} = w_{c,g,t} - w^e_{c,g}$ is the growth of the average log earnings of cohort c between year t and labor-market entry (e).

Let $w_{g,t}$ represent the average log earnings of all workers of gender g in year t and let $a\,(c,t)$ denote the age of cohort c in year t. The change between year t and year $t' > t$ in the average log earnings of gender group g is:

$$\begin{aligned} w_{g,t'} - w_{g,t} \;=\; & \underbrace{\sum_{c:a(c,t')\in[25,64]} s_{c,g,t'} \cdot w^e_{c,g} - \sum_{c:a(c,t)\in[25,64]} s_{c,g,t} \cdot w^e_{c,g}}_{\text{Between-Cohort Change}} \\ & + \underbrace{\sum_{c:a(c,t')\in[25,64]} s_{c,g,t'} \cdot \Delta w_{c,g,t'} - \sum_{c:a(c,t)\in[25,64]} s_{c,g,t} \cdot \Delta w_{c,g,t}}_{\text{Within-Cohort Change}}, \end{aligned} \tag{2}$$

where $s_{c,g,t}$ is the share of workers of gender g from cohort c in year t over the total number of workers of the same gender. This share operates as a weight for a given cohort of active employees (between 25 and 64 years old) in a given year.

[18]The timing of fertility may have played a role: in the past four decades, the average maternal age at the birth of a child has moved from 26 to 29 in the US and from 28 to 32 in Italy (OECD, 2022).

The first two terms in Equation (2) quantify the *between-cohort* change between t and t'. This between-cohort component keeps the earnings of each cohort fixed over time and equal to their value at labor-market entry while allowing the weight of each cohort to change over time. Hence, it captures variation in the mean log earnings of gender group g that stems from (i) cross-cohort differences in early-career earnings and (ii) changes in the distribution of workers across birth cohorts over time.

In contrast, the last two terms of Equation (2) measure the *within-cohort* change between t and t'. This within-cohort component isolates variation in the mean log earnings of gender group g that originates from (i) changes in the life-cycle earnings growth for the average cohort active in the labor market between t and t', and (ii) changes in the age composition of workers over time. If we subtract Equation (2) for women from Equation (2) for men, we can quantify the contribution of the *between-cohort* and *within-cohort* components to the aggregate trend in the gender pay gap.

To implement this decomposition, we need to know the earnings of each cohort c at the time of labor-market entry e. Unfortunately, this information is available only for Italy, the country for which we have access to panel data. Hence, we opt for a different definition of entry earnings that we can measure in all of our datasets. In our baseline results, we compute $w^e_{c,g}$ as the mean earnings of each cohort-gender pair in the year in which the workers in that birth cohort were 25 years old. We then show the robustness of our results to an array of alternative definitions, including the actual entry earnings available in the Italian longitudinal data.

Another practical issue in the implementation of Equation (2) concerns the definition of $w^e_{c,g}$ for older cohorts in the early years of the sample. For example, the mean earnings at age 25 are not observable for any worker who is at least 26 years old in the first sample year. More generally, denote with $\underline{t}$ the first sample year in each dataset; we will not observe the entry earnings of all workers older than $t - \underline{t} + 25$ in year t. Therefore, when applying Equation (2) to the data, we assign the average earning in year $\underline{t}$ to each cohort-gender pair comprising workers older than 25 in year $\underline{t}$. Then, the resulting *between-cohort* component represents a lower bound of the true variation between worker cohorts if the life-cycle growth in the gender pay gap has increased across cohorts, and vice versa. In our robustness tests, we also show the results using a much shorter time series in which we observe the mean log earnings at age 25 for all cohorts, including the older ones.

Figure 3 plots the total change in the gender pay gap together with the counterfactual change that isolates the *between-cohort* component of Equation (2). These results highlight the importance of cross-cohort effects in driving the trend in the overall gender pay gap. In both the United States and Italy, the between-cohort decline in the gender pay gap is at least as large in magnitude as the total decline. In the US, the between-cohort change in the gender gap in log weekly earnings is equal to 116 percent of the aggregate decline in 1980, 110 percent in 1990, 130 percent in 2000, and 128 percent in 2010 (Figure 3, Panel A). Similarly, in Italy, the between-cohort component accounts for up to 139 percent (in 2010) of the total decline (Figure 3, Panel B). These findings hold for different earnings measures and in other high-income countries (Figure A6). The fact that, in most sample years, the between-cohort component accounts for more than 100 percent of the total decline in the gender pay gap indicates that the gender gap has been increasing over the life cycle of the average worker cohort.

4.3 Robustness Checks and Compositional Changes

Entry age definition. We first show that our findings are robust to alternative definitions of early-career earnings ($w^e_{c,g}$): earnings at age 25 (the baseline), earnings at age 28, earnings at age 30, and mean earnings between age 25 and age 30 (Table 1, Columns 3 to 6). Results remain very similar across all these specifications. Specifically, by the last sample year, the between-cohort component accounts for more than 100 percent of the total decline in the gender pay gap across all definitions of "early career" and in all countries.

Except for the US, the extent to which the counterfactual exceeds the aggregate trend decreases as the age threshold increases. For example, in Italy, the between-cohort component accounts for 127 percent of the total decline in the gender pay gap between 1976 and 2019 when early-career earnings are computed at age 25, 115 percent when early-career earnings are computed at age 28, and 104 percent when early-career earnings are computed at age 30. This comparison across age thresholds suggests that the gender pay gap starts increasing within a cohort soon after entry into the labor market so that early-career earnings observed at higher age thresholds incorporate a larger share of the within-cohort growth in the gender pay gap.

Next, we show that the results are robust to computing the decomposition in Equation (2) on two shorter time series. First, we assess our results' sensitivity to assigning mean earnings in the first sample year, rather than at age 25, to older cohorts (Table 1, Columns 7 to 8). To this end, we repeat the analysis starting from 2015, the first year we observe earnings at age 25 for all worker cohorts in our sample. Second, we study whether the choice of not measuring wages exactly at the time of labor-market entry influences our results. To address this concern, we exploit the longitudinal dimension of the Italian data to define $w^e_{c,g}$ as the earnings observed during the first year in the labor market for each worker who started working after 1976 (Table 1, Columns 9 to 10). In both cases, the between-cohort component continues to account for more than the whole drop in the gender pay gap.

Labor force participation. Next, we explore whether changes in the observable characteristics of new entrants in the labor market can explain why most of the decline in the gender pay gap has happened between worker cohorts. We start by studying trends in the share of women active in the labor market (Figure A7). In the US, female labor force participation at age 25 increased by 8 percentage points between 1976 and 1986 and has remained stable ever since. Albeit slightly noisier due to the smaller sample size, the Italian data indicate that women's labor-force participation was fairly constant in the 1980s and the 1990s and slightly declined in the 21st century.[19]

In short, if the features of the marginal female worker have changed over time, shifts in labor-force participation do not seem to be responsible. In the next paragraphs, we will control more directly for cross-cohort changes in observable characteristics.

Sectoral shifts. In the next set of robustness checks, we test whether the decline in the gender pay gap across cohorts is related to variation in the sorting of men and women across different sectors. To do so, we consider the between-cohort component in Equation (2) and further decompose it between and within sectors. Sectors are defined using 1-digit NACE Rev. 2 codes in the United States and 2-digit codes in Italy.[20]

In both countries, the data indicate that most of the between-cohort decline in the gender pay gap has taken place within sectors (Figure A8, Panel A). Specifically, 18 percent of the total between-cohort change between 1976 and 2019 occurred across 1-digit sectors in the US, while none of the decline between 1976 and 2019 took place between 2-digit sectors in Italy. Therefore, the data suggest that a loss of employment in economic sectors where men historically received high wages, such as manufacturing (Charles, Hurst, and Schwartz, 2019), does not appear responsible for a meaningful portion of the cross-cohort decline in the gender pay gap.

Changes in educational attainment. In another set of tests, we investigate whether an increase in the college graduation rate among women can explain the cross-cohort shrinkage of the gender pay gap. Using an approach similar to the one described for studying sectoral shifts, we decompose both the total decline in the gender pay gap and just its between-cohort component within and between two education levels: workers with and without a college degree (Figure A8, Panel B). In this decomposition, the between-college component quantifies how much of the gender pay gap's decline stems from the fact that women have become more likely to graduate from college over time (or men have become less likely to

[19] Given that the Italian Social Security data include only employees, this analysis is based on data from the Bank of Italy's Survey of Household and Income Wealth (SHIW).

[20] Although the information on 2-digit codes is available for the United States, we use 1-digit codes in order to have enough observations in each cohort-year-sector cell.

graduate). In the US, the between-college dimension accounts for only a minor share of the decline in both the total (14 percent) and the between-cohort (14 percent) gender gap between 1976 and 2019.[21]

Child penalty. We test whether a progressive decrease in the child penalty borne by mothers in the labor market could explain the importance of the between-cohort convergence in earnings. Since childbirth tends to happen in a woman's earlier career stages, a progressively lower child penalty could disproportionately benefit women in younger cohorts.

For each year, we estimate how much of the between-cohort change in the gender pay gap is explained by disparities in the negative consequences of having children on the careers of new mothers and fathers. Following Kleven, Landais, and Søgaard (2019) and Casarico and Lattanzio (2023), we first calculate estimates of the child penalty in weekly earnings between mothers and fathers in the United States and between women with and without children in Italy, using a subset of data for which we have fertility information from maternity leave applications. We then create counterfactual weekly earnings by multiplying women's weekly earnings by the product of the child penalty and the fraction of mothers in each year and cohort. This variable is an estimate of women's earnings if we were to eliminate the adverse effects of parenthood on women's careers.[22]

As expected, we find that accounting for the child penalty decreases the contribution of the between-cohort component to the overall decline in the gender gap, but this effect tends to be small in magnitude (Figure A8, Panels C and D). Even after adjusting for the child penalty in the United States and Italy, the main takeaway of Section 4.2 still holds: the decline in the gender pay gap stems entirely from differences across worker cohorts.

Changes in the share of full-time women. We then move to investigate whether a disproportionate increase in the proportion of full-time female workers could serve as an explanation for the between-cohort convergence in mean earnings. To this end, we repeat the main decomposition in Equation (2), limiting the sample to only full-time workers (Figure A8, Panel E). In this more restricted sample, the main findings remain similar: the decline in the gender pay gap between cohorts is at least as large in magnitude as the aggregate decline.

Residual earnings. Finally, we control for multiple observable characteristics at once. We regress the log of weekly earnings on a dummy for part-time workers, a dummy for temporary workers (only in Italy), a dummy for domestic-born workers (only in Italy), several dummies for race and ethnicity (only in the US), and a dummy for workers with children (only in the

[21] The Italian Social Security data do not include information on completed education for most workers in the sample.

[22] Appendix C provides more details about this procedure.

US). We estimate these regressions separately by year and country to allow the coefficients to vary over time and across different datasets. Based on these coefficients, we compute residual earnings and use them to repeat the main decomposition in Equation (2) (Figure A8, Panel F). The between-cohort component still accounts for more than the entire decline in the gender pay gap.

5 The Positions of Younger Workers in the Wage Distribution

Section 4 has documented that the decline in the overall gender gap can be traced back to gradual changes across cohorts, which is consistent with Prediction 1. Specifically, the data indicate that until 1995, each new cohort of younger workers faced progressively lower gender pay gaps at the time of labor-market entry (Figure A4). Section 5.1 and Section 5.2 below now present the empirical evidence in support of Prediction 2, which states that the gender pay gap among younger workers has closed because younger men have fallen toward younger women in the pay distribution and in firms' hierarchies. Section 5.3 then describes how the gender pay gap has shrunk within higher- and lower-paying firms (Prediction 3) and between them (Prediction 4).

5.1 The Pay Rank Gap at Labor-Market Entry

We first investigate changes in the rank of new entrants within the wage distribution that existed at the time of their entry in the labor market. Let $F_t(w)$ represent the distribution of weekly earnings for all workers in year t.[23] For each new entrant i with weekly earnings w^E_{it}, we compute where they rank in the overall distribution: $p_{it} \equiv F_t(w^E_{it})$. We then compute the average rank among all entrants of gender g in year t as follows: $\bar{p}_{gt} = \frac{1}{N_{gt}} \sum_{i \in \{g,t\}} p_{it}$, where N_{gt} is the number of labor market entrants of gender g in year t. As discussed in Section 4.2, we fix the time of labor-market entry in the data at age 25 so that the variable $\bar{p}_{gt}$ measures the mean position in the wage distribution of men and women at age 25.

Building upon the analysis of the racial wage gap in Bayer and Charles (2018), we also consider an alternative metric focusing on the *median* entrant, rather than the average. Let $F_t(w)$ represent the distribution of weekly earnings for all workers in year t, while $F^E_{gt}(w)$ represents the distribution of weekly earnings at entry among gender g in year t. The gender-specific quantile-q entry pay is given by w^E_{qgt}, defined so that $F^E_{gt}(w^E_{qgt}) = q$. Let P_{qgt} quantify where the qth percentile of weekly earnings at entry of gender g ranks in the overall distribution: $P_{qgt} = F_t(w^E_{qgt})$ with $q = 0.5$.

[23] Using the position in the population's pay distribution allows a close mapping between the model and the results. In fact, Prediction 2 concerns the *rank* of younger men and women in the pay distribution, rather than their average wage. In addition, by focusing on ranks, we avoid any potential confounders in the level of wages that stem from trends in wage inequality, among other factors.

Consistent with Prediction 2, the gender pay gap at labor-market entry has shrunk because the mean position of younger men in the wage distribution has declined substantially more than that of younger women (Figure 4). In 1976, the average rank of younger men at age 25 was equal to the 50th percentile of all weekly earnings in the United States and to the 47th percentile in Italy. By 1995, the position of younger men had fallen to the 39th percentile in the US and to the 36th percentile in Italy. During the same period, the mean position of women at age 25 remained fairly stable around the 30th percentile in both countries.[24]

We reach the same conclusions if we consider the median position, rather than focusing on the mean rank (Figure A9, Panels A-B). Compared to the median woman at age 25, the median man at age 25 experienced a much larger decline in his rank in the distribution of weekly earnings. Similarly, the results are robust to using (i) survey data from Canada and the United Kingdom (Figure A9, Panels C-D), (ii) hourly earnings for the US (Figure A10), and (iii) residualized weekly earnings (Figure A11). Moreover, the findings hold if we control for potential on-the-job experience using education data (for the US only; Figure A12). The steeper fall experienced by younger men until the mid-1990s does not depend on differential trends in years of completed education.

Finally, this analysis highlights one additional insight on the labor-market conditions of new entrants. Figure 4 shows two distinct phases of convergence at labor-market entry. Between 1976 and the mid-1990s, the gender gap in pay rank at entry quickly closed due to the large positional losses experienced by younger men. From the mid-1990s onwards, the ranks of both younger men and women have stabilized, indicating that there has not been any further convergence at labor-market entry. Section 6 will investigate this result further.

5.2 The Probability of Holding Higher-Ranked Positions

The section shows that the main takeaway of Section 5.1 holds if we study the trend in the probability of younger men and women holding higher-ranked jobs within firms' hierarchies rather than their rank in the overall pay distribution. In the CPS data, we classify higher-ranked positions as all jobs with (i) 2-digit Standard Occupational Classification (SOC) code 11 (management occupations) and (ii) associated annual earnings in the top quartile of the year-specific distribution.[25]

We first establish that the likelihood of a higher-ranked position being filled by a young worker decreased from 13 percent in 1976 to 7 percent in 2019, even though higher-ranked

[24] Fukui, Nakamura, and Steinsson (2023) finds that increases in female employment rates do not crowd out male workers in the US labor market, alleviating the concern that an initial increase in the participation of younger women between 1976 and 1986 might have had negative effects on the opportunities of younger men.

[25] Unlike the findings on pay rank in Section 5.1, this analysis can be performed only in the US because detailed occupation variables become available in the Italian data only for recent years.

jobs accounted for a slightly increasing share of all jobs in the economy (from 6.5 percent in 1976 to 7 percent in 2019) during the same period (Figure A13, Panel A).

This crowding out of younger workers from the top of firms' hierarchies has affected younger men more than younger women, which aligns with Prediction 2 (Figure A13, Panel B). The share of higher-ranked jobs filled by men between the ages of 25 and 30 fell from 12 percent in 1976 to 5 percent in 1995, and then to 4 percent in 2019. Meanwhile, the corresponding share for women in the same age group and job category increased from 1 percent in 1976 to 2 percent in 1995, and to 3 percent in 2019. This finding holds even when the definition of higher-ranked positions is extended to managerial jobs with above-median pay (Figure A13, Panels C and D).

5.3 The Gender Pay Gap Between and Within Firms

This section documents how the gender pay gap among new entrants has changed within and between firms. For this analysis, we must use data matching workers to firms. Therefore, we can show evidence only from the Italian administrative dataset. Moreover, this section focuses on the years between 1976 and 1995 because we have already established that the gender pay gap among new entrants stopped closing after the mid-1990s.

We first study how the positions of younger men and women changed within different types of firms. To do so, we divide workers into one hundred percentiles based on their employer's mean wage in each sample year so that these firm groups have the same number of workers but varying wage levels. Then, we compute the mean percentile of men and women at age 25 using the wage distribution within each percentile of mean firm pay, rather than the overall wage distribution from all firms and workers in the labor market.

The empirical findings support Prediction 3, which states that younger men experience career spillovers that are more negative than younger women do in higher- and lower-paying firms, and that these losses are larger for younger workers of both genders in higher-paying firms. We expect to find larger effects in higher-paying firms because older workers have become increasingly entrenched in these firms, as Bianchi and Paradisi (2023) established. Compared to younger women, younger men have experienced larger declines in their mean percentile within ninety-three out of one hundred firm groups. Therefore, the gender pay gap has shrunk among new entrants within both lower- and higher-paying firms (Figure 5, Panel A). Moreover, the results indicate that the positional losses were more pronounced within higher-paying firms for all younger workers. Between 1976 and 1995, the mean positional change within above-median firm groups was equal to -10 percentiles for men and -6 percentiles for women, while the same change within below-median firm groups was -7 percentiles for men and +0.5 percentiles for women.

Next, consistent with Prediction 4, the share of younger men in higher-paying firms has declined more than that of younger women (Figure 5, Panel B): between 1976 and 1995, the probability of 25-year-old men working in the top decile of firm groups decreased on average by 6 percentage points (a 62 percent decline from the 1976 level), while the same probability for 25-year-old women fell by only 2 percentage points (a 28 percent decline from the 1976 level). At baseline, younger men were fairly equally distributed across lower-paying and higher-paying firms, while women were overrepresented among lower-paying firms. Over time, the distribution of younger men across different firm groups moved closer to that of younger women.

6 Two Phases of Convergence in the Gender Pay Gap

Section 4.1, Section 5.1, and Section 5.2 have shown that the gender pay gap of new entrants into the labor market rapidly shrank until the mid-1990s and then stabilized until 2019. Section 6.1 now quantifies the contribution of the convergence in entry earnings to the overall between-cohort change. Section 6.2 then studies the implications of the post-mid-1990s halt in the convergence at labor-market entry for the future of the aggregate gender pay gap. Finally, Section 6.3 provides an explanation for the fact that the convergence at entry stopped (Prediction 5).

6.1 Convergence from Inflows and Outflows of Worker Cohorts

In this section, we quantify to what extent the decline in the between-cohort gender pay gap has stemmed from the fact that men and women have been entering the labor market with progressively more similar mean earnings (*convergence at entry*), or, instead, from the fact that older worker cohorts, who had higher-than-average gender disparity in earnings, have naturally retired over time (*convergence through exit*).

To this end, starting from the between-cohort component in Equation (2), we further neutralize any cross-cohort convergence in the mean earnings of men and women that happened at the time of entry into the labor market. Specifically, considering a baseline year t_b, our new counterfactual measure of pay fixes the earnings at labor-market entry (i.e., at age 25) of all cohorts who enter in that year *or later* to the average earnings at entry computed between t_b and the following two years. Hence, this new between-cohort component measures what would have happened if the convergence in the early-career earnings of men and women had stopped at t_b and if only the natural turnover of older cohorts had affected the overall level of the gender pay gap.

Under this counterfactual scenario, in which we freeze convergence at entry in the first sample year, the data indicate that both sources of convergence (at entry and through exit) are important drivers of the decline in the gender gap that took place between 1976 and 2019 (Figure 6, Panels A and B). In 2019, convergence at entry accounted for 36 percent of the total between-cohort shrinkage in the gender gap in the US and 43 percent in Italy.

Consistent with our prior findings, we observe that the importance of the convergence at labor-market entry wanes as t_b increases, while the opposite is true for convergence through labor-market exit (Figure 6, Panel C). The latter begins to consistently explain at least 100 percent of the total decline in the between-cohort gap when the benchmark year t_b is equal to 2001 in the US and 2003 in Italy. In other words, the entire decline in the gender pay gap between the early 2000s and 2019 originated from the fact that older cohorts, who had higher levels of the gender pay gap, progressively retired, therefore reducing the mean pay differential between men and women who were still active in the labor market.[26]

6.2 Consequences for Future Convergence

The ongoing importance of convergence through exit carries relevant implications for the future of the gender pay gap. Given that retirees with higher-than-average gaps are currently the only source of convergence, the gender pay gap is *not* projected to close. In the absence of further changes in the labor market, it can, at best, converge to the level of the gender pay gap currently observed among younger worker cohorts at the time of labor-market entry.

Strikingly, if we were to predict future convergence by looking at the recent trends in the *aggregate* gender pay gap computed on the whole workforce, we would reach vastly different conclusions. Since the aggregate gender pay gap has been decreasing, any future projection would predict full convergence in a few decades in most high-income economies. For example, in its 2023 Global Gender Gap Report, the World Economic Forum predicts that Europe will reach gender parity in 67 years, while North America will get there in 95 years (World Economic Forum, 2023).

To better appreciate the differences between the two predictions, we estimate the rate of earnings convergence at early career stages across subsequent worker cohorts and over time. Denote with g_s^e the gender pay gap at age 25 of the cohort who entered the labor market in year s. Assuming that the convergence rate is linear, we model the gender pay gap at entry in year t as follows:

$$g_s^e = \alpha_t - \beta_t \left(s - \underline{s}_t\right), \tag{3}$$

where $\underline{s}_t$ is the entry year of the cohort with the maximum age used to estimate convergence in year t (age 45 in the baseline analysis), $s \in [\underline{s}_t, t]$, α_t is the gender gap at age 25 for the cohort $\underline{s}_t$, and the coefficient β_t measures the rate of convergence in the gender pay gap at

[26] All these findings hold if we also use data from Canada and the United Kingdom (Figure A14).

age 25 observed between year $\underline{s}_t$ and t. A higher value of β_t implies faster convergence across subsequent cohorts.

Starting from a given year t, if the convergence continued at the same rate in the following cohorts and the demographic composition of men and women remained the same, the gender pay gap would close for the first time for the cohort who entered the labor market in year $s^* = \frac{\alpha_t}{\beta_t} + \underline{s}_t$. Notice that $\lim_{\beta \to 0^+} s^* = +\infty$ shows that if the cross-cohort convergence stops, the gender pay gap at labor-market entry never closes.

In addition to estimating Equation (3), we can estimate the linear trend in the aggregate gender pay gap for each year t using the prior twenty years of data. We can then use the estimated coefficients from these regressions to predict the first year in which the total gap would close if its linear path continued without modifications, a scenario we already know is incorrect.

As expected, the between-cohort gender pay gap at age 25 is not bound to converge. We first show the estimated linear function in Equation (3) in the US and Italy for two years: the year 2000 and the last sample year (Figure 7, Panels A and D). Both countries' convergence rate β_t dramatically decreased from 2000 to 2019. For example, in the United States, β_t declined from 0.008 in 2000 to 0.0003 in 2019. In addition to highlighting two different years, we show the evolution of β_t across all years in the sample (Figure 7, Panels B and E): at least from 1995, the convergence rate rapidly decreased until it reached zero in the second half of the 2010s.

Next, we show how the projected year of closure of the gender pay gap at age 25 changed over time (Figure 7, Panels C and F). Specifically, we plot the predicted s^*, or year of entry into the labor market for the first cohort with zero gap at age 25, for all years t in our sample. The first year with no gender gap at age 25 has been following the opposite trend of the convergence rate β_t. In 1995, the projected first worker cohort without a positive gender pay gap at age 25 was slated to enter the market in 2022 in the United States and 2028 in Italy. By 2019, the entry year of this same cohort had slipped after year 2300 in the US, while no cohort was projected to have a zero gap in Italy.

In contrast, simply extrapolating the current negative trend in the total gender gap suggests that convergence should be attainable in 2073 in the US and 2062 in Italy.[27]

6.3 Halt in the Convergence From Inflows of Worker Cohorts

This section examines why the convergence at entry stopped after the mid-1990s. Our stylized framework predicts that when younger men and women are fairly equally concentrated in higher-paying jobs, a larger supply of older workers cannot shrink the gender pay gap by

[27] Data from Canada and the United Kingdom allow us to reach the same conclusions (Figure A15).

pushing more younger men toward lower-ranked positions. Consistent with the theory, we have shown that the differences between younger men and women in mean pay rank (Section 5.1) and in the share of top jobs (Section 5.2) were small by 1995.

After the mid-1990s, the effects of a further increase in the number of older workers on the gender pay gap depends on the sources of the remaining earning differential between younger men and women. If the outstanding gap depends on education outcomes that predate their labor-market entry, such as their choice of a college major, Prediction 5 states that a larger stock of older workers will not contribute to closing the gap because younger women are not more likely than younger men to acquire high-return skills (Black et al., 2008; Bertrand, 2020; Huneeus et al., 2021; Sloane, Hurst, and Black, 2021; and Bovini, De Philippis, and Rizzica, 2023).

To provide empirical support for the theoretical prediction, this section documents that gender differences in college majors have accounted for a substantial share of the gender pay gap at labor-market entry since the mid-1990s. For this analysis, we focus on college graduates and assess the portion of the remaining entry gender pay gap explained by their college major choices. To this end, we use the American Community Survey for the United States and the Italian Quarterly Labor Force Survey for Italy. Building upon Bertrand (2018), we start from the population of full-time native-born male employees and compute average residual weekly earnings for each college major from regressions that net out year fixed effects and a quadratic polynomial of age. We then quantify the share of individuals in each cohort and gender group that graduated in a specific major. By interacting average residual weekly earnings with the shares of graduates in each college major, we compute the major-predicted average weekly earnings of each cohort-gender combination at labor-market entry.[28]

In both countries, the gender pay gap at entry predicted by younger workers' major choices slightly declined until the mid 1980s (Figure 8).[29] After this period, it has remained remarkably stable for nearly three decades. Notably, the major-predicted gap has been stable since the convergence at entry stopped, constituting approximately 63 percent of the entry gap for college graduates in the United States and 51 percent in Italy.[30] These two large shares are likely to be a lower bound of the importance of pre-labor-market choices for the

[28]Given that we observe the combination of wages and college majors only in more recent years, the major-predicted earnings for older cohorts implicitly assume that the relative wages between majors have been stable over time. In support of this assumption, we show that college major choices have been stable after 1995 in both countries (Figure A16). In addition, our results show a stable major-predicted gender pay gap, which is unlikely to be driven by changes in average wages that perfectly offset each other.

[29]This effect is quantitatively small relative to the observed convergence at entry in average wages (for the US, -0.04 log points compared to -0.25 log points).

[30]Our results based on Italian survey data are slightly smaller than estimates based on Italian administrative data (60 percent from Bovini, De Philippis, and Rizzica (2023)).

gender pay gap at entry, given that we are considering only a single predetermined factor (major choice).

In short, the empirical evidence supports Prediction 5. The gap predicted by major choices has always favored younger men, and started accounting for a very large share of the gender pay gap at labor-market entry by the mid-1990s. Under these conditions, our framework suggests that further increases in the stock of older workers should not keep shrinking the earning differential between younger men and women, as evidenced by the flat trend followed by the gender pay gap at labor-market entry.

7 Conclusion

Our examination of the evolution of the gender pay gap over the past four decades reveals a complex story marked by both progress and weakening in worker conditions. The conventional wisdom attributes the steep narrowing of the gender pay gap observed between the 1970s and 1990s to substantial improvements in women's earnings and labor market outcomes. A halt in such improvements is then generally considered to be the primary reason behind the more recent slowdown in the closing of the gender gap. Instead, our research suggests that both the decline and the halt in the convergence of the gender pay gap are linked to a worsening in the careers of younger workers, particularly younger men, who have faced widening age pay gaps, lower promotion rates, and declining lifetime earnings.

To explore these dynamics, we developed a conceptual framework linking an increased supply of older workers in higher-paying jobs to a smaller gender pay gap among younger workers. Consistent with the importance of cross-cohort effects, we show that the entirety of the decline in the gender pay gap can be attributed to newer cohorts who entered the labor market with smaller gender differentials and to older cohorts who exited the labor market with larger differentials.

Importantly, our study challenges the assumption that the decline in the gender pay gap primarily reflects improved opportunities for women. Instead, we argue that the aging of the workforce has disproportionately and negatively affected younger men because, at baseline, they were more likely to find employment in the type of jobs occupied by an increasing stock of older workers. The data confirm that it was a significant decline in opportunities for younger men, rather than substantial gains for younger women, that drove the convergence in labor-market outcomes that occurred in the 1970s through the mid-1990s. After this point, the remaining gender gap at labor-market entry reflected mostly gender differences in educational choices, rather than differences in initial job allocations. Therefore, further increases in the number of older workers has continued to create bottlenecks to the careers of all younger workers, but not differentially between men and women.

In conclusion, our findings paint a nuanced picture of the dynamics of the gender pay gap. Despite the many prior policies that intended to boost women's position in the wage distribution, our findings caution against interpreting the progressive shrinkage of the gender pay gap as a sign of improved opportunities for women. The design of future gender-responsive policies should strongly consider the outcomes of younger women at or immediately after labor-market entry, because those outcomes represent a largely untapped source for improving gender equality.

References

Acemoglu, Daron and Pascual Restrepo. 2023. "Automation and Rent Dissipation: Implications for Inequality, Productivity, and Welfare." working paper, MIT.

Altonji, Joseph G. and Rebecca M. Blank. 1999. "Race and Gender in the Labor Market." *Handbook of Labor Economics*, , ed. David Card and Orley Ashenfelter Vol. 3, Chapter 48, 3143–3259. Elsevier B.V.

Antwi, James and David C. Phillips. 2013. "Wages and health worker retention: Evidence from public sector wage reforms in Ghana." *Journal of Development Economics*, 102: 101–115.

Arellano-Bover, Jaime. 2024. "Career consequences of firm heterogeneity for young workers: First job and firm size." *Journal of Labor Economics*, 42(2): 549–589.

Azmat, Ghazala and Barbara Petrongolo. 2014. "Gender and the Labor Market: What Have We Learned From Field and Lab Experiments?" *Labour economics*, 30: 32–40.

Bailey, Martha J., Thomas Helgerman, and Bryan A. Stuart. forthcoming. "How the 1963 Equal Pay Act and 1964 Civil Rights Act Shaped the Gender Gap in Pay." *The Quarterly Journal of Economics*.

Banca d'Italia. 2023. "Survey of Household Income and Wealth."

Bayer, Patrick and Kerwin Kofi Charles. 2018. "Divergent paths: A new perspective on earnings differences between black and white men since 1940." *Quarterly Journal of Economics*, 133(3): 1459–1501.

Berger, Mark C. 1985. "The Effect of Cohort Size on Earnings Growth: A Reexamination of the Evidence." *Journal of Political Economy*, 93(3): 561–573.

Bertoni, Marco and Giorgio Brunello. 2021. "Does a Higher Retirement Age Reduce Youth Employment." *Economic Policy*, 36(106): 325–372.

Bertrand, Marianne. 2018. "Coase Lecture–The Glass Ceiling." *Economica*, 85(338): 205–231.

Bertrand, Marianne. 2020. "Gender in the Twenty-First Century." *AEA Papers and Proceedings*, 110: 1–24.

Bianchi, Marco, Björn R. Gudmundsson, and Gylfi Zoega. 2001. "Iceland's natural experiment in supply-side economics." *American Economic Review*, 91(5): 1564–1579.

Bianchi, Nicola and Matteo Paradisi. 2023. "Countries for Old Men: An Analysis of the Age Wage Gap." working paper, Northwestern University.

Bianchi, Nicola, Giulia Bovini, Jin Li, Matteo Paradisi, and Michael Powell. 2023. "Career Spillovers in Internal Labor Markets." *Review of Economic Studies*, 90(4): 1800–1831.

Black, Dan A., Amelia M. Haviland, Seth G. Sanders, and Lowell J. Taylor. 2008. "Gender Wage Disparities Among the Highly Educated." *Journal of Human Resources*, 43(3): 630–659.

Blair, Peter Q and Benjamin Posmanick. 2023. "Why Did Gender Wage Convergence in the United States Stall?" National Bureau of Economic Research Working Paper.

Blau, Francine D. and Lawrence M. Kahn. 2017. "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature*, 55(3): 789–865.

Blundell, Richard, Hugo Lopez, and James P. Ziliak. 2024. "Labour Market Inequality and the Changing Life Cycle Profile of Male and Female Wages." *Working Paper, Institute for Fiscal Studies.*

Boeri, Tito, Pietro Garibaldi, and Espen R. Moen. 2022. "In Medio Stat Victus: Labor Demand Effects of an Increase in the Retirement Age." *Journal of Population Economics*, 35(2): 519–556.

Bovini, Giulia, Marta De Philippis, and Lucia Rizzica. 2023. "The Origins of the Gender Pay Gap: The Role of School-to-Work Transition." *Temi di discussione, Rome: Bank of Italy.*

Card, David, Ana Rute Cardoso, and Patrick Kline. 2016. "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women." *Quarterly Journal of Economics*, 131(2): 633–686.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. "Firms and labor market inequality: Evidence and some theory." *Journal of Labor Economics*, 36(S1): S13–S70.

Card, David and Thomas Lemieux. 2001. "Can falling supply explain the rising return to college for younger men? A cohort-based analysis." *Quarterly Journal of Economics*, 116(2): 705–746.

Carlana, Michela. 2019. "Implicit Stereotypes: Evidence from Teachers' Gender Bias." *The Quarterly Journal of Economics*, 134(3): 1163–1224.

Casarico, Alessandra and Salvatore Lattanzio. 2023. "Behind the child penalty: understanding what contributes to the labour market costs of motherhood." *Journal of Population Economics*, 36(3): 1489–1511.

Casarico, Alessandra and Salvatore Lattanzio. 2024. "What Firms Do: Gender Inequality in Linked Employer-Employee Data." *Journal of Labor Economics*, 42(2): 325–355.

Charles, Kerwin Kofi, Erik Hurst, and Mariel Schwartz. 2019. "The Transformation of Manufacturing and the Decline in US Employment." *NBER Macroeconomics Annual*, 33(1): 307–372.

Charles, Kerwin Kofi, Jonathan Guryan, and Jessica Pan. forthcoming. "The Effects of Sexism on American Women: The Role of Norms Vs. Discrimination." *Journal of Human Resources.*

Cross-National Data Center. 2023. "Luxembourg Income Study (LIS) Database."

Dabla-Norris, Era, Carlo Pizzinelli, and Jay Rappaport. 2023. "Job Polarization and the Declining Wages of Young Female Workers in the United Kingdom." *Oxford Bulletin of Economics and Statistics.*

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2014. "The Role of Entrepreneurship in US Job Creation and Economic Dynamism." *Journal of Economic Perspectives*, 28(3): 3–24.

De Quinto, Alicia and Libertad González. 2024. "The Short-and Long-Term Effects of Family-Friendly Policies on Women's Employment." *BSE Working Paper 1434.*

Eissa, Nada and Hilary Williamson Hoynes. 2004. "Taxes and the labor market participation of married couples: the earned income tax credit." *Journal of Public Economics*, 88(9-10): 1931–1958.

England, Paula, Andrew Levine, and Emma Mishel. 2020. "Progress toward gender equality in the United States has slowed or stalled." *Proceedings of the National Academy of Sciences*, 117(13): 6990–6997.

Fernández, Raquel, Alessandra Fogli, and Claudia Olivetti. 2004. "Mothers and Sons:

Preference Formation and Female Labor Force Dynamics." *The Quarterly Journal of Economics*, 119(4): 1249–1299.

Ferrari, Irene, Jan Kabátek, and Todd Morris. 2023. "Longer careers: A barrier to hiring and coworker advancement." IZA Discussion Paper No. 16098.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, and Michael Westberry. 2022. "Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]."

Foliano, Francesca, Alex Bryson, Heather Joshi, Bozena Wielgoszewska, and Dave Wilkinson. 2023. "Gender wage gap among young adults: a comparison across British cohorts." IZA Discussion Paper 15973.

Freeman, Richard B. 1979. "The Effect of Demographic Factors on Age-Earnings Profiles." *The Journal of Human Resources*, 14(3): 289.

Fukui, Masao, Amy Nakamura, and Jon Steinsson. 2023. "Women, Wealth Effects, and Slow Recoveries." *American Economic Journal: Macroeconomics*, 15(1): 269–313.

Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review*, 104(4): 1091.

Goldin, Claudia and Joshua Mitchell. 2017. "The New Life Cycle of Women's Employment: Disappearing Humps, Sagging Middles, Expanding Tops." *Journal of Economic Perspectives*, 31(1): 161–182.

Guaitoli, Gabriele and Roberto Pancrazi. 2023. "Global Trends in Income Distribution Across Age Groups?" working paper, University of Warwick.

Guvenen, Fatih, Greg Kaplan, Jae Song, and Justin Weidner. 2022. "Lifetime Earnings in the United States over Six Decades." *American Economic Journal: Applied Economics*, 14(4): 446–79.

Huneeus, Federico, Conrad Miller, Christopher Neilson, and Seth Zimmerman. 2021. "Firm Sorting, College Major, and the Gender Earnings Gap." *Industrial Relations Section Working Paper Series*, 648.

Istituto Nazionale della Previdenza Sociale. 2021. "Database UNIEMENS."

Istituto Nazionale di Statistica. 2023. "Rilevazione Sulle Forze di Lavoro."

Kahn, Lisa B. 2010. "The Long-Term Labor Market Consequences of Graduating From College in a Bad Economy." *Labour Economics*, 17(2): 303–316.

Karademir, Sencer, Jean-William P. Laliberté, and Stefan Staubli. 2024. "The Multigenerational Impact of Children and Childcare Policies." *NBER Working Paper 32204*.

Kleven, Henrik. 2023. "The Geography of Child Penalties and Gender Norms: A Pseudo-Event Study Approach." *NBER Working Paper*, 30176.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard. 2019. "Children and Gender Inequality: Evidence from Denmark." *American Economic Journal: Applied Economics*, 11(4): 181–209.

Kleven, Henrik, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller. 2024. "Do Family Policies Reduce Gender Inequality? Evidence From 60 Years of Policy Experimentation." *American Economic Journal: Economic Policy*, 16(2): 110–149.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler. 2019. "Imperfect Competition, Compensating Differentials and Rent Sharing in the U.S. Labor Market." BFI Working Paper 2019-84.

Lazear, Edward P., Kathryn L. Shaw, and Christopher T. Stanton. 2018. "Who Gets Hired? The Importance of Competition Among Applicants." *Journal of Labor Economics*, 36(S1): S133–S181.

Manoli, Dayanand and Andrea Weber. 2011. "Nonparametric evidence on the effects of retirement benefits on labor force participation decisions." *SSRN Working Paper*.

Mohnen, Paul. forthcoming. "The Impact of the Retirement Slowdown on the U.S. Youth Labor Market." *Journal of Labor Economics*.

Morchio, Iacopo and Christian Moser. 2023. "The Gender Pay Gap: Micro Sources and Macro Consequences." *CEPR Discussion Paper 16383*.

OECD. 2022. "Age of mothers at childbirth and age-specific fertility." *OECD Strategic Foresight 2.3*.

Olivetti, Claudia and Barbara Petrongolo. 2016. "The Evolution of Gender Gaps in Industrialized Countries." *Annual Review of Economics*, 8(1): 405–434.

Oreopoulos, Philip, Till von Wachter, and Andrew Heisz. 2012. "The Short- and Long-Term Career Effects of Graduating in a Recession." *American Economic Journal: Applied Economics*, 4(1): 1–29.

Porter, Catherine and Danila Serra. 2020. "Gender Differences in the Choice of Major: The Importance of Female Role Models." *American Economic Journal: Applied Economics*, 12(3): 226–54.

Rosolia, Alfonso and Roberto Torrini. 2007. "The generation gap: relative earnings of young and old workers in Italy." Bank of Italy, Economic Research and International Relations Area Temi di discussione (Economic working papers) 639.

Ruffini, Krista. 2022. "Worker earnings, service quality, and firm profitability: Evidence from nursing homes and minimum wage reforms." *Review of Economics and Statistics*, 1–46.

Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Renae Rogers, and Megan Schouweiler. 2023. "IPUMS USA: Version 14.0 [dataset]."

Sloane, Carolyn M., Erik G. Hurst, and Dan A. Black. 2021. "College Majors, Occupations, and the Gender Wage Gap." *Journal of Economic Perspectives*, 35(4): 223–248.

U.S. Census Bureau. 2023. "American Community Survey."

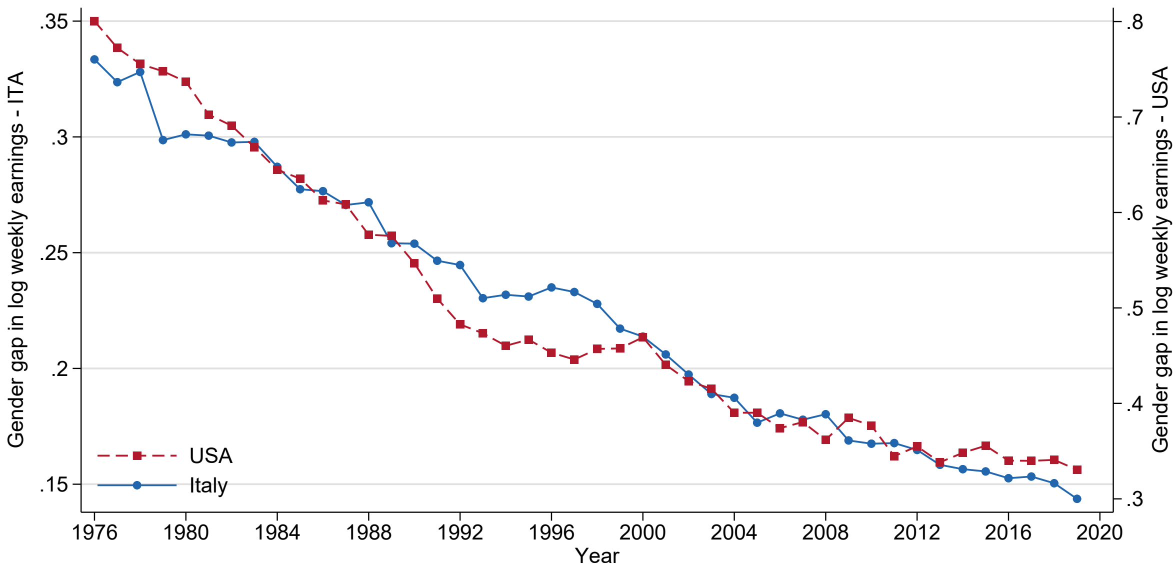
Welch, Finis. 1979. "Effects of Cohort Size on Earnings: The Baby Boom Babies' Financial Bust." *Journal of Political Economy*, 87(5): S65–S97.

World Economic Forum. 2023. "Global Gender Gap Report."

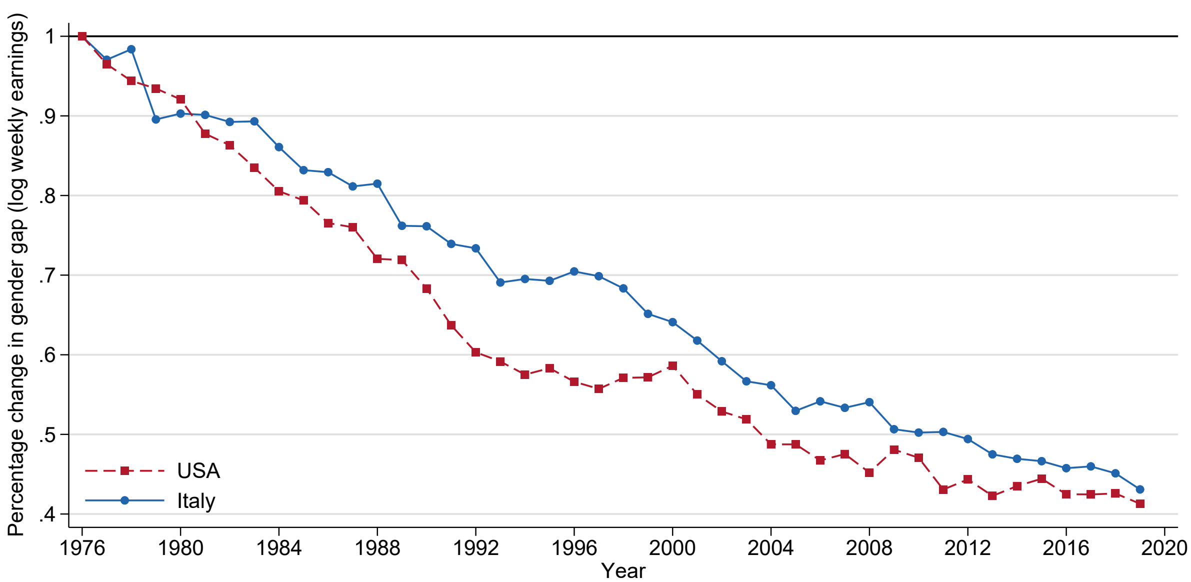
Figures and Tables

Figure 1: Gender Gap in Weekly Earnings

Panel A: Raw gap



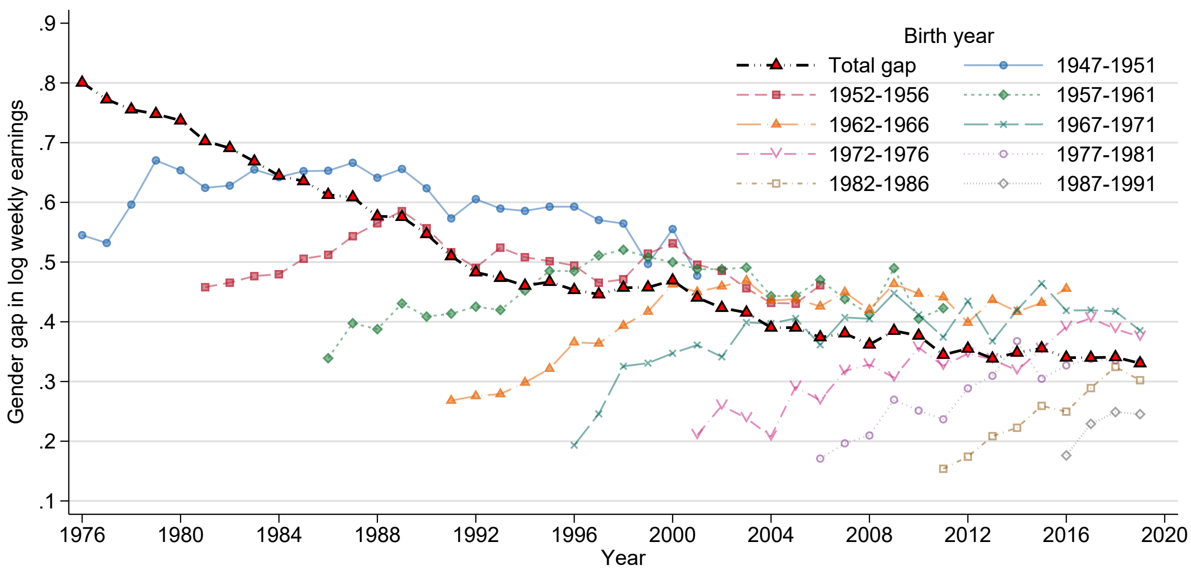
Panel B: Percentage deviation from first year



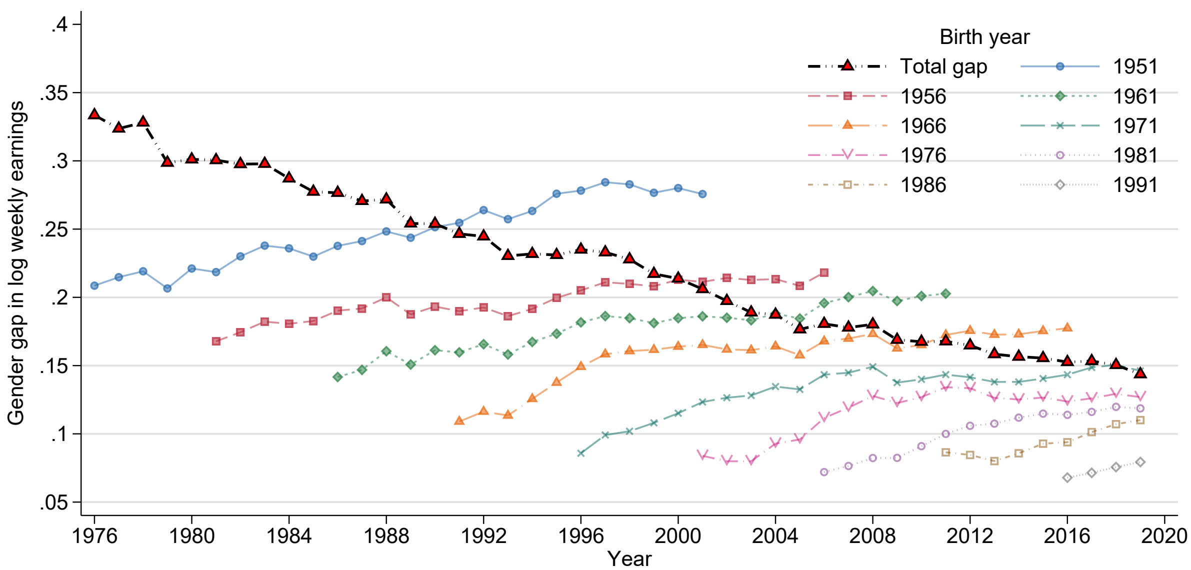
Notes: Panel A plots the trend in the raw mean gender pay gap (log weekly earnings of men - log weekly earnings of women) in Italy and the United States. Panel B shows the percentage deviation from the first year (1976). In each year, the data encompass information about all workers who were between 25 and 64 years old, had worked in the private sector for at least 24 weeks, had earned strictly positive earnings, and (only in Italy) had not retired by December 31.
Source for Italy: UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Source for the United States*. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Figure 2: Gender Pay Gap Between and Within Cohorts

Panel A: USA



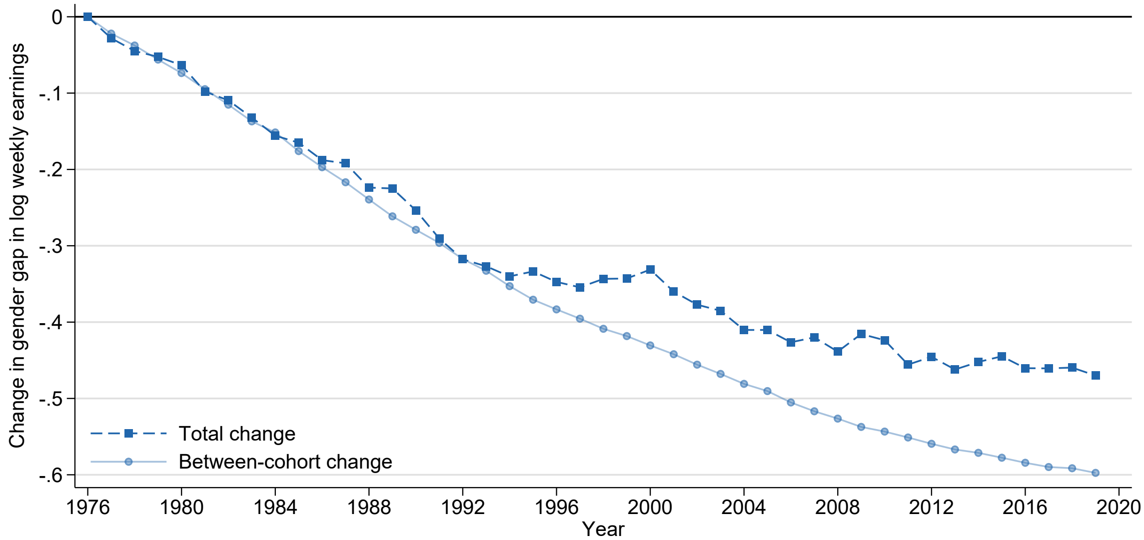
Panel B: Italy



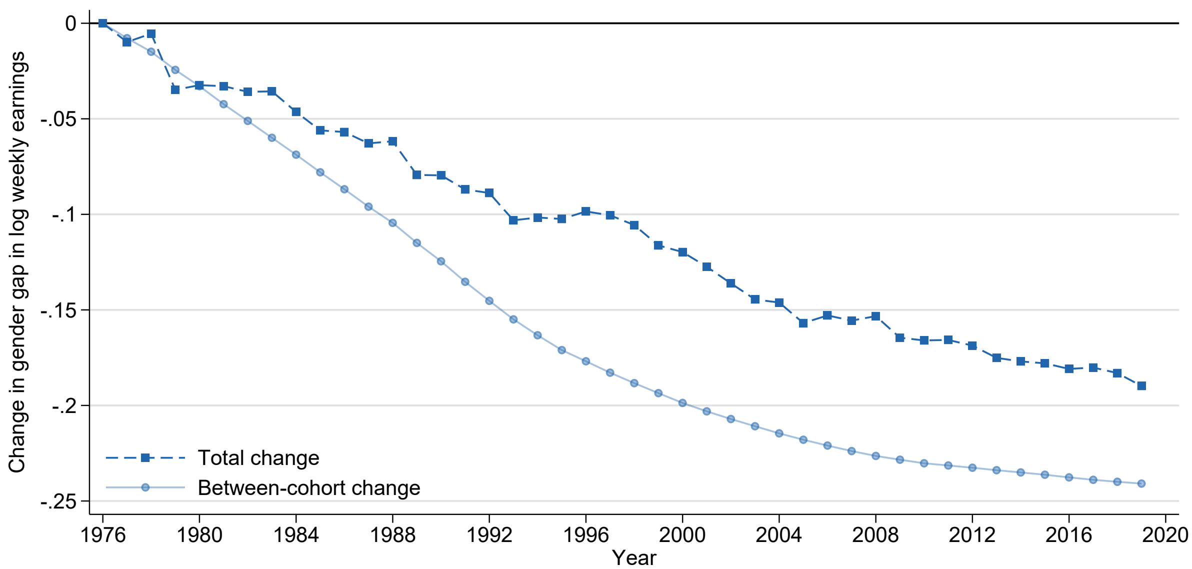
Notes: Panel A (Panel B) shows the trend in the mean gender pay gap (log weekly earnings) across different groups of birth cohorts in the United States (single birth cohorts in Italy). The red triangles depict the trend in the mean gender pay gap across all cohorts active in the labor market in each year. This analysis includes only workers aged 50 or younger to limit the influence of cross-cohort changes in selection into retirement. In each year, the data encompass information about all workers who were between 25 and 50 years old, had worked in the private sector for at least 24 weeks, had earned strictly positive earnings, and (only in Italy) had not retired by December 31.
Source for Italy: UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Source for the United States*. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Figure 3: Between-Cohort Decline in Gender Pay Gap

Panel A: USA



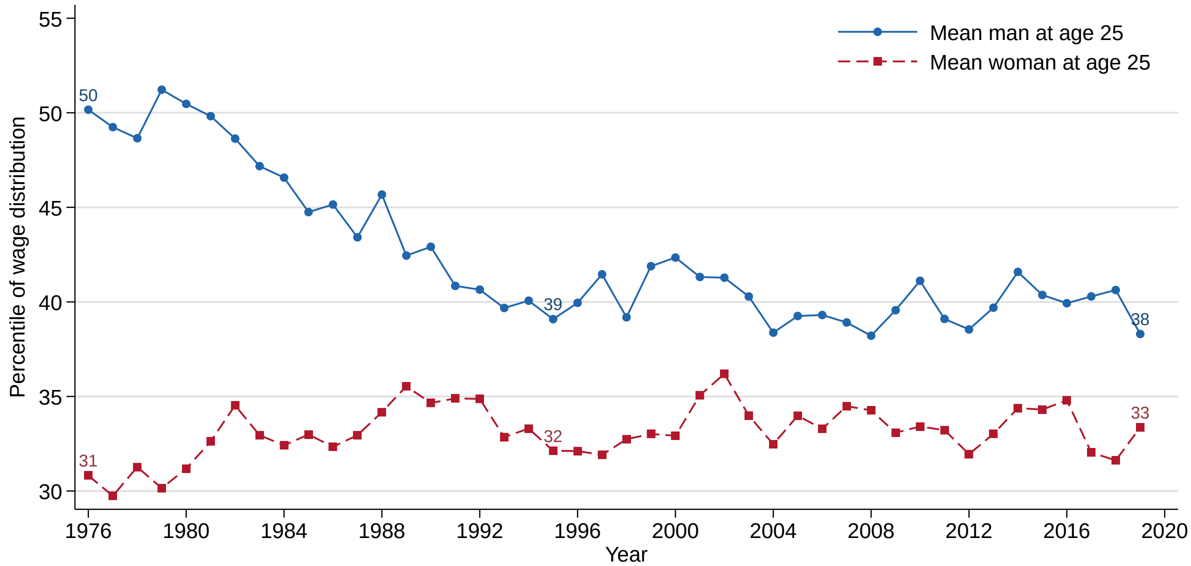
Panel B: Italy



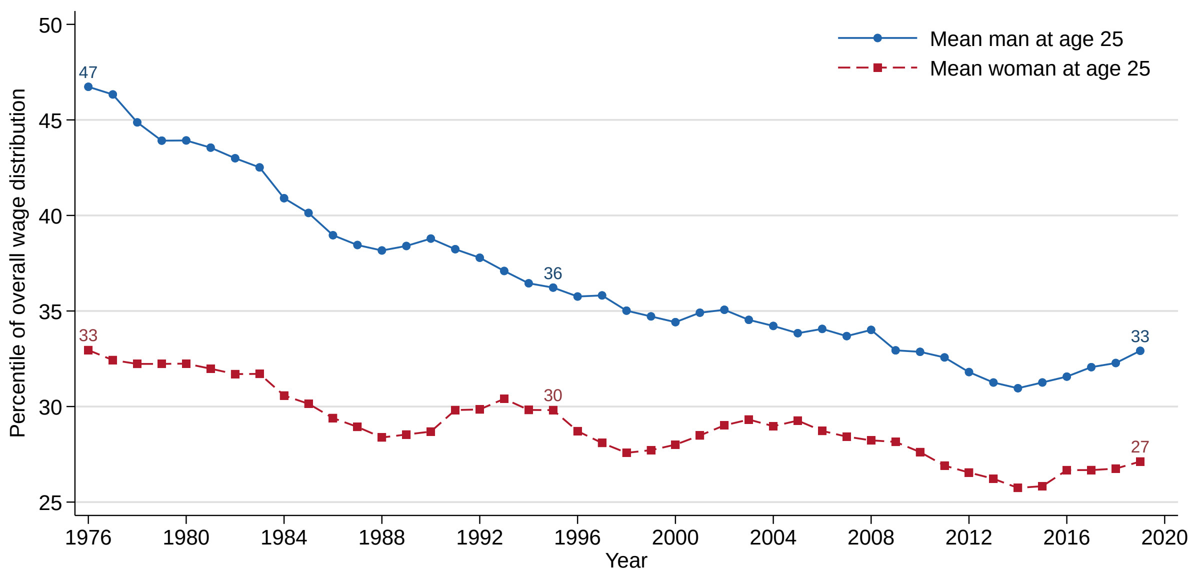
Notes: Panels A and B show the change in the total gender pay gap and its between-cohort component in the United States and Italy, respectively, for log weekly earnings. To compute the between-cohort component, we assign to each cohort (defined as a combination of birth year and gender) its mean log weekly earnings at labor-market entry in each year (Equation (2)). In the baseline analysis, entry into the labor market corresponds with the year in which workers in each cohort were 25 years old. We assign to cohorts who were older than 25 in the first sample year their mean weekly earnings in the first sample year. In each year, the data encompass information about all workers who were between 25 and 64 years old, had worked in the private sector for at least 24 week, had earned strictly positive earnings, and (only in Italy) had not retired by December 31.
Source for Italy: UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). Source for the United States. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Figure 4: Positions in Pay Distribution at 25 Years Old

Panel A: USA



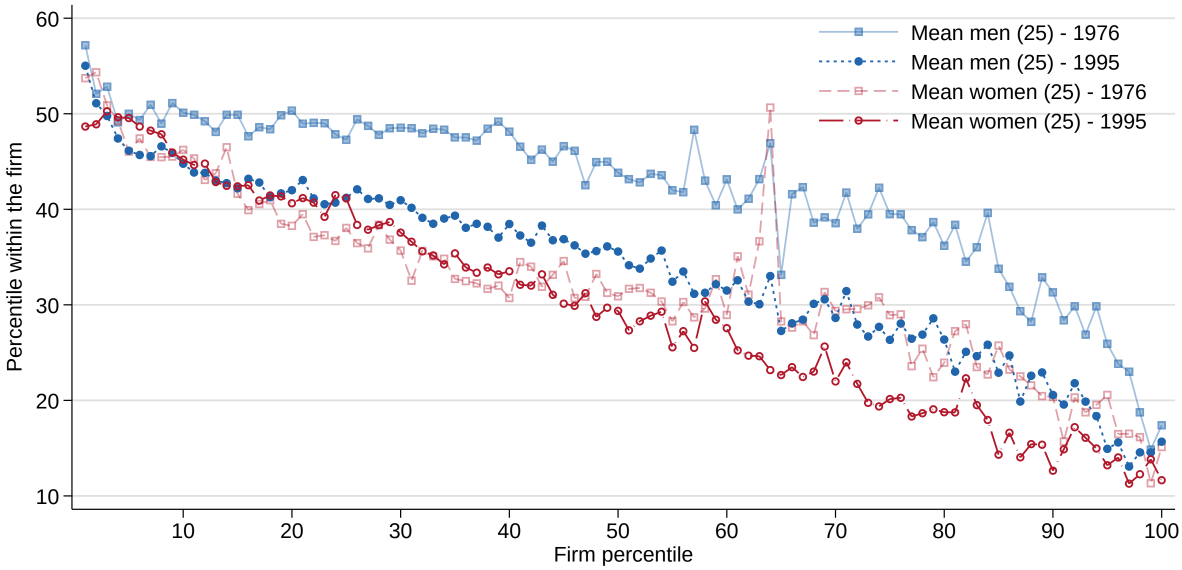
Panel B: Italy



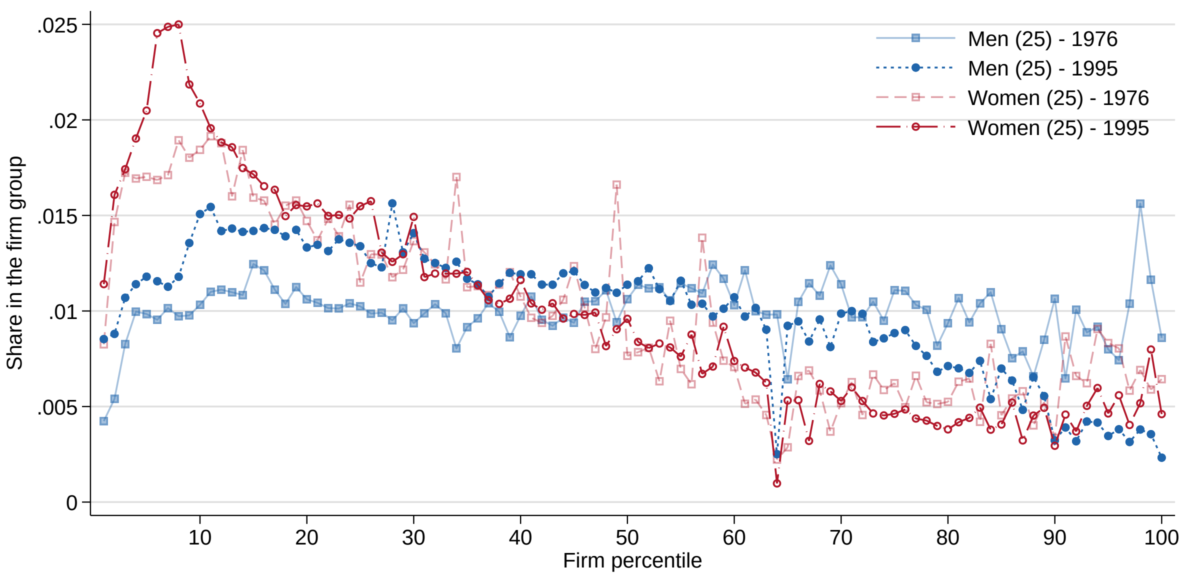
Notes: Panels A and B show the average earning percentile of men and women at 25 years old in the United States and Italy, respectively. In each year, the data encompass information about all workers who were between 25 and 64 years old, had worked in the private sector for at least 24 weeks, had earned strictly positive earnings, and (only in Italy) had not retired by December 31.
Source for Italy: UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Source for the United States*. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Figure 5: Distribution Within and Between Firms for Workers Age 25—Italy

Panel A: Average position within firm earnings percentiles, 1976 and 1995



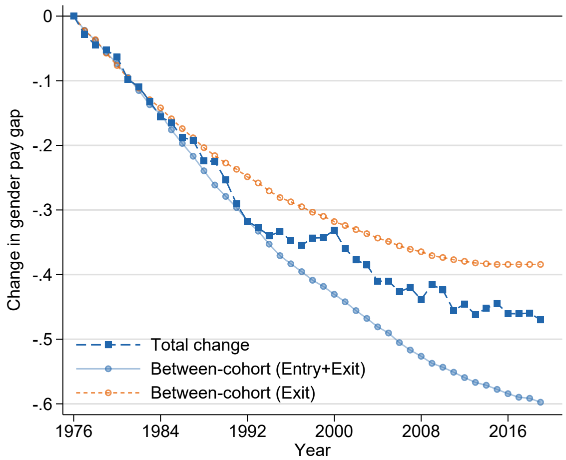
Panel B: Distribution across firm earnings percentile, 1976 and 1995



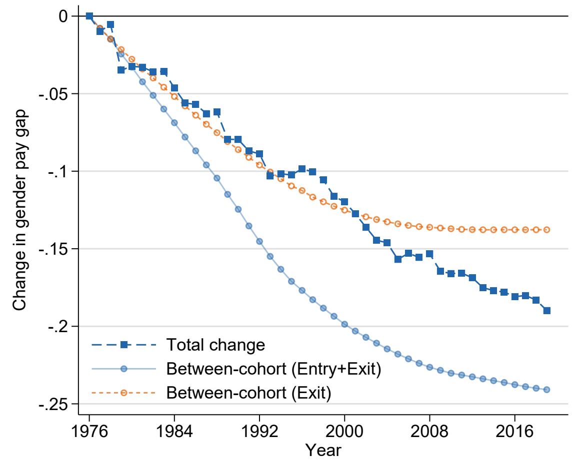
Notes: Panel A shows the average percentile, in the distribution of weekly earnings, of men and women at age 25 across percentiles of firm mean pay in 1976 and 1995. Panel B shows the share of men and women at age 25 across percentiles of firm average pay in 1976 and 1995. In each year, the data encompass information about all workers who were between 25 and 64 years old, had worked in the private sector for at least 24 weeks, had earned strictly positive earnings, and had not retired by December 31.
Source: UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS).

Figure 6: Between-Cohort Change Without Convergence in Entry Earnings

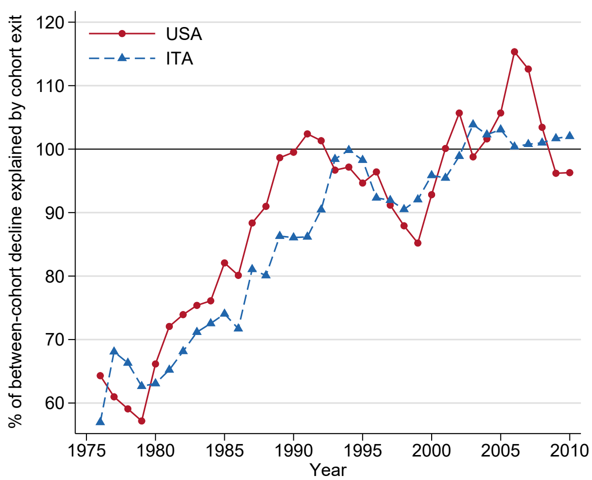
Panel A: United States



Panel B: Italy

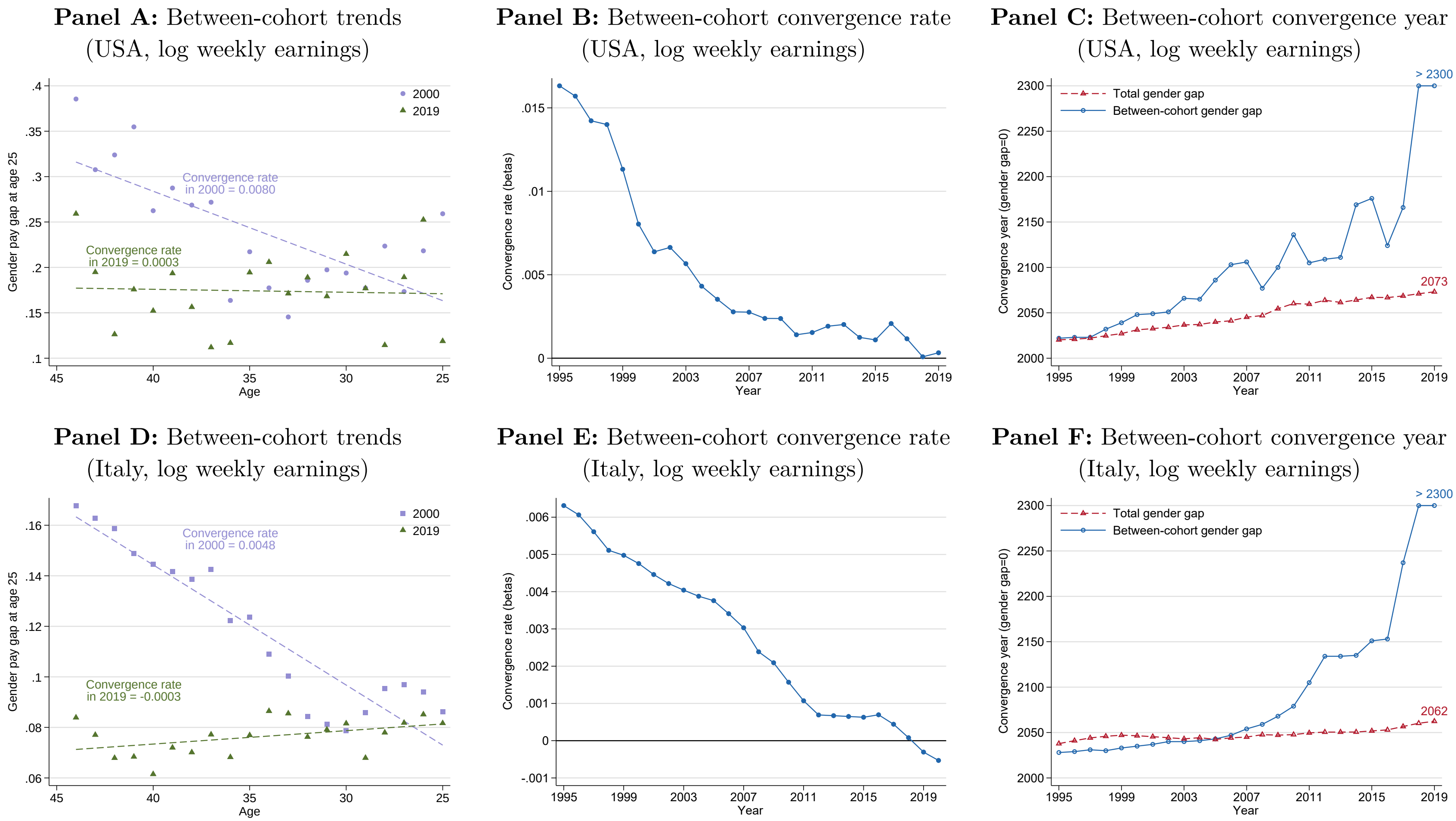


Panel C: Share of convergence at labor-market exit
over total between-cohort convergence



Notes: Panels A and B compute for the United States and Italy, respectively, what would have happened to the gender pay gap if the cross-cohort convergence in the earnings of men and women at age 25 had stopped in benchmark year $t_b = 1976$. In this counterfactual exercise, the weekly earnings of cohorts who entered the labor market after 1976 are set equal to the average weekly earnings of cohorts who entered the labor market between 1976 and the following two years. Panel C shows the ratio between the change (last year - first year) in the wage gap predicted by this new counterfactual scenario and the change in the gender gap predicted by the total between-cohort component from Equation (2) when the benchmark year t_b moves between 1976 and 2010. The ratio is such that 100 implies that the decline in the gender pay gap predicted by the new counterfactual scenario accounts for the entire between-cohort change. In each year, the data encompass information about all workers who were between 25 and 64 years old, had worked in the private sector for at least 24 weeks, had earned strictly positive earnings, and (only in Italy) had not retired by December 31.
Source for Italy: UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Source for the United States*. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

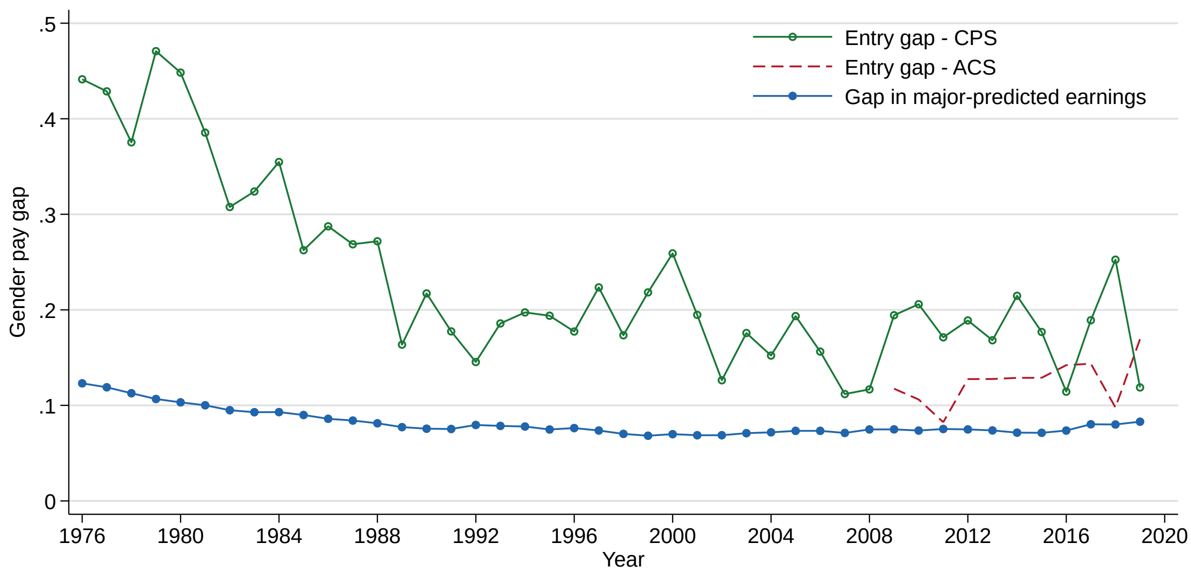
Figure 7: Projected Convergence in the Gender Pay Gap



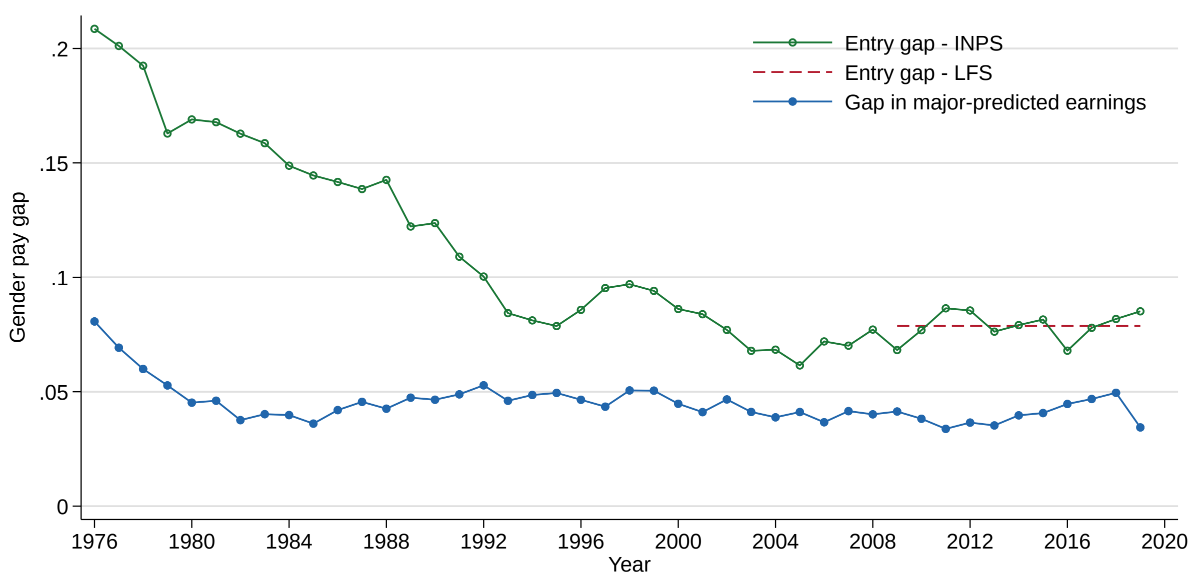
Notes: In each year t, we compute the gender gap at an early career stage for workers in age group a using their earnings at age 25. Then, we estimate the linear relationship between the mean gender gap at labor-market entry and age (Equation (3)). Panels A and D show the best fit line in 2000 and 2019 for the United States and Italy, respectively. Panels B and E show the coefficients of age (β_t) for each year t. Panels C and F show the first year of convergence in the gender gap (first cohort with gap at most equal to zero at age 25) predicted by Equation (3) for each year between 1976 and 2019. In each year, the data encompass information about all workers who were between 25 and 64 years old, had worked in the private sector for at least 24 weeks, had earned strictly positive earnings, and (only in Italy) had not retired by December 31.
Source for Italy: UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Source for the United States:* Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Figure 8: Major-Predicted Gender Pay Gap

Panel A: USA



Panel B: Italy



Notes: The figure shows the gender gap in major-predicted residual weekly earnings computed from the American Community Survey in the United States (Panel A) and from the Labor Force Survey in Italy (Panel B). Predicted residual weekly earnings are computed by averaging across college majors the residual weekly earnings of native-born male employees working full-time in the private sector. Residual weekly earnings are obtained from OLS log wage regressions that control for a quadratic polynomial in age and time fixed effects over the period 2009-2019 in both countries. Then, we multiply these major-specific residuals by cohort-gender shares in each major. There are 32 college majors in the data for Italy and 176 for the United States. The gender gap at entry (age 25) is measured with both CPS and ACS (for college graduates) data in the United States and with both INPS and LFS (for college graduates) data in Italy. In the last case, we report the average between 2009 and 2019 because yearly averages are too noisy.
Sources for Italy: Quarterly Labour Force Survey, Istituto Nazionale di Statistica (ISTAT); UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Sources for the United States*: Integrated Public Use Microdata Series, American Community Survey; Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Table 1: Alternative Definitions of Labor-Market Entry

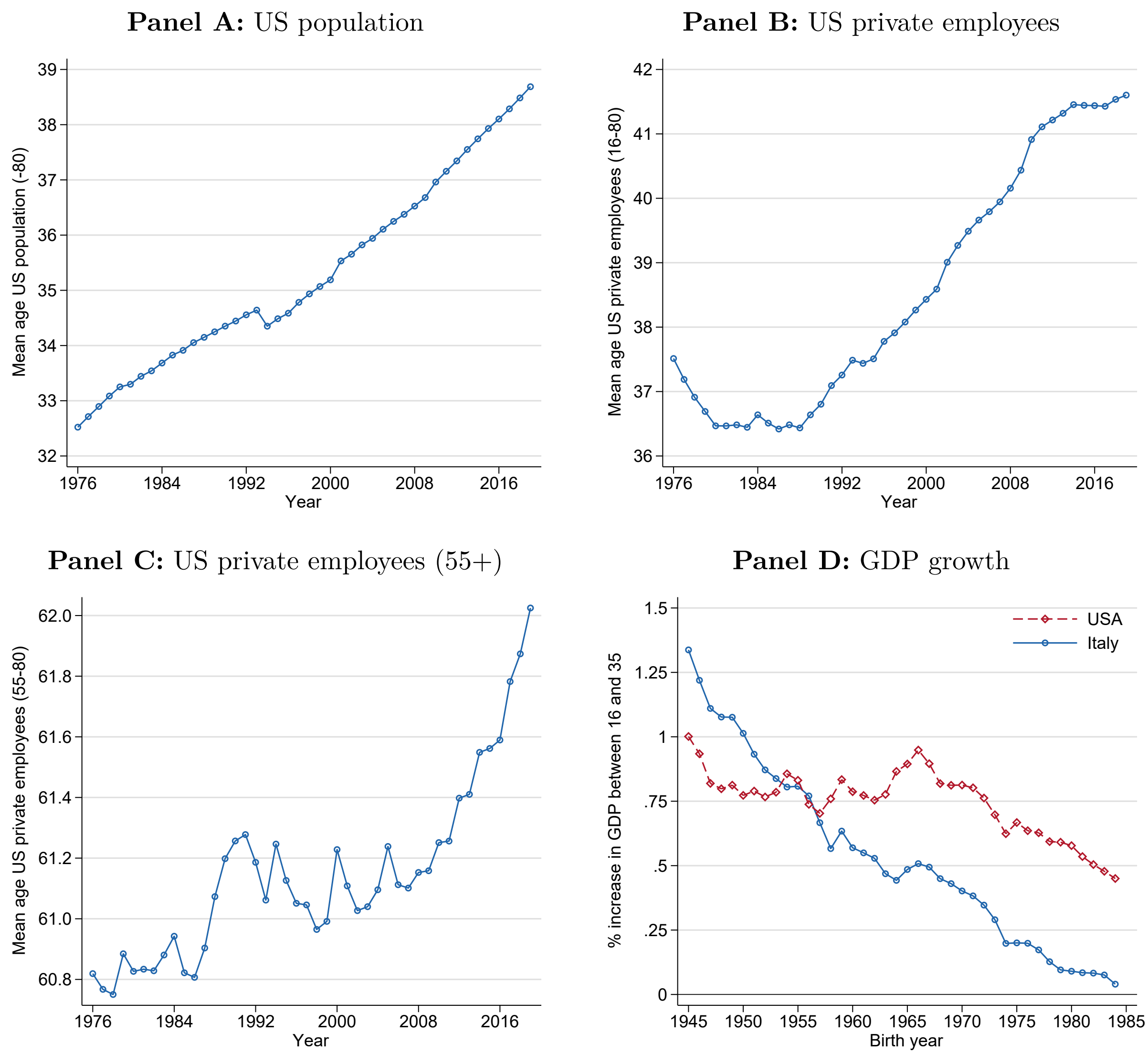
	Total change in gender gap (last year - first year)		Between-cohort change in gender gap (last year - first year)				Between-cohort change in gender gap (Shorter time series)			
	First year	Change	Earnings at age 25	Earnings at age 28	Earnings at age 30	Earnings b/w age 25 and age 30	Total change (last y. - 2015)	Earnings at age 25	Total change (last y. - 1980)	True entry earnings
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Log weekly earnings										
United States (1976-2019)	0.800	-0.470	-0.598	-0.606	-0.615	-0.575	-0.025	-0.034	-	-
Italy (1976-2019)	0.333	-0.190	-0.241	-0.218	-0.197	-0.218	-0.012	-0.014	-0.157	-0.280
Canada (1973-2011)	0.672	-0.288	-0.430	-0.507	-0.326	-0.335	-	-	-	-
Panel B: Log yearly earnings										
United States (1976-2019)	0.835	-0.495	-0.621	-0.645	-0.657	-0.598	-0.029	-0.035	-	-
Italy (1976-2019)	0.350	-0.092	-0.204	-0.165	-0.127	-0.166	-0.012	-0.016	-0.061	-0.270
Canada (1973-2011)	0.695	-0.308	-0.460	-0.621	-0.397	-0.359	-	-	-	-
United Kingdom (1976-2019)	1.058	-0.627	-0.750	-0.629	-0.622	-0.669	-0.065	-0.038	-	-
Panel C: Log hourly earnings										
United States (1976-2019)	0.568	-0.361	-0.452	-0.464	-0.458	-0.441	-0.012	-0.024	-	-

Notes: Columns 1 and 2 show the change in gender gap between the first and last available years for each country. Columns 3 to 6 show the between-cohort component of the change in the gender gap between the first and last available years (Equation (2)). In this counterfactual scenario, we assign to each cohort (defined as a combination of birth year and gender) its mean earnings (weekly, yearly, or hourly) at labor-market entry in every year. These columns differ in the definition of labor-market entry: mean earnings at age 25 (Col. 3), at age 28 (Col. 4), at age 30 (Col. 5), and between age 25 and age 30 (Col. 6). Cohorts who are above these age thresholds in the first sample year receive their mean earnings in the first sample year. In Column 8, we assign to each cohort its mean earnings at age 25, starting from the year in which they are observable for all cohorts, including older ones (2015). This analysis can be performed only with data from Italy, the United States, and the United Kingdom. Column 7 shows the total change in gender gap between 2015 and the last available year. In Column 10, we assign to each cohort its true earnings at labor-market entry, rather than its mean earnings at age 25. This analysis is available only for Italy and only from 1980. Column 9 shows the total change in gender gap between 1980 and the last available year. In this analysis. the time series for Canada stops in 2011, the last sample year in which the exact age of each worker is available. In each year, the data encompass information about all workers who were between 25 and 64 years old, had worked in the private sector for at least 24 weeks (not available in the UK), had earned strictly positive earnings, and (only in Italy) had not retired by December 31 (only in Italy). *Source for Italy:* UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Source for the United States*. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS. *Source for LIS data:* Luxembourg Income Study (LIS) Database, last accessed on 06/01/2023 at https://www.lisdatacenter.org/.

Online Appendix

A Additional Figures and Tables

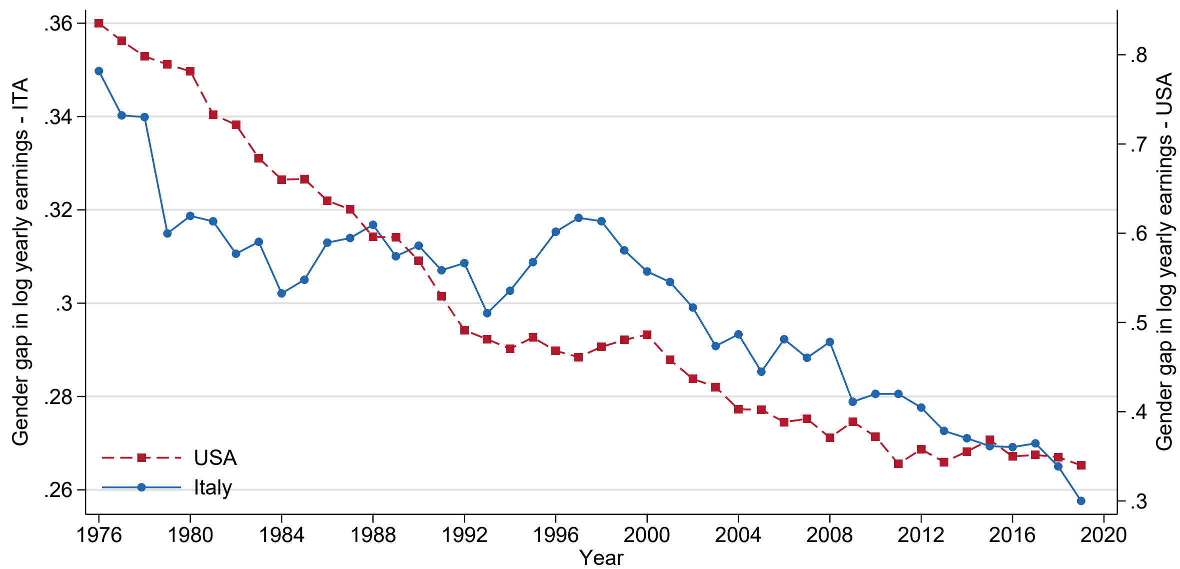
Figure A1: Workforce Aging and GDP Growth



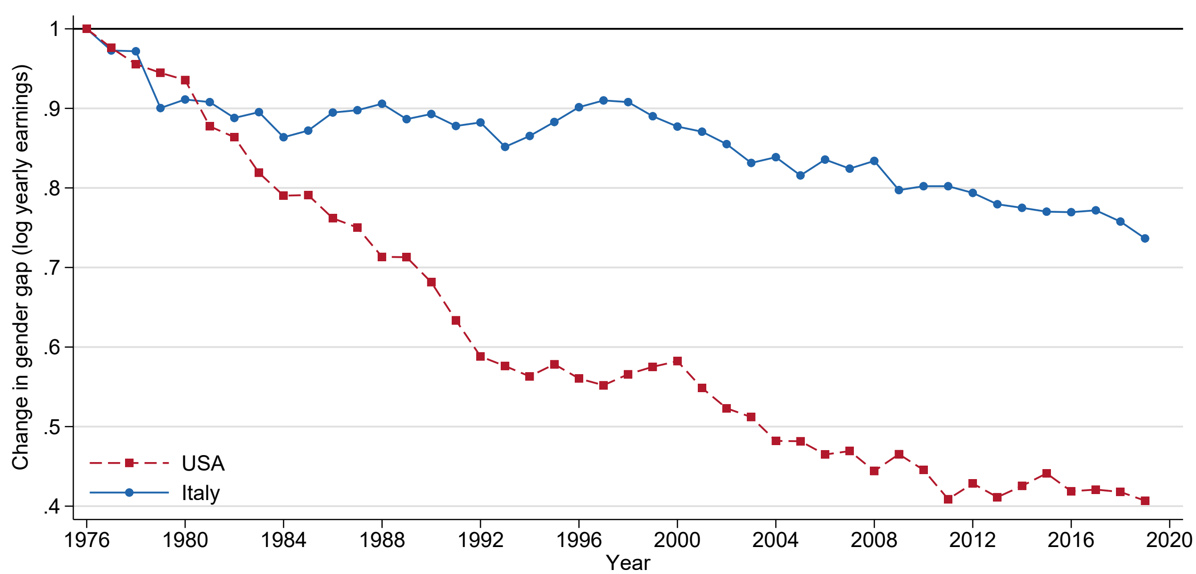
Notes: Panels A to C show the mean age of the US population, private-sector employees in the US (16+), and older private-sector employees in the US (55+). All age variables are winsorized at age 80. Panel D computes the cumulative percentage change in GDP (in 2010 USD) over the first years in the labor market for individuals born in different years. For example, the data point for the birth year "1945" computes the percentage growth in GDP between 1961 (when individuals born in 1945 were 16 years old) and 1980 (when individuals born in 1945 were 35 years old).
Source for Italy: UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Source for the United States*: Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS. *Source for GDP data*: World Development Indicators by the World Bank, last accessed on 04/21/2023 at `https://databank.worldbank.org/reports.aspx?source=2&series=NY.GDP.MKTP.CD&country=`.

Figure A2: Gender Gap in Yearly Labor Earnings

Panel A: Raw gap



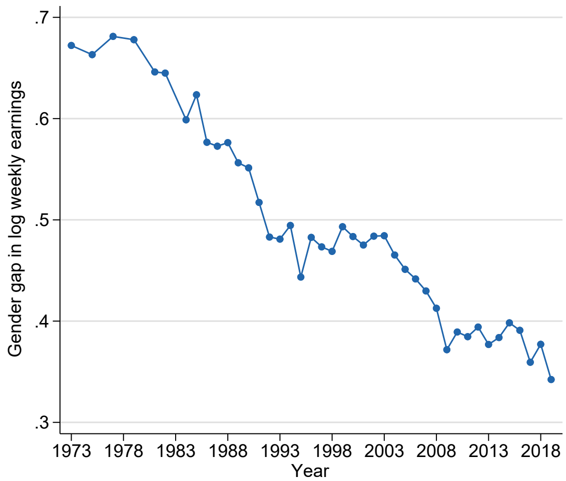
Panel B: Percentage deviation from first year



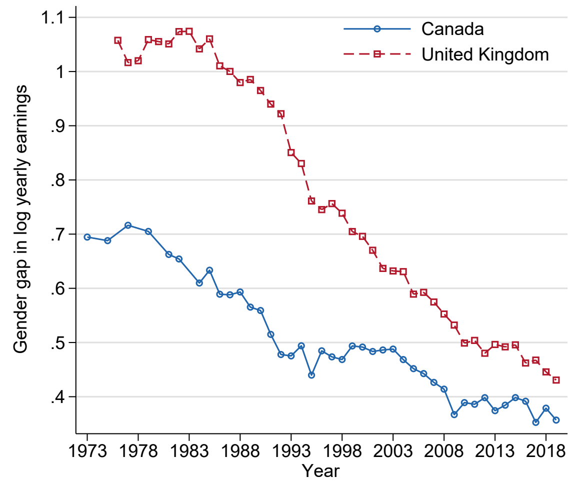
Notes: Panel A plots the trend in the raw mean gender gap (log yearly earnings of men - log yearly earnings of women) in Italy and the United States. Panel B shows the percentage deviation from the first year (1976). In each year, the data encompass information about all workers who were between 25 and 64 years old, had worked in the private sector for at least 24 weeks, had earned strictly positive earnings, and (only in Italy) had not retired by December 31.
Source for Italy: UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Source for the United States*. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Figure A3: Gender Pay Gap in Other High-Income Countries

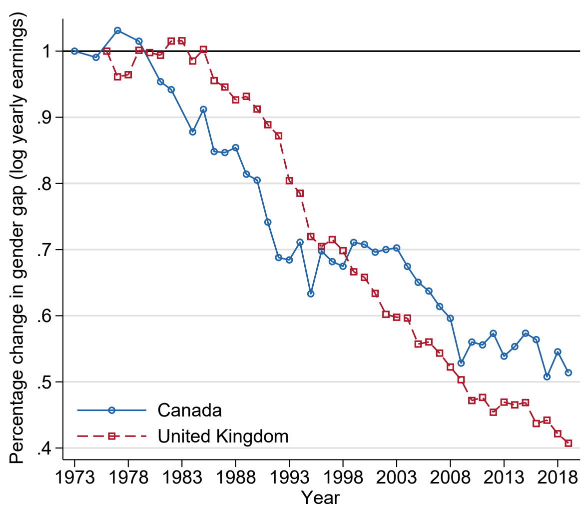
Panel A: Log weekly earnings (Canada)



Panel B: Log yearly earnings



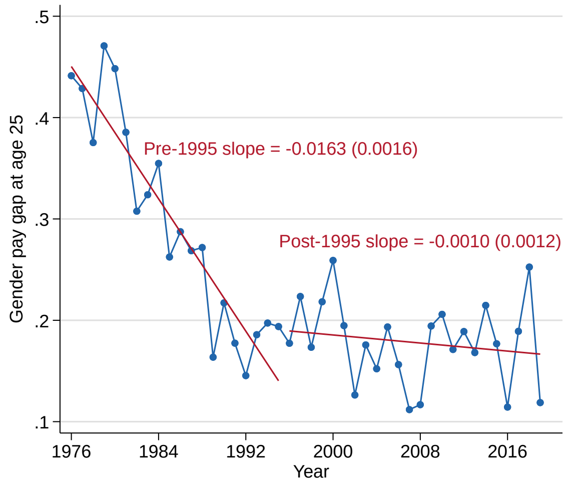
Panel C: Percentage deviation from first year
(log yearly earnings)



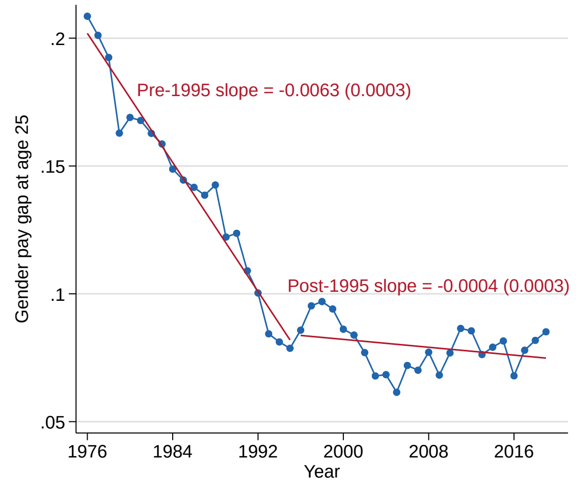
Notes: Panel A plots the trend in the raw mean gender gap in log weekly earnings for Canada. Panel B plots the trend in the raw mean gender gap in log yearly earnings, which are available in Canada and the United Kingdom. Panel C shows the percentage deviation from the first in the gender gap in log yearly earnings. In each year, the data encompass information about all workers who were between 25 and 64 years old, had worked in the private sector for at least 24 weeks (only in Canada), and had earned strictly positive earnings. *Source for LIS data:* Luxembourg Income Study (LIS) Database, last accessed on 06/01/2023 at https://www.lisdatacenter.org/.

Figure A4: Gender Pay Gap at Age 25

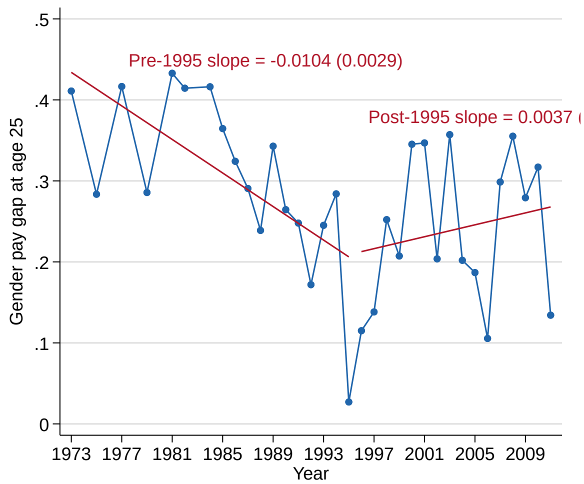
Panel A: USA
(log weekly earnings)



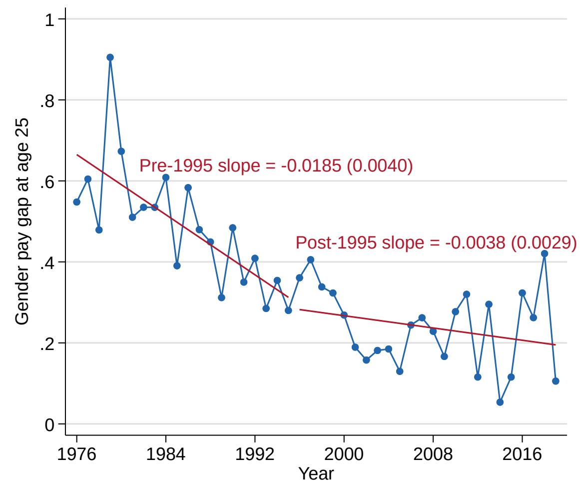
Panel B: Italy
(log weekly earnings)



Panel C: Canada
(log weekly earnings)



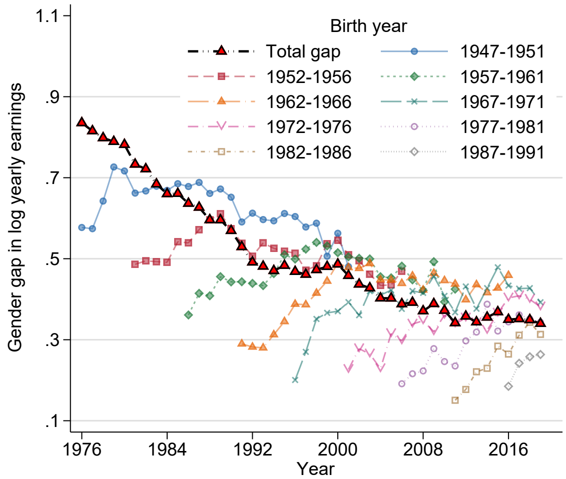
Panel D: United Kingdom
(log yearly earnings)



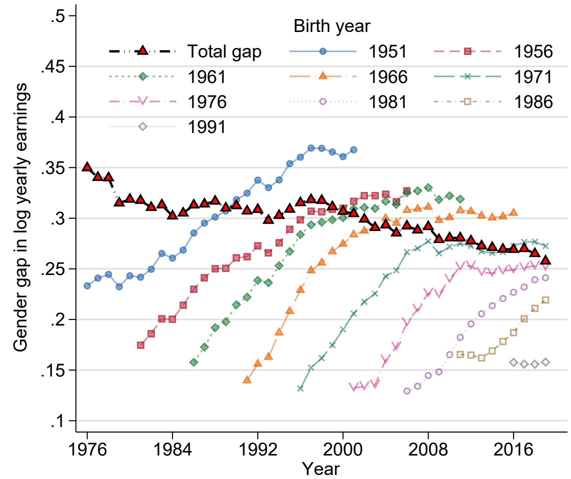
Notes: Panels A to D show the trend in the mean gender gap in log earnings at age 25 in the United States, Italy, Canada, and the United Kingdom, respectively. In each year, the data encompass information about all workers who were 25 years old, had worked in the private sector for at least 24 weeks (not available in the UK), had earned strictly positive earnings, and (only in Italy) had not retired by December 31.
Source for Italy: UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Source for the United States.* Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS. *Source for LIS data:* Luxembourg Income Study (LIS) Database, last accessed on 06/01/2023 at https://www.lisdatacenter.org/.

Figure A5: More Results on Cross-Cohort Trends in Gender Pay Gap

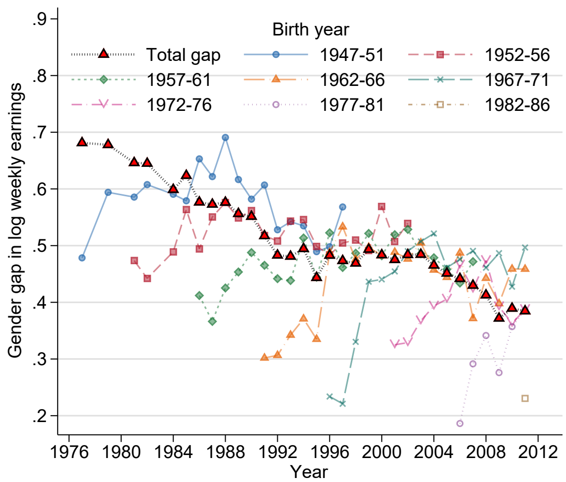
Panel A: USA
(log yearly earnings)



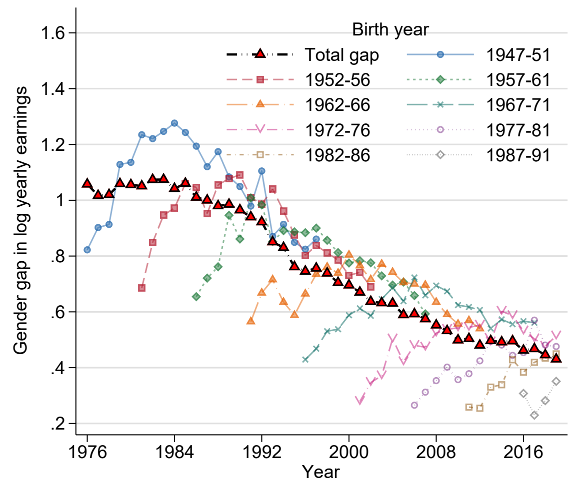
Panel B: Italy
(log yearly earnings)



Panel C: Canada
(log weekly earnings)

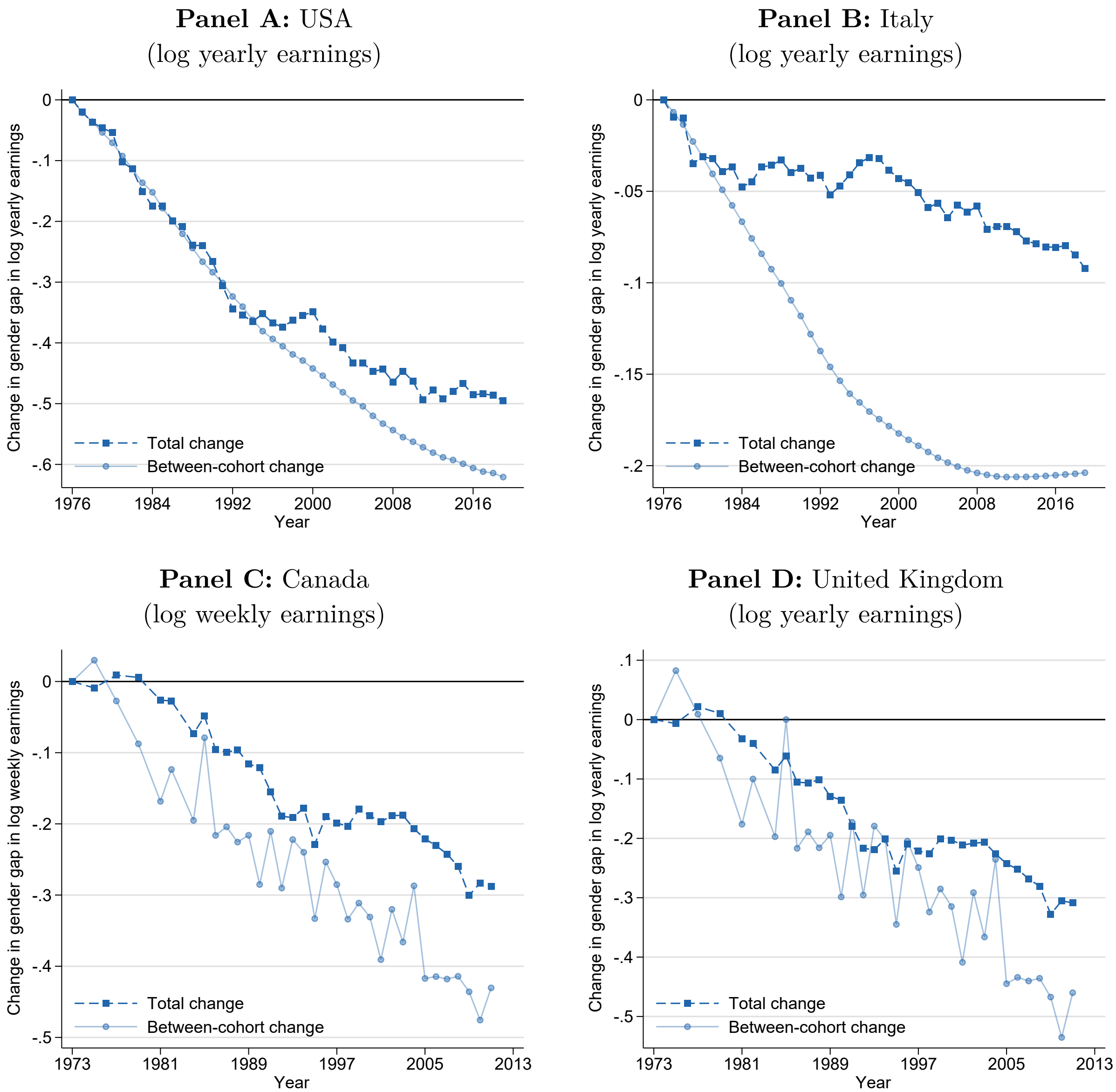


Panel D: United Kingdom
(log yearly earnings)



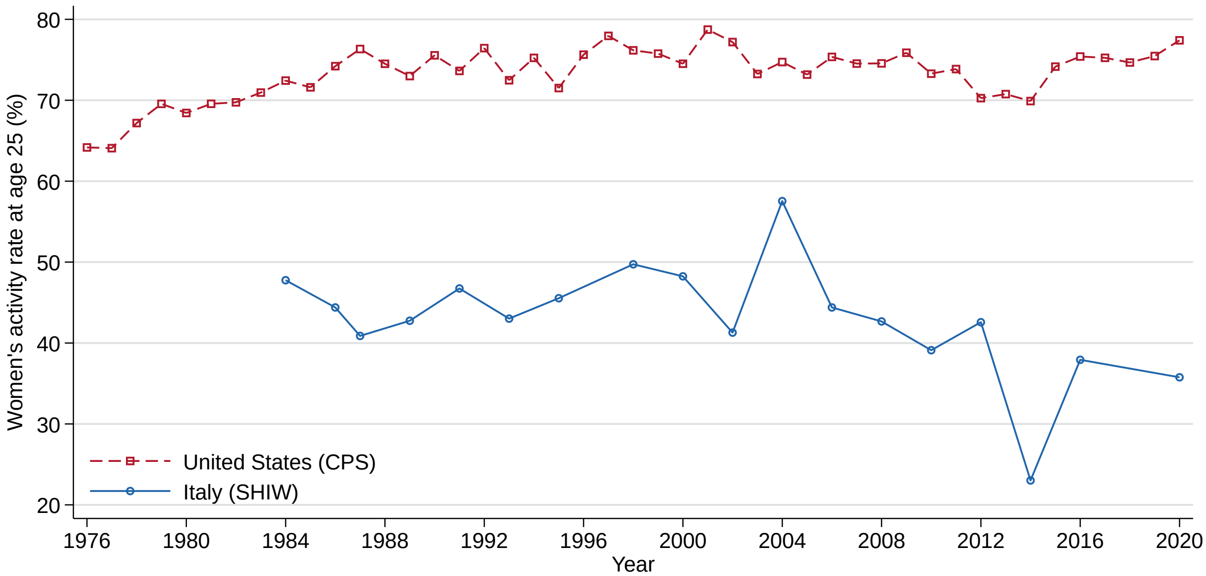
Notes: Panels A, B, and D show the trend in the mean gender gap in log yearly earnings across different birth cohorts in the United States, Italy, and the United Kingdom, respectively. Panel C shows the cross-cohort trend in the mean gender gap in log weekly earnings in Canada. The red triangles depict the trend in the mean gender pay gap across all cohorts active in the labor market in each year. This analysis includes only workers aged 50 or younger to limit the influence of cross-cohort changes in the selection into retirement. In each year, the data encompass information about all workers who were between 25 and 50 years old, had worked in the private sector for at least 24 weeks (not available in the UK), had earned strictly positive earnings, and (only in Italy) had not retired by December 31.
Source for Italy: UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Source for the United States.* Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS. *Source for LIS data:* Luxembourg Income Study (LIS) Database, last accessed on 06/01/2023 at https://www.lisdatacenter.org/.

Figure A6: More Results on Between-Cohort Decline in Gender Pay Gap



Notes: Panels A, B and D show the change in the total gender gap and its between-cohort component in the United States, Italy, and the United Kingdom, respectively for log yearly earnings. Panel C shows the change in the total gender gap and its between-cohort component in Canada for log weekly earnings. To compute the between-cohort component, we assign to each cohort (defined as a combination of birth year and gender) its mean log (yearly or weekly) earnings at labor-market entry in each year (Equation (2)). In the baseline analysis, entry in the labor market corresponds with the year in which workers in each cohort were 25 years old. We assign to cohorts who were older than 25 in the first sample year their mean weekly earnings in the first sample year. In each year, the data encompass information about all workers who were between 25 and 64 years old, had worked in the private sector for at least 24 weeks (not available in the UK), had earned strictly positive earnings.
Source for LIS data: Luxembourg Income Study (LIS) Database, last accessed on 06/01/2023 at https://www.lisdatacenter.org/.

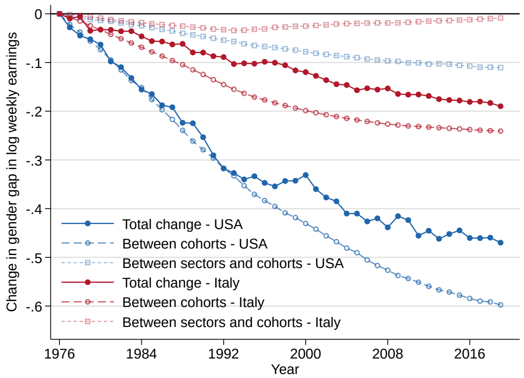
Figure A7: Participation Rates at Age 25



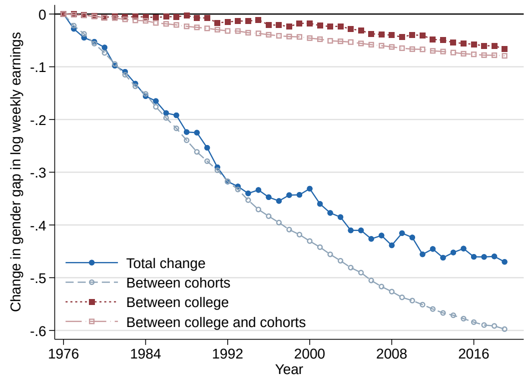
Notes: The figure shows participation rates of women at age 25 in the United States and Italy. *Source for Italy*: Survey of Household Income and Wealth, Bank of Italy. *Source for the United States*: Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Figure A8: Between-Cohort Decline in Gender Pay Gap—Controlling for Compositional Changes

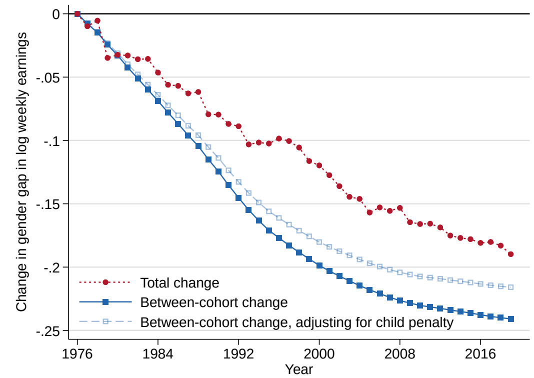
Panel A: Between cohorts and sector (log weekly earnings)



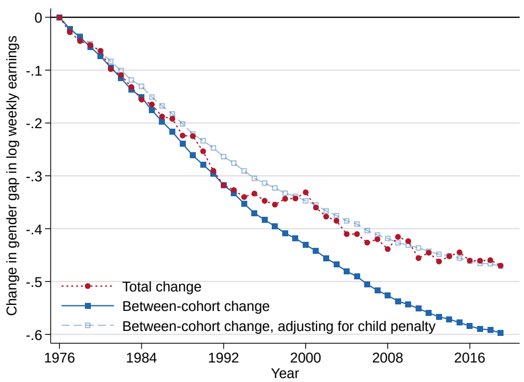
Panel B: Between cohorts and college (USA, log weekly earnings)



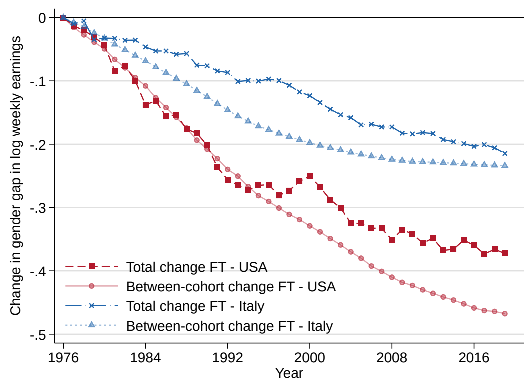
Panel C: Adjusting for child penalty (Italy, log weekly earnings)



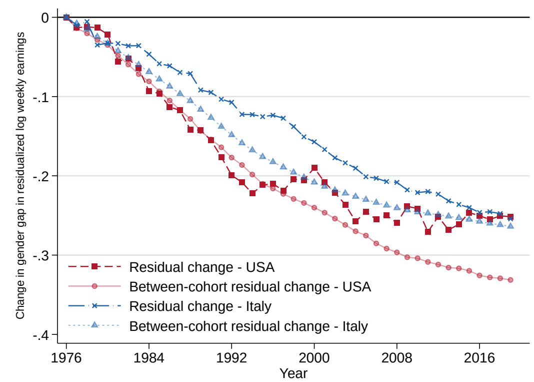
Panel D: Adjusting for child penalty (USA, log weekly earnings)



Panel E: Full-time workers (log weekly earnings)

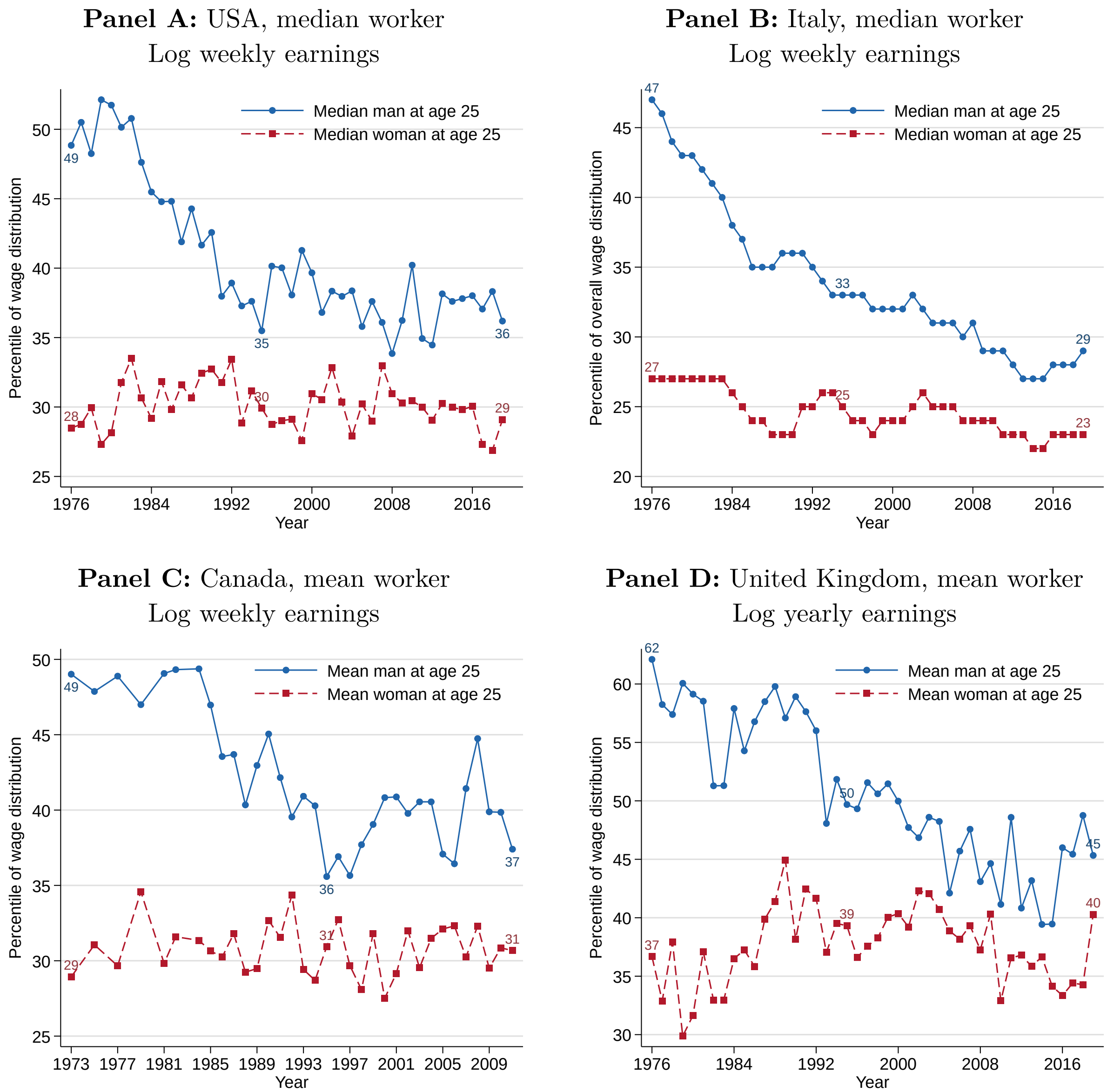


Panel F: Residual earnings (log weekly earnings)



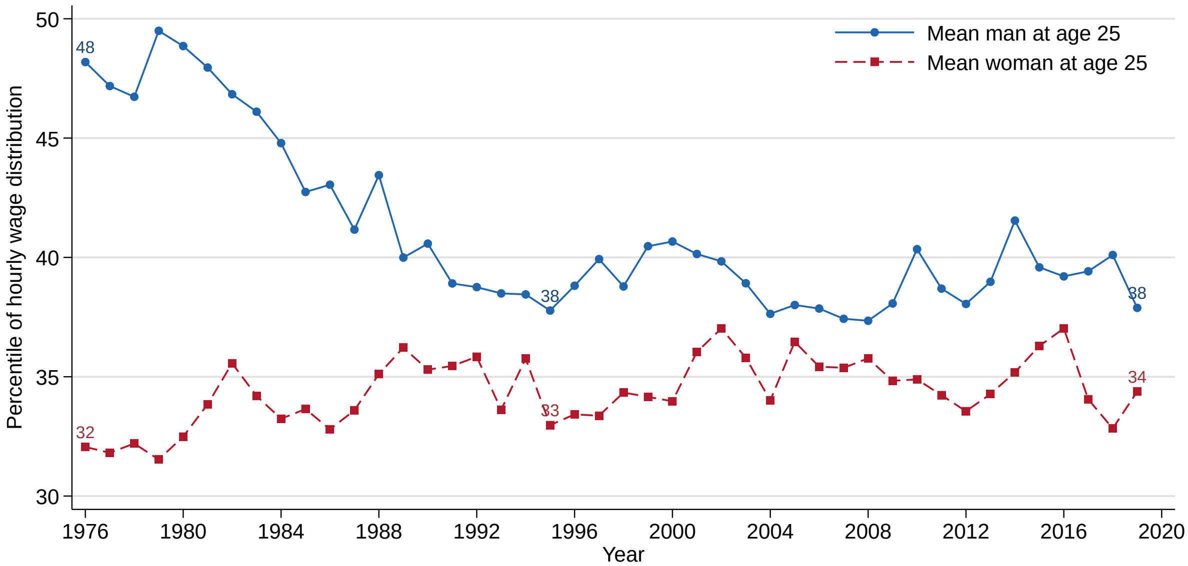
Notes: Panel A decomposes both the total change in gender pay gap and just its between-cohort component between and within sector (1- and 2-digit NACE Rev. 2 in the United States and Italy, respectively. Panel B decomposes both the total change in gender pay gap and just its between-cohort component between and within college graduation. Panels C and D adjust the between-cohort component of the decrease in the gender gap for the child penalty borne by women in the labor market (Kleven, Landais, and Søgaard, 2019). Panel E replicates the main decomposition with only full-time workers. Panel F uses residualized wages after regressing log weekly earnings, separately for each year and country, on a dummy for part-time workers, a dummy for temporary workers (only Italy), a dummy for domestic-born workers (only Italy), dummies for race and ethnicity (only USA), and a dummy for workers with children (only USA). In each year, the data encompass information about all workers who were between 25 and 64 years old, had worked in the private sector for at least 24 weeks, had earned strictly positive earnings, and (only in Italy) had not retired by December 31. *Source for Italy*: UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Source for the United States*: Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Figure A9: More Results on Positions at Labor-Market Entry



Notes: Panels A and B show the median earning percentile of men and women at age 25 in the United States and Italy, respectively. Panels C and D show the average earning percentile of men and women at age 25 in Canada and the United Kingdom, respectively. In each year, the data encompass information about all workers who were 25 years old, had worked in the private sector for at least 24 weeks (not available in the UK), had earned strictly positive earnings, and (only in Italy) had not retired by December 31.
Source for Italy: UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Source for the United States*. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS. *Source for LIS data:* Luxembourg Income Study (LIS) Database, last accessed on 06/01/2023 at https://www.lisdatacenter.org/.

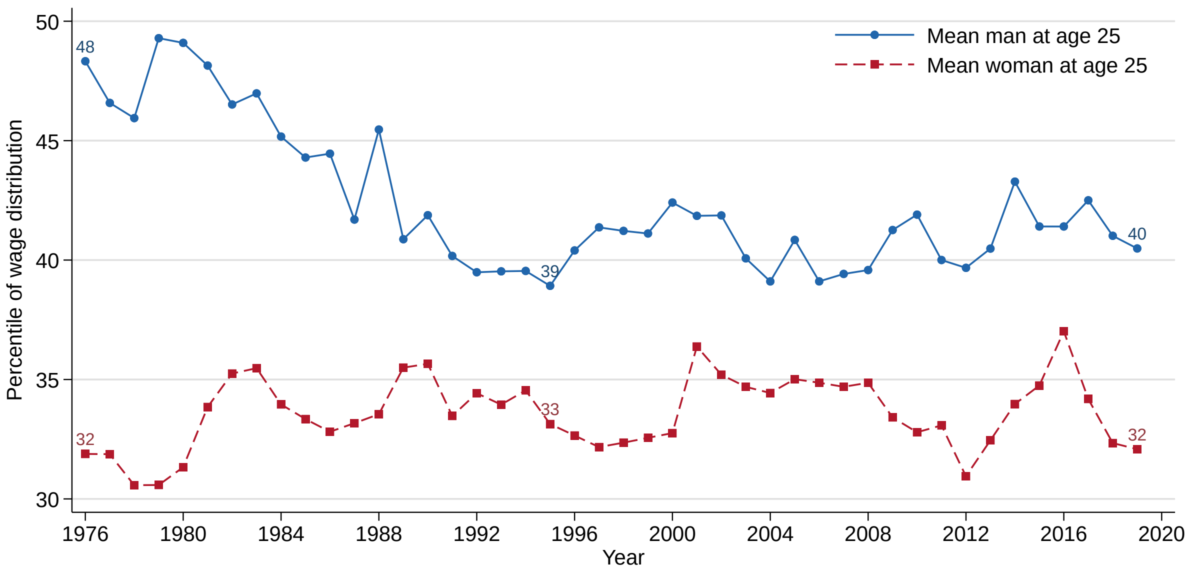
Figure A10: Positions in Distribution of Hourly Earnings (US)



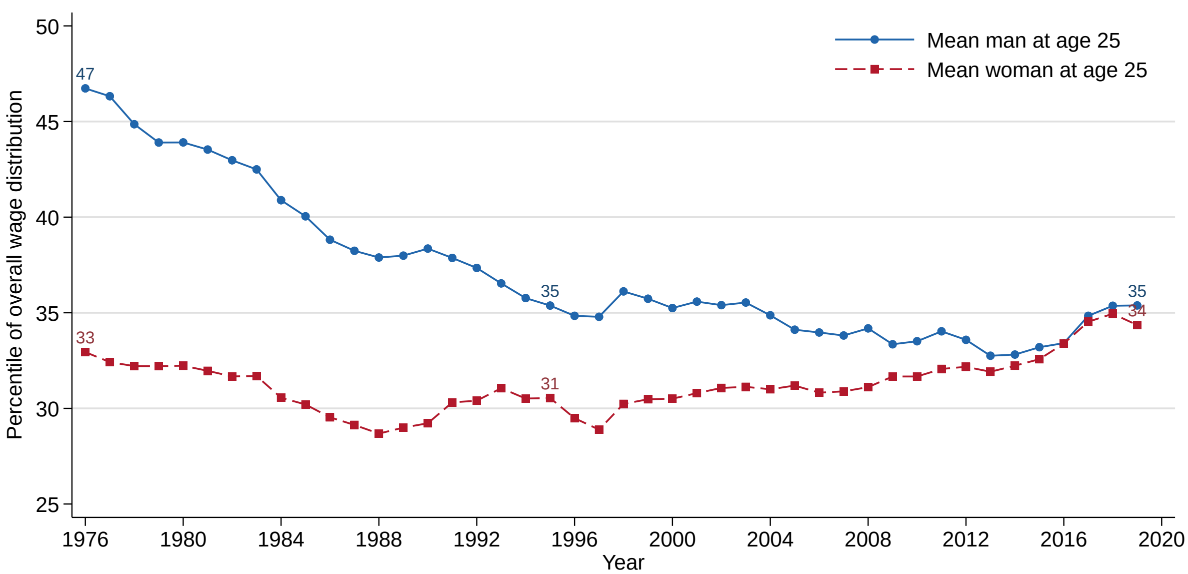
Notes: The figure shows the average percentile of men and women at age 25 in the distribution of hourly earnings in the United States. In each year, the data encompass information about all workers who were between 25 and 64 years old, had worked in the private sector for at least 24 weeks, had earned strictly positive earnings.
Source for the United States. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Figure A11: Positions in Distribution of Residualized Weekly Earnings

Panel A: USA

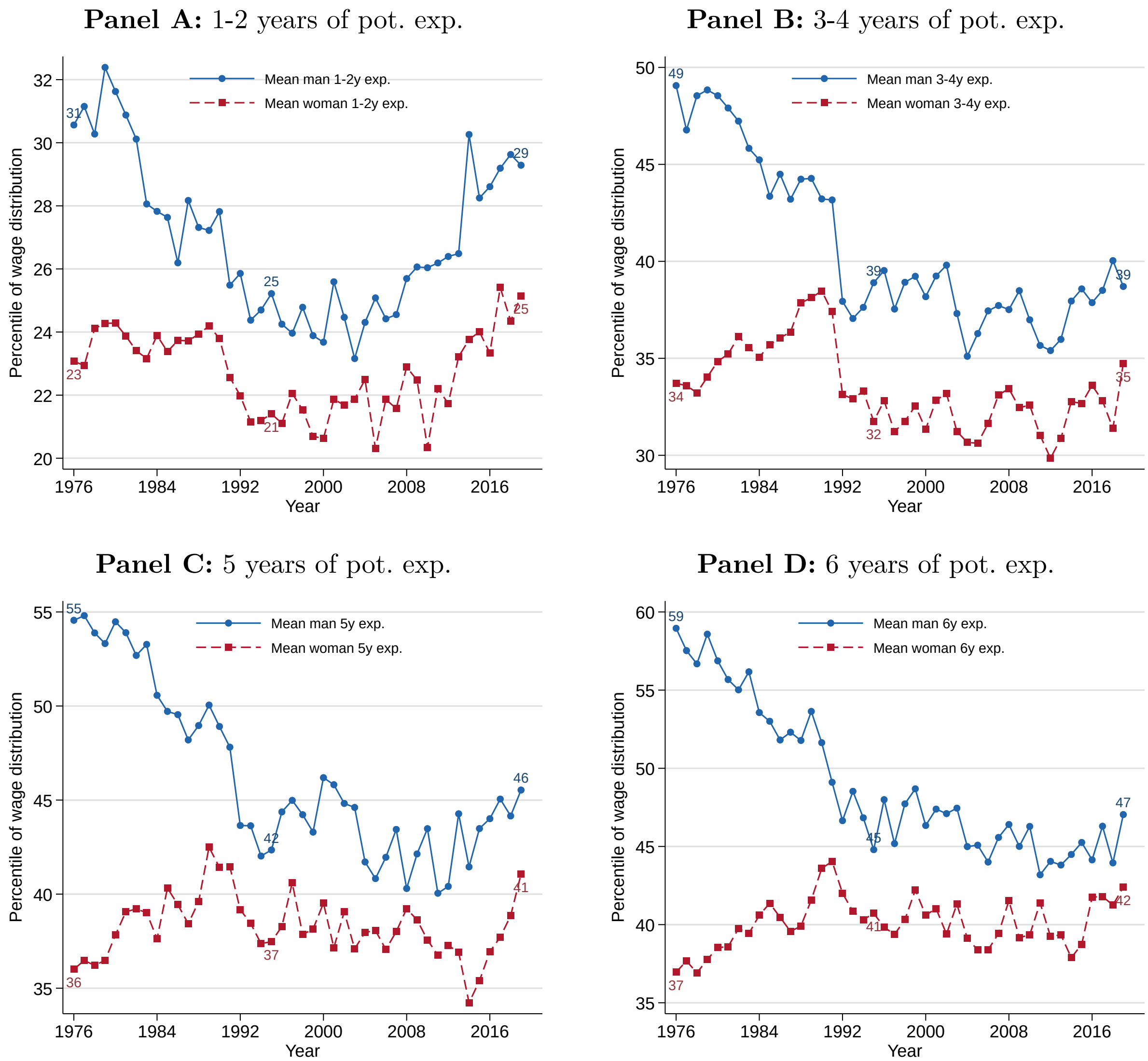


Panel B: Italy



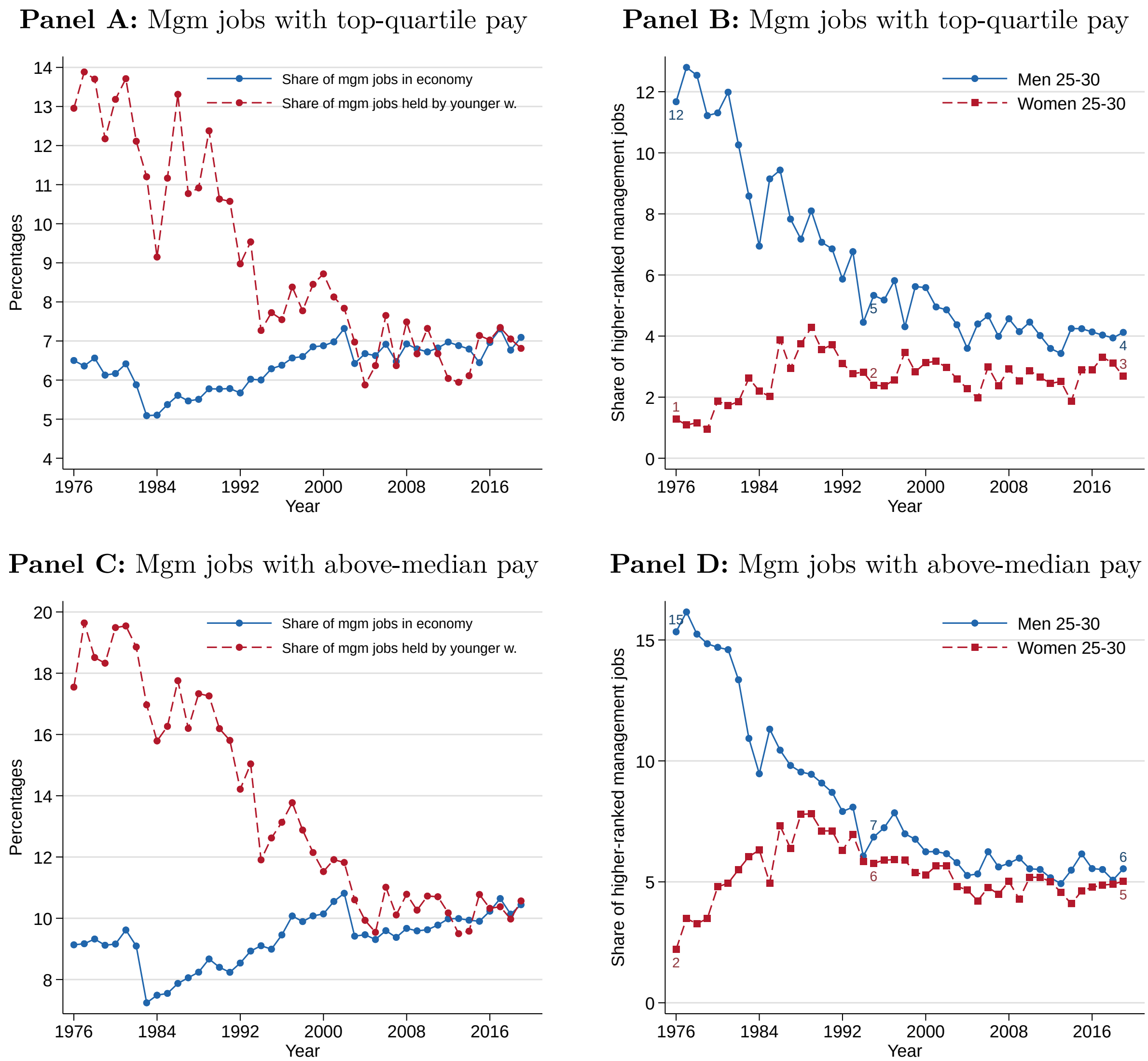
Notes: Panels A and B show the average earning percentile of men and women at age 25 in the distribution of residualized weekly earnings in the United States and Italy, respectively. In each year, the data encompass information about all workers who were between 25 and 64 years old, had worked in the private sector for at least 24 weeks, had earned strictly positive earnings, and (only in Italy) had not retired by December 31. Residualized earnings are computed by regressing log weekly earnings, separately for each year and country, on a dummy for part-time workers, a dummy for temporary workers (only Italy), a dummy for domestic-born workers (only Italy), dummies for race and ethnicity (only USA), and a dummy for workers with children (only USA).
Source for Italy: UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Source for the United States*. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Figure A12: Positions in Wage Distribution Based on Potential Experience



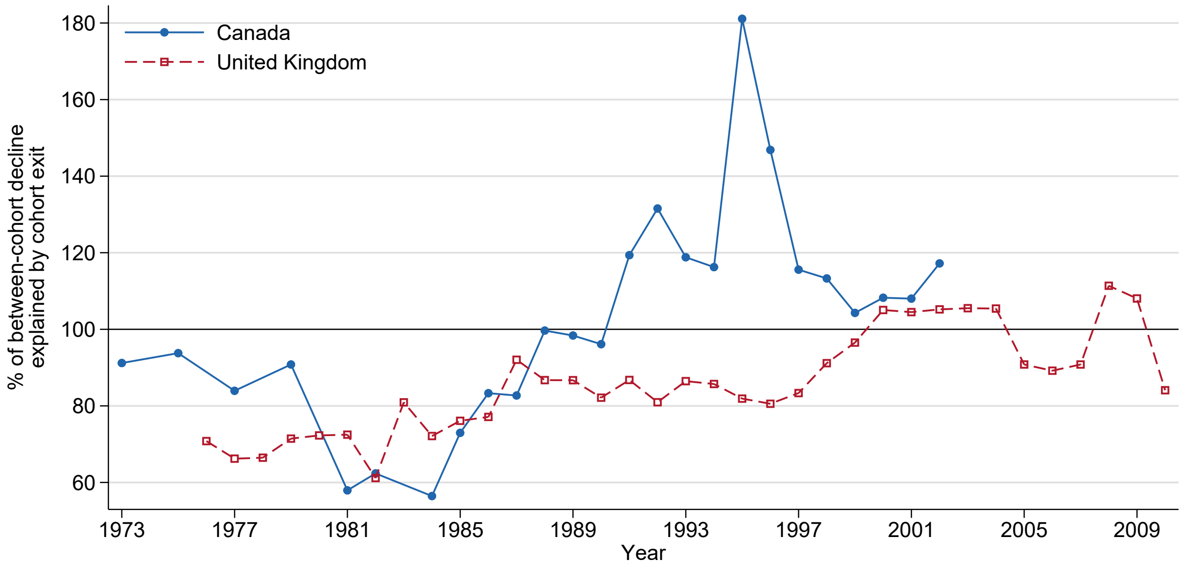
Notes: Panels A to D show the average earning percentile of men and women at different years of potential experience in the United States. Potential experience is calculated as age − years of education − 6. In each year, the data encompass information about all workers who were between 25 and 64 years old, had worked in the private sector for at least 24 weeks, had earned strictly positive earnings.
Source for the United States. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Figure A13: Higher-Ranked Managerial Positions



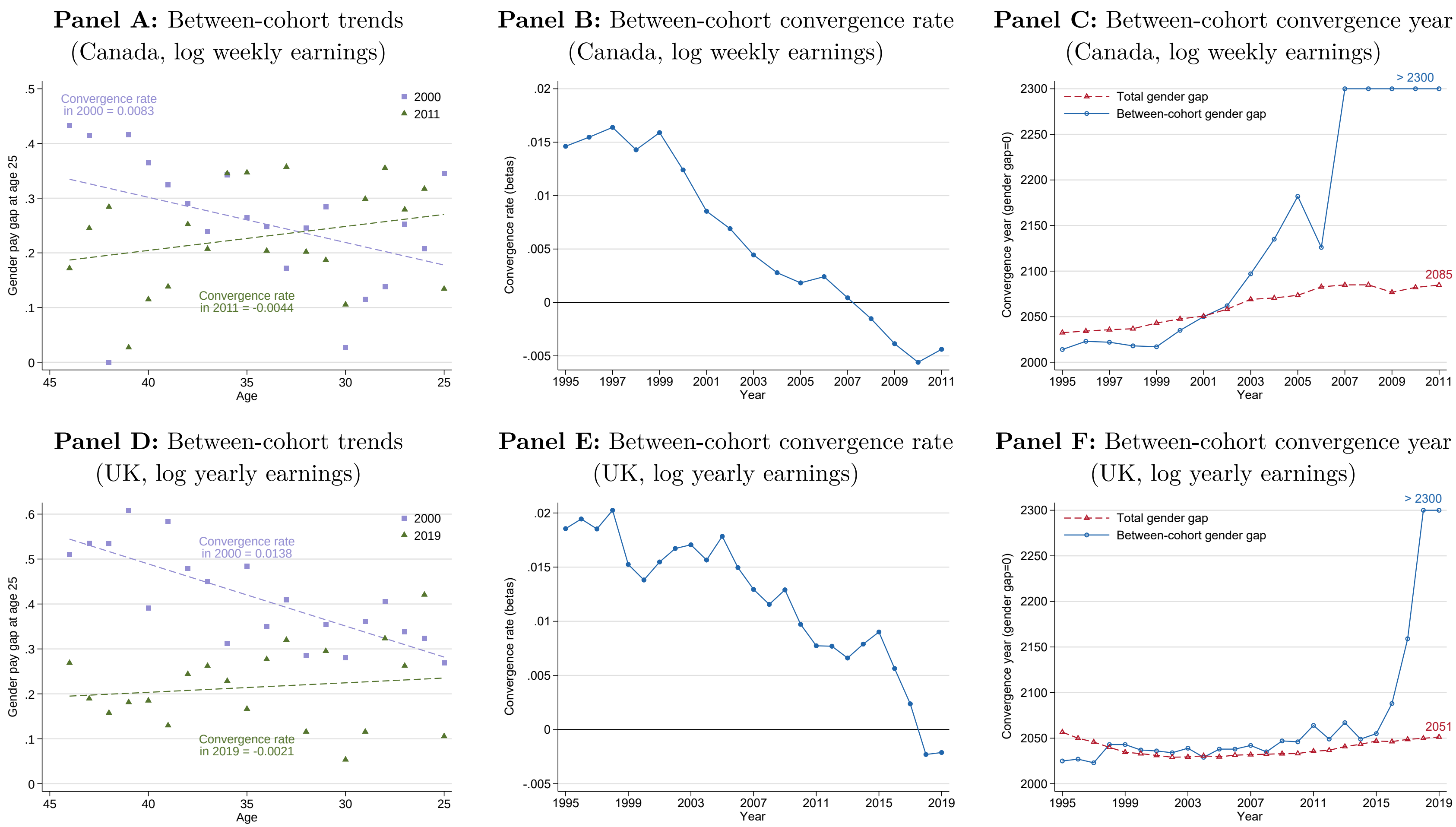
Notes: In Panels A and B, we define as higher-ranked managerial jobs all occupations with (i) 2-digit Standard Occupational Classification (SOC) code 11 (management occupations) and (ii) annual earnings in the top quartile of the year-specific distribution of annual earnings. Based on this categorization, Panel A shows the share of high-ranked managerial jobs out of all jobs in the economy and the share of high-ranked managerial jobs held by workers between 25 and 30 years old. Panel B shows the share of high-ranked managerial jobs held by men and women between 25 and 30 years old. In Panels C and D, we define as higher-ranked managerial jobs all occupations with (i) 2-digit Standard Occupational Classification (SOC) code 11 (management occupations) and (ii) annual earnings above the median of the year-specific distribution of annual earnings. Based on this categorization, Panels C and D replicate the same results included in Panels A and B, respectively.
Source for the United States. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Figure A14: No Convergence in Entry Earnings in Other High-Income Countries



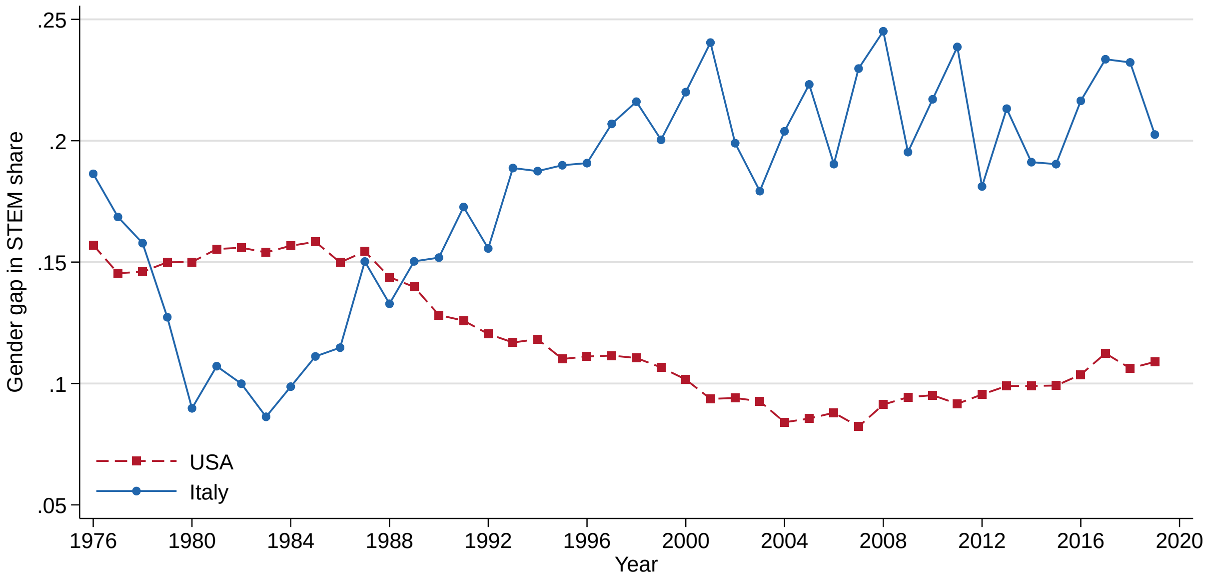
Notes: "Convergence at exit" computes what would have happened to the gender pay gap if the cross-cohort convergence in the earnings of men and women at age 25 had stopped in different benchmark years for Canada and the United Kingdom, respectively. The figure shows the ratio between the change (last year - first year) in the wage gap predicted by this new counterfactual scenario and the change in the gender gap predicted by the total between-cohort component from Equation (2) when the benchmark year t_b moves from the first sample year to 2002 for Canada and to 2010 for the UK. In each year, the data encompass information about all workers who were between 25 and 64 years old, had worked in the private sector for at least 24 weeks (not available in the UK), had earned strictly positive earnings. *Source for LIS data:* Luxembourg Income Study (LIS) Database, last accessed on 06/01/2023 at https://www.lisdatacenter.org/.

Figure A15: Between-Cohort Convergence in Other High-Income Countries



Notes: In each year t, we compute the gender pay gap at an early career stage for workers in age group a using their (weekly for Canada and yearly for the UK) earnings at age 25. Then, we estimate the linear relationship between the mean gender gap at labor-market entry and age (Equation (3)). Panels A and D show the best fit line in two different years for Canada and the United Kingdom, respectively. Panels B and E show the coefficients of age (β_t) for each year t. Panels C and F show the first cohort with gap at most equal to zero at age 25 predicted by Equation (3) for each year between 1976 and 2019 (1973 and 2011 for Canada). In each year, the data encompass information about all workers who were between 25 and 64 years old, had worked in the private sector for at least 24 weeks (not available in the UK), had earned strictly positive earnings. *Source for LIS data:* Luxembourg Income Study (LIS) Database, last accessed on 06/01/2023 at https://www.lisdatacenter.org/.

Figure A16: Gender Gap in the Share of STEM Graduates



Notes: The figure shows the gender gap (men − women) in the share of graduates in STEM subjects at age 25 in the United States and Italy. STEM subject areas are: Natural Sciences, Physics, Mathematics, Statistics, Computer Science, Information Engineering, Industrial Engineering, Architecture and Civil Engineering. *Source for Italy:* Quarterly Labour Force Survey, Istituto Nazionale di Statistica (ISTAT). *Source for the United States*: Integrated Public Use Microdata Series, American Community Survey. Minneapolis, MN: IPUMS.

Table A1: Characteristics of Data Sources

	# available years	# observations	# workers	# firms	Yearly earnings	Weekly earnings	Hourly earnings	Restrict working weeks
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Current Population Survey								
United States (1976-2019)	44	2,053,131	-	-	Yes	Yes	Yes	Yes
Panel B: Social Security data								
Italy (1976-2019)	44	373,117,856	32,112,786	5,174,323	Yes	Yes	No	Yes
Panel C: Survey data from the Luxembourg Income Study (LIS) Database								
Canada (1973-2019)	42	1,078,555	-	-	Yes	Yes	No	Yes
United Kingdom (1976-2019)	44	565,117	-	-	Yes	No	No	No

Source for Italy: Database UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Source for United States:* Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS. *Sources for Canada:* Survey of Consumer Finances (1973-1995); Survey of Labour and Income Dynamics (1996-2011); Canadian Income Survey (2012 and later). After 2011, data from Canada include a coarser categorization of age that does not allow us to study the outcomes of younger workers at a specific age (for example, at age 25). Hence, while we can plot the aggregate gender pay gap until 2019, the rest of the analysis can be performed only between 1973 and 2011. *Sources for United Kingdom:* Family Expenditure Survey (1991 and earlier); Family Resources Survey (1994 and later).

B A Model of Career Spillovers

Solving the firm problem. The firm problem is

$$\max_{m_{y,t},f_{y,t},m_{y,b},f_{y,b}} AY\left(L_y, L_o\right) - \sum_{g\in\{m,f\}}\sum_{a\in\{y,o\}}\sum_{j\in\{t,b\}}\left(w^g_{a,j}g_{a,j}\right) - \frac{\kappa}{2}K^2 - \sum_{g\in\{m,f\}}\left(\frac{c_g}{2}g^2_{y,t}\right).$$

The FOCs are

$$\begin{aligned}
\{f_{y,t}\}: &\quad AY_{L_y}\theta_{y,t} - w^f_{y,t} - \kappa K - c_f f_{y,t} &= 0;\\
\{m_{y,t}\}: &\quad AY_{L_y}\theta_{y,t} - w^m_{y,t} - \kappa K - c_m m_{y,t} &= 0;\\
\{f_{y,b}\}: &\quad AY_{L_y}\theta_{y,b} - w^f_{y,b} &= 0;\\
\{m_{y,b}\}: &\quad AY_{L_y}\theta_{y,b} - w^m_{y,b} &= 0.
\end{aligned}$$

The last two first order conditions indicate that the marginal revenue products of labor of younger men and women in the bottom job b are the same, hence their wages in the bottom job are also the same:

$$w^f_{y,b} = w^m_{y,b} = w_{y,b} = AY_{L_y}\theta_{y,b}.$$

Given that the wage in the bottom job pays a constant wedge over the wage in the bottom job, it follows that the wages of younger men and women in the top job t are also the same:

$$w^f_{y,t} = w^m_{y,t} = w_{y,t} = \mu_y w_{y,b}.$$

From the two initial first order conditions, we can pin down the optimal employment of younger men and women in the top job:

$$\begin{aligned}
f^*_{y,t} &= \frac{AY_{L_y}\left(\theta_{y,t} - \mu_y\theta_{y,b}\right) - \kappa K}{c_f};\\
m^*_{y,t} &= \frac{AY_{L_y}\left(\theta_{y,t} - \mu_y\theta_{y,b}\right) - \kappa K}{c_m}.
\end{aligned}$$

Given that $c_f > c_m$, we can conclude that the optimal number of younger women in top jobs is lower than the optimal number of men in top jobs: $f_{t,y} < m_{t,y}$. Furthermore, in equilibrium, the firm keeps the ratio of younger women and men in the top job constant:

$$\frac{f_{y,t}}{m_{y,t}} = \frac{c_m}{c_f} = \delta_f < 1.$$

Next, we consider an increase in the number of older workers in top jobs from period -1. The bottom wage of younger workers (for both gender groups) responds as follows:

$$\frac{\partial w_{y,b}}{\partial l^{-1}_{o,t}} = A\theta_{y,b}\left(Y_{L_yL_y}\frac{\partial L_y}{\partial l^{-1}_{o,t}} + Y_{L_yL_o}\frac{\partial L_o}{\partial l^{-1}_{o,t}}\right).$$

The derivatives of the efficient units of younger and older labor with respect to $l_{o,t}^{-1}$ are:

$$\begin{aligned}\frac{\partial L_y}{\partial l_{o,t}^{-1}} &= \frac{\partial\left[\theta_{y,t}\left(K-l_{o,t}\right)+\theta_{y,b}\left(l_y-K+l_{o,t}\right)\right]}{\partial l_{o,t}^{-1}} = \left(\theta_{y,t}-\theta_{y,b}\right)\left(\frac{\partial K}{\partial l_{o,t}^{-1}}-\rho_t\right)\\ \frac{\partial L_o}{\partial l_{o,t}^{-1}} &= \frac{\partial\left[\theta_{o,t}\rho_t l_{o,t}^{-1}+\theta_{o,b}\rho_b l_{o,b}^{-1}\right]}{\partial l_{o,t}^{-1}} = \theta_{o,t}\rho_t.\end{aligned}$$

We can rewrite the change in the bottom wage as follows:

$$\begin{aligned}\frac{\partial w_{y,b}}{\partial l_{o,t}^{-1}} &= A\theta_{y,b}\left(Y_{L_yL_y}\frac{\partial L_y}{\partial l_{o,t}^{-1}}+Y_{L_yL_o}\frac{\partial L_o}{\partial l_{o,t}^{-1}}\right)\\ &= A\theta_{y,b}\left[Y_{L_yL_y}\left(\theta_{y,t}-\theta_{y,b}\right)\left(\frac{\partial K}{\partial l_{o,t}^{-1}}-\rho_t\right)+Y_{L_yL_o}\theta_{o,t}\rho_t\right].\end{aligned}$$

An increase in the supply of older workers causes negative career spillovers (or crowding out of younger workers from top spots) if $\frac{\partial l_{y,t}}{\partial l_{o,t}^{-1}} = \frac{\partial m_{y,t}}{\partial l_{o,t}^{-1}} + \frac{\partial f_{y,t}}{\partial l_{o,t}^{-1}} = \frac{\partial K}{\partial l_{o,t}^{-1}} - \rho_t < 0$, which implies that the total number of top jobs grows less that the number of older workers at the top. The key driver for the sign of this derivative is the change in the number of top jobs. We can write the number of top slots as follows:

$$K = \frac{AY_{L_y}\left(\theta_{y,t}-\mu_y\theta_{y,b}\right)-c_m m_{y,t}}{\kappa}.$$

Next, we write the derivative of K with respect to $l_{o,t}^{-1}$:

$$\frac{\partial K}{\partial l_{o,t}^{-1}} = \frac{A\left(\theta_{y,t}-\mu_y\theta_{y,b}\right)}{\kappa}\left[Y_{L_yL_y}\left(\theta_{y,t}-\theta_{y,b}\right)\left(\frac{\partial K}{\partial l_{o,t}^{-1}}-\rho_t\right)+Y_{L_yL_o}\theta_{o,t}\rho_t\right]-\frac{c_m}{\kappa}\frac{\partial m_{y,t}}{\partial l_{o,t}^{-1}}.$$

We use the fact that the ratio of younger men and women in top jobs is constant in equilibrium to rewrite the derivative of $m_{y,t}$ as a function of the derivative of K:

$$\begin{aligned}\text{[Step 1]}\quad \frac{\partial f_{y,t}}{\partial l_{o,t}^{-1}} &= \frac{c_m}{c_f}\frac{\partial m_{y,t}}{\partial l_{o,t}^{-1}}.\\ \text{[Step 2]}\quad \frac{\partial m_{y,t}}{\partial l_{o,t}^{-1}} &= \frac{\partial K}{\partial l_{o,t}^{-1}}-\frac{\partial f_{y,t}}{\partial l_{o,t}^{-1}}-\rho_t\\ &= \frac{\partial K}{\partial l_{o,t}^{-1}}-\frac{c_m}{c_f}\frac{\partial m_{y,t}}{\partial l_{o,t}^{-1}}-\rho_t\\ &= \frac{c_f}{c_f+c_m}\left(\frac{\partial K}{\partial l_{o,t}^{-1}}-\rho_t\right).\end{aligned}$$

Going back to the derivative of K, we can rewrite it as follows:

$$\frac{\partial K}{\partial l_{o,t}^{-1}} = \frac{1}{\kappa}\left\{A\left(\theta_{y,t}-\mu_y\theta_{y,b}\right)\left[Y_{L_yL_y}\left(\theta_{y,t}-\theta_{y,b}\right)\left(\frac{\partial K}{\partial l_{o,t}^{-1}}-\rho_t\right)+Y_{L_yL_o}\theta_{o,t}\rho_t\right]-\frac{c_m c_f}{c_f+c_m}\left(\frac{\partial K}{\partial l_{o,t}^{-1}}-\rho_t\right)\right\}.$$

We simplify the notation:

$$
\begin{aligned}
B &= A\left(\theta_{y,t}-\mu_y\theta_{y,b}\right)>0\\
D &= \left(\theta_{y,t}-\theta_{y,b}\right)>0\\
E &= \frac{c_m c_f}{c_f+c_m}>0.
\end{aligned}
$$

Then, we can further simplify the derivative of K as follows:

$$
\begin{aligned}
\frac{\partial K}{\partial l_{o,t}^{-1}} &= \frac{1}{\kappa}\left\{B\left[Y_{L_yL_y}D\left(\frac{\partial K}{\partial l_{o,t}^{-1}}-\rho_t\right)+Y_{L_yL_o}\theta_{o,t}\rho_t\right]-E\left(\frac{\partial K}{\partial l_{o,t}^{-1}}-\rho_t\right)\right\}\\
&= \underbrace{\frac{1}{\kappa-Y_{L_yL_y}BD-E}\left\{B\left[Y_{L_yL_o}\theta_{o,t}-Y_{L_yL_y}D\right]-E\right\}}_{\text{Negative career spillovers if }<1}\rho_t.
\end{aligned}
$$

If the term multiplying ρ_t is less than 1, an increase in the number of older workers in top jobs decreases the number of top slots available to younger workers. This scenario happens when the following condition holds:

$$
\begin{aligned}
1 &> \frac{1}{\kappa-Y_{L_yL_y}BD-E}\left\{B\left[Y_{L_yL_o}\theta_{o,t}-Y_{L_yL_y}D\right]-E\right\}\\
\kappa-Y_{L_yL_y}BD-E &> B\left[Y_{L_yL_o}\theta_{o,t}-Y_{L_yL_y}D\right]-E\\
\kappa &> \bar{\kappa}=BY_{L_yL_o}\theta_{o,t}=A\left(\theta_{y,t}-\mu_y\theta_{y,b}\right)Y_{L_yL_o}\theta_{o,t}>0.
\end{aligned}
$$

This inequality indicates that the cost parameter κ needs to be higher than the productivity gains that younger workers experience from their complementarity with older workers. The term on the right-hand side is greater than zero because $(\theta_{y,t}-\mu_y\theta_{y,b})>0$. The latter follows from the condition that guarantees a positive K (see the formula of the equilibrium K) and from the fact that $f_{y,t}=K-m_{y,t}-l_{t,o}$.

So, when $\kappa>\bar{\kappa}$, we can draw several conclusions. First, as already pointed out, there is crowding out of younger workers in top jobs: $\frac{\partial K}{\partial l_{o,t}^{-1}}-\rho_t<0$. Second, all younger workers become less likely to hold top jobs, but younger men lose more top spots than younger women because they are more likely to hold top jobs at baseline: $\left|\frac{\partial f_{t,y}}{\partial l_{o,t}^{-1}}\right|=\delta_f\left|\frac{\partial m_{t,y}}{\partial l_{o,t}^{-1}}\right|<\left|\frac{\partial m_{t,y}}{\partial l_{o,t}^{-1}}\right|$ because $\delta_f<1$. Third, a larger supply of older workers at the top raises the bottom wage of younger workers:

$$
\frac{\partial w_{y,b}}{\partial l_{o,t}^{-1}}=A\theta_{y,b}\left[Y_{L_yL_y}\left(\theta_{y,t}-\theta_{y,b}\right)\left(\frac{\partial K}{\partial l_{o,t}^{-1}}-\rho_t\right)+Y_{L_yL_o}\theta_{o,t}\rho_t\right]>0,
$$

because $Y_{L_yL_y}<0$, $\left(\frac{\partial K}{\partial l_{o,t}^{-1}}-\rho_t\right)<0$, and $Y_{L_yL_o}>0$.

Next, we address a change in the mean wage of younger men and women. The mean wages of younger men and women are:

$$
\begin{aligned}
\{f_y\}\quad \overline{w}_{y,f} &= \frac{f_{y,t}\cdot w_{y,t}^f+f_{y,b}\cdot w_{y,b}^f}{f_y};\\
\{m_y\}\quad \overline{w}_{y,m} &= \frac{m_{y,t}\cdot w_{y,t}^m+m_{y,b}\cdot w_{y,b}^m}{m_y},
\end{aligned}
$$

where $w^f_{y,b} = w^m_{y,b} = w_{y,b}$, $w^f_{y,t} = w^m_{y,t} = w_{y,t} = \mu_y w_{y,b}$, and $g_y = g_{y,t} + g_{y,b}$ for each $g \in \{m, f\}$. Therefore, we can rewrite them as follows:

$$\begin{aligned} \{f_y\} \quad \overline{w}_{y,f} = & \quad \frac{f_{y,t} \cdot \mu_y w_{y,b} + (f_y - f_{y,t}) \cdot w_{y,b}}{f_y} = \frac{1}{f_y} (\mu_y - 1) f_{y,t} w_{y,b} + w_{y,b}; \\ \{m_y\} \quad \overline{w}_{y,m} = & \quad \frac{1}{m_y} (\mu_y - 1) m_{y,t} w_{y,b} + w_{y,b}. \end{aligned}$$

Starting from younger men, we consider the change in the mean wage that stems from a marginal increase in the number of older workers who held top jobs in period -1 ($l^{-1}_{t,o}$). Under the empirically relevant scenario of $\kappa > \bar{\kappa}$, we find that:

$$\frac{\partial \bar{w}_{y,m}}{\partial l^{-1}_{o,t}} = \underbrace{\frac{1}{m_y} (\mu_y - 1) \frac{\partial m_{y,t}}{\partial l^{-1}_{o,t}} w_{y,b}}_{\text{career spillovers} <0} + \underbrace{\left(\frac{1}{m_y} (\mu_y - 1) m_{y,t} + 1 \right) \frac{\partial w_{y,b}}{\partial l^{-1}_{o,t}}}_{\text{wage level} >0}.$$

The first component is negative because $\frac{\partial m_{y,t}}{\partial l^{-1}_{o,t}} < 0$ and $\mu_y - 1 > 0$. In contrast, the second component is positive because $\frac{\partial w_{b,y}}{\partial l^{-1}_{t,o}} > 0$. Due to the complementarity between the two age groups, a larger number of older workers makes younger men more productive, contributing to raise their mean wage.

The same decomposition applies to the mean wage of younger women:

$$\frac{\partial \bar{w}_{y,f}}{\partial l^{-1}_{o,t}} = \frac{1}{f_y} (\mu_y - 1) \frac{\partial f_{y,t}}{\partial l^{-1}_{o,t}} w_{y,b} + \left(\frac{1}{f_y} (\mu_y - 1) f_{y,t} + 1 \right) \frac{\partial w_{y,b}}{\partial l^{-1}_{o,t}}.$$

Next, we compare the magnitude of the negative career spillovers between younger men and women. As long as the share of women in the top job is smaller than the equivalent share of men, the career spillovers have a more negative effect on the mean wage of younger men:

$$\begin{aligned} \frac{1}{m_y} (\mu_y - 1) \frac{\partial m_{y,t}}{\partial l^{-1}_{o,t}} w_{y,b} & < \frac{1}{f_y} (\mu_y - 1) \frac{\partial f_{y,t}}{\partial l^{-1}_{o,t}} w_{y,b} \\ \frac{1}{m_y} (\mu_y - 1) \frac{\partial m_{y,t}}{\partial l^{-1}_{o,t}} w_{y,b} & < \frac{1}{f_y} (\mu_y - 1) \delta_f \frac{\partial m_{y,t}}{\partial l^{-1}_{o,t}} w_{y,b} \\ \frac{1}{m_y} & > \frac{1}{f_y} \delta_f \\ \frac{m_{y,t}}{m_y} \frac{1}{m_{y,t}} & > \frac{f_{y,t}}{f_y} \frac{1}{f_{y,t}} \delta_f \\ \tfrac{m_{y,t}}{m_y} / \tfrac{f_{y,t}}{f_y} & > \frac{m_{y,t}}{f_{y,t}} \delta_f \\ \tfrac{m_{y,t}}{m_y} / \tfrac{f_{y,t}}{f_y} & > 1. \end{aligned}$$

The main takeaway from this exercise is that when an increase in the supply of older workers decreases the mean wages of younger men and women, the gender pay gap closes as long as men are more concentrated in top jobs at baseline.

Finally, we can show that an increase in the retention rate of older workers at the top and a decrease in the economy-wide level of economic growth generate similar consequences on the gender pay gap. Starting from an increase in the retention rate of older workers at the top, the bottom wages of younger men and women change as follows:

$$\frac{\partial w_{y,b}}{\partial \rho_t} = A\theta_{y,b}\left[Y_{L_yL_y}\left(\theta_{y,t}-\theta_{y,b}\right)\left(\frac{\partial K}{\partial \rho_t}-l_{o,t}^{-1}\right)+Y_{L_yL_o}\theta_{o,t}l_{o,t}^{-1}\right].$$

We simplify the notation:

$$\begin{aligned} B &= A\left(\theta_{y,t}-\mu_y\theta_{y,b}\right)>0 \\ D &= \left(\theta_{y,t}-\theta_{y,b}\right)>0 \\ E &= \frac{c_mc_f}{c_f+c_m}>0. \end{aligned}$$

Then, we can rewrite the derivative of K as follows:

$$\begin{aligned} \frac{\partial K}{\partial \rho_t} &= \frac{1}{\kappa}\left\{B\left[Y_{L_yL_y}D\left(\frac{\partial K}{\partial l_{o,t}^{-1}}-l_{o,t}^{-1}\right)+Y_{L_yL_o}\theta_{o,t}l_{o,t}^{-1}\right]-E\left(\frac{\partial K}{\partial l_{o,t}^{-1}}-l_{o,t}^{-1}\right)\right\} \\ &= \underbrace{\frac{1}{\kappa-Y_{L_yL_y}BD-E}\left\{B\left[Y_{L_yL_o}\theta_{o,t}-Y_{L_yL_y}D\right]-E\right\}}_{\text{Negative career spillovers if }<1}l_{o,t}^{-1}. \end{aligned}$$

Therefore, there are negative career spillovers when $\kappa > \bar{\kappa} = A_f\left(\theta_{y,t}-\mu_y\theta_{y,b}\right)Y_{L_yL_o}\theta_{o,t}$, which is the same condition we found for an increase in the number of older workers at the top. All subsequent results follow through.

Furthermore, we can model κ as a decreasing function of the economic economic growth rate: $\kappa(g)$ with $\kappa'(g)<0$. Bianchi and Paradisi (2023) shows that the condition on the third derivatives of the production function that is needed for a larger κ to generate more negative career spillovers. A decline in g increases κ and, therefore, lowers the response of K to a larger supply of older workers at the top, leading to more crowding out of younger workers in top positions. This result does not change the fact that older men are more affected than younger women as long as they are more concentrated in top jobs at baseline.

Gender gap in productivity. Instead of assuming that the firm faces different costs for employing younger men and women in the top job, we could assume that (i) women are less productive than men and (ii) men and women are imperfect substitutes in the top job. The first assumption could still be micro-founded based on either taste-based or statistical discrimination against women.

In this context, the firm chooses the number of younger men and women in both the top and bottom job in order to maximize its profits,

$$\max_{m_{y,t},f_{y,t},m_{y,b},f_{y,b}} AY\left(L_y,L_o\right)-\sum_{g=m,f}\sum_{a=y,o}\sum_{j=t,b}\left(w_{a,j}^g g_{a,j}\right)-\frac{\kappa}{2}K^2.$$

The efficient units of younger and older labor become $L_a=\theta_{a,t}\left(m_{a,t}^{\frac{\sigma-1}{\sigma}}+\delta_f^{\frac{1}{\sigma}}f_{a,t}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}+\theta_{a,b}\left(m_{a,b}+f_{a,b}\right)$, where $a\in\{y,o\}$ and $\delta_f<1$.

The first order conditions for the employment level in the bottom jobs are unchanged. Therefore, in equilibrium, we still have that $w_{y,b}^f=w_{y,b}^m=w_{y,b}=AY_{L_y}\theta_{y,b}$. Moreover, $w_{y,t}^f=w_{y,t}^m=w_{y,t}=$

$\mu_y w_{y,b}$. In top jobs, we find that

$$\begin{aligned}
\{f_{y,t}\}: \quad & AY_{L_y}\theta_{y,t}\left(m_{y,t}^{\frac{\sigma-1}{\sigma}} + \delta_f^{\frac{1}{\sigma}} f_{y,t}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1}{\sigma-1}} \delta_f^{\frac{1}{\sigma}} f_{y,t}^{-\frac{1}{\sigma}} - w_{y,t}^f - \kappa K & = 0; \\
\{m_{y,t}\}: \quad & AY_{L_y}\theta_{y,t}\left(m_{y,t}^{\frac{\sigma-1}{\sigma}} + \delta_f^{\frac{1}{\sigma}} f_{y,t}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1}{\sigma-1}} m_{y,t}^{-\frac{1}{\sigma}} - w_{y,t}^m - \kappa K & = 0.
\end{aligned}$$

Combining the two first order conditions, we find that:

$$\begin{aligned}
\delta_f^{\frac{1}{\sigma}} f_{y,t}^{-\frac{1}{\sigma}} &= m_{y,t}^{-\frac{1}{\sigma}} \\
\frac{f_{y,t}}{m_{y,t}} &= \delta_f < 1.
\end{aligned}$$

In equilibrium, the firm wants to keep a fixed ratio of younger men and women in top jobs. Moreover, compared with younger women, younger men are more concentrated at the top. All the main results from the baseline specification follow.

Resource constraint. Instead of assuming that there is a given number of costly top jobs, we assume that the firm faces a constraint on the amount of resources that it can spend for top jobs. The firm problem becomes:

$$\begin{aligned}
\max_{m_{y,t}, f_{y,t}, m_{y,b}, f_{y,b}} AY\left(L_y, L_o\right) - & \sum_{g\in\{m,f\}} \sum_{a\in\{y,o\}} \sum_{j\in\{t,b\}} \left(w_{a,j}^g g_{a,j}\right) \\
& - \sum_{g\in\{m,f\}} \left(\frac{c_g}{2} g_{y,t}^2\right) - \kappa \cdot \left(l_{o,t} + m_{y,t} + f_{y,t}\right),
\end{aligned}$$

subject to the organizational resource constraint on top jobs ($\kappa \cdot (l_{o,t} + m_{y,t} + f_{y,t}) \leq K$). Here, K is the (exogenous) maximum amount of resources that can be spent on maintaining top jobs.

The first order conditions for bottom jobs are unchanged. Therefore, the wages of younger men and women in both bottom and top jobs are the same. The first order conditions for employment in top jobs are:

$$\begin{aligned}
\{f_{y,t}\} \quad & AY_{L_y}\theta_{y,t} - w_{y,t}^f - c_f f_{y,t} - (1+\lambda)\cdot\kappa & = 0; \\
\{m_{y,t}\} \quad & AY_{L_y}\theta_{y,t} - w_{y,t}^m - c_m m_{y,t} - (1+\lambda)\cdot\kappa & = 0.
\end{aligned}$$

In equilibrium, we conclude that:

$$\frac{f_{y,t}}{m_{y,t}} = \frac{c_m}{c_f} = \delta_f < 1.$$

All the main results from the baseline specification follow.

Endogenous labor force participation. In this extension, we drop the assumption of fixed labor supply and directly model the choice of participating in the labor market. Specifically, we assume that younger workers work whenever their expected wage is above their reservation wage. Workers draw reservation wages from the following cumulative distribution function:

$$P_g\left(w\right) = \left(\frac{w - w_{r,g}^{LB}}{w_{r,g}^{UB} - w_{r,g}^{LB}}\right)^{\eta_g},$$

for each $g \in \{m, f\}$, where $w_{r,g}^{LB}$ is the lower bound of reservation wages for workers of gender g, and $w_{r,g}^{UB}$ is the corresponding upper bound.

Specifically, we assume that younger workers choose whether to work or not for the representative firm based on the mean wage that the firm offers to younger workers ($\bar{w}_{y,g}$). They base their choice on the mean wage, rather than the actual wage they are going to receive when employed by the firm, because they are randomly assigned to either the top or bottom job after they join the company until the marginal product of labor equates to its cost. In this context, the number of employed younger men is equal to $P_m(\bar{w}_{y,m}) m_y$, while the number of employed younger women is $P_f(\bar{w}_{y,f}) f_y$.

The rest of the problem is unchanged. First, the firm receives the legacy older workers from period -1. Then, given a set of wages, the firm decides how many younger men and women to slot in the top and bottom jobs by equating the marginal revenue products of younger labor in the two positions to their marginal costs. In equilibrium, the market clears so that the demand for younger workers equates younger workers' supply: $g_y^d = P_g(\bar{w}_{y,g}) g_y$. Then, the firm allocates the younger workers randomly between the top and bottom jobs until its labor demands in the two positions are satisfied. Finally, the production is realized, and the firm pays all workers.

The labor-supply response to a change in the mean wage is:

$$\begin{aligned}
\frac{\partial P_g(\bar{w}_{y,g}) g_y}{\partial \bar{w}_{y,g}} &= \eta_g \frac{1}{w_{r,g}^{UB} - w_{r,g}^{LB}} \left(\frac{\bar{w}_{y,g} - w_{r,g}^{LB}}{w_{r,g}^{UB} - w_{r,g}^{LB}} \right)^{\eta_g - 1} g_y \\
&= \eta_g \frac{1}{w_{r,g}^{UB} - w_{r,g}^{LB}} P_g(\bar{w}_{y,g}) \left(\frac{\bar{w}_{y,g} - w_{r,g}^{LB}}{w_{r,g}^{UB} - w_{r,g}^{LB}} \right)^{-1} g_y \\
&= \eta_g \frac{P_g(\bar{w}_{y,g})}{\bar{w}_{y,g} - w_{r,g}^{LB}} g_y.
\end{aligned}$$

We assume that the lower-bound reservation wage is zero, which is akin to assuming that there is a positive share of workers with no fixed cost from participating in the labor market, to further simplify the derivative:

$$\frac{\partial P_g(\bar{w}_{y,g}) g_y}{\partial \bar{w}_{y,g}} = \eta_g \frac{P_g(\bar{w}_{y,g})}{\bar{w}_{y,g}} g_y.$$

The FOCs are unchanged. In equilibrium, wages of younger men and women are the same in both jobs. Younger men are more represented at the top. Moreover, the firm wants to keep a constant ratio $\delta_f < 1$ between younger women and men in top jobs.

Next, we are going to focus on the change in the mean wage of younger men and women when the number of older workers in the top job increases. The mean wage for younger workers of gender g can be written as follows:

$$\begin{aligned}
\overline{w}_{y,g} &= \frac{g_{y,t} \cdot w_{y,t}^{g} + g_{y,b} \cdot w_{y,b}^{g}}{P_g(\bar{w}_{y,g}) g_y} \\
&= \frac{g_{y,t} \cdot \mu_y w_{y,b} + (P_g(\bar{w}_{y,g}) g_y - g_{y,t}) \cdot w_{y,b}}{P_g(\bar{w}_{y,g}) g_y} \\
&= \frac{(\mu_y - 1) g_{y,t} \cdot w_{y,b}}{P_g(\bar{w}_{y,g}) g_y} + w_{y,b},
\end{aligned}$$

because $w_{y,b}^{f} = w_{y,b}^{m} = w_{y,b}$, $w_{y,t}^{f} = w_{y,t}^{m} = \mu_y w_{y,b}$, and $P_g(\bar{w}_{y,g}) g_y = g_{y,t} + g_{y,b}$ for each $g \in \{m, f\}$.

An increase in the number of older workers in top jobs changes the mean wage of younger workers of gender g as follows:

$$
\begin{aligned}
\frac{\partial \bar{w}_{y,g}}{\partial l_{o,t}^{-1}} &= \frac{\left(P_g\left(\bar{w}_{y,g}\right) g_y\right)\left[\left(\mu_y-1\right) \frac{\partial g_{y,t}}{\partial l_{o,t}^{-1}} w_{y,b}+\left(\mu_y-1\right) g_{y,t} \frac{\partial w_{y,b}}{\partial l_{o,t}^{-1}}\right]}{\left(P_g\left(\bar{w}_{y,g}\right) g_y\right)^2} \\
&- \frac{\frac{\eta_g P_g(\bar{w}_{y,g}) g_y}{\bar{w}_{y,g}} \frac{\partial \bar{w}_{y,g}}{\partial l_{o,t}^{-1}}\left(\mu_y-1\right) g_{t,y} w_{y,b}}{\left(P_g\left(\bar{w}_{y,g}\right) g_y\right)^2}+\frac{\partial w_{y,b}}{\partial l_{o,t}^{-1}} \\
&= \frac{\left[\left(\mu_y-1\right) \frac{\partial g_{y,t}}{\partial l_{o,t}^{-1}} w_{y,b}+\left(\mu_y-1\right) g_{y,t} \frac{\partial w_{y,b}}{\partial l_{o,t}^{-1}}\right]}{\left(P_g\left(\bar{w}_{y,g}\right) g_y\right)} \\
&- \frac{\frac{\eta_g}{\bar{w}_{y,g}} \frac{\partial \bar{w}_{y,g}}{\partial l_{o,t}^{-1}}\left(\mu_y-1\right) g_{y,t} w_{y,b}}{\left(P_g\left(\bar{w}_{y,g}\right) g_y\right)}+\frac{\partial w_{y,b}}{\partial l_{o,t}^{-1}}.
\end{aligned}
$$

Further simplifying and rewriting this formula leads to the usual separation of the total effect into a crowding-out component and a relative labor supply effect:

$$
\begin{aligned}
\frac{\partial \bar{w}_{y,g}}{\partial l_{t,o}} &= \left(\frac{P_g(\bar{w}_{y,g}) g_y}{P_g(\bar{w}_{y,g}) g_y+\frac{\eta_g}{\bar{w}_{y,g}}(\mu_y-1) g_{y,t} w_{y,b}}\right) \left\{\frac{\left(\mu_y-1\right) w_{y,b}}{\left(P_g\left(\bar{w}_{y,g}\right) g_y\right)} \frac{\partial g_{y.t}}{\partial l_{o,t}^{-1}}\right. \\
&\left.+\left(\frac{(\mu_y-1) g_{y,t}}{(P_g(\bar{w}_{y,g}) g_y)}+1\right) \frac{\partial w_{y,b}}{\partial l_{o,t}^{-1}}\right\}.
\end{aligned}
$$

The crowding-out component for $g \in \{m, f\}$ becomes:

$$
\left(\frac{1}{P_g\left(\bar{w}_{y,g}\right) g_y+\frac{\eta_g}{\bar{w}_{y,g}}\left(\mu_y-1\right) g_{y,t} w_{y,b}}\right)\left(\mu_y-1\right) w_{y,b} \frac{\partial g_{y,t}}{\partial l_{o,t}^{-1}}.
$$

When there is crowding out of younger workers in top jobs, this component is negative because $\frac{\partial g_{y,t}}{\partial l_{o,t}^{-1}}<0$. Moreover, the negative crowding-out effect on the mean wages of younger men is larger in magnitude than that for younger women when the following inequality holds:

$$
\begin{aligned}
\delta_f &< \frac{1+\frac{\eta_f}{\bar{w}_{y,f}}\left(\mu_y-1\right) \frac{f_{y,t}}{P_f\left(\bar{w}_{y,f}\right) f_y} w_{y,b}}{1+\frac{\eta_m}{\bar{w}_{y,m}}\left(\mu_y-1\right) \frac{m_{y,t}}{P_m(\bar{w}_{y,m}) m_y} w_{y,b}}, \\
1-\delta_f &> \left(\mu_y-1\right) w_{y,b}\left(\delta_f \frac{\eta_m}{\bar{w}_{y,m}} \frac{m_{y,t}}{P_m\left(\bar{w}_{y,m}\right) m_y}-\frac{\eta_f}{\bar{w}_{y,f}} \frac{f_{y,t}}{P_f\left(\bar{w}_{y,f}\right) f_y}\right) \\
1-\delta_f &> \left(\mu_y-1\right) w_{b,y} \frac{f_{y,t}}{P_f\left(\bar{w}_{y,f}\right) f_y}\left(\frac{\eta_m}{\bar{w}_{y,m}} \frac{P_f\left(\bar{w}_{y,f}\right) f_y}{P_m\left(\bar{w}_{y,m}\right) m_y}-\frac{\eta_f}{\bar{w}_{y,f}}\right).
\end{aligned}
$$

This inequality is more likely to hold when (i) the share of younger women in top jobs is lower, (ii) younger men are more likely to be employed than younger women, and (iii) the employment elasticity of younger women is larger than that of younger men. All these conditions are likely to find empirical support. For example, the idea that younger women's employment is more elastic than that of younger men is supported by several prior papers (for example, Bianchi, Gudmundsson, and Zoega (2001), Eissa and Hoynes (2004), and Manoli and Weber (2011)).

No exogenous rents. It is possible to drop the assumption that top-job wages pay an exogenous rent $\mu_y > 1$ over bottom-job wages. However, dropping this assumption requires further modifications to the baseline model.

The key takeaway from a model without exogenous rents is that the difference in the top wages of younger men and women reflects the difference in hiring costs between the two genders. However, since the firm can price discriminate between the two groups, it does not necessarily need to choose difference quantities of younger men and women in top jobs. In this scenario, there are multiple equilibria that satisfy the following condition:

$$w_{y,t}^m - w_{y,t}^f = c_f f_{y,t} - c_m m_{y,t}.$$

We can restore a positive gender gap in job allocations in favor of younger men by making the labor supply endogenous to the level of wages, as outlined in the section above. Together with a linear cost of hiring, this assumption ensures that a positive difference in the wages of men and women translates into a positive difference in the concentration of younger men and women in top jobs. We further explore this scenario when we discuss an extension with heterogeneous firms.

Introducing skills. When all top jobs are already occupied by older workers, an increase in the supply of older workers has the following effect on the mean wage of younger men:

$$\begin{aligned}
\frac{\partial \bar{w}_{y,m}}{\partial l_{o,t}^{-1}} &= \sum_{s\in\{h,l\}} \left(\frac{1}{m_y}\left(\mu_y - 1\right) m_{y,t,s} + 1\right) \frac{\partial w_{y,b,s}}{\partial l_{o,t}^{-1}} \\
&= \sum_{s\in\{h,l\}} \left(\frac{m_{y,s}}{m_y}\right) A\theta_{y,b,s} Y_{L_{y,s}L_o} \theta_{o,t,s}\rho_t,
\end{aligned}$$

where $\frac{1}{m_y}\left(\mu_y - 1\right) m_{y,t,s} + 1 = \frac{\mu m_{y,t,s} + m_{y,b,s}}{m_y} = \frac{m_{y,s}}{m_y}$ because all top jobs are occupied by older workers and, therefore, $m_{y,t,s} = 0$, and $m_{y,b,s} = m_{y,s}$. Moreover, $\frac{\partial w_{y,b,s}}{\partial l_{o,t}^{-1}} = A\theta_{y,b,s} \cdot \left(Y_{L_{y,s}L_{y,s}} \frac{\partial L_{y,s}}{\partial l_{t,o}^{-1}} + Y_{L_{y,s}L_o} \frac{\partial L_{o,s}}{\partial l_{t,o}^{-1}}\right) = A\theta_{y,b,s} Y_{L_{y,s}L_o} \theta_{o,t,s}\rho_t$ because $\frac{\partial L_{y,s}}{\partial l_{o,t}^{-1}} = 0$ for all skills/tasks s. Moreover, we assume that the retention rate of older workers in top jobs (ρ_t) is the same across all tasks.

We repeat the same computation for younger women:

$$\frac{\partial \bar{w}_{y,f}}{\partial l_{o,t}^{-1}} = \sum_{s\in\{h,l\}} \left(\frac{f_{y,s}}{f_y}\right) A\theta_{y,b,s} Y_{L_{y,s}L_o} \theta_{o,t,s}\rho_t.$$

For both younger men and women, the change in mean wages is positive. If we compare the two derivatives, we find that an increase in the number of older workers in top jobs shrinks the gender pay gap if the wage increase in larger for younger women:

$$\sum_{s\in\{h,l\}} \left(\frac{f_{y,s}}{f_y}\right) A\theta_{y,b,s} Y_{L_{y,s}L_o} \theta_{o,t,s}\rho_t > \sum_{s\in\{h,l\}} \left(\frac{m_{y,s}}{m_y}\right) A\theta_{y,b,s} Y_{L_{y,s}L_o} \theta_{o,t,s}\rho_t.$$

This inequality crucially depends on the distribution of younger men and women across different skills/tasks. We can further show this point by assuming that the cross-complementarity between younger and older workers is proportional to a task's marginal product: $Y_{L_{y,s}L_o} = Y_{L_{y,s}} \cdot C\left(\mathbf{L_y}, L_o\right)$, where $C\left(\cdot,\cdot\right)$ is the same across tasks.[31] In this case, the change in mean wages for younger women

[31] This is true if the production function is a Cobb-Douglas or a nested CES. Indeed, for the Cobb-Douglas

is larger than that for younger men if:

$$\frac{f_{y,h}}{f_y} Y_{L_{y,h}} \theta_{y,b,h} \theta_{o,t,h} + \frac{f_{y,l}}{f_y} Y_{L_{y,l}} \theta_{y,b,l} \theta_{o,t,l} > \frac{m_{y,h}}{m_y} Y_{L_{y,h}} \theta_{y,b,h} \theta_{o,t,h} + \frac{m_{y,l}}{f_y} Y_{L_{y,l}} \theta_{y,b,l} \theta_{o,t,l}$$

$$\left(\frac{f_{y,h}}{f_y} - \frac{m_{y,h}}{m_y} \right) \left(Y_{L_{y,h}} \theta_{y,b,h} \theta_{o,t,h} - Y_{L_{y,l}} \theta_{y,b,l} \theta_{o,t,l} \right) > 0.$$

Assuming that high-skill tasks have higher marginal product than low-skill tasks ($Y_{L_{y,h}} \theta_{y,b,h} \theta_{o,t,h} - Y_{L_{y,l}} \theta_{y,b,l} \theta_{o,t,l} > 0$), the inequality holds if women are overrepresented in high-skill tasks: $\frac{f_{y,h}}{f_y} > \frac{m_{y,h}}{m_y}$.

Heterogeneous firms. First, we replace the representative firm with N firms, but each firm is small and does not internalize the consequences of its actions on other firms. We further assume that $\rho_{j,n}$ increases with firm-level productivity A_n of firm n (Antwi and Phillips, 2013; Ruffini, 2022). Second, firms set wages for the bottom and top jobs, instead of taking them as given. Third, the ratio of top and bottom wages is not equal to a fixed rent and is not necessarily constant across firms. Fourth, we assume that $\kappa(K_n)$ is 0 up to a threshold level $\overline{K}_n$ and then is ∞ beyond $\overline{K}_n$. In practice, this is equivalent to a binding constraint on the number of top jobs: $l_{o,t,n} + m_{y,t,n} + f_{y,t,n} = \bar{K}_n$. Fifth, to make the computations more tractable, we assume that the cost of hiring younger workers in the top job is linear, instead of quadratic.

The timing of the model is as follows. First, each firm receives legacy older workers from period -1. Then, each firm posts wage offers for its bottom and top jobs, and each younger worker joins the firm and job that maximizes her utility. Finally, the production is realized, and the firm makes payments to all workers. In this scenario, the firm problem is to choose the wages of younger men and women in the top and bottom job that maximize its profits,

$$\max_{w^m_{y,t,n}, w^f_{y,t,n}, w^m_{y,b,n}, w^f_{y,b,n}} A_n Y \left(L_{y,n}, L_{o,n} \right) - \sum_{g \in \{m,f\}} \sum_{a \in \{y,o\}} \sum_{j \in \{t,b\}} \left(w^g_{a,j,n} g_{a,j,n} \right) - \sum_{g \in \{m,f\}} c_g g_{y,t,n},$$

subject to

$$l_{o,t,n} + m_{y,t,n} + f_{y,t,n} \leq \bar{K}_n.$$

On the worker side, we assume that a worker i of age group $a(i)$ and gender $g(i)$ derives the

---

we have $Y = L_o^{\alpha_o} \cdot \prod_{s \in \{h,l\}} L_{a,s}^{\alpha_{a,s}}$ and

$$Y_{L_{y,s} L_o} = \alpha_{y,s} \alpha_o \frac{Y}{L_{y,s} L_o} = Y_{L_{y,s}} \cdot \frac{\alpha_o}{L_o},$$

where $\frac{\alpha_o}{L_o}$ is the same across all skills. In the nested CES case $Y = \left[\beta_y \left(L_{y,h}^{\rho_y} + L_{y,l}^{\rho_y} \right)^{\frac{\rho}{\rho_y}} + \beta_o L_o^{\rho} \right]^{\frac{1}{\rho}}$ and

$$\begin{aligned} Y_{L_{y,s} L_o} &= \beta_y L_{y,s}^{\rho_y - 1} \left(L_{y,h}^{\rho_y} + L_{y,l}^{\rho_y} \right)^{\frac{\rho}{\rho_y} - 1} \left[\beta_y \left(L_{y,h}^{\rho_y} + L_{y,l}^{\rho_y} \right)^{\frac{\rho}{\rho_y}} + \beta_o L_o^{\rho} \right]^{\frac{1}{\rho} - 2} \left(\frac{1}{\rho} - 1 \right) \beta_o \rho L_o^{\rho - 1} \\ &= \underbrace{\beta_y L_{y,s}^{\rho_y - 1} \left(L_{y,h}^{\rho_y} + L_{y,l}^{\rho_y} \right)^{\frac{\rho}{\rho_y} - 1} \left[\beta_y \left(L_{y,h}^{\rho_y} + L_{y,l}^{\rho_y} \right)^{\frac{\rho}{\rho_y}} + \beta_o L_o^{\rho} \right]^{\frac{1}{\rho} - 1}}_{= Y_{L_{y,s}}} \underbrace{\left(\frac{1}{\rho} - 1 \right) \beta_o \rho L_o^{\rho - 1} \left[\beta_y \left(L_{y,h}^{\rho_y} + L_{y,l}^{\rho_y} \right)^{\frac{\rho}{\rho_y}} + \beta_o L_o^{\rho} \right]^{-1}}_{= C(\mathbf{L_y}, L_o)}, \end{aligned}$$

which proves the result.

following utility when working in job j and firm n:

$$U_{i,a,j,n} = \log\left(w^g_{a,j,n}\right) + \xi_{i,a,j,n},$$

where $\xi_{i,a,j,n}$ represents the idiosyncratic preference of worker i over job j of firm n. We assume that $\xi_{i,a,j,n}$, which is unobserved by firms, follows a type-1 extreme distribution with a parameter σ that captures the degree of substitutability across jobs and firms in workers' preferences. In this context, firm n faces the following labor supply function for its job j from younger workers of gender g:

$$g_{y,j,n} = \frac{\left(w^g_{y,j,n}\right)^{\frac{1}{\sigma}}}{\sum_{n=1}^{N}\sum_{j\in\{t,b\}}\left(w^g_{y,j,n}\right)^{\frac{1}{\sigma}}} g_y.$$

The marginal change in the wage of job j and firm n has the following effect on the labor supply of younger workers of gender g anticipated by the firm:

$$\begin{aligned}\frac{\partial g_{y,j,n}}{\partial w^g_{y,j,n}} &= \frac{1}{\sigma}\frac{\left(w^g_{y,j,n}\right)^{\frac{1}{\sigma}-1}}{\sum_{n=1}^{N}\sum_{j\in\{t,b\}}\left(w^g_{y,j,n}\right)^{\frac{1}{\sigma}}} g_y \\ &= \frac{1}{\sigma}\frac{g_{y,j,n}}{w^g_{y,j,n}}.\end{aligned}$$

In the first row, the denominator is left unchanged due to the assumption that firms do not anticipate the effect of a change in their wage on the other wages in the economy (Card et al., 2018; Lamadon, Mogstad, and Setzler, 2019).

Firms choose optimal wages as follows:

$$\left\{w^g_{y,b,n}\right\} : A_n Y_{L_{y,n}} \theta_{y,b,n} \frac{\partial g_{y,b,n}}{\partial w^g_{y,b,n}} - w^g_{y,b,n}\frac{\partial g_{y,b,n}}{\partial w^g_{y,b,n}} - g_{y,b,n} = 0$$

$$\left\{w^g_{y,t,n}\right\} : A_n Y_{L_{y,n}} \theta_{y,t,n} \frac{\partial g_{y,t,n}}{\partial w^g_{y,t,n}} - w^g_{y,t,n}\frac{\partial g_{y,t,n}}{\partial w^g_{y,t,n}} - g_{y,t,n} - c_g\frac{\partial g_{y,t,n}}{\partial w^g_{y,t,n}} - \lambda_n\frac{\partial g_{y,t,n}}{\partial w^g_{y,t,n}} = 0,$$

where λ_n is the multiplier on the constraint on the quantity of top jobs. Therefore, wages in the bottom job are equal to:

$$w^g_{y,b,n} = \underbrace{\frac{1}{1+\sigma}}_{\text{Markdown}} A_n Y_{L_{y,n}} \theta_{y,b,n},$$

so that bottom wages pay a markdown below the marginal product of labor, and $w^f_{y,b,n} = w^m_{y,b,n}$ for all firms.

Instead, wages in the top job are equal to:

$$w^g_{y,t,n} = \underbrace{\frac{1}{1+\sigma}}_{\text{Markdown}} \left(A_n Y_{L_{y,n}} \theta_{y,t,n} - c_g - \lambda_n\right).$$

Again, wages in the top job pay a markdown below the marginal product of labor in the top job. Moreover, as we discussed in the version with a representative firm, younger men are paid more than younger women in top jobs ($c_m < c_f$) and, therefore, are more likely to hold these positions. Finally, all firms have $\overline{K_n}$ top jobs.

Next, we consider the effect of an increase in the economy-wide number of older workers on wages and the number of top slots. Specifically, we study a marginal increase in $l_{o,t}^{-1}$, the total number of older workers in top jobs in period -1. We assume that this increase affects all firms proportionately to the share of the total number of older workers they employ in top jobs. So, in firm n, a marginal increase in $l_{o,t}^{-1}$ increases period-0 older workers in top jobs by $\rho_{t,n} l_{o,t,n}^{-1}/l_{o,t}^{-1}$.

Wages at the bottom change as follows (by the same amount for both genders):

$$\begin{aligned}
\frac{\partial w_{y,b,n}}{\partial l_{o,t}^{-1}} &= \frac{1}{1+\sigma} A_n \theta_{y,b,n} \left(Y_{L_{y,n}L_{o,n}} \theta_{o,t,n} \rho_{t,n} \frac{l_{o,t,n}^{-1}}{l_{o,t}^{-1}} \right. \\
&\quad \left. + Y_{L_{y,n}L_{y,n}} \left(\frac{1}{\sigma} \frac{(m_{y,b,n} + f_{y,b,n})}{w_{y,b,n}} \frac{\partial w_{y,b,n}}{\partial l_{o,t}^{-1}} \theta_{y,b,n} + \left(\frac{\partial m_{y,t,n}}{\partial l_{o,t}^{-1}} + \frac{\partial f_{y,t,n}}{\partial l_{o,t}^{-1}} \right) \theta_{y,t,n} \right) \right) \\
&= \frac{\frac{1}{1+\sigma} A_n \theta_{y,b,n} \left(Y_{L_{y,n}L_{o,n}} \theta_{o,t,n} - Y_{L_{y,n}L_{y,n}} \theta_{y,t,n} \right)}{1 - \frac{1}{1+\sigma} \frac{1}{\sigma} A_n \theta_{y,b,n}^2 Y_{L_{y,n}L_{y,n}} \frac{(m_{y,b,n} + f_{y,b,n})}{w_{y,b,n}}} \rho_{t,n} \frac{l_{o,t,n}^{-1}}{l_{o,t}^{-1}} > 0,
\end{aligned}$$

Wages at the top change as follows (again, by the same amount for both genders):

$$\frac{\partial w_{y,t,n}}{\partial l_{o,t}^{-1}} = \frac{\partial w_{y,b,n}}{\partial l_{o,t}^{-1}} \frac{\theta_{y,t,n}}{\theta_{y,b,n}} - \frac{1}{1+\sigma} \frac{\partial \lambda_n}{\partial l_{o,t}^{-1}}.$$

If the marginal productivities of younger workers in top and bottom jobs are not too dissimilar, an increase in the number of older workers increases wages in the top job less than wages in the bottom job. This is due to the fact that $\frac{\partial \lambda_n}{\partial l_{o,t}^{-1}} > 0$ because the shadow value of relaxing the quantity constraint increases with the number of older workers in the economy.

Next, we show that the restricted access to top jobs leads to larger limitations in the number of younger men who hold these positions. We start by considering the ratio of younger workers of gender g between the top and bottom job in firm n :

$$\frac{g_{y,t,n}}{g_{y,b,n}} = \left(\frac{w_{y,t,n}^g}{w_{y,b,n}^g} \right)^{\frac{1}{\sigma}}.$$

The derivative of this ratio with respect to the number of older workers is as follows:

$$\begin{aligned}
\frac{\partial \frac{g_{y,t,n}}{g_{y,b,n}}}{\partial l_{o,t}^{-1}} &= \frac{1}{\sigma} \left(\frac{w_{y,t,n}^g}{w_{y,b,n}^g} \right)^{\frac{1}{\sigma}-1} \frac{\frac{\partial w_{y,t,n}^g}{\partial l_{o,t}^{-1}} w_{y,b,n}^g - \frac{\partial w_{y,b,n}^g}{\partial l_{o,t}^{-1}} w_{y,t,n}^g}{\left(w_{y,b,j}^g \right)^2} \\
&= \frac{1}{\sigma} \frac{g_{y,t,n}}{g_{y,b,n}} \left(\frac{\partial w_{y,t,n}^g}{\partial l_{o,t}^{-1}} \frac{1}{w_{y,t,n}^g} - \frac{\partial w_{y,b,n}^g}{\partial l_{o,t}^{-1}} \frac{1}{w_{y,b,n}^g} \right).
\end{aligned}$$

When $\frac{\partial \lambda_n}{\partial l_{o,t}^{-1}}$ is between $\frac{\theta_{y,t,f} - \theta_{y,b,f}}{\theta_{y,b,f}} (1+\sigma) \frac{\partial w_{y,b,f}}{\partial l_{o,t}^{-1}}$ and $\frac{\theta_{y,t,f}}{\theta_{y,b,f}} (1+\sigma) \frac{\partial w_{y,b,f}}{\partial l_{o,t}^{-1}}$, the change in the top wage is positive but lower than the change in the bottom wage. Hence, the difference within the parentheses is negative and the employment of younger workers in the top job decreases relative to their employment in the bottom job.

Moreover, it is possible to compare $\partial \frac{m_{y,t,n}}{m_{y,b,n}} / \partial l_{o,t}^{-1}$ to $\partial \frac{f_{y,t,n}}{f_{y,b,n}} / \partial l_{o,t}^{-1}$. The wage change at the top $\frac{\partial w_{y,t,n}^g}{\partial l_{o,t}^{-1}}$

is the same for both genders, but younger men are paid more in top jobs. Therefore, $\frac{\partial w^m_{y,t,n}}{\partial l^{-1}_{o,t}} \frac{1}{w^m_{y,t,n}}$ is smaller than $\frac{\partial w^f_{y,t,n}}{\partial l^{-1}_{o,t}} \frac{1}{w^f_{y,t,n}}$. In contrast, the percentage change at the bottom is the same for both younger men and women. Therefore, the difference in parentheses is more negative for younger men. In addition, it is also true that $\frac{m_{y,t,n}}{m_{y,b.n}} > \frac{f_{y,t,n}}{f_{y,b,n}}$ at baseline because the wages in top jobs are higher for younger men (due to the higher cost of hiring women in top jobs). Therefore, it follows that, if top wages increase in response to a larger supply of older workers (due to positive relative supply effects), then the share of younger men in top jobs decline more than the share of women in the same positions.

In short, workforce aging can shrink the gender pay gap by blocking more younger men from reaching higher-paying positions. And the main reason why younger men are more affected is that they are more represented in top jobs at baseline, and, thus, have more to lose from congested access at the top of firms' hierarchies.

Finally, we consider whether more older workers can change the distribution of younger workers across different types of firms. Under the assumption that the retention rates are higher in higher-productivity firms, an increase in the total number of older workers in the economy leads to larger increases in the number of older workers in the top jobs of higher-productivity firms. The fact that younger workers (both men and women) experience more negative career spillovers in higher-productivity firms does not necessarily mean that they find employment in the bottom jobs of the same set of firms. Bianchi and Paradisi (2023) shows that an increase in the number of older workers induces younger workers to move toward firms with higher percentage increases in the bottom wages (the increase is again a result of complementarity with older workers):

$$\frac{\partial \frac{g_{y,b,n}}{g_{y,b,n'}}}{\partial l^{-1}_{o,t}} = \frac{1}{\sigma} \frac{g_{y,b,n}}{g_{y,b,n'}} \left(\frac{\partial w^g_{y,b,n}}{\partial l^{-1}_{o,t}} \frac{1}{w^g_{y,b,n}} - \frac{\partial w^g_{y,b,n'}}{\partial l^{-1}_{o,t}} \frac{1}{w^g_{y,b,n'}} \right).$$

The derivative is positive if $\frac{\partial w^g_{y,b,n}}{\partial l^{-1}_{o,t}} \frac{1}{w^g_{y,b,n}} > \frac{\partial w^g_{y,b,n'}}{\partial l^{-1}_{o,t}} \frac{1}{w^g_{y,b,n'}}$. Bianchi and Paradisi (2023) also derives under what conditions a larger supply of older workers in top jobs coincides with younger workers moving toward lower-productivity firms, a fact that finds empirical support.

However, to better understand the consequences of this migration toward lower-productivity firms for the gender pay gap, we need to compare the formula above for younger men and women. The two derivatives $\partial \frac{m_{y,b,n}}{m_{y,b,n'}} / \partial l^{-1}_{o,t}$ and $\partial \frac{f_{y,b,n}}{f_{y,b,n'}} / \partial l^{-1}_{o,t}$ are the same because all their components are the same between younger men and women. This result is due to the fact that younger men and women always earn the same wages in the bottom job when they are employed by the same firm. Thus, the number of younger workers who migrate across firms is entirely driven by how many younger workers are blocked from accessing top jobs. Given that (i) younger men are always more negatively impacted than younger women, and (ii) there is more congestion within higher-productivity firms, there are more younger men who move from higher-productivity to lower-productivity firms. In short, workforce aging can contribute to shrinking the gender pay gap by making the distributions of younger men and women across higher- and lower-paying firms more even.

C Adjusting Weekly Earnings by the Contribution of the Child Penalty

In this section, we describe the procedure to compute the adjustment of weekly earnings for the child penalty in both the United States and Italy. The methodology follows closely Kleven, Landais, and Søgaard (2019), Kleven (2023), and Casarico and Lattanzio (2023). It relies on estimating the effects of childbirth on women and men in the United States and women without children in Italy. Specifically, we are interested in the counterfactual weekly earnings of women, absent the child penalty, in each birth cohort and calendar year:[32]

$$\widetilde{w} = \frac{w}{(1 - p\pi)}, \tag{4}$$

where w are the average weekly earnings of women, p is the cohort- and year-specific child penalty, and π is the average fraction of mothers in each cohort-year pair. We describe in the following paragraphs how we calculate each of these components.

C.1 United States

For each mother i in our CPS sample with year-t weekly earnings w_{it}, our goal is to construct $\tilde{w}_{it}$, the counterfactual weekly wage we would expect if that person were not a mother:

$$\tilde{w}_{it} = \frac{w_{it}}{1 - p_{jc(i,t)}}, \tag{5}$$

where p_{jc} is the child penalty, which is specific to birth cohort c and time-since-childbirth j. We estimate the set of child penalties p_{jc} following Kleven (2023) very closely, with the key difference that we estimate the weekly wage child penalty conditional on working.

Data. Following Kleven (2023), we estimate the child penalties using repeated cross sections of the March CPS 1968–2020 and American Community Survey (ACS) 2000–2019. We follow his sample restrictions and further limit the sample to include workers employed for at least half of the weeks of the preceding year.

Empirical Approach. We follow the pseudo-event-study approach of Kleven (2023) in order to estimate child penalties using repeated cross sections. The initial step in the approach involves matching to generate a pseudo-panel of men and women before and after the birth of their first child (see Kleven (2023) for more details). We then estimate the following earnings equation, separately for men and women:

$$w_{it}^{g} = \boldsymbol{\alpha^{g} D_{it}^{Event}} + \boldsymbol{\beta^{g} D_{it}^{Age}} + \boldsymbol{\gamma^{g} D_{it}^{Year}} + \nu_{it}^{g}, \tag{6}$$

where w_{it}^{g} is the weekly earnings for (pseudo-)individual i of gender $g = w, m$ at event time t. On the right-hand side, boldface denotes vectors. The first term includes dummies for each event time t, omitting a base year before child birth. The event time coefficients α^{g} measures the impact of child birth on gender g in event year t, relative to the base year. The second and third terms include a full set of age and year dummies.

The estimated level effects are transformed into percentage effects by calculating:

$$P_{t}^{g} = \frac{\alpha_{t}^{g}}{E\left[\tilde{w}_{it}^{g}|t\right]}. \tag{7}$$

Here, $\tilde{w}_{it}^{g}$ is the predicted counterfactual weekly wage in the absence of children.

Finally, the child penalty p is defined as the average effect of having children on women relative to men over a specified event time horizon:

$$p = E\left[P_{t}^{m} - P_{t}^{w}|t \geq 0\right] - E\left[P_{t}^{m} - P_{t}^{w}|t < 0\right] \tag{8}$$

[32] To simplify notation, we ignore here time and cohort subscripts.

Results. Figure C1 presents the pooled event study for men and women surrounding the birth of their first child. The event study horizon displayed in this and ensuing graphs ranges from $t = -5$ years to $t = 10$ years. The average child penalty on weekly earnings conditional on working, across this event time spectrum, is calculated to be 17%. Without conditioning on employment, Kleven (2023) finds an annual earnings child penalty of 33%.

Crucially for our purposes—studying the gender gap dynamics over a long period—we allow the child penalties that feed into counterfactual wages (5) to vary over time. To this end, we estimate child penalties separately across nine birth cohort groups that span the women in our sample: birth cohorts 1930–44, 1945–49, 1950–54, 1955–59, 1960–64, 1965–69, 1970–74, 1975–1979, and 1980–95.[33] Figure C2 shows the estimated child penalties for each of these nine birth cohort groups. The child penalty on weekly earnings conditional on working fluctuates between 11%–23%. Even though the pattern is not monotonic, there is a decreasing trend of the child penalty size across birth cohorts.

The event-time estimates 0–10, one for each of the nine birth cohort groups, represent our adjustment factors p_{jc} in Equation (5). That is, we assign each of the mothers in our CPS analysis sample to one of 99 cells based on her birth year and the age of her eldest son.[34] This cell assignment determines the adjustment factor p_{jc} assigned to each mother and her resulting counterfactual weekly earnings net of the child penalty.

C.2 Italy

Defining the Sample to Estimate Child Penalties. In Italy, we first recover childbirth episodes from the "contribution archive," which reports the full history of workers' Social Security contributions from their first employment spell. This archive not only records actual contributions paid by employers but also imputed contributions related to leaves of absence, sick leaves, unemployment benefit receipt and, crucially, maternity leave. The latter allow us to identify childbirth episodes based on the first month of maternity leave, which has a mandatory duration of five months and can be taken one to two months before the expected childbirth and lasts until three to four months after. The contribution archive is only available for a sample of workers born after 1950. We, therefore, restrict the data to workers included in such sample and extrapolate our estimates of the child penalty to the full set of workers. We further restrict the sample to women who had their first child (their first maternity leave episode) between 25 and 45. As we follow them for at most 10 years after childbirth, our sample comprises women between 25 and 55 years old.

As there is no information on fathers, the first step is to recover a suitable control group of non-mothers. In our sample of 25 to 55-year-old women, we first focus on those born between 1950 (the first birth year available in the contribution archive) and 1974, who were not yet 45 years old by 1976 (the first year in our sample) and turned 45 by 2019. Among these women, we identify mothers from maternity leave take-up both during our observation period (1976-2019) and before, as we have workers' full Social Security contribution histories. Women born between 1950 and 1974 who do not have a child enter the group of never-mothers. Women born after 1974 are subject to right-censoring, as they were not yet 45 by the end of the observation period (2019) and might have had a child after. We solve this issue by assigning a birth probability to the truncated cohort. Specifically, we estimate a linear probability model in the non-truncated cohorts 1950-1974 by regressing a dummy taking value one for never-mothers on the following set of dummy controls:

[33] The first and last birth cohort groups span a larger number of birth years because these are the groups with fewer observations.

[34] We apply the 10-year child penalty to mothers whose eldest son is older than 10.

quartiles of the cohort-specific log weekly earnings distribution and province of residence.[35] We then assign all women in the truncated birth cohorts the predicted probability of giving birth based on the coefficients estimated in the linear probability model. We then sort women born after 1974 based on such predicted probability and, starting from the largest value, assign them to the control group up to the point in which the fraction of "predicted" never-mothers in the truncated cohort post-1974 equals the fraction of actual never-mothers in the non-truncated cohorts 1950-1974.[36] The final sample consists of three groups of women: actual mothers, actual never-mothers from non-truncated birth cohorts and predicted never-mothers from truncated birth cohorts. The latter two groups constitute the control group.

The second step is to assign a placebo year of birth to the control group of never-mothers. We do so by assigning a placebo age at childbirth to non-mothers, drawing from the actual distribution of age at childbirth for mothers. We distinguish again between actual and predicted never-mothers. For actual never-mothers, we assume that the distribution of age at childbirth $A_{c,q}$ follows a log-normal distribution within cells of birth cohort c and quartiles of log weekly earnings q, $A_{c,q} \sim \mathcal{LN}\left(\widehat{\mu}_{c,q}, \widehat{\sigma}_{c,q}\right)$, where mean $\widehat{\mu}_{c,q}$ and variance $\widehat{\sigma}_{c,q}$ are obtained from the actual within-cell distribution for mothers. We assign a random draw from this distribution to actual never-mothers. For predicted never-mothers, we use random draws from a distribution with same variance $\widehat{\sigma}_{c,q}$ but different mean $\widetilde{\mu}_{c,q}$, which is obtained by predicting age at childbirth from the estimation of a regression on a quadratic time trend for actual mothers to allow women born after 1974 to have their first child at an older age.[37]

Estimating Child Penalties by Year and Birth Cohort. Our goal is to estimate the child penalty in each year and birth cohort. To do so, we run the following event study, separately for mothers and non-mothers:

$$w^g_{ist} = \sum_y \sum_{j \neq -1} \alpha^g_{yj} \mathbf{I}[\mathbf{j} = \mathbf{t}] \mathbf{I}[y = s] + \sum_k \beta^g_k \mathbf{I}[k = \text{age}_{is}] + \sum_y \gamma^g_y \mathbf{I}[y = s] + \varepsilon^g_{ist},$$

where we interact event time dummies (year to childbirth) with year dummies in order to estimate the year-specific coefficients α^g_{yj} (which are equivalent to birth-cohort-specific coefficients, as birth cohort $c = s - t$). The equation also controls for age and year dummies. Figure C3 reports the average child penalty estimates (the average difference in coefficients for mothers and non-mothers) in the first 10 years after childbirth for different birth cohorts. The estimates hover around 6-8%. We assume that the relative stability in the child penalty also applies to the earlier birth cohorts. To this end, we fit a linear trend and assign to the birth cohorts before 1990 the predicted child penalty from later cohorts.[38]

Fraction of Mothers in Each Cohort and Year. The second element needed to correct the weekly earnings of women is an estimate of the fraction of mothers at each event time and for

[35] In other words, we estimate the following regression $\text{Never-Mother}_{iT} = \alpha + X'_{it}\beta + \epsilon_{it}$, where Never-Mother_{iT} is a dummy equal to 1 for never-mothers in birth cohorts 1950-1974, and X_{it} includes the dummy controls indicated in the text.

[36] The assumption that the fraction of never-mothers is constant contrasts with the secular reduction in fertility rates. However, this assumption is rather innocuous, as it only marginally affects the size of the child penalty, as highlighted in Casarico and Lattanzio (2023).

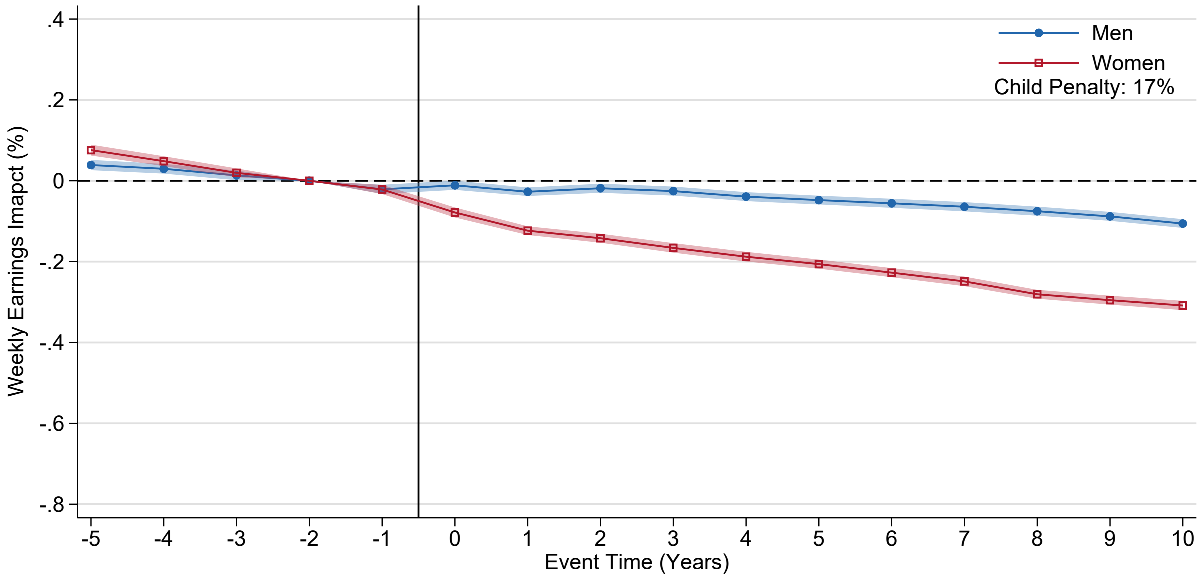
[37] This adjustment is necessary because of the truncation issue. As we do not observe completed fertility for truncated cohorts, age at childbirth would be skewed to the right if we did not make any adjustment.

[38] We choose 1990 as a threshold as this is the first birth cohort for which we observe enough observations per mother.

each cohort. For the years in which we have observations, we compute the share of new mothers and the total share of mothers in the data at each age. Figure C4 shows the share of new mothers by year and different age groups in Panel A. As expected, the share of new mothers displays an inverse U-shape relationship with age: it is small at earlier ages, peaks around 30 years old, and then declines. The peak in age at childbirth has increased over time, as can be seen by the upward-sloping trend in the share of new mothers at age 28, especially after the 1990s, mirrored by a decline in the share of new mothers at age 25. For the years in which we do not have enough observations to compute the share of new mothers, we fit a quadratic time trend by age and assume the share of new mothers equals the predicted values of the fit, reported as lines in Figure C4. We also estimate the share of total mothers in a given year at each age. Again, we fit a quadratic trend to retrieve the fraction of mothers by age in the years in which we have no observations.

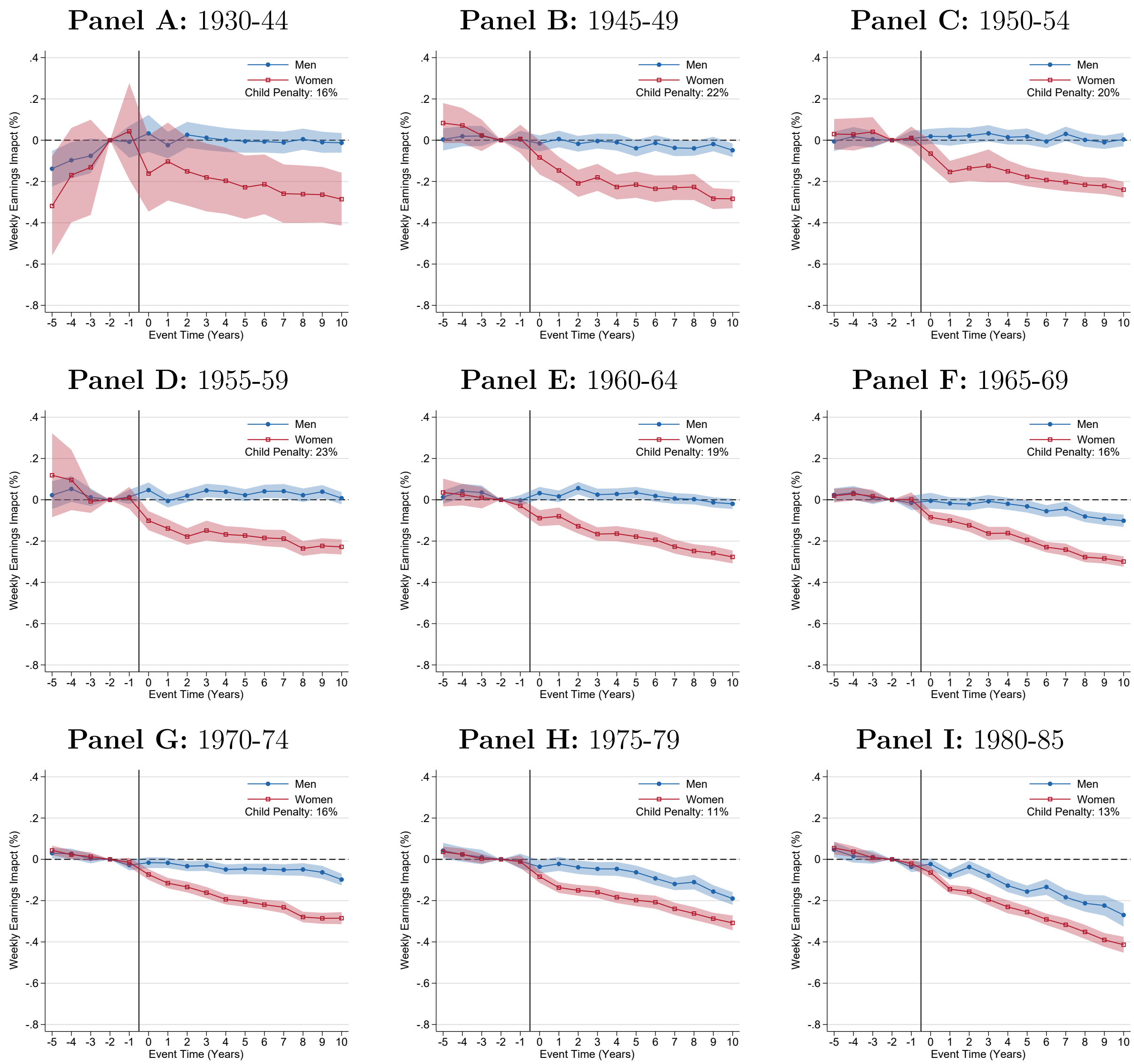
Correcting Weekly Earnings. We now have all the elements to perform the correction of weekly earnings in Equation 4. Figure C5 reports the life cycle profile, averaged over calendar years, of the gender gap in weekly earnings with and without the adjustment for the child penalty. The unadjusted gender wage gap starts at around 0.12 log points at age 25 and increases to 0.16 log points by age 40. Correcting for the child penalty removes most of the life cycle wage growth in the gap: the gender gap would be 0.13 log points at 40 years old, close to 20% lower than that observed in the data.

Figure C1: Child Penalty Event Study: Weekly Earnings Conditional on Working



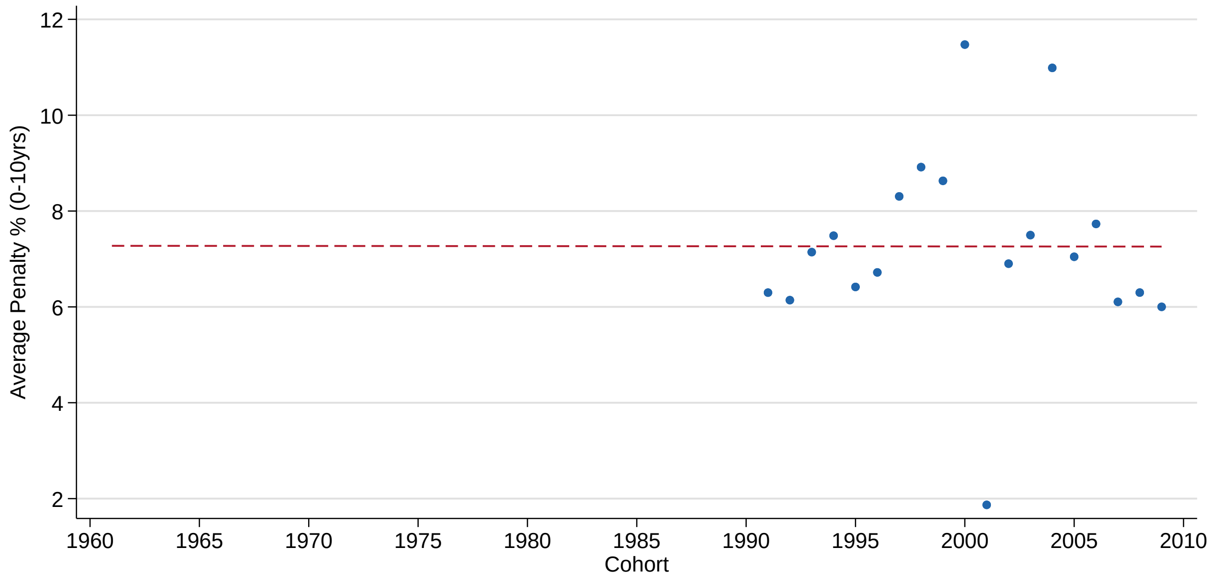
Notes: Event study for weekly earnings for men and women around the birth of their first child at $t = 0$. The series show the percentage impact of child birth on men and women at each event time t, i.e., $\hat{P}_t^m$ and $\hat{P}_t^w$ estimated from Equations (6)-(7). The figure also displays the average child penalty over event times $0 - 10$ defined as in Equation (8). Age at first birth is restricted to be between ages 25–45. The 95% confidence intervals are based on robust standard errors. *Source:* Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Figure C2: Child Penalty by Birth Cohort: Weekly Earnings Conditional on Working



Notes: Event studies of first child birth for weekly earnings conditional on working across birth cohort groups. The sample of parents is split by birth cohort group and the event study specification (6) is estimated separately for cohort group. Each panel displays the average child penalty over event times 0-10 (defined in Equation 8) for the time period in question. The 95% confidence intervals are based on robust standard errors. *Source:* Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Figure C3: Child Penalty Estimates by Birth Cohort

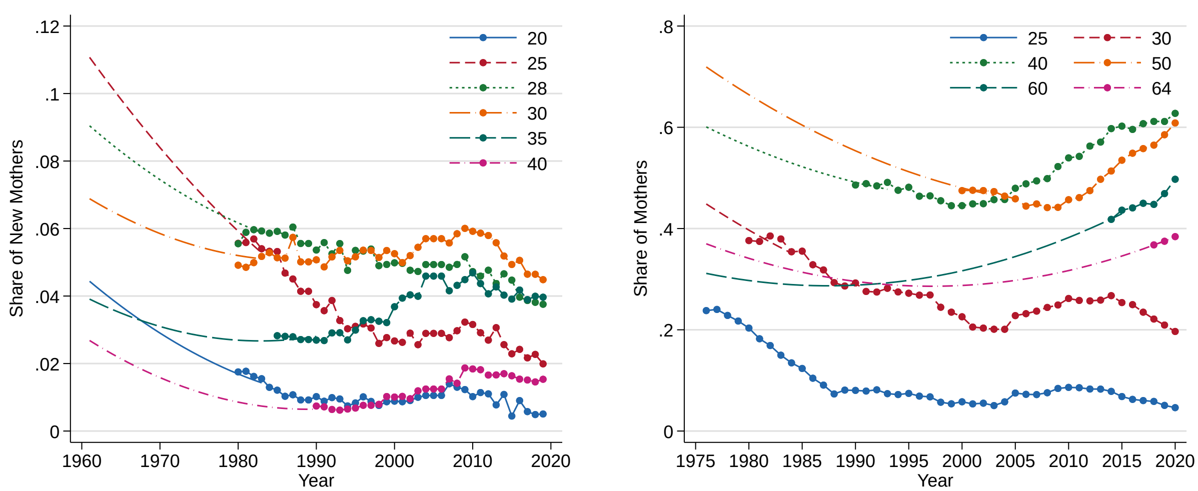


Notes: Average child penalty estimates in the first 10 years following childbirth (dots) and linear trend (dashed lines) by birth cohort. *Source:* Database UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS).

Figure C4: Share of Mothers by Year and Cohort

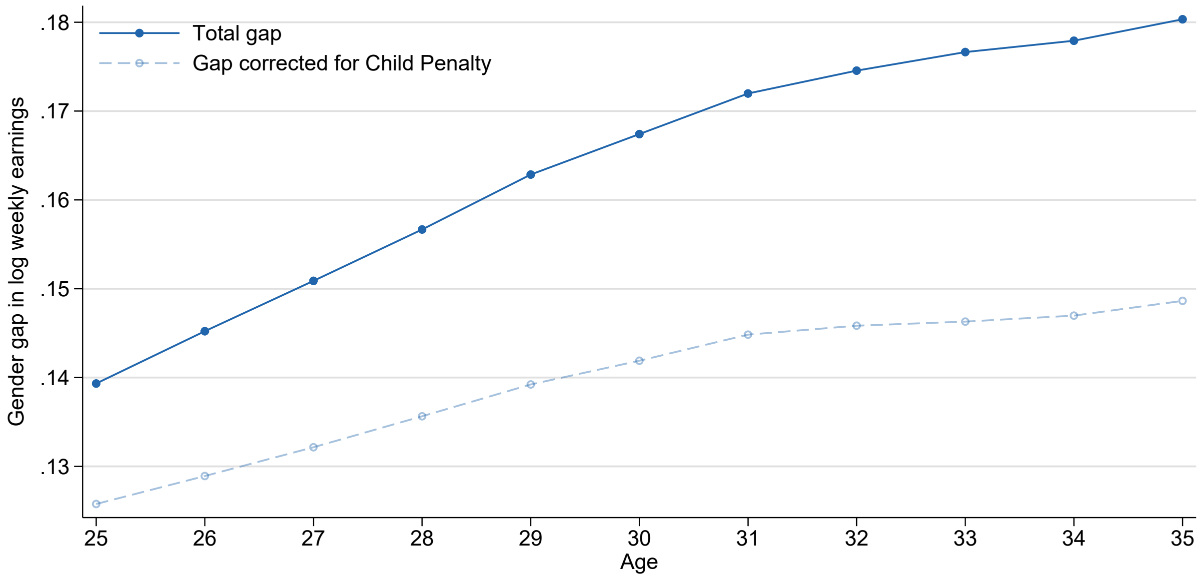
Panel A: Share of New Mothers

Panel B: Share of Total Mothers



Notes: Panel A shows the observed share of new mothers in any given year (dots), computed as the fraction of women having children over total women. Panel B shows the observed share of total mothers in any given year, computed as the share of women with children over total women. In both panels dashed lines are quadratic trends by age. *Source:* Database UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS).

Figure C5: Raw and Child-Penalty-Adjusted Weekly Earnings Over the Life Cycle



Notes: The figure reports the gender gap in log weekly earnings over the life cycle averaged over time and cohorts. The solid line is the observed gap. The dashed line is the gap obtained after correcting women's weekly earnings by the estimated child penalty. The sample includes the first ten years after labor-market entry for all cohorts who entered starting from 1995. *Source:* Database UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS).