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

Jaime Arellano-Bover, Nicola Bianchi, Salvatore Lattanzio, Matteo Paradisi

Preliminary draft, comments welcome

June 16, 2023

Abstract

We provide new insights on the decline of the gender pay gap from the 1970s until today, using data from multiple high-income countries. We find support for a cohort-driven interpretation of the shrinking pay gap: newer worker cohorts with smaller gender gaps gradually replaced older cohorts with larger gaps. We show that the cross-cohort decline in the gender gap in entry wages among younger workers accounts for the overall reduction in the gender gap at most until 2000. After 2000, the exit of older cohorts with larger gaps becomes the prominent driver of cross-cohort convergence. Next, we further examine the progressive reduction in the entry-pay gender gap among younger workers that stopped in 2000. Rather than stemming from younger women's improved outcomes, it mainly originates from younger men's absolute and relative (to women) earning losses. Finally, we show that changes across cohorts in initial allocations to firms help explain the decline in the overall gap, but this is not true of initial sector allocations. Overall, we conclude that the convergence achieved over the last two decades is mostly the slow-moving consequence of equality gains achieved decades ago, a realization that provides a pessimistic outlook for the future trajectory of the gender gap.

JEL Classification: J16, J31, J11.
Keywords: gender gap, wage growth, age wage gap, older workers, labor market entry, entry wages, initial conditions.

*Contact information: Jaime Arellano-Bover, Tor Vergata University of Rome, EIEF, and IZA, via Columbia 2, 00133 Rome, j.arellano.bover@gmail.com; Nicola Bianchi, Kellogg School of Management, Northwestern University, and NBER, 2211 Campus Drive, 60201 Evanston IL, nicola.bianchi@kellogg.northwestern.edu; Salvatore Lattanzio, Bank of Italy, via Nazionale 91, 00184 Rome, salvatore.lattanzio@bancaditalia.it; Matteo Paradisi, EIEF, via Sallustiana 62, 00187 Rome, matteo.paradisi@eief.it. We thank Sergey Abramenko, Carolina Bussotti, Sean Chen, and Chengmou Lei for outstanding research assistance. The realization of this article was possible thanks to the the sponsorship of the "VisitINPS Scholars" program. The views expressed in this paper are those of the authors only and should not be attributed to the Bank of Italy nor to the Eurosystem.

1 Introduction

The gender pay gap has decreased in all high-income economies over recent decades. The median gender wage gap in OECD countries was 12 percent in 2020, down from 30 percent in 1975.[1] A large literature (see Blau and Kahn (2017) for a survey) has documented the evolution of the gender gap over time and has examined different explanations for its existence, including differences in educational attainment and field of study, sorting across firms and occupations, child bearing, compensating differentials for long and flexible work hours, and discrimination.

The literature has devoted much efforts to studying the drivers of gender pay gap *levels*, but less so towards understanding its downward *trend*. Generally speaking, the decreasing gap might be explained by two types of explanations. The first type is a *market-wide* convergence in pay, unfolding one year at a time, and broadly impacting new and experienced workers alike. In the second type of explanation, newer *cohorts* may have progressively entered the labor market with lower gender gaps and, together with the retirement of older cohorts, brought down the overall gap one cohort at a time. Understanding which is the dominant explanation is important for at least three reasons. First, improving our basic understanding of the shrinking gender pay gap, a key trend across the rich world. Second, different explanations point to different root causes: *market-wide* convergence would point to societal changes impacting broad cross-sections of women (e.g., changes in discrimination and gender stereotypes, or equal employment opportunity laws); *cohort*-driven convergence would instead be consistent with changes in policies or attitudes that impact young women's formation and early labor-market experiences (e.g., education choices or family-formation policies). Third, the prospects of future convergence clearly depend on the type of explanation. E.g., even conditional on a *cohort*-driven explanation, recent declines that are due to *active* convergence (continued inflows of lower-gap cohorts) vs. *passive* convergence (outflows of higher-gap cohorts) have dramatically different implications.

In this paper, we use administrative data from Italy, as well as survey data from Canada, France, Germany, and the United Kingdom, and the United States, to provide a new perspective on the decline of the gender pay gap. We carry out a novel set of analyses focusing on cross-cohort differences, generational change, and initial labor market entry conditions. These exercises result in four key takeaways. First, our evidence supports a *cohort-driven* interpretation of the shrinking of the gender gap, as opposed to a *market-wide* convergence interpretation. Second, the cross-cohort decline in the *entry-pay gap*, coupled with the gradual replacement of older cohorts, almost perfectly explain the reduction in the overall gap.

[1] See the OECD Family Database at https://www.oecd.org/els/family/database.htm.

Third, the decline in the entry-pay gap we observe since the mid 1970s came to a halt by the late 1990s. As a result, during the past two decades, only the exit of older cohorts with greater gaps has contributed to shrink the overall gap. Fourth, when using older (white) men as a benchmark, the observed decline in the entry-pay gap up until the 2000s is explained in large part due to *worsening* outcomes for younger men.

The paper proceeds in four steps. First, we show descriptively that the gender pay gap has consistently narrowed in all countries in our sample for at least the last four decades, with the decline stemming from reductions happening mostly across cohorts rather than within any given cohort. In both Italy and the US, despite some differences in the within-cohort patterns, the cross-cohort trend is very different from the aggregate one, suggesting that the convergence between the pay of men and women is not affecting all cohorts simultaneously. Moreover, we show that the gap at entry in the labor market decreased across cohorts until the 2000s, meaning that younger cohorts started from lower levels of gender pay gaps compared to older ones. Therefore, the intergenerational turnover favors the decreasing trend in the overall gap.

Second, we precisely quantify how much of the decreasing gender pay gap over time stems for cross-cohort convergence between genders. Specifically, we decompose the aggregate gender pay gap into a *between-cohort* and *within-cohort* components. The former captures how the overall gap would evolve if there was no change in the life-cycle wage growth across cohorts and genders. Hence, it measures the effect of a pure compositional change in the age demographics of men and women. In both countries the entire trend in the gender gap in weekly earnings can be accounted for by a progressive decrease in the between-cohort component, i.e., in the entry gap across cohorts.[2] Therefore, our analysis reveals that the secular trend in the gender pay gap comes from younger cohorts with smaller pay gaps replacing older ones with large gaps. However, since the convergence in entry gaps has stalled during the 21st century, we show that the exit of older cohorts has nowadays become the most important driver of the declining overall gap. Subsequently, in the absence of any structural breaks, a projection exercise predicting the timing of the closure of the overall gap results in a never-convergence result.

Third, for the Italian case, we investigate the potential role of sectors and firms in determining the cross-cohort convergence in the gender pay gap documented above. Specifically, we measure how much of the decline in the gender pay gap is explained by (i) a decline in the gender gap in sector- or firm-level average aggregate wages, and (ii) a decline in sector- or firm-level average *entry* wages. Our analysis reveals that sectors only marginally contribute to the shrinking wage gap. In contrast, better sorting across firms explain a larger portion

[2] We confirm these results also hold in datasets from Canada, France, Germany, and the UK.

of such decrease.[3]

Finally, we ask whether the 1970s–2000 decline in the entry-pay gap was a reflection of improved prospects for young women or, instead, worse ones for young men. Answering this question requires moving beyond the entry male-female gap itself and, instead, comparing young women and young men to a common benchmark. To this end, we compare how young men and women aged 25–29 have ranked, over time, in the pay distribution of (white) men older than 55. In Italy and in the US, we find that the narrowing of the entry gap until the 2000s was mostly driven by younger men doing *worse* in absolute terms, rather than by younger women experiencing disproportionately meaningful improvements.

All in all, our findings imply that, both in Italy and US, pay convergence has manifested itself one cohort at a time; initially spurred by a decline in entry-pay gaps but, during the past two decades, merely driven by the outflow of older cohorts featuring greater gaps. Even more pessimistically, the original catalysts of these dynamics—male-female convergence in entry outcomes until the 2000s—were not driven by improved prospects for young women but, rather, for disproportionately worsening outcomes for young men. This interpretation leads to a cautionary note against interpreting narrowing gaps as undoubted gains for the less-privileged group.

There is a vast literature documenting different aspects of gender gaps in the labor market (Altonji and Blank, 1999; Olivetti and Petrongolo, 2016; Blau and Kahn, 2017; Bertrand, 2020). Convergence in the pay gap across richer economies is a well documented fact (e.g., Olivetti and Petrongolo, 2016; Blau and Kahn, 2017), yet our analyses focusing on cohorts, initial conditions, and relative vs. absolute opportunities provide new insights into the dynamics and potential sources of the observed convergence. Goldin (2014) documents how, among US college graduates, the declining gap masks cohort-specific gaps that rise with age. Our analyses, rather than focusing on age, clearly illustrates the importance of cohort effects in initial conditions towards determining the time trend in the overall gap.

Recent work has documented gender gaps in initial conditions (Bovini, De Philippis and Rizzica, 2023; Foliano et al., 2023). We complement and add to these works by showing how, in two distinct advanced economies, trends in entry-wage gaps and the replacement of older cohorts can almost perfectly explain the trend in the overall pay gap. Our evidence on cohorts-based convergence and the importance of initial conditions is consistent with changes in policies, attitudes, or allocations that impact young women's formative and labor-market-entry experiences and materialize one cohort at a time. Potential examples include gender

[3] Interestingly, the improvements in the relative sorting of women across firms are more evident in the first years in the labor market (between 25 and 30 years old) than in the aggregate, indicating that, over the life cycle, women then tend to sort to worse firms compared to men.

stereotypes in education (Carlana, 2019), field of study choice (Porter and Serra, 2020), family-leave policies (Blair and Posmanick, 2023), or first-job sorting across heterogeneous employers and industries (Casarico and Lattanzio, 2023; Arellano-Bover, 2024; Olivetti and Petrongolo, 2016).

We additionally provide a new way of interpreting gender gaps in entry wages, building on recent works on the age gap (Dabla-Norris, Pizzinelli and Rappaport, 2023; Bianchi and Paradisi, 2023) and using older (white) men as an absolute benchmark comparison group. This allows us to innovate relative to the existing literature, going beyond relative comparisons, towards an understanding of how much of the shrinking gap in entry wages is due to increased opportunities for young women vs. worsening outcomes for young men (race-to-the-bottom effects). Finally, our analysis on the degree to which starting firms and sectors explain the trend in the aggregate gap contributes to studies that link differential sorting between firms and sectors as underlying explanations of pay gap levels (e.g., Card, Cardoso and Kline, 2016; Casarico and Lattanzio, 2023; Huneeus et al., 2021; Blau and Kahn, 2017).

The remainder of this paper is structured as follows. Section 2 describes our data sources and presents aggregate trends. Section 3 carries out the within- and between-cohort gap decomposition and related analyses, including the role of differential sorting across firms and sectors. Section 4 zooms into entry wages and compares the outcomes of young men and women to those of older (white) men. Section 5 concludes.

2 The Gender Gap: Data and Aggregate Trends

2.1 Italian Social Security

Our analysis mainly relies on confidential administrative data provided by the Italian Social Security Institute (INPS). This dataset comprises 44 years (1976-2020) of matched employer–employee records for the whole population of private-sector, nonagricultural firms with at least one salaried employee. The INPS data contain worker-level demographic and economic variables, as well as basic firm-level information.

The worker-level information on labor contracts includes the annual gross earnings—this earning variable considers all forms of labor compensation gross of labor income taxes and social security contributions—, the number of working weeks, whether the schedule is part-time or full-time, whether the contract is fixed-term or open-ended (this differentiation started in 1998), a coarse measure of job level—apprentice, blue-collar, white-collar, middle manager, manager—, and the starting and ending date of the contract. The demographic information includes gender, the birth year, and the region of residence. Moreover, we gather information on each firm's economic sector, based on the statistical classification of economic

activities of the European Community (NACE Rev. 2).

The original INPS dataset is structured at the level of worker, labor contract, and year. In other words, workers can appear multiple times in a year if they experienced significant variations in their employment status. In order to simplify the data to have unique worker-year pairings, we retain the contract information only for each worker's main job, defined as the one with highest earnings. However, before dropping excess information, we compute each worker's total annual earnings, summing across all available contracts. Next, we convert working weeks in full-time equivalent (FTE) units by multiplying the number of actual working weeks by the ratio between the number of hours worked in a month and the number of contractual hours for the full-time equivalent position. We then derive the full-time equivalent weekly earnings, our favorite measure of labor compensation, as the ratio between total annual earnings and FTE working weeks. All measures of labor earnings are expressed in 2015 euros, using the CPI provided by the OECD.[4] Moreover, they are winsorized at the 99.9^{th} percentile to limit the influence of extreme outliers.

Finally, we impose a few sample restrictions to exclude workers without meaningful labor-force participation. In each year, we focus on workers who were between 25 and 64 years old, had worked at least 24 weeks (corresponding to 50 percent of a standard working year), had earned strictly positive earnings, and did not retire.

2.2 US Current Population Survey

Our analysis on US data relies on repeated cross sections from the Current Population Survey between 1976 and 2019, accessed through IPUMS (Flood et al., 2022). Our main analyses are carried out using the CPS's Annual Social and Economic Supplement (ASEC). In line with the restrictions imposed on the Italian data, we limit our sample to individuals who were between 25 and 64 years old, had worked in the private sector for at least 27 weeks during the past year, and had earned strictly positive wages.

The CPS data allow us to construct three compensation measures. First, we observe the annual wage and salary income earned during the previous year. Second, we obtain a measure of weekly earnings, computed by dividing annual wage and salary income over the number of weeks worked during the previous year. Third, we leverage a measure of hourly earnings, constructed by dividing weekly earnings over the usual number of hours worked per week. The main compensation measure in our analysis is weekly earnings, although we also provide additional results considering annual wage and salary income or hourly earnings. All compensation measures are expressed in 2019 USD, using the CPI provided by the BLS.

We complement the ASEC results by carrying out robustness tests using data from the

[4] The tables can be downloaded from https://data.oecd.org/price/inflation-cpi.htm.

Outgoing Rotation Group/Earner Study between 1982 and 2019. This sample is restricted to those who were employed as wage or salaried workers at the time of the interview. In this set of results, we focus on private-sector workers and utilize weekly earnings at the time of the interview as pay measure.

2.3 Luxembourg Income Studies

In addition to the confidential Social Security data for Italy and the large-scale CPS data for the United States, we leverage survey data from the Luxembourg Income Survey (LIS) database for four additional countries: Canada (1973-2018), France (2002-2019), Germany (1994-2018), and the United Kingdom (1979-2019). We compute weekly earnings, our favorite pay variable, for Canada and Germany, the two datasets that include information on working weeks, while we observe total yearly labor earnings in all four countries. Both earnings variables are expressed in real terms after converting them to 2011 purchasing-power-parity US dollars and are winsorized at the 99.9^{th} percentile.

Whenever possible, we apply the same sample limitations that we used for the Italian and US data. Specifically, in all four countries, we keep workers who were between 25 and 64 years old and had strictly positive labor earnings. However, as already discussed, we can focus on workers who had worked at least 24 weeks only in Canada and Germany, because the other two datasets do not have any information on the number of working weeks. In total, the four LIS datasets comprise 2.4 million observations across 125 survey years.

Finally, it is worth noting that both the CPS data and the LIS database lack informational depth compared to the more comprehensive Italian data. For example, unlike the Social Security dataset, the survey data do not have information about employers and tend to have much smaller sample sizes. Therefore, we mainly leverage the survey data to show the robustness of a subset of all empirical tests conducted with the Italian INPS data.

2.4 Aggregate Trends

In the last four decades, both Italy and the United States experienced a steep decline in the gender gap in weekly earnings (Figure 1). In Italy, the gender gap decreased by 0.21 log points (or 62 percent with respect to the baseline year), from 0.33 log points in 1976 to 0.13 log points in 2020. The United States experienced an even larger decline in absolute value. Here, the gender gap fell by 0.46 log points (or 59 percent with respect to the baseline year), from 0.78 log points in 1976 to 0.32 log points in 2019. These findings hold if we replace weekly earnings with the total yearly earnings. In this case, the gender gap decreased by 0.12 log points in Italy and by 0.47 log points in the United States (Figure A1).

The decline in the gender gap extends to other countries, appearing to be a generalized

trend within high-income economies (Figure A2). Specifically, we find that the gender gap in yearly earnings decreased by 0.32 log points in Canada (1973-2018), 0.12 log points in Germany (1994-2019), 0.06 log points in France (2002-2018), and 0.63 log points in the United Kingdom (1979-2019). If we consider only the period in which all data sources overlap (2002-2018), the gender gap in yearly earnings decrease by 0.03 log points in Italy, 0.09 log points in the United States, 0.11 log points in Canada, 0.06 log points in France, 0.20 log points in Germany, and 0.19 log points in the United Kingdom.

To control for changes over time in the composition of the workforce, we compute the trend in the gender gap using residualized earnings. Specifically, we regress log earnings on a series of variables, whenever these are available in one of our datasets (Table A2): a male dummy, a dummy for nonimmigrant workers (replaced by a dummy for white workers in the United States), a dummy for temporary contracts, a dummy for college education, and a dummy for disability status. We estimate separate regressions in each year and country, therefore allowing each coefficient to vary over time and geography. We then use the residuals from these regressions to compute the gender gap. In almost all countries, the gender gap's decline remains substantial. For example, the conditional gender gap decreased by 0.25 log points in Italy, while the unconditional one declined by 0.21 log points. Similarly, in the United States, the gender gap experienced a 0.46-log-point decline in raw weekly earnings and a xx-log-point decline in residual weekly earnings. The sole exception is represented by France; its conditional gender gap decreased by only 0.002 log points, down from a decline in the unconditional gender gap of 0.06 log points.

3 Cross-Cohort Trends in the Gender Gap

3.1 Decomposing the Gender Gap Within and Between Cohorts

In this section, we evaluate the contribution of workers' cohorts to the evolution of the gender pay gap in the last decades. We consider two alternative possibilities about how the gap has evolved. First, the gap might have closed because of a labor market-wise convergence in the earnings of male and female employees. If this was the case, the pay gap should have closed for *all* cohorts over time at a similar rate. Alternatively, the decrease in the gender pay gap might stem from a progressive reduction of the pay gap *at entry*. In the latter case, the observed reduction in the population gap would be the result of a compositional effect: new cohorts with lower gaps would have progressively replaced older cohorts with larger gaps.

To test which of the two alternatives is more consistent with the data, we develop a decomposition of the change in the gender pay gap into a *within-* and *between-cohorts* components. We implement this decomposition for a measure of weekly earnings, and we show yearly labor earnings as a robustness. First, we write the average log-earnings outcome of a

birth year cohort c of gender g at time 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 entry earnings, and $\Delta w_{c,g,t} = w_{c,g,t} - w^e_{c,g}$ is the growth of the average log-earnings over the life-cycle of cohort c. Denote with $a\,(c,t)$ the age of cohort c in year t. The change between year t and year $t' > t$ in the average log-earnings of a given gender group g is:

$$\begin{aligned} w_{g,t'} - w_{g,t} = & \underbrace{\sum_{c:a(c,t')\in[25,64]} s_{c,g,t'} \cdot w^e_{c,g} - \sum_{c:a(c,t)\in[25,64]} s_{c,g,t} \cdot w^e_{c,g}}_{\text{Between-Cohort Change}} \\ & + \underbrace{\sum_{c:a(c,t')\in[25,64]} s_{c,g,t'} \cdot \Delta w_{c,g,t'} - \sum_{c:a(c,t)\in[25,64]} s_{c,g,t} \cdot \Delta w_{c,g,t}}_{\text{Within-Cohort Change}} \end{aligned} \tag{1}$$

defining $s_{c,g,t}$ as the share of workers of gender g from cohort c in year t over the total number of employees in the same gender group. This share operates as a weight for a given cohort of active employees (between 25 and 64 years old) in a given year. The first two terms in (1) quantify the *between-cohort* over time change. It keeps the earnings of each cohort fixed over time and equal to their entry earnings, but it lets the weight of that cohort change according to the observed demographic trend. Hence, it captures the impact of changes in the age composition of the pool of employees of a certain gender group in case different cohorts have heterogeneous entry earnings. On the other hand, the latter two terms of equation (1) measure the *within-cohort* over time change. This component quantifies changes in the life-cycle earnings growth for the average cohort active in the labor market. Subtracting the overall change for women from the one for men, we can quantify the contribution of the *between-* and *within-cohort* components to the change in the gender pay gap in a double-difference version of equation (1).

To implement this exercise, we need a precise definition of $w^e_{c,g}$, *i.e.* the entry earnings of cohort c. Because we have access to panel data only for Italy, it is not possible to measure the exact earnings in the entry year for the workers in other countries' data. Hence, we opt for a more homogeneous definition of entry earnings that we can measure in all samples. In our baseline results, we define as entry earnings for a cohort-gender combination their average earnings in the 25th year of age. We then show the robustness of our results to an array of alternative definitions, including the average before the 30th years of age, the earnings in other years of age (e.g. 28 and 30), and – for Italy only – the actual entry earnings as measured in longitudinal data.

Another practical issue that arises in the implementation of (1) concerns the definition of $w^e_{c,g}$ for older cohorts in the early years of the sample. As an example, in the first available calendar year the entry earnings are not observable for any worker who is 26 or older. In general, denote with $\underline{t}$ the first calendar year in the sample, in a given year t we will not observe the entry earnings of those who are older than $t - \underline{t} + 25$ years of age. We therefore implement a constrained version of (1) where we assign to each cohort-gender combination older than 25 in year $\underline{t}$ the average earnings that they earned in that year. The resulting *between-cohort* component will represent an upper-bound to the true component in equation (1) if the life-cycle growth in the gender pay gap has decreased across cohorts, while it will be a lower-bound if the opposite is true. In our robustness tests, we exploit the long time-series for Italy and the US to show the ideal decomposition in (1) on a shorter time-series, where we use the excluded years at the beginning of the sample to quantify the entry earnings for all older cohorts.

3.2 Aggregate Trend and Cohort Trends

Figure 2 contrasts the aggregate trend in the gender pay gap with the cohort-specific evolution. For a set of cohorts we show how the gap in the average weekly earnings changes over the life-cycle from the year in which the cohort turned 25 years old to their 50th year of age. In the US the gap evolves in a similar way across cohorts showing an inverted-U-shape with an increase over the first ten years in the labor market, followed by a decrease that brings the gap back to its initial level by the time the cohort is over 40. Italian cohorts instead show a more stable gap over the life-cycle with a slight increase as the cohort ages. Increases in the gap over a cohort life-cycle are consistent with many of the typical explanations for the existence of a gender gap in the first place. For instance, the child penalty or the propensity of women to move towards more flexible jobs are likely to emerge during the lifetime of a cohort, increasing the gender gap within cohort.[5] The decrease observed in the US at later ages is consistent with changes in the selection into work, and a progressive increase in women's average earnings due to the fact that lower-earnings women are more likely to leave the workforce as they get older (Goldin and Mitchell, 2017). Despite some difference in the within-cohort patterns, in both countries the cohort trend is very different from the aggregate one, suggesting that the convergence between the earnings of men and women is not affecting all cohorts simultaneously. At the same time, in both countries the gap at 25 years old is decreasing across cohorts so that younger cohorts start from lower gender pay

[5] The mother's average age at *first* birth has increased from 25 to 27 between 2000 and 2020 in the US. Since the 1970s the average age at birth has moved from 26 to 29 in the US and from 27 to 31 in Italy (OECD, 2022). In both countries – and for all the other countries included in our analysis – the age at first birth has always been strictly above 20 years old.

gaps compared to older ones and the intergenerational turnover favors the decreasing trend in the gap. In both countries, the decreasing trend in entry earnings seems to have slowed down among the younger cohorts in the sample.

3.3 The Importance of the Between-Cohort Component

The decomposition in equation (1) can quantify the extent to which the convergence across cohorts is crucial for the observed reduction of the gender pay gap, complementing the descriptive evidence provided in Figure 2. We implement the decomposition using the earnings at 25 to calibrate the entry earnings of a cohort, and setting all the entry earnings for older cohorts in the baseline year of the time series to their average earnings in that year. Figure 3 shows the result of this exercise for Italy and the US. For each country we plot the observed change in the gender pay gap and the counterfactual change measured by the *between-cohort* component of equation (1). The latter captures what would happen if there was no change in the life-cycle earnings growth across cohorts and genders. Hence, it measures the effect of a pure compositional change in the age demographics of men and women. In both countries the predicted drop in the gender gap is larger than the observed one: -0.59 against -0.46 log-points in the US and -0.24 against -0.21 log-points in Italy. This is because – as already discussed in Figure 2 – the gap slightly increases over time within the average cohort, pushing the aggregate gap in the opposite direction to the one observed in the secular trend. It follows that the entire trend in the gender gap in weekly earnings can be accounted for by a progressive decrease in the entry gap across cohorts. We conclude that the secular trend in the gender pay gap stems from the fact that younger cohorts with smaller pay gaps replace older cohorts retiring with large gaps, rather than from an overall convergence of the earnings of men and women in the labor market.

We also replicate the analysis on yearly labor earnings in Table A3. Results are very similar. The *between-cohort* component predicts more than 100 percent of the observed drop in the yearly earnings gender gap for both countries.

Robustness: We test the robustness of our finding by running a series of alternative specifications of the decomposition in (1), and we collect their results in Table A3. The first set of tests changes the age at which we define the entry earnings for younger cohorts. We try three alternative ages: 28, 30, and the average before the 30th year of age. Results remain very similar across all these specifications so that the counterfactual always explains all of the observed change in the gender pay gap for both countries, except in one case in which it explains 96% of the total. The extent to which the counterfactual exceeds the observed trend decreases in both countries with the entry age that we choose. In Italy, for instance, the decomposition using 28 years of age to calibrate the entry earnings is

matching almost perfectly the observed change in the gap, while the one using 25 overshoots the change. This comparison across entry ages suggests that the gender pay gap starts increasing within a cohort at very early ages, so that when we assign to each cohort a gender gap that incorporates the growth in the years immediately after the 25th we predict a smaller reduction in the population gap. As an additional robustness of the definition of entry earnings, for Italy only we exploit the longitudinal dimension of the data to measure the actual entry earnings as those earned by a worker in the first year in the labor market. Using this definition, we find that the between-cohort component explains even a larger portion of the drop compared to our baseline specification.

Then, we assess the sensitivity of our result to the implicit assumption that we made by assigning to older cohorts the average earnings they had in the baseline year of the time series. We show for both Italy and the US a shorter time series where we calibrate the entry earnings at 25 for all cohorts of workers from the data, exploiting the calendar years excluded at the beginning of the sample. This time series starts from 2015.[6] Since that year and using the earnings at 25 for all cohorts, the between-cohort component still explains the majority of the drop in both Italy (65%) and the US (above 100%).

External validity: To provide external validity to our result, we run the same analysis using Germany, France, Canada and the UK (Table A3). We observe weekly earnings in Canada and Germany only, while we use yearly labor earnings for the other two countries. For all these countries and both measures of earnings, the *between-cohort* counterfactual overshoots the observed decrease in the gender gap confirming the finding in Italy and the US. We also implement the battery of robustness discussed above and the finding remains robust.

3.4 Across-Cohort Convergence in the Gender Pay Gap

We have established that the entire drop in the gender pay gap is the result of a compositional change driven by younger cohorts with progressively lower gender gaps replacing older ones. A direct implication of this finding is that entry earnings are sufficient to capture the observed trend in the gender pay gap. We therefore focus on the convergence rate of entry pay gaps across cohorts and relate it to the pace of convergence in the population gap. Denote with g_s^e the gender pay gap at entry of the cohort who entered the labor market in year s. Suppose the convergence rate is linear, we model the gender pay gap in year t as

$$g_s^e = \alpha_t - \beta_t (s - \underline{s}) \tag{2}$$

[6] Year 2015 is the result of the following sum 1976+64-25.

where $\alpha_t = g^e_{\underline{s}}$ and $\underline{s}$ is the entry year of the cohort with the maximum age at year t (*i.e.* the 64 years old cohort). The parameter β_t measures the rate of convergence observed between year $\underline{s}$ and t. Higer β_t implies higher convergence. 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 in $T_t = \alpha_t/\beta_t$ years. Notice that $\lim_{\beta \to 0^+} T_t = +\infty$, which shows that if the cross-cohort convergence stops, the gender gap never closes.

Figure 4 Panel A shows the relationship in equation (2) in Italian data when we consider the youngest 20 cohorts active in the labor market in a given year (hence, we set $\underline{s} = t - 20$). We show data for 1976 and 2020, the initial and final years in our sample. The convergence rate decreased from $\beta_{1976} = 0.009$ to $\beta_{2020} = -0.0005$. While in 1976 the entry gap was reducing by 0.009 log-points per year, it has been stable for the past 20 years. If anything, the small negative coefficient suggests that the entry earnings mildy diverged in the most recent period. Panel B shows the evolution of β_t across all years in the sample. Starting around 1980, the convergence rate decreased at an almost linear rate and the β_t moved from a level of around 0.12 to just below 0 in the last years.

To quantify the impact on the population gender pay gap of the documented reduction in convergence, we show how the projected year of closure of the gender pay gap changed over time. In Figure 4 Panel C we exploit the definition of T_t to illustrate for any calender year when the gender gap would have closed if the convergence continued at the same rate observed in the previous 20 years. In the early years of the sample, the gap was projected to close before the end of our data. For instance, if the rate of convergence had stayed constant at the level we observe in 1980, the population gender pay gap would have closed in 2012. By 2000 the projection had already moved forward by more than 20 years, with an expected closure in 2035. However, the steady decrease in convergence starts impacting in a dramatic way on the projection starting around 2010 when the expected closure of the gender pay gap moves towards the end of this century. Finally, in the last two years of the sample – when β_t turns positive – the projection reaches a no convergence result.

3.5 Importance of Cohorts Turnover and Convergence in Entry Earnings

Motivated by the fact that the convergence in the entry gender gap decreased over the last decades, we quantify to what extent the gender pay gap dynamic in the population stemmed from such convergence, and how much instead can be accounted for by the natural turnover of older cohorts.

To this end, we implement an exercise where we neutralize the convergence in entry earnings of men and women and we compare this counterfactual to the between-cohort

component that we quantified in Section 3.3. Our new counterfactual starts from a given baseline year t_b and fixes the entry earnings of all cohorts who enter in that year or later to the average entry earnings computed between t_b and the following two years. Older cohort entry earnings are instead measured as in the previous exercise. Hence, the counterfactual of this exercise measures what would have happened if the convergence in the entry earnings of men and women had stopped at t_b (i.e. $\beta_q = 0$ for $q > t_b$), and only the turnover of older cohorts with larger gaps could have affected the population gender pay gap. By comparing this counterfactual to the one of Section 3.3, we can establish how much the convergence at entry from t_b on has contributed to the observed trend in the time period after t_b.

Figure 5 shows the results for Italy and US over the full period (Panel A), and starting from $t_b = 2000$ (Panel B). In the US, the new counterfactual accounts for 2/3 of the one that allows for convergence in entry earnings in the period after 1976, and predicts a drop in the gender pay gap just below 0.4 log-points and almost equal to the observed one. Hence, we conclude that convergence in entry earnings after 1976 is responsible for a 0.2 log-points decrease in the gender pay gap, but it was not the main driver of the trend. Italy displays a similar qualitative result: the new counterfactual is about half of the previous one after 1976, and accounts for a drop in the pay gap that equates the observed one. Notice that if we were to start the US counterfactual from 1985, the two countries would look even more similar: in the US the new counterfactual would be 80 percent of the old one, and would account for just a little more than the observed drop. We conclude that the convergence in entry earnings from 1985 on can only account for about 20-25 percent of the observed drop in the gender pay gap from that year on.

We then focus on a shorter time span. Because the entry earnings of men and women did not converge in the last 20 years of our sample (Figure 4), we focus on the time period from year 2000 on. In both countries the two counterfactuals almost coincide. As expected, the convergence in entry earnings from the 2000 entry cohort on had no impact on the gender pay gap. Hence, the observed trend in the past 20 years is purely driven by demographic turnover, and is the mechanical reflection of a convergence in entry earnings that happened decades before.

Panel C and D of Figure 5 provide a more detailed description of the ratio between the two counterfactuals. In both countries, starting from the mid 90s the counterfactual that neutralizes the cross-cohort convergence in entry earnings is almost equivalent to 100 percent of the between-cohort component in the old counterfactual. This implies that in both countries the last couple of decades of convergence in average men and women earnings were not the result of contemporaneous improvements in equality of opportunities, but the long-lasting consequences of a convergence that happened much earlier in time.

3.6 The Role of Sectors and Firms

One natural question arises: what is the role of sectors and firms behind the cross-cohort convergence in male and female earnings? Specifically, the reduction in the gender pay gap across cohors may stem from women being increasingly employed by higher paying sectors or firms relative to men. We therefore measure how much of the decline in the gender pay gap is explained by (i) a decline in the gender gap in sector- or firm-level average aggregate weekly earnings, and (ii) a decline in sector- or firm-level average *entry* earnings. We can perform this exercise only with the Italian data, which record the identifier of the firm and the sector where each worker is employed. For each firm and sector, we compute the average earnings across all employees in any given year. Figure 6 reports the contribution of sectors to the decline in the gender pay gap in panel A and that of firms in panel B. Both graphs show the decline in the aggregate gender gap in weekly wages and two counterfactuals, representing the decline that would be observed if we assign to male and female workers: (1) the average weekly earnings in their sector or firm; (2) the average weekly earnings in the sector or firm at age 25. The first counterfactual captures how much the pay gap would change because of changes in the allocation of men and women across sectors or firms, reflected in changes in the average earnings they receive in the sectors or firms which employ them. The second counterfactual show by how much the gender pay gap would change if we assign to men and women their entry sector- or firm-level earnings in the labor market. By comparing the latter with the former we also get an indirect indication of how sorting changes over the life cycle.

Panel A shows that sectors only marginally contribute to the shrinking pay gap. On the one hand, gender differences in average sector earnings are substantially stable over time, meaning that the convergence in male-female earnings is a within-sector phenomenon. On the other hand, we observe a small decline stemming from convergence in entry sector earnings, which is stronger until the early 1990s, and it reverses in the last two decades (being equal to -0.01 log-points by 2020). Hence, we interpret this evidence as higlighting a small contribution of the first sector of employment at entry in the labor market to the gender pay gap. In addition, women tend to systematically sort into lower-paying sectors over life cycle, as such contribution is decreasing if we focus on average sector-level weekly earnings across all ages.

Panel B shows that sorting across firms explain a larger portion of the decreasing gender pay gap. Gender differences in firm-level earnings shrink by 0.14 log-points (or by approximately 70 percent of the overall decrease in the aggregate gender pay gap). The gender gap in *entry* average firm-level weekly earnings explains almost the entirety of the decline in the gender gap in *current* firm earnings. In 2020 the decline in entry earnings amounts to -0.12

log-points, or 58 percent of the overall decline in the gender pay gap when fixing the entry age at 25. The reduction in the gender gap in entry earnings has been decelerating since the late 1990s, while the gender pay gap kept declining. Combining this evidence with that reported in Figure 3, which shows how between cohort convergence is the main driver behind the gender pay gap reduction, we can interpret our findings as an indication that until the late 1990s convergence between cohorts mainly resulted from relative improvements in women's sorting across firms relative to men's. Since then, the between-cohort convergence is instead primarily the result of within-firm convergence, as there is a stall in the decline of the between-firm between-cohort gender gap. Overall, we conclude that the decline in the gender pay gap between cohorts is mainly a within-sector between-firm phenomenon.

4 The Entry Gap and Absolute Gains and Losses: Older (White) Men as a Benchmark

We have so far documented how the decline in the overall gender gap over time can be traced back to gradual changes occurring one cohort at a time. In particular, we have seen that the decline in the gender pay gap over the last few decades can be explained by two forces: i) a decline in the *entry-pay gap* up until the 2000s and ii) the subsequent demographic replacement of older cohorts that featured greater gaps. We now focus our attention on entry pay, to better understand the catalyst of these two forces. Motivated by recent work documenting a growing age gap (Bianchi and Paradisi, 2023; Rosolia and Torrini, 2007) we ask the following question: Was the decline in the entry-pay gap documented in Figure 2 due to young women seeing their opportunities improve over time or, instead, young men seeing theirs worsen? This question cannot be answered by inspection of the gender gap in entry pay alone, requiring instead a common benchmark group to study the separate evolution, in absolute terms, of male and female entry pay.

The benchmark group we use to gauge the progress, in absolute terms, of young men and women are male workers aged 55 and older; in the US, we further narrow the benchmark group to *white* males over 55. The motivation for this choice is intuitive. Pre-retirement older (white) men have traditionally represented the dominant group in the labor market. Moreover, their own labor market outcomes are not impacted by direct effects of the labor market integration of women (and non-whites in the US) that has played out over the last decades. As such, (white) men over 55 represent a suitable absolute benchmark against which to compare male and female labor market entrants.

4.1 The Pay Rank Gap

We compute the pay rank gap between the gender-specific entry pay distribution and the pay distribution of (white) older men. Let $F_t^{O55}(w)$ represent the distribution of weekly earnings for (white) males over 55 in year t. For each entrant i with weekly earnings w_{it}^E, we compute where they would rank in the older (white) males distribution with $g_{it} \equiv F_t^{O55}(w_{it}^E)$. We then compute the average over-55-rank among all entrants of a given gender J in a given year t:

$$\bar{g}_t^J = \frac{1}{N_t^J} \sum_{i \in \{J,t\}} g_{it},$$

where N_t^J is the number of entrants of gender J in year t. Our baseline measure of the evolution of male and female entrants' absolute progress over time is $\bar{g}_t^J$. It allows us to understand how have young male and female entrants fared over time without directly comparing them to each other.

Building upon the analysis of the US white-black gap of Bayer and Charles (2018), we also consider an alternative metric that focuses on the median entrant rather than the average. As before, let $F_t^{O55}(w)$ represent the distribution of weekly earnings for (white) males over 55 in year t, while $F_t^{E,J}(w)$ represents the distribution of entry weekly earnings among gender J in year t. The quantile q entry pay is given by $w_{q,t}^{E,J}$, defined so that $F_t^{E,J}(w_{q,t}^{E,J}) = q$. Let $G_{q,t}^J$ quantify where the qth entry weekly earnings of gender J would rank in the older (white) males distribution:

$$G_{q,t}^J = F_t^{O55}(w_{q,t}^{E,J}).$$

$G_{q,t}^J$ when evaluated at the median $q = 0.5$ is our additional measure of the evolution of male and female entrants' absolute progress over time.

4.2 A Declining Entry Gap Due to Young Men's Larger Absolute Losses

Figure 7 shows the evolution of $\bar{g}_t^M$ and $\bar{g}_t^F$ for Italy (Panel A) and US (Panel B). In both countries, we can see two distinct phases. During the first phase, up until 1996, there is a narrowing of the entry pay gender gap; from then onwards, the weekly earnings of young men and women evolve in parallel. The stabilization of the entry pay gap in the 2000s mirrors the patterns documented in Figure 2.

More importantly, the comparison to older (white) men unveils an important fact: the narrowing of outcomes between young men and women up until 1996 was mostly driven, not by young women doing better in absolute terms, but by young men doing *worse*. In Italy, young men in 1985 were earning on average as the 39th percentile older man, while

in 1996 they were down to the 32nd percentile. Young Italian women instead went from the 29th percentile in 1985 to the 26th percentile in 1996. A starker picture arises in the US: in 1976, young men were earning on average as the 43rd percentile older man, while in 1996 they were down to the 30th percentile; young women instead went from the 21st percentile in 1976 to the 23rd percentile in 1996. As such, in both countries, the narrowing up until 1996 was driven by young men's large losses overcoming women's modest losses (in Italy) or modest gains (in the US).[7]

Robustness: Appendix Figure A3 provides additional evidence on the absolute trends of young entrants, showing our estimates of $G^{M}_{0.5,t}$ and $G^{F}_{0.5,t}$ for Italy and US. The main takeaways when focusing on median entrants are quite similar as with the average, with the main difference being that the gains of the US median female entrant are larger than the gains on average (they both reach similar percentiles in 1996, but the median started from a lower 1976 baseline).

Overall, this section provides a new and somewhat pessimistic perspective on the declining entry pay gender gap in Italy and the US. Equipped with the absolute benchmark of (white) older men, we show that the declines in entry gender gaps between the mid 1970s and 2000s were mostly due to race-to-the-bottom effects driven by large male losses.[8]

5 Conclusion

This paper presents a series of facts that shed new light on the reduction in the gender pay gap, one of the most consequential changes in labor markets across high-income countries. The key takeaway is that cross-cohort differences, generational change, and initial conditions all play important roles in the dynamics underlying the shrinking of the overall gender gap. We also provide a pessimistic interpretation of the underlying dynamics that have driven gender convergence since the mid 1970s until today.

Both in Italy and the US, pay convergence between men and women over the last few decades has taken place one cohort at a time; through the entry of newer cohorts featuring lower gaps since entry, and the exit of older cohorts featuring greater gaps. Crucially, the entry-pay gap stopped shrinking with the turn of the century and the gains towards convergence achieved over the last two decades are mostly the slow-moving consequences of equality gains achieved a long time ago. Moreover, we also paint a rather pessimistic picture

[7] The absolute benchmark of older (white) men also reveals that, in the US, young men and women made significant gains from 1996 up until the turn of the century and, with some fluctuations, their prospects have remained rather constant until 2019. More dramatically, Italian young men and women have been doing persistently worse since 2002 and lost, over the course of 16 years, seven (men) and six (women) percentiles.

[8] While young men have lost ground in absolute terms during our sample periods in both countries, in Italy, additionally, young women have also seen worsened absolute opportunities.

regarding the gender convergence in entry pay that did occur up until the mid-1990s: in both Italy and the US, on average, young men and women indeed became more similar to each other—yet such convergence was not driven by improvements in young women's prospects but, rather, by the worsening prospects of young men.

Overall, our findings are useful to guide theories of the ultimate causes of the shrinking wage gap, and can also inform policies aiming to further reduce the gap. More broadly, we see this paper as improving our basic understanding of the occurrence of a key shift among high-income societies, providing a more nuanced interpretation of the reasons to be more or less optimistic about true convergence of labor market opportunities for men and women.

References

Altonji, Joseph G., and Rebecca M. Blank. 1999. "Race and Gender in the Labor Market." 3: 3143–3259.

Arellano-Bover, Jaime. 2024. "Career consequences of firm heterogeneity for young workers: First job and firm size." *Journal of Labor Economics*.

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.

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

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

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.

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.

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. "What Firms Do: Gender Inequality in Linked Employer-Employee Data." *Journal of Labor Economics*.

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*.

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*.

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.

Huneeus, Federico, Conrad Miller, Christopher Neilson, and Seth Zimmerman. 2021. "Firm Sorting, College Major, and the Gender Earnings Gap." *working paper, Princeton University.*

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

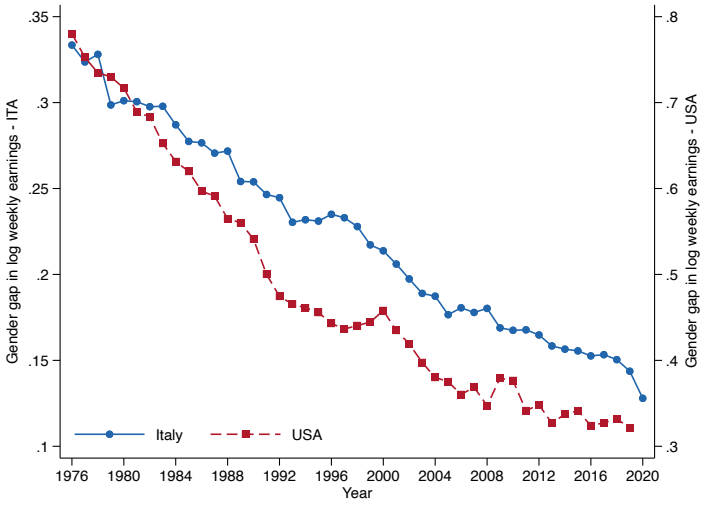
Olivetti, Claudia, and Barbara Petrongolo. 2016. "The Evolution of Gender Gaps in Industrialized Countries." *Annual Review of Economics*, 8(1): 405–434.

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.

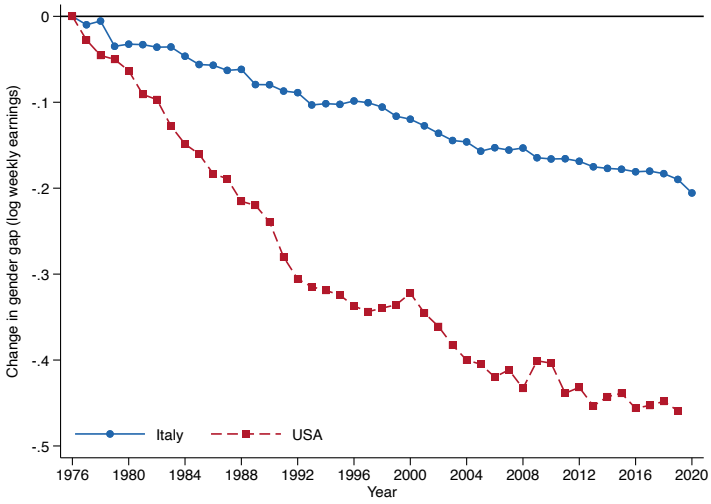
Rosolia, Alfonso, and Roberto Torrini. 2007. "The generation gap: relative earnings of young and old workers in Italy." Bank of Italy, Economic Research and International Relations Area Temi di discussione (Economic working papers) 639.

Figures and Tables

Figure 1: Gender Gap in Weekly Earnings



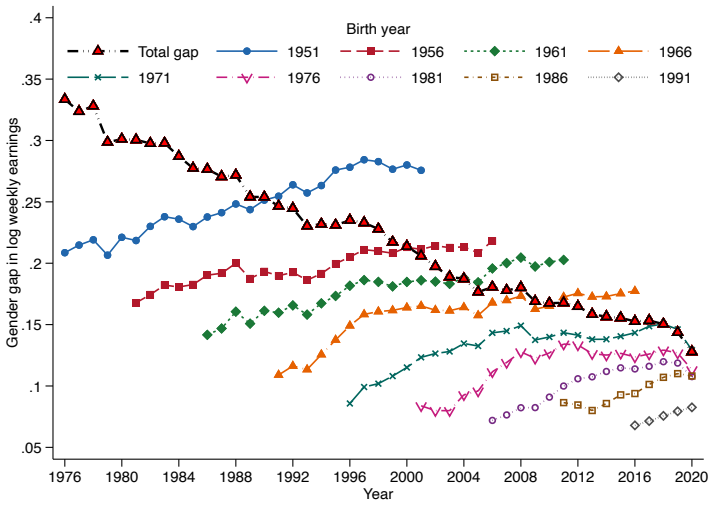
Panel A: Raw gap



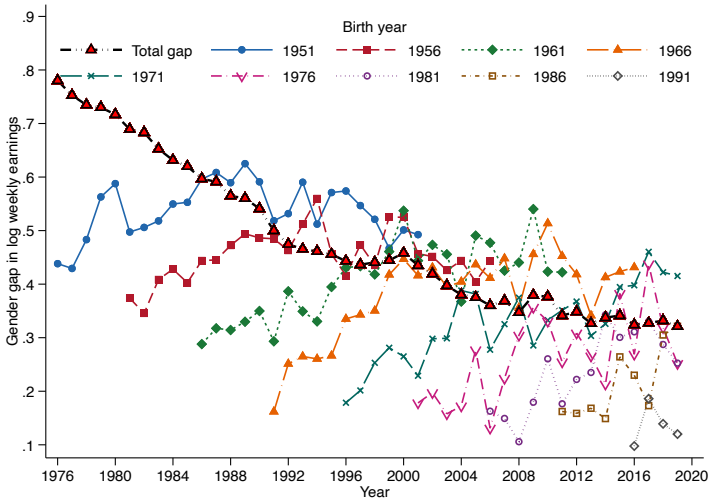
Panel B: Deviation from first year

Notes: Panel A plots the trend in the raw mean gender gap (log weekly earnings of men - log weekly earnings of women) in Italy and the United States. Panel B shows the deviation from the first year (1976). *Sources:* In each year, the data pools information about all workers who were between 25 and 64 years old, worked in the private sector for at least 27 weeks year, earned strictly positive earnings, and did not retire. *Sources for Italy:* UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Sources for the United States.* Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Figure 2: Gender Gap Between and Within Cohorts



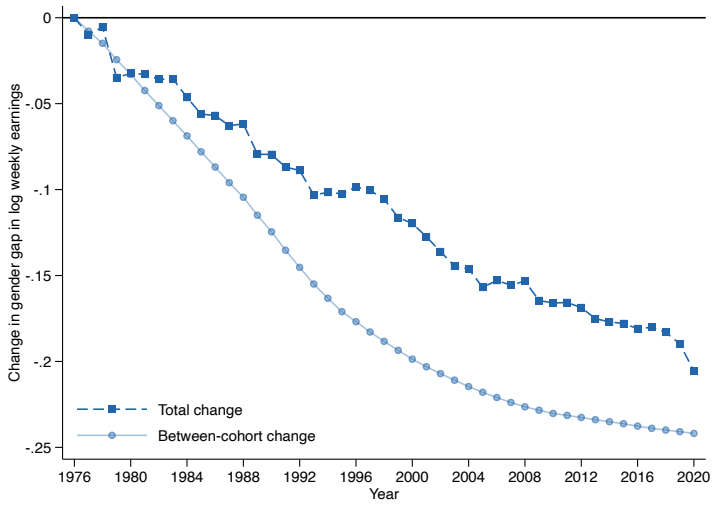
Panel A: Italy



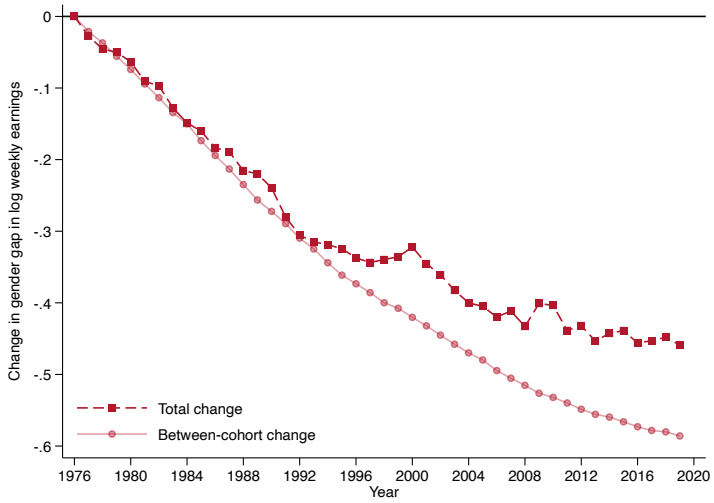
Panel B: USA

Notes: Panel A (Panel B) shows the trend in the mean gender gap across different birth cohorts in Italy (the United States). The red triangles depict the trend in the mean gender 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. *Sources:* In each year, the data pools information about all workers who were between 25 and 50 years old, worked in the private sector for at least 27 weeks year, earned strictly positive earnings, and did not retire. *Sources for Italy:* UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Sources for the United States*. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Figure 3: Between-Cohort Change in the Gender Gap



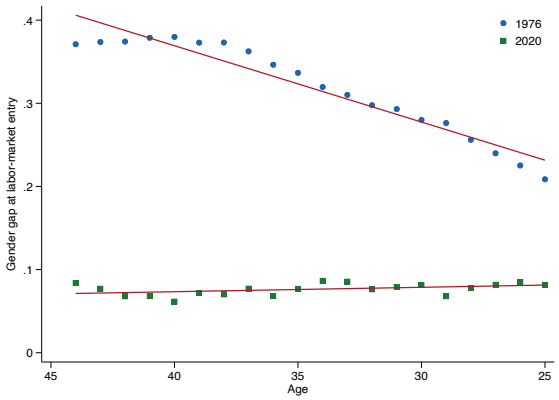
Panel A: Italy



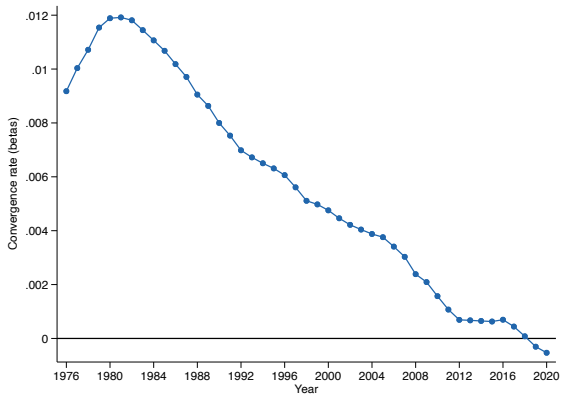
Panel B: USA

Notes: Panel A (Panel B) shows the change in the total gender gap and its between-cohort component in Italy (the United States). To compute the between-cohort component, we assign to each cohort (defined as a combination of year of entry in the labor market and gender) its mean weekly earnings at labor-market entry in each year (Equation (1)). In the baseline analysis, entry in the labor market corresponds with the year in which each cohort was 25 years old. We assign to cohorts who were older than 25 in 1976 their mean weekly earnings in 1976. *Sources:* In each year, the data pools information about all workers who were between 25 and 64 years old, worked in the private sector for at least 27 weeks year, earned strictly positive earnings, and did not retire. *Sources for Italy:* UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Sources for the United States.* Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

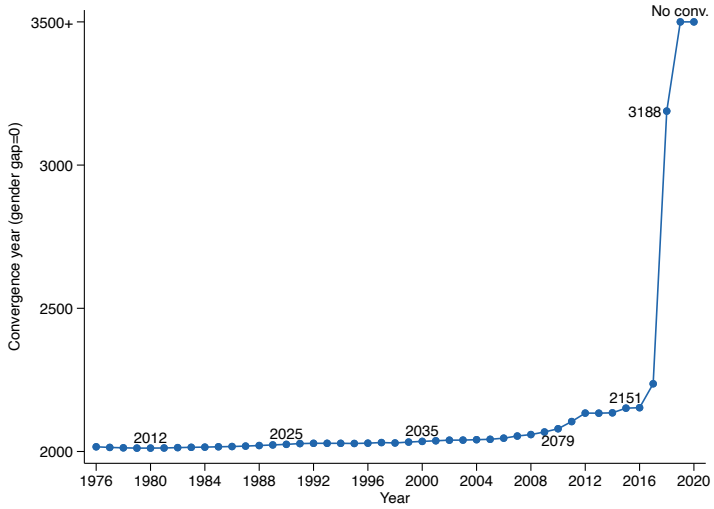
Figure 4: Between-Cohort Convergence in the Gender Gap



Panel A: Between-cohort trend



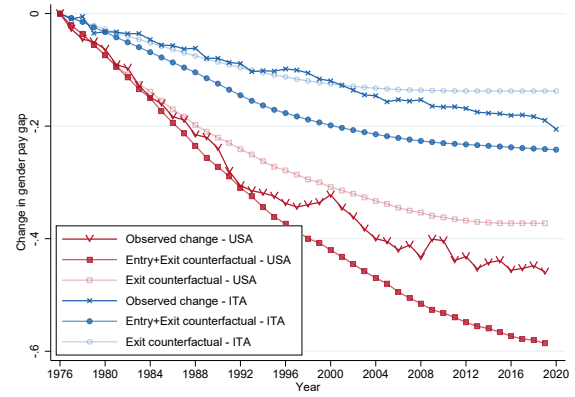
Panel B: Convergence rate



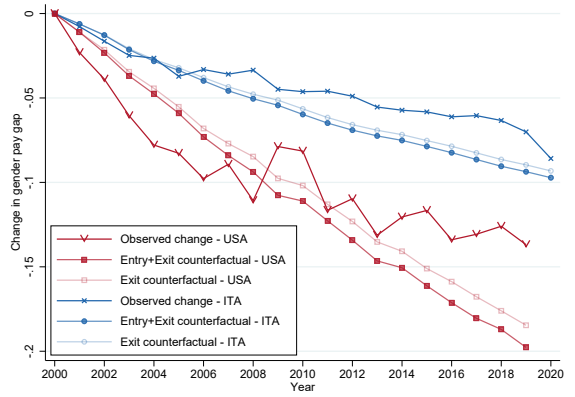
Panel C: Convergence year

Notes: Using Italian data, Panel A shows the trend in gender gap at labor-market entry across cohorts in 1976 and 2020. In year t, we compute the gender gap at labor-market entry for age group a using their earnings at 25 years old. Based on the linear relationship between the mean gender gap at labor-market entry and birth cohorts in each year between 1976 and 2020, we can compute the variation over time in the predicted year when the mean gender gap is going to close (Equation xx). Panel B shows the slope of this linear relationship over time, while Panel B shows the first year of predicted convergence in the gender gap for each year between 1976 and 2020. *Sources:* In each year, the data pools information about all workers who were between 25 and 64 years old, worked in the private sector for at least 24 weeks year, earned strictly positive earnings, and did not retire. *Sources for Italy:* UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS).

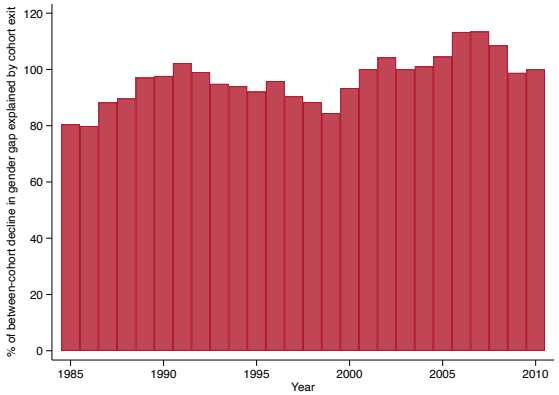
Figure 5: Between-Cohort Change in the Gender Gap Without Convergence in Entry Earnings



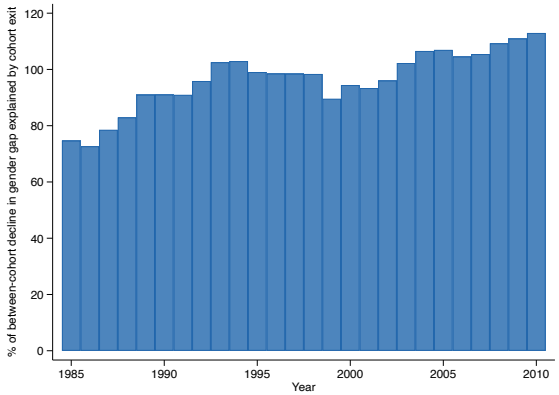
Panel A: Benchmark Year = 1976



Panel B: Benchmark Year = 2000



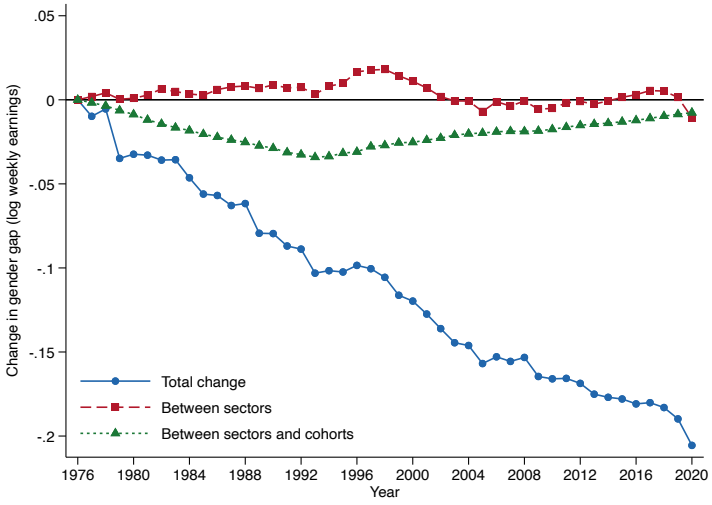
Panel C: Contribution by t_b - USA



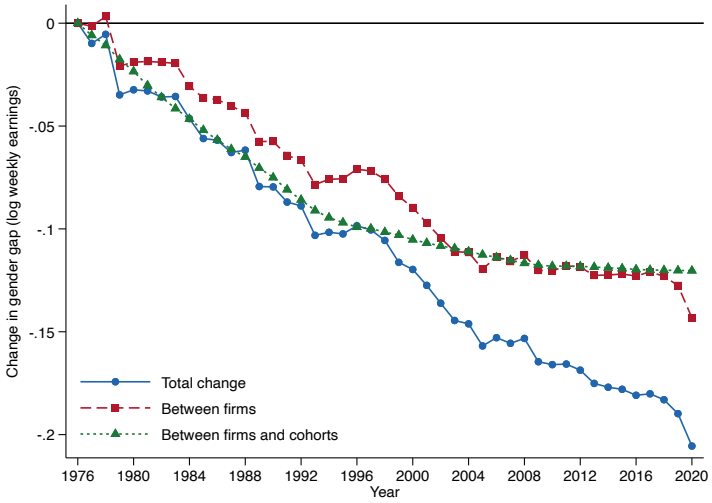
Panel D: Contribution by t_b - Italy

Notes: These figures computes a counterfactual scenario that shows what would have happened to the gender pay gap if the convergence in the entry earnings of men and women had stopped in year t_b. In this counterfactual exercise, the weekly earnings of cohorts who entered the labor marker after year t_b correspond to the average weekly earnings of cohorts who entered the labor market between t_b and the following two years. The panels differ in the definition of the benchmark year t_b: in Panel A, $t_b = 1976$, while in Panel B, $t_b = 2000$. Entry in the labor market is observed at 25 years old. Panel C and D show the ratio between the change in the wage gap predicted by this new counterfactual scenario and the one predicted by the old counterfactual scenario from Figure 3 when the benchmark year t_b moves between 1985 and 2010. The ratio is such that 100 implies that the decline in the gender gap predicted by the new counterfactual scenario perfectly matches the one predicted by the previous counterfactual scenario from Figure 3. *Sources:* In each year, the data pools information about all workers who were between 25 and 64 years old, worked in the private sector for at least 24 weeks year, earned strictly positive wages, and did not retire. Italy. Time period: 1976-2020. UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). USA. Time period: 1976-2019. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Figure 6: Change in the Gender Gap Between Sectors and Firms



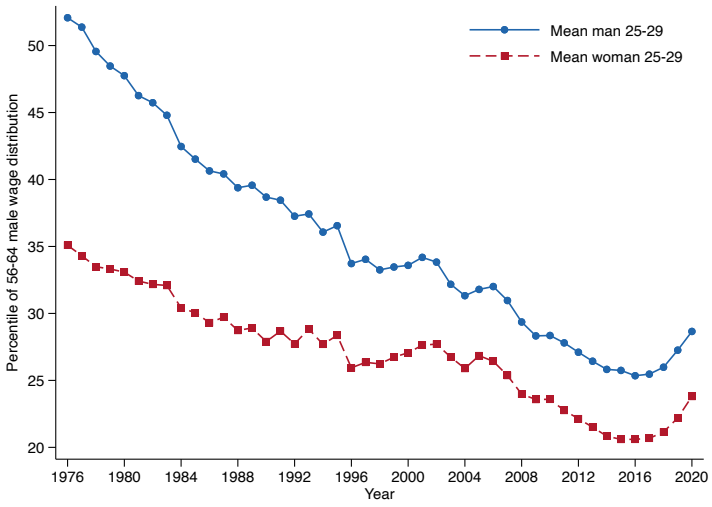
Panel A: Sectors



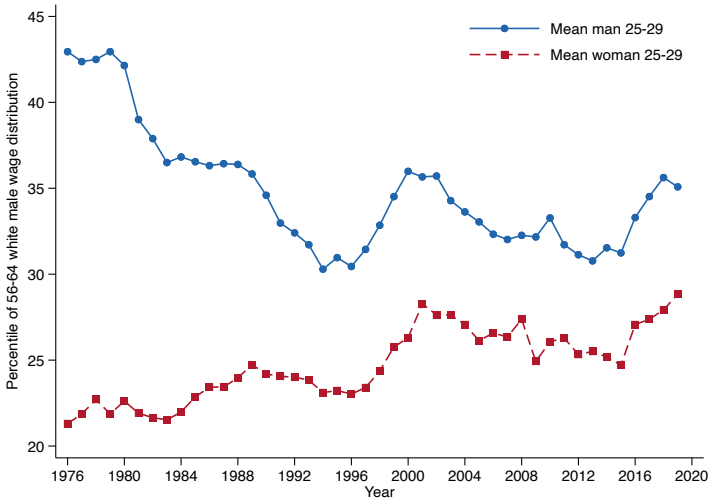
Panel B: Firms

Notes: Using Italian data, this figure shows how much of the total change in the gender gap took place between sectors and firms. In Panel A, we compute the between-sector component by assigning to workers the mean log weekly earnings in their 2-digit sector. Moreover, we compute the between-sector-and-cohort component by assigning to workers the mean log weekly earnings in their 2-digit sector observed at the time of entry in the labor market (25 years old). Panel B replicates this analysis replacing 2-digit sectors with firms. *Sources:* In each year, the data pools information about all workers who were between 25 and 64 years old, worked in the private sector for at least 24 weeks year, earned strictly positive earnings, and did not retire. *Sources for Italy:* UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS).

Figure 7: Positional Gains and Losses at Labor-Market Entry



Panel A: Italy



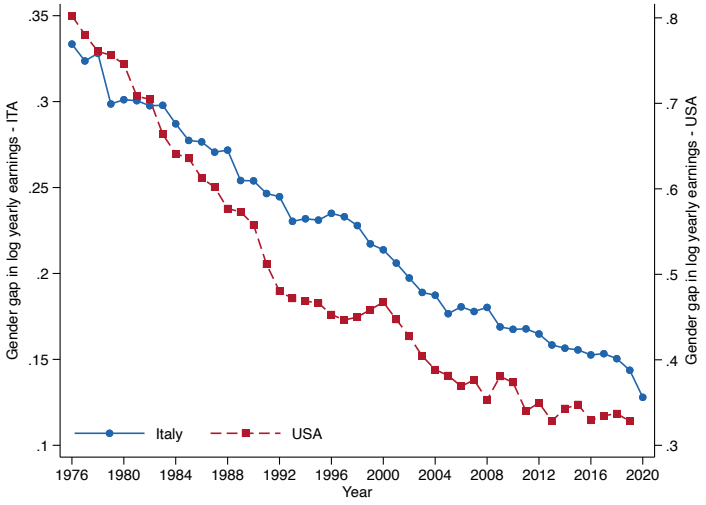
Panel B: USA

Notes: These figures show the average earning percentile of men and women at the time of entry in the labor market (between 25 and 29 years old), using the distribution of log weekly earnings of over-55-years-old (in the US, white) men as benchmark. *Sources:* In each year, the data pools information about all workers who were between 25 and 64 years old, worked in the private sector for at least 24 weeks year, earned strictly positive earnings, and did not retire. *Sources for Italy:* UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Sources for the United States*. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

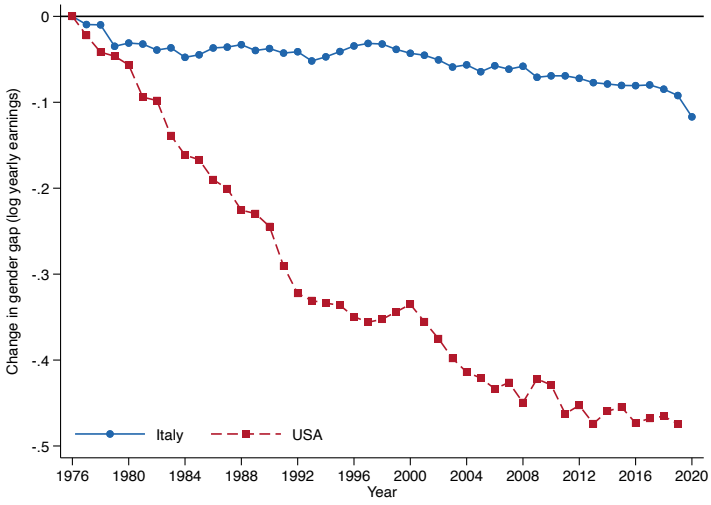
Online Appendix

A Additional Figures and Tables

Figure A1: Gender Gap in Yearly Labor Earnings



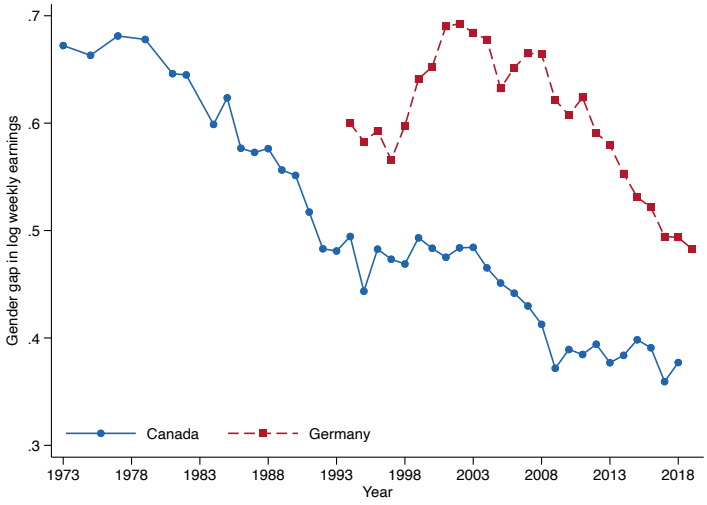
Panel A: Raw gap



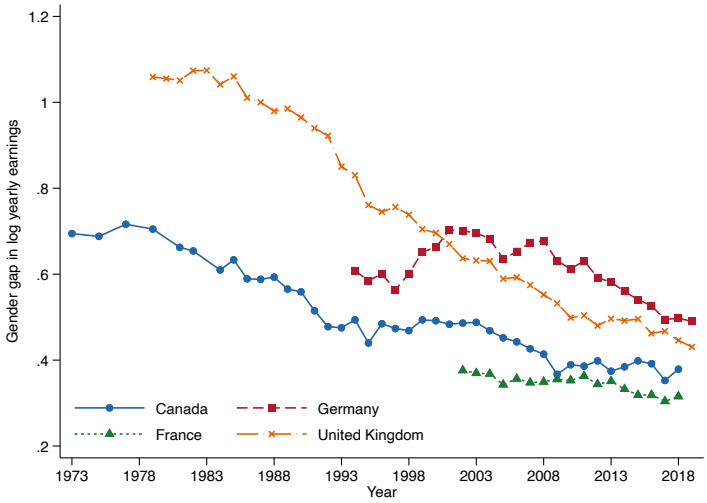
Panel B: Deviation from first year

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 deviation from the first year (1976). *Sources:* In each year, the data pools information about all workers who were between 25 and 64 years old, worked in the private sector for at least 24 weeks year, earned strictly positive earnings, and did not retire. *Sources for Italy:* UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Sources for the United States.* Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS.

Figure A2: Gender Gap in Other High-Income Countries



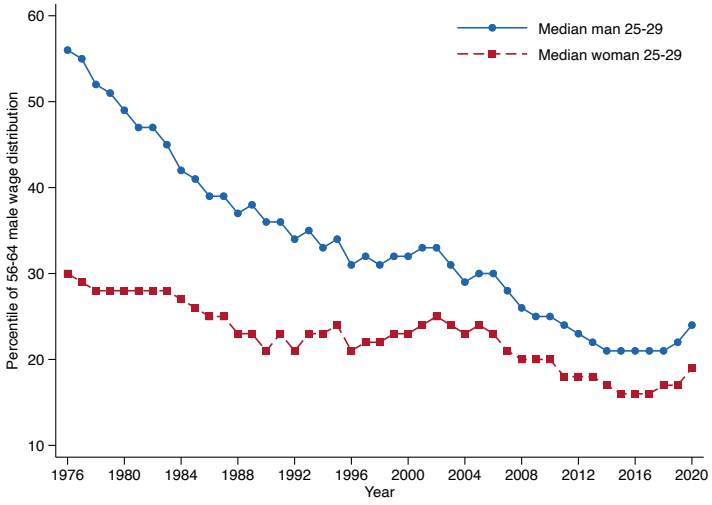
Panel A: Log weekly earnings



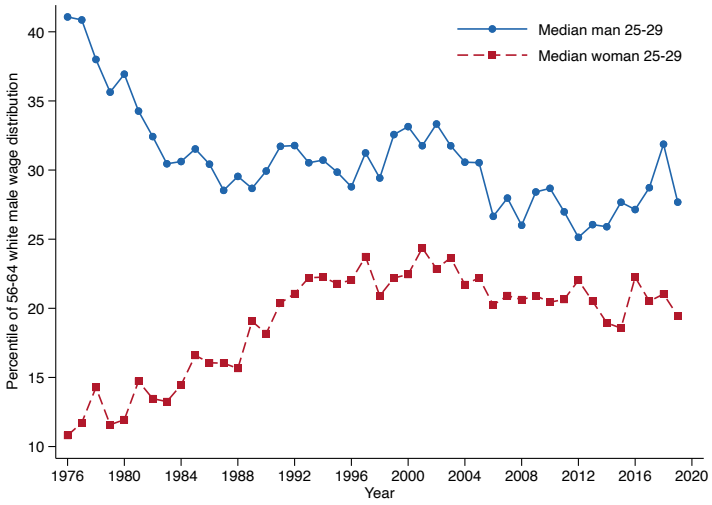
Panel B: Log yearly earnings

Notes: Panel A plots the trend in the raw mean gender gap in log weekly earnings, available in Canada and Germany. Panel B plots the trend in the raw mean gender gap in log yearly earnings, available in Canada, Germany, France, and the United Kingdom. *Sources:* Whenever possible, the data pools information about all workers who were between 25 and 64 years old, worked in the private sector for at least 24 weeks year, earned strictly positive earnings, and did not retire. *Sources for LIS data:* Luxembourg Income Study (LIS) Database, last accessed on 06/01/2023 at https://www.lisdatacenter.org/.

Figure A3: Additional Results on Positional Gains and Losses at Labor-Market Entry



Panel A: Italy, weekly earnings
Median new entrant



Panel B: USA, weekly earnings
Median new entrant

Notes: Panels A and B show the median earning percentile of men and women at the time of entry in the labor market (between 25 and 29 years old), using the distribution of log weekly earnings of over-55-years-old (in the US, white) men as benchmark. *Sources:* In each year, the data pools information about all workers who were between 25 and 64 years old, worked in the private sector for at least 24 weeks year, earned strictly positive earnings, and did not retire. *Sources for Italy:* UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Sources for the United States*. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. 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: Social Security data								
Italy (1976-2020)	45	383,256,256	32,544,238	5,243,930	Yes	Yes	No	Yes
Panel B: Current Population Survey								
United States (1976-2019)	44	1,996,548	-	-	Yes	Yes	Yes	Yes
Panel B: Survey data from the Luxembourg Income Study (LIS) Database								
Canada (1973-2018)	41	1,053,226	-	-	Yes	Yes	No	Yes
France (2002-2018)	26	255,671	-	-	Yes	No	No	No
Germany (1994-2019)	17	551,373	-	-	Yes	Yes	No	Yes
United Kingdom (1979-2019)	41	545,896	-	-	Yes	No	No	No

Sources for Italy: Database UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Sources 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). *Sources for France:* Tax and Social Incomes Survey. *Sources for Germany:* German Socio-Economic Panel. *Sources for United Kingdom:* Family Expenditure Survey (1991 and earlier); Family Resources Survey (1994 and later).

Table A2: Gender Gap and Workforce Composition

	Controls							Total change in gender gap	
	Age	PT	Edu	Child	Temp	Nat	Dis	Baseline	Residuals
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Log weekly earnings									
Italy (1976-2020)	Yes	Yes	No	No	Yes	Yes	No	-0.206	-0.256
United States (1976-2019)	Yes	Yes	Yes	Yes	No	Yes	No		
Canada (1973-2018)	Yes	Yes	Yes	Yes	No	No	No	-0.295	-0.188
Germany (1994-2019)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-0.117	-0.098
Panel B: Log yearly earnings									
Italy (1976-2020)	Yes	Yes	No	No	Yes	Yes	No	-0.117	-0.254
United States (1976-2019)	Yes	Yes	Yes	Yes	No	Yes	No		
Canada (1973-2018)	Yes	Yes	Yes	Yes	No	No	No	-0.316	-0.202
Germany (1994-2019)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-0.118	-0.095
France (2002-2018)	Yes	Yes	Yes	Yes	Yes	Yes	No	-0.061	-0.002
United Kingdom (1979-2019)	Yes	Yes	No	Yes	No	No	Yes	-0.628	-0.176

Notes: We regress log weekly and yearly earnings on several worker characteristics, separately in each country and available year. Then, we use the earning residuals from these regressions to compute the change in the gender gap between the first and last available year. "Age" denotes a quadratic polynomial of age. "PT" is a dummy variable for part-time workers. "Edu" is a dummy for college education. "Child" is a dummy for workers with at least one child. "Temp" is a dummy for temp workers. "Nat" denotes a dummy for nonimmigrant workers (white workers in the United States to control for race, instead). "Dis" controls for disability status. *Sources:* In each year, the data pools information about all workers who were between 25 and 64 years old, worked in the private sector for at least 24 weeks year, earned strictly positive earnings, and did not retire. Italy. *Sources for Italy:* UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Sources for the United States*. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS. *Sources for LIS data:* Luxembourg Income Study (LIS) Database, last accessed on 06/01/2023 at https://www.lisdatacenter.org/.

Table A3: 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	Last year	Change	Earnings at 25 y.o.	Earnings at 28 y.o.	Earnings at 25-29 y.o.	Earnings at 30 y.o.	Total change (last y. - 2015)	Earnings at 25 y.o.	Total change (last y. - 1980)	True entry earnings
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A: Log weekly earnings											
Italy (1976-2020)	0.333	0.128	-0.206	-0.242	-0.219	-0.219	-0.198	-0.028	-0.018	-0.173	-0.281
United States (1976-2019)	0.780	0.321	-0.459	-0.586	-0.591	-0.562	-0.607	-0.020	-0.031	-	-
Canada (1973-2018)	0.672	0.377	-0.295	-0.395	-0.525	-0.345	-0.327	-	-	-	-
Germany (1994-2019)	0.600	0.483	-0.117	-0.377	-0.277	-0.286	-0.228	-	-	-	-
Panel B: Log yearly earnings											
Italy (1976-2020)	0.350	0.233	-0.117	-0.204	-0.165	-0.166	-0.128	-0.037	-0.020	-0.086	-0.269
United States (1976-2019)	0.803	0.328	-0.474	-0.501	-0.595	-0.545	-0.655	-0.020	-0.028	-	-
Canada (1973-2018)	0.695	0.379	-0.316	-0.425	-0.644	-0.372	-0.402	-	-	-	-
France (2002-2018)	0.608	0.490	-0.118	-0.385	-0.276	-0.286	-0.218	-	-	-	-
Germany (1994-2019)	0.376	0.316	-0.061	-0.144	-0.115	-0.120	-0.100	-	-	-	-
United Kingdom (1979-2019)	1.059	0.431	-0.628	-0.751	-0.618	-0.656	-0.598	-	-	-	-

Notes: Columns 1 to 3 show the change in gender gap between the first and last available year for each country. Columns 3 to 7 show the between-cohort component of the change in the gender gap between the first and last available year (Equation (1)). In this counterfactual scenario, we assign to each cohort (defined as a combination of year of entry in the labor market and gender) its mean weekly earnings at labor-market entry in each year. Columns 3 to 7 differ in the definition of earnings at labor-market entry: mean weekly earnings at 25 years old, at 28 years old, between 25 and 29 years old, and at 30 years old. Cohorts who are above these age thresholds in the first available year are assigned their mean weekly earnings in the first year of the sample. In column 9, we assign to each cohort its mean weekly earnings at 25 years old, starting when it is observable for all cohorts (2015). This analysis can only be performed with data from Italy and the United States. Column 8 shows the total change in gender gap between 2015 and the last available year. In column 11, whenever available, we assign to all cohorts their true weekly earnings at entry, rather than their mean weekly earnings at 25 years old. This analysis is available only for Italy and starts in 1980. Column 10 shows the total change in gender gap between 1980 and the last available year. *Sources:* In each year, the data pools information about all workers who were between 25 and 64 years old, worked in the private sector for at least 24 weeks year, earned strictly positive earnings, and did not retire. Italy. *Sources for Italy:* UNIEMENS, Istituto Nazionale della Previdenza Sociale (INPS). *Sources for the United States*. Integrated Public Use Microdata Series, Current Population Survey: Version 10.0 [dataset]. Minneapolis, MN: IPUMS. *Sources for LIS data:* Luxembourg Income Study (LIS) Database, last accessed on 06/01/2023 at https://www.lisdatacenter.org/.