ISSN 2279-9362

Collegio Carlo Alberto

Beneath the Surface: the Decline in Gender Injury Gap

Tiziano Razzolini
Roberto Leombruni
Giovanni Mastrobuoni
Mario Pagliero

No. 286
December 2012

Carlo Alberto Notebooks

www.carloalberto.org/research/working-papers

Beneath the Surface: the Decline in Gender Injury Gap

Tiziano Razzolini*, Roberto Leombruni†,
Giovanni Mastrobuoni,‡and Mario Pagliero§

Abstract

There is little evidence on the joint evolution of gender differences in wages and other job amenities. We analyze gender differences in wages *and* workplace safety using 9 years of Italian administrative micro-level data. We document that a decline in the gender wage gap was accompanied by a narrowing of the gender gap in injury risk and an increased concentration of injuries among low-skilled female workers. A decomposition *a la* DiNardo et al. (1996) suggests that the main driver of the reduction in the wage gender gap is the sorting of workers across sectors and occupations, while the reduction in the injury gender gap and the increased concentration of injuries among low skilled female workers can be attributed to changes in unobservable job tasks and worker skills.

Keywords: Gender gap, workplace injury, job amenities, wage differentials, polarization of hazardous tasks

JEL classification codes: J16, J28, J31

*University of Siena, tiziano.razzolini@unisi.it.
†University of Turin and Labor.
‡Collegio Carlo Alberto and CeRP, giovanni.mastrobuoni@ccarloalberto.org.
§University of Turin and Collegio Carlo Alberto, mario.pagliero@carloalberto.org.

1 Introduction

The gender gap measures systematic differences in outcomes that men and women experience in the labor market. Several developed countries are witnessing a secular trend toward a reduced gender wage gap (see among others Blau, 1998; Blau and Kahn, 1997; O'Neill and Polachek, 1993), including Italy (Bono and Vuri, 2011). However, whereas wages are quite readily observed and quantified, other job amenities are not so easily captured by labor force surveys or administrative data and remain beneath the surface, unobserved (see for an overview Altonji and Blank, 1999; Goldin, 1994).

The main contribution of this paper is the joint analysis of gender differences in wages *and* workplace safety, the latter of which remains one of the most important, yet understudied, job amenities. In 2009, over 2.8 million accidents involving more than three days lost occurred in the EU (Eurostat 2012). During the same year, 3.8 thousand workers died as a consequences of workplace accidents in the EU, about 4.5 thousand in the U.S. (Bureau of Labor Statistics 2012). Our investigation is based on matched employer-employee administrative data containing detailed information about wages and working histories in Italy over the period 1994-2002.

A novel feature of this data set is the availability of individual-level information on the occurrence of workplace injuries, which allows simultaneous study of the distribution of wages and workplace injuries.

We document that: i) there is a significant gender gap both in wages and in work-related injuries. In 1994, wages for men were on average 20 percent higher than for women and men experienced approximately three times as many work-related injuries as women; ii) both these gender gaps diminished over time. Between 1994 and 2002, the difference in wages between men and women decreased by 21 percent, while the difference in the probability of injury decreased by 37 percent; iii) convergence in the probability of injury results from a greater decline for men (-29 percent) than for women (-11 percent); iv) compensating wage differentials do not explain the reduction in the gender wage gap, as female work injuries are increasingly concentrated among low paid workers, for whom gender wage gaps tend to be more persistent.

In particular, women are becoming more active in riskier jobs, and overall their wages are converging toward male wages. Looking at the aggregate data, this may seem to be driven by compensating wage differentials. However, variations in the concentration of injuries along the wage distribution of male and female workers show that this is not the case. Male injuries are concentrated among low paid workers, and this concentration is fairly stable over time. For women, the concentration of injury among low skilled workers only became apparent towards the end of the sample period, when the majority of injuries affecting women were also concentrated in the lowest part of the wage distribution. Only in the top two quartiles of the wage distribution did female workers experience a clear decline in injuries. This is consistent with women branching out from more traditional jobs and sorting into more diverse jobs. On the one hand, low-skilled (and lower-paid) women have not enjoyed a decline in work injury rate, thus reducing the gender gap in the average probability of injury. On the other hand, highly-skilled women, who are less at risk of injury, have experienced a substantial increase in wages, accounting for the overall reduction in the gender gap in average wages.

Our results are in line with Hamermesh (1999), who argues that, despite a positive risk premium, unobserved differences in human capital may lead more skilled workers to earn more and to trade-off part of their earnings for workplace safety, a normal good. Hence, these unobserved skills create a negative correlation between wages and injuries. Our results are also in line with the transitioning of women from traditionally "female" professions (e.g. blue collar workers in the textile sector), which are usually concentrated in the middle of the wage distribution (Blau and Kahn, 2000), toward "male" professions at both the lowest and highest ends of the earnings distribution. The entry of new cohorts of more educated and skilled women has allowed some of them to advance in the job hierarchy. Moreover, changes related to international trade and especially to the digital revolution may have influenced the price of skills, inducing a stronger polarizing effect for women (see Autor et al., 2003, Bacolod and Blum, 2010 and Black and Spitz-Oener, 2010). The increased price of cognitive skills relative to the price of routine and manual tasks, partly driven by computerization, are believed to influence those in the central part of the wage distribution, and such workers are often women. The drop in the price of low skilled labor led to a reduction in all job amenities, including workplace safety, and to a concentration of risk at the bottom of the wage distribution. Low-paid women bear the brunt of the costs of the decline in the gender injury gap.

We use the DiNardo et al. (1996) (hereafter DFL) decomposition to identify the determinants of the changes in gender gaps between 1994 and 2002, generating counterfactuals.

Compositional changes (changes in observable characteristics) capture part of these changes, and are mainly due to the switching of jobs across sectors and occupations. Changes in unobserved characteristics led to a polarization of female jobs much like that of males'. On the one hand, the lowest paid female workers now bear most of the risk of injury. On the other hand, the highest paid female workers face better prospects in terms of both wages and injury risk.

The paper is organized as follows. The next section describes the data and provides some descriptive statistics. Section 3 presents the methodological framework. Section 4 reports the empirical results, and Section 5 concludes.

2 The data

Data availability has been a long-standing issue in the literature on job amenities, and injury risk in particular. We overcome this issue by using administrative data on a 1 : 90 random sample of Italian workers, the Work Histories Italian Panel (WHIP), linked with administrative records from the Italian Workers' Compensation Authority (INAIL), covering the years 1994-2002 (Bena et al., 2012). Overall, the data set includes about 120,000 individual records for each of the 9 years in the sample. This matched employer-employee data set provides information on worker and job characteristics (age, sex, place of birth, type of occupation, sector, size of firm, number of weeks worked in a year, part-time job, earnings), the number of work-related injuries and their exact description, and the days of work lost due to such accidents. Physicians report and certify the diagnosis and prognosis of the accidents. Hence, our data set provides an exceptionally rich source of information which we use to analyze the joint distribution of (deflated) weekly earnings and workplace injuries.

Despite this wealth of information we are upfront about three limitations of our data. First, a precise estimation of injury risk is only possible for employees in the non-agricultural private sector. This is because employees in the the public and the agricultural and fishing sectors are not included in the database, while information about the self-employed, although included, is not adequate both for measuring the risk-exposure (worked weeks are not recorded) and the outcome (self employed without employees in the trade sector are not required to enroll in public insurance).

Second, information on workplace injuries is currently available only for the 9-year period between 1994 and 2002. However, although relatively short, this period witnessed important economic and social changes and provides significant variability in wages, injury probabilities, and injury index.[1] Another reason why 9 years might not be such a short period when studying workplace injuries is that most injuries happen at the beginning of a job spell (see Table 1) and the Italian workforce changed significantly during the 1994-2002 period. Female labor force participation increased significantly, and new cohorts of workers entered the labor market. These changes were so pronounced that 42 percent of women and 35 percent of men working as employees in 2002 were still out of employment in 1994.

Third, like many administrative records that are used to compute social security benefits, our data set has no information on education, since education does not enter the benefit formula directly. Fortunately, the data does include information on whether the worker is a blue or white collar worker or a manager, which tends to be highly correlated with education.

Tables 2 and 3 describe the evolution of our main controls, while the rest of this section focuses on the two main outcomes of interest, namely wages and injuries. Table 2 describes

[1]In 1992 a severe recession forced the Italian government to devalue the lira (the currency lost almost 40 percent of its value with respect to the German Mark) and abandon the European Exchange Rate Mechanism. The old political parties disappeared in the midst of widespread corruption charges against their leaders, and a government of technocrats was formed. New parties and new leaders entered the new political scene, called "2nd Republic," in 1996. Despite these changes, economic growth saw an unprecedented low, with average yearly growth rates close to 1.5 percent.

the participation rates in industries in 1994 and 2002. There are several changes in labor force participation across sectors. The most evident are the 3 percentage point increase for males and the almost 6 percentage point increase for females in the financial intermediation sector. Given that temporary help agencies (agencies that rent out workers on a non-permanent basis to firms) are recorded as financial intermediaries, part or even most of this increase might be attributed to these agencies. Most of these workers seem to come from traditional sectors like vehicle manufacturing for men (-1 percentage point) and the textile industry for women(-6 percentage points). Changes in the textile industry generate the largest convergence between male and female occupation across sectors. This industry almost exclusively employs women and has been constantly declining in size since the 1990s.

Table 3 shows that there were no major changes between 1994-2002 in the proportion of white and blue collar employees or in the geographical location of jobs, either for female or for male workers. Both men and women experienced a substantial increase in part-time jobs. Interestingly, the participation rate of male workers in large firms declined over time, whereas the female participation rate increased. The last two columns show that the gender differences in the average age of workers has been declining, mainly due to the increase in the age of retirement for women. Although retiring much later than in the past, women are increasingly employed in part-time jobs, so their tenure in office is decreasing in comparison to men.

Next we describe the evolution of our first main variable of interest, wages. Figure 1 displays the evolution of the 10th, 25th, 50th 75th and 90th percentiles of the log of deflated weekly earnings (annual earnings divided by the number of weeks worked per year) for male and female workers over the period 1994-2002. Wages include overtime payments and are deflated to constant 1995 values. The changes in female wage distribution are more pronounced than the variations for males, especially at the top and bottom of the distributions.

The numbers are shown in Table 4, where we focus on the distribution of log weekly wages in 1994 and in 2002 (the 2002 "prime," year represents a counterfactual year that we will discuss later). Overall, wages have been stagnant in the middle and increasing at the tails of the wage distribution. Women in the top percentiles of the distribution underwent the largest increase in wages, although women in all quantiles saw some improvement. This was not the case for men. Men between the 10th and 75th percentile received lower wages in 2002 than 9 years earlier.

The Gini index and the Theil index show that the relatively more stagnant evolution of wages in the middle of the wage distribution has led to increasing inequality.

Regarding our second main variable of interest, workplace injuries, we use two slightly different measures. To account for differences in exposure to workplace risk the weekly probability of injury is computed by dividing the likelihood that a worker was injured at least once during the year by the number of weeks worked during that year.[2] We also compute an injury index, defined as the annual number of days lost due to an injury divided by the number of weeks worked during the year. This index depends not only on the occurrence of an injury but also on its severity. To facilitate interpretation of the results, we report the probability of injury and the injury index on a yearly basis, i.e., the probability of injury in a year and the number of days lost per year.

Figure 2 and Table 5 show the evolution of injuries over time and also by wage quartile. The overall risk of injury for male workers dropped from 8.7 percent in 1994 to 6.2 percent in 2002. The corresponding decline among female workers was considerably smaller, from 2.8 percent

[2]The proportion of workers experiencing more than one injury within any given year is negligible (0.2 percent). The number of weeks worked is adjusted to account for part-time in order to provide an accurate measure of exposure to workplace injury risk.

to 2.5 percent. A similar picture emerges from the injury index, which fell from 2.8 to 1.8 for men and from 0.9 to 0.7 for women. Despite the noise, Figure 2 shows that the injury index is decreasing for the lowest male quartile and the highest female quartile. Table 6 describes the changes in the average gender wage and injury gap. The wage gender gap decreased from 0.201 to 0.158 (-21 percent) and the gender gap in injury probability from 0.059 to 0.037 (-37 percent). The gap in the injury index also declined during the same period (from 1.878 to 1.165, a reduction of 38 percent). The next section introduces a semi-parametric strategy for testing whether the evolution of observable characteristics can explain these patterns in wages and injuries.

3 Methodological framework

Given the well-known heterogeneity in risk across sectors, the pattern in injury risk and wages may simply be a result of changes in the participation rates across sectors (as documented earlier, for example, in Table 2). Figure 3 plots the change in the percentage of workers in each sector between 1994 and 2002 against the change in the average injury index in each sector (normalized to sum up to one). Male and female workers' participation increased in higher risk sectors in 2002. The slope of the fitted values for women is twice as large as that for men, indicating that this phenomenon is particularly strong for female workers. However, sectorial composition is only one of many potential driving forces.

More generally, changes over time in the distribution of injuries and wages and their joint distribution can be ascribed to changes in several observable characteristics of the workers (age, tenure, part-time work, industrial sector, occupation, region of work, region of birth), or unobservable factors, such as the price of skills, the workers' attitudes towards risk and the assignment of hazardous tasks and wages among workers with the same characteristics and within a specific occupation. Hence, the observed changes in the gender gaps (in wages and workplace injuries) can also be decomposed into these two components. To do so, we estimate the counterfactual densities (indicated by a prime ') of wages and injuries that would have been observed at the end of the sample period (2002) had the observed characteristics been distributed as they were at the beginning of the period (1994).

The difference in the statistics based on the counterfactual and the observed density in 1994 represents the impact of changes in unobservable factors affecting wages and workplace injury risk (for a given type of worker and job), but also changes in workers' attitudes towards risk and changes in the assignment of tasks within a specific type of occupation. The difference between the observed density in 2002 and the counterfactual one (denoted by 2002') indicates the impact of changes in observable characteristics. This reveals how changes in wages and injuries are driven by variations in labor force participation and by the sorting decisions of workers into different types of activities. To compute the counterfactual density, we adopt the reweighting approach of DiNardo et al. (1996). In our analysis, we apply this procedure to create counterfactual population measures of a single outcome variable (i.e. wages or injury indexes, *separately*) or counterfactual measures based on the joint distribution of wages and workplace injuries.

In the construction of the counterfactual measure for a single variable of interest the notation is as follows. Consider a sample of N workers, with $n = 1, ..., N$. Each worker is characterized by individual observed characteristics (y, x), where the vector x includes socioeconomic variables to condition on in the construction of the counterfactual. We define a statistic $I(t)$ at time t, based on the function $g(y)$ as

$$I(t) = \int g(y)dF(y|t) = \tag{1}$$

$$= \int \int_{x\in\Omega} g(y)dF(y,x|t), \tag{2}$$

where $F(y,x|t)$ represents the conditional joint distribution of the outcome variable y and characteristics x at time t. As shown by DiNardo et al. (1996), counterfactual scenarios can be constructed by adopting an appropriate reweighting procedure. One can build the counterfactual population measure $I(2002, 1994)$, that is, the population measure over an outcome of interest y that would have occurred had the conditional distribution $F(y|x)$ been the same as in 2002, and had the distribution of x been the same as in 1994. [3] The counterfactual

[3]We have also created counterfactual scenarios focusing on a subset of covariates and changing the order in the decomposition as in DFL. In all these specifications the differences between the actual and counterfactual 2002 distributions were mainly driven by changes in the distribution of sectors and occupations. Since the inclusion of other covariates did not affect our results, we only report the specification in which all observations are reweighted to replicate the distribution of covariates in 1994.

$I(2002, 1994)$ can then be computed using a reweighting procedure:

$$I(2002, 1994) = \int \int_{x \in \Omega} g(y) f(y|t_y = 2002) dF(x|t_x = 1994) = \tag{3}$$

$$= \int \int_{x \in \Omega} g(y) f(y|t_y = 2002) \Phi_x(2002, 1994) dF(x|t_x = 2002) \tag{4}$$

where the weights are:

$$\Phi_x(2002, 1994) = \frac{dF(x|t_x = 1994)}{dF(x|t_x = 2002)}. \tag{5}$$

The sample weights can be derived and estimated using Bayes' rule,

$$\hat{\Phi}_x(2002, 1994) = \frac{\hat{Pr}(t_x = 1994|x)\hat{Pr}_{2002}}{\hat{Pr}(t_x = 2002|x)\hat{Pr}_{1994}}, \tag{6}$$

where $\hat{Pr}(t_x = 1994|x)$ and $\hat{Pr}(t_x = 2002|x)$ are the probabilities for an observation to belong to 1994 or 2002, conditional on the covariates x; $\hat{Pr}_{1994}$ and $\hat{Pr}_{2002}$ are the unconditional probabilities for an observation to belong to 1994 or 2002, respectively.[4] This procedure can easily be extended to compute statistics based on the joint distribution of two outcomes of

[4]As noted by DiNardo (2002), the reweighting procedure is a function of the propensity score and two constants.

interest, namely wages and injury indexes. Equation 1 can be modified by specifying $g(.)$ as a function of both wages w and injury index inj and by replacing the conditional density function $f(y|t_y = 2002)$ in equation 3 with the conditional joint density function $f(w, inj|t_y = 2002)$. We can thus compute the counterfactual joint distributions of wages and injury probabilities and the related statistics such as concentration curves (O'Donnell et al., 2008) and concentration indexes (Wagstaff et al., 1991 and Kakwani et al., 1997), which are discussed later in Section 4.3. The weights from equation 6 can be used to construct these counterfactuals, as both measures can be computed using sampling weights (O'Donnell et al., 2008).

4 Evidence of gender differences

4.1 The magnitude of the gender *wage* gap

We can now turn from our description of counterfactual outcomes to the description of Tables 4 and 5. A comparison of the actual 2002 wage distribution and its counterfactual with covariates distributed as in 1994 reveals that the (observed) compositional effect has penalized male workers in every quantile of the wage distribution. Changes in unobserved factors alone (age, tenure, part-time work, industrial sector, occupation, region of work, region of birth) would have led to much higher wages than observed in 2002, especially at the tails of the distribution. Moreover, changes in unobserved factors alone would have led to a more homogenous wage distribution than observed in 2002 in terms of Gini and Theil indices.

Female workers experienced an average wage increase of 2.5 percent between 1994 and 2002. The wages of women in the 90th and 95th percentiles rose by much more: 5 and 9 percent respectively. When disentangling the effect due to unobservables from the effect of changes in the distribution of job and worker attributes, we find two distinct effects for the portions of the wage distribution below and above the median. The changes in the distribution of unobservables would have led, in the counterfactual 2002 distribution, to a larger increase in wages below the median. Variations in the distribution of the observed covariates have partly limited this growth. Despite improvement in female wages due to unobservables, sectors and occupations where females were more likely to be employed and better paid may have been forced out of the market by internationalization and global competition. The textile sector is the clearest example of this phenomenon. Conversely, the wage increase in the upper part of the distribution can be explained by both changes in unobserved and observed characteristics. Figures 4 to 6 provide a more complete picture of the different impact of observables and unobservables on the 2002 wage distribution for male and female workers. Figure 4 shows that observables reduced wages of male workers at each percentile, and that changes in wages were positive only at the lowest and at the highest percentiles. Women also experienced larger increases in the tails of the wage distribution but, unlike men, increases occurred across the whole spectrum. The other difference is that changes in observable characteristics partially explain the rise among high-skilled women and the lower rise among low-skilled ones. Figure 5 describes the difference in percentiles of the wage distribution between men and women. The gender gap is more pronounced at the lowest and the highest wage percentiles, with changes in observables explaining most of the almost parallel downward shift in the gap. Figure 6 shows that, apart from the very extremes, this shift is almost uniform; however, the most skilled women experienced the largest reduction in the gap, while the lowest skilled women actually experienced an increase in the wage gap.

Overall, the actual and counterfactual values of the Gini and Theil indices in Table 4 indicate that these variations in the wage distribution of women have led to a higher level of inequality to be ascribed with almost equal weights to compositional effects and to changes in unobserved characteristics. Table 6 provides exact figures for the evolution of the gender wage gap. The unconditional average gender wage gap was 20 percent in 1994 and approximately 4 percentage points lower in 2002 (a 20 percent reduction). The gender gap narrowed substantially in the upper part of the wage distribution, where the gap was considerably larger. At the 95th (5th) percentile, it fell from 36 to 30 (17 to 15) percent. On average, the decline in the gender wage gap was mainly determined by changes in observed characteristics (2002 vs. 2002').

4.2 The magnitude of the gender *injury* gap

Wages are not the only important job characteristic; job safety is certainly another. Counterfactuals will tell us whether the large decline in injury risk among men is driven by compositional changes or by other unobserved variables. Table 5 shows a 29 percent reduction in injury risk for male workers and an 11 percent reduction for females. The counterfactual results show that unobservables drive these changes. Similarly, Table 6 shows that gender differences in the probability of injury and the severity of the accidents fell mainly because of unobserved variables. Changes in the composition of jobs, occupations and sectors has counteracted the effects of unobserved variables only to a very limited degree.

Figure 7 shows that the same is true regarding the severity of injuries. Figure 8 makes this even more explicit by showing the differences between the male and female distribution of days lost due to workplace injuries. Finally, Figure 9 reports the difference-in-difference between 2002 and 1994, and the male and female distribution of days lost due to injury. The narrowing of the injury gap is clearly driven by changes in injuries that lead to short term absences from work.

4.3 Gender wage and injury gap together

After describing the marginal distribution of wages and injuries, together with their actual and counterfactual evolution, we now describe their joint distribution. The purpose of joint analysis is to investigate whether the observed reductions in the wage and the injury gap can be interpreted as the consequence of reductions in gender differences in compensating wage differentials (conditional on the observable characteristics).

Figure 2 shows the evolution of the average injury index within each wage quartile for male and female workers. The pattern described in the figure reveals the increased concentration of injuries among low wage female workers. To determine whether these patterns are driven by changes in observable characteristics, the probability of injury is computed for male and female workers by wage quartile in 1994, 2002 and in the 2002 counterfactual scenario (Table 7). The risk of injury declined for male workers in all quartiles, with the biggest difference in the first quartile (-37 percent). The probability of injury among female workers slightly increased for the first quartile, but remained stable for the second and considerably decreased for the 3rd and 4th. Observables explain little about the evolution of male injuries across all quartiles, but sometimes predict even larger changes in injury risks for women.

The corresponding evolution of the injury index (taking severity into account) is shown in Table 8. The 30 percent decline in injury probability in the first quartile of the male wage distribution is accompanied by a 42 percent decline in the injury index. This indicates that the severity of injuries decreased considerably in the lowest part of the male wage distribution. For female workers, there is almost no change in injury probability in the first wage quartile, while there is a 16 percent increase in the mean injury index, indicating again that the severity of injuries is driving the changes.

The severity of workplace injuries decreased for men and women in the upper part of the wage distribution. The 20 and 43 percent decrease in probability of injury among women in the 3rd and 4th wage quartiles was accompanied by a reduction in the mean injury index of about 61 and 40 percent, respectively. It appears that when fixing the distribution of covariates as in 1994, male and female workers in 2002 would have experienced fewer and less serious injuries in every quartile. In contrast, the compositional change between 1994 and 2002 in the distribution of covariates seems to have only slightly increased the number of workers subject to injury risk.

The last three columns of Table 7 and 8 describe the gender gap in workplace injuries by wage quartile. For the injury probabilities, the gap is always positive and, except for the 4th quartile, decreasing over time. The gender gap tends to be larger among low skilled workers. Similarly, the gender gap in the severity of injuries is generally positive, decreasing over time, and larger for low skilled workers. In 1994, women in the 4th wage quartile had a larger injury index then men, but the gap is not statistically significant and had shrunk to zero by 2002. The gender gap in the injury index decreased significantly (from 5.2 to 2.4) for workers in the

1st quartile of the wage distribution. The 2002 counterfactual values are very close to the 2002 observed ones, suggesting that observables are not the main driving force of these changes.

Quartile-specific measures of injury risk neglect intra-quartile differences in injury risk. A more comprehensive measure of inequality in the distribution is provided by the concentration curve (Wagstaff et al., 1991 and Kakwani et al., 1997), which describes the proportion of injuries that are attributable to the cumulative percentage of the sample ranked by wage. If wages and the probability of injury were uncorrelated, the concentration curve would correspond to the 45-degree line. If, instead, low wage earners bear more risk of injury, the concentration curve will fall above the 45-degree line.[5]

Figure 10 presents the concentration curves of injury probability by wage level for men and women in panel (a) and (b) respectively. The position of the concentration curves of men relative to the 45° line shows a permanent concentration of hazardous tasks among low paid workers, and a marginal reduction in concentration between 1994 and 2002. This decrease in

[5] A quantitative measure of the distribution of the injury risk y over income levels is provided by the concentration index that represents the area between the concentration curve and the 45° line. For a sample of n individuals, this index is

$$C = \frac{2}{n\mu} \sum_{i=1}^{n} y_i R_i - 1 \tag{7}$$

where μ is the mean injury index and R_i is the fractional rank of the *i-th* individual in the income distribution.

concentration cannot be explained by observable factors. Table 9 attests that such changes are statistically significant and that the counterfactual scenario in 2002 is identical to the actual one. The opposite is true for women. Wages and the probability of injury are as good as uncorrelated in 1994.

However, by 2002, the concentration curve resembles that of men, indicating a substantial shift in the probability of injury towards lower paid female workers, but again, changes cannot be attributed to observable factors. The *cfactual* line represents the counterfactual concentration curve that would have occurred in 2002 if all observed attributes were distributed as in 1994. The increase in the concentration of injury risk among low wage female workers seems to be mainly driven by changes in unobservable attributes (see also Table 9). This could be driven by an increased unobserved heterogeneity between high skilled and low skilled women that may have disproportionately increased wages and workplace safety at the top of the wage distribution. Similarly, a decrease in the relative price of low skills may have induced low skilled women to bear the risk associated with more hazardous tasks.

5 Concluding Remarks

Using a unique dataset, we study the narrowing of the gender gap in wages and workplace injury risk in Italy between 1994 and 2002. This decrease in gender differences is associated with rising inequality within the female wage distribution and an increased concentration of injuries among low-skilled female workers. While the reduction in the gender wage gap is largely explained by changes in the observed characteristics of workers and jobs, the increased concentration of injuries for female workers is largely due to changes in unobservable attributes.

Two possible explanations for the changes in observable characteristics of workers and jobs are skill-biased technological progress and the falling of barriers to career advancement for women. Over the past few decades, women have increasingly been moving from traditionally female professions to professions previously dominated by males. The entry of new, better skilled and educated cohorts has contributed to growing numbers of women in more highly paid positions (Blau and Kahn, 2000) and allowed some of them to advance in the job hierarchy. However, changes related to skill biased technological change might be driving increased wage inequality for women (Autor et al., 2003, Bacolod and Blum, 2010, Black and Spitz-Oener, 2010).

The falling of institutional and social barriers and the change in attitudes towards work-related risk within a specific occupation likely explain part of the wage increases in the lowest part of the distribution for females. Weichselbaumer and Winter-Ebmer (2007) show that the elimination of government regulations prohibiting women from working in unpleasant jobs or tasks requiring physical strength has lowered the gender wage gap. The computer revolution and growing international trade that characterized the 90s likely increased the price of cognitive skills relative to physical skills. Increased remuneration of skills might have induced more highly-skilled women to trade-off part of their salary for workplace safety, whereas low-skilled female workers, whose skills have been penalized by the IT revolution, had to change job and sector (observable changes in our data set), or to sort into different tasks within the same job and sector (unobservable changes), and bear most of the cost deriving from injury risk.

Changes in (unobserved) risk preferences or preferences for overtime and part-time work may also have affected the allocation of hazardous tasks within jobs. Epidemiological studies indicate that overtime work is highly correlated with the probability of injury. Hence, changes in female risk preferences may affect overtime work and the allocation of tasks within firms. Similarly, the fall in the number of children in Italian families may also have affected women's preferences for overtime and part-time work, and hence the allocation of hazardous tasks.

References

J.G. Altonji and R.M. Blank. Race and Gender in the Labor Market. *Handbook of labor economics*, 3:3143–3259, 1999. ISSN 1573-4463.

David H. Autor, Frank Levy, and Richard J. Murnane. The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333, November 2003.

Marigee P. Bacolod and Bernardo S. Blum. Two sides of the same coin: U.s.; inequality and the gender gap. *Journal of Human Resources*, 45(1), 2010.

Antonella Bena, Roberto Leombruni, Massimiliano Giraudo, and Giuseppe Costa. A new italian surveillance system for occupational injuries: Characteristics and initial results. *American Journal of Industrial Medicine*, 2012.

Sandra E Black and Alexandra Spitz-Oener. Explaining women's success: Technological change and the skill content of women's work. *The Review of Economics and Statistics*, 92(1):187–194, 04 2010.

F.D. Blau. Trends in the Well-Being of American Women, 1970-1995. *Journal of Economic Literature*, 36(1):112–165, 1998.

F.D. Blau and L.M. Kahn. Swimming Upstream: Trends in the Gender Wage Differential in the 1980s. *Journal of Labor Economics*, 15(1):1–42, 1997.

Francine D. Blau and Lawrence M. Kahn. Gender differences in pay. *Journal of Economic Perspectives*, 14(4):75–99, Fall 2000.

Emilia Del Bono and Daniela Vuri. Job mobility and the gender wage gap in italy. *Labour Economics*, 18(1):130–142, January 2011.

J. DiNardo, N.M. Fortin, and T. Lemieux. Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica*, 64(5):1001–1044, 1996.

John DiNardo. Propensity score reweighting and changes in wage distributions. Mimeo, University of Michigan, 2002.

Claudia Goldin. Understanding the Gender Gap: An Economic History of American Women. *Equal Employment Opportunity: Labor Market Discrimination and Public Policy*, 1994.

Daniel S. Hamermesh. Changing inequality in markets for workplace amenities. *The Quarterly Journal of Economics*, 114(4):1085–1123, November 1999.

Nanak Kakwani, Adam Wagstaff, and Eddy van Doorslaer. Socioeconomic inequalities in health: Measurement, computation, and statistical inference. *Journal of Econometrics*, 77 (1):87–103, March 1997.

Owen O'Donnell, Adam Wagstaff, Magnus Lindelow, Eddy Van Doorslaer, Adam Wagstaff, and Magnus Lindelow. *Analyzing Health Equity Using Household Survey Data: A Guide to Techniques and Their Implementation.* The World Bank, 2008.

J. O'Neill and S. Polachek. Why the Gender Gap in Wages Narrowed in the 1980s. *Journal of Labor Economics*, 11(1):205–228, 1993.

Adam Wagstaff, Pierella Paci, and Eddy van Doorslaer. On the measurement of inequalities in health. *Social Science & Medicine*, 33(5):545–557, January 1991.

Doris Weichselbaumer and Rudolf Winter-Ebmer. The effects of competition and equal treatment laws on gender wage differentials. *Economic Policy*, 22:235–287, 04 2007.

Figure 1: Percentiles of log of deflated weekly earnings for male and female workers in 1994-2002.

(a) Men (b) Women

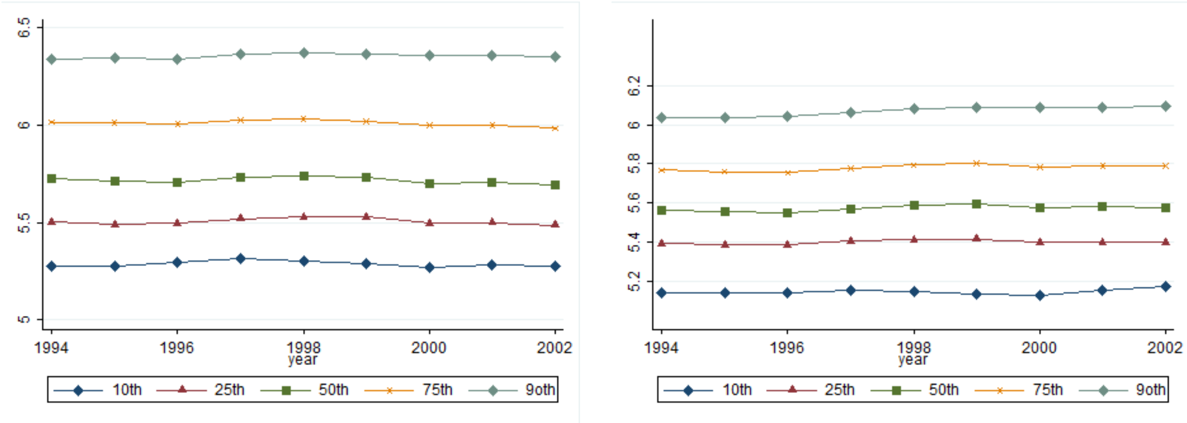
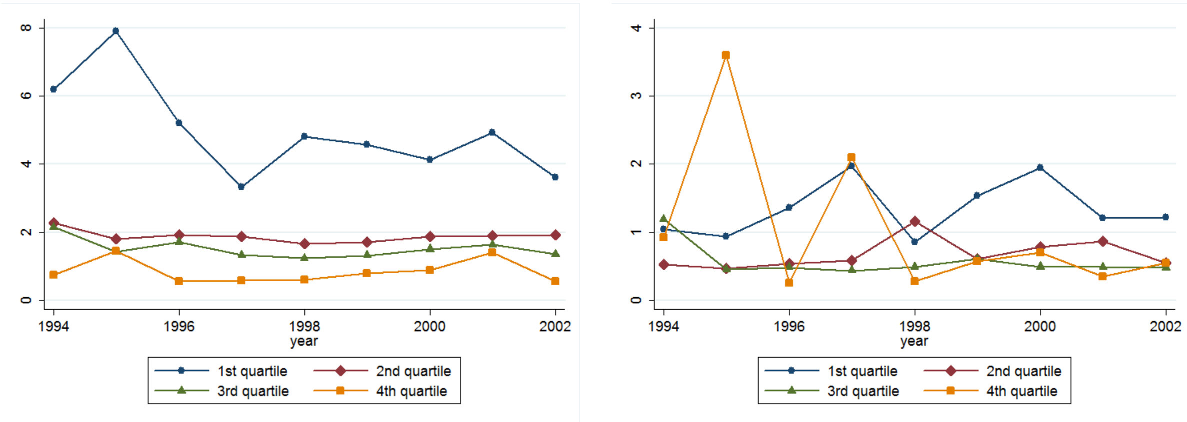


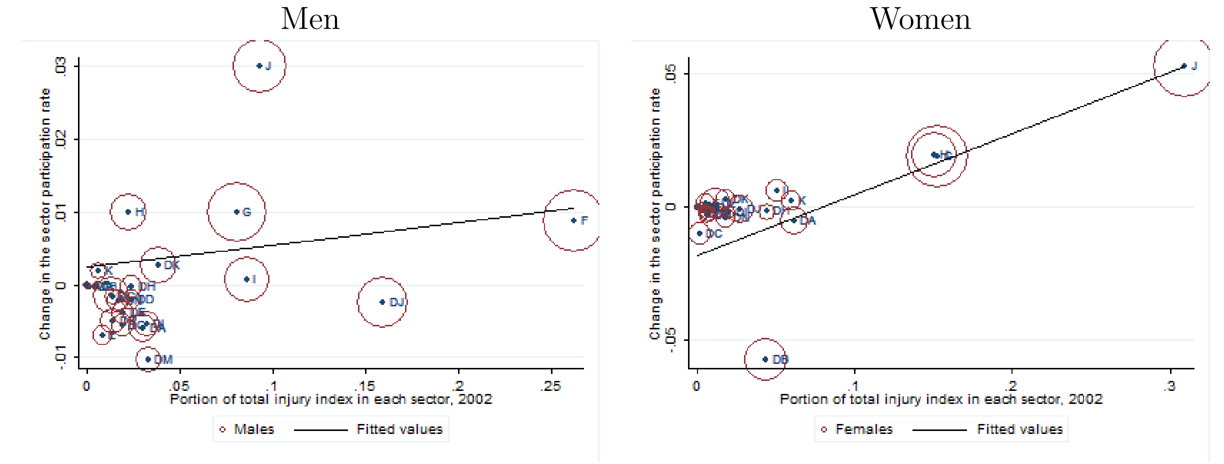
Figure 2: The average injury index by wage quartile.

(a) Men (b) Women



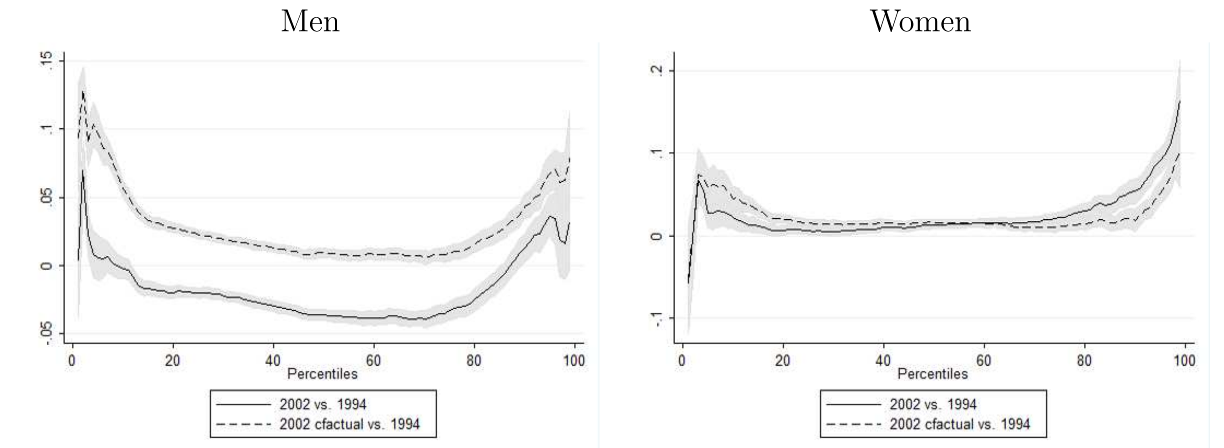
The series represent the evolution of the average injury index within each wage quartile for male and female workers.

Figure 3: Changes in workforce participation in each sector and injury risk in 2002.



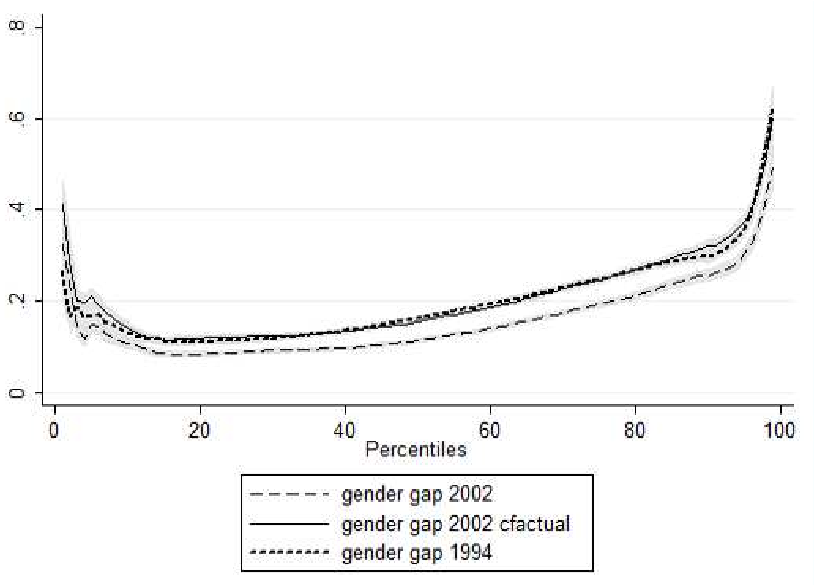
Note: The vertical axis represents changes in workers' participation rate in each sector. The horizontal axis represents the proportion of days lost due to injury in each sector. The diameter of the circles is proportional to the number of workers in each sector and the regression line is weighted based on that number.

Figure 4: Differences in the percentiles of the log wage distribution for male and female workers.



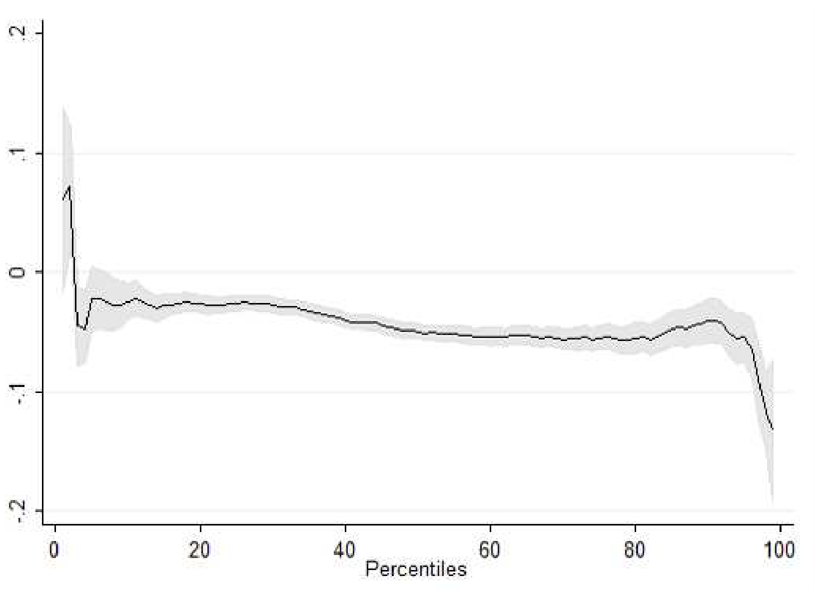
Note: The solid line shows the difference between the percentiles of the 2002 log wage distribution and the percentiles of the 1994 log wage distribution. The dotted line shows the difference between the percentiles of the 2002 counterfactual log wage distribution (computed assuming the same distribution of observable characteristics as in 1994) and the percentiles of the 1994 log wage distribution. Grey areas represent 95 percent confidence intervals.

Figure 5: Gender gap in log wage distribution for male and female workers.



Note: The dotted line shows the gender gap at every percentile of the 1994 wage distribution. Similarly, the dashed line represents the gender gap in 2002 and the solid line the gender gap in the 2002 counterfactual distribution. Grey areas represent 95 percent confidence intervals.

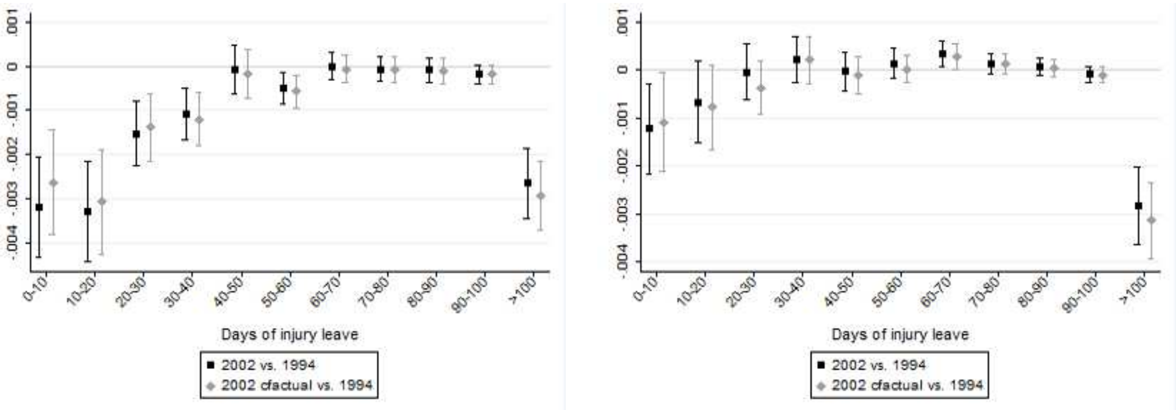
Figure 6: Difference between the gender gap in 2002 and 1994.



Note: The solid line represents the difference between the two gender gaps, computed in 2002 and 1994. The grey area represents 95 percent confidence intervals.

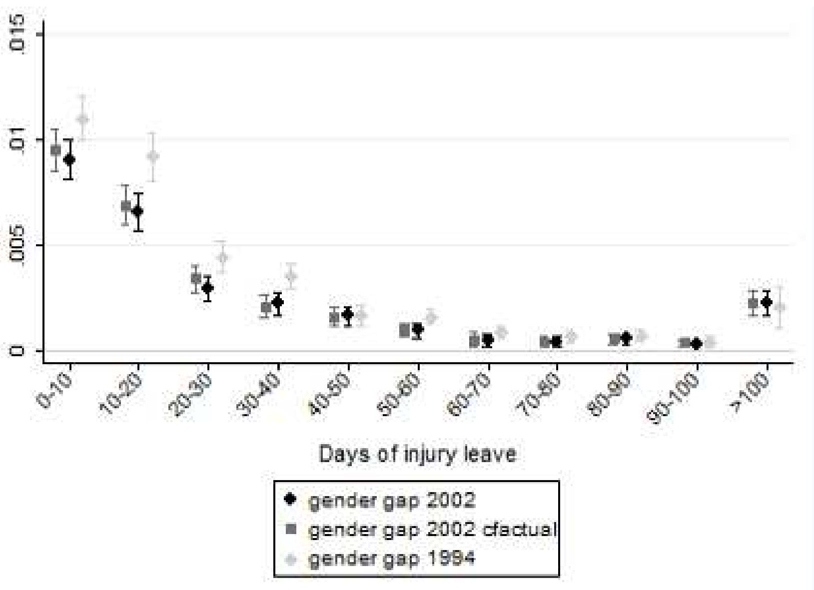
Figure 7: Changes in the distribution of days lost due to workplace injury by gender

Men Women



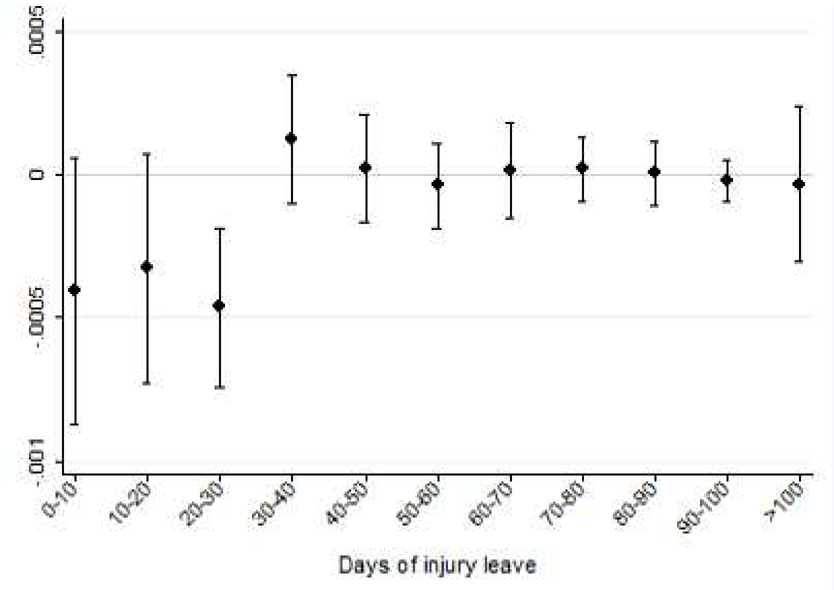
Note: The black squares represent the difference of frequencies between the distribution of days lost due to workplace injury in 2002 and in 1994. The grey circles represent the difference of frequencies between the counterfactual distributions of days lost due to workplace injury in 2002 (computed assuming the same distribution of observable characteristics as in 1994) and the distribution in 1994. The capped spikes represent 95 percent confidence intervals.

Figure 8: The gender gap in the frequency of days lost due to workplace injury.



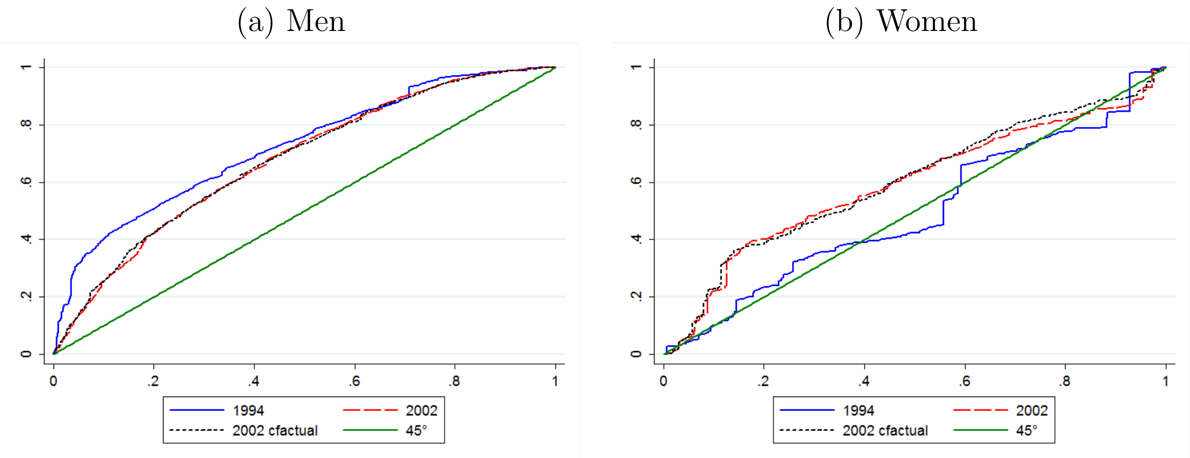
Note: The black circles represent the gender gap in the frequency of days lost due to workplace injury in 2002. The grey squares represent the gender gap in 2002 counterfactual distribution. The light diamonds represent the gender gap in 1994. The capped spikes represent 95 percent confidence intervals.

Figure 9: Change in the gender gap in the frequency of days lost due to workplace injury between 1994 and 2002 .



Note: The black circles represent the difference between the gender gap in the frequency of days lost due to injury in 2002 and in 1994. The capped spikes represent 95 percent confidence intervals.

Figure 10: Concentration curves for men and women in 1994 and 2002.



The concentration curves describe the proportion of injuries that are attributable to the cumulative percentage of the sample ranked by wage. The curve denoted by 2002 cfactual represents the counterfactual concentration curve for 2002, assuming the same distribution of observable characteristics as in 1994.

Table 1: Descriptive statistics on the probability of injury and the number of working days lost due to injury by age and tenure.

	Males		Females	
	injury probability	injury index	injury probability	injury index
16 $\leq$age $<$ 30	.092	2.250	.029	.623
30 $\leq$age $<$ 40	.069	1.898	.022	.678
40 $\leq$age $<$ 50	.059	2.081	.026	.809
50 $\leq$age $<$ 65	.062	2.992	.034	1.811
tenure $\leq$ 1 year	.110	3.549	.043	1.367
1 $<$ tenure $<$ 2 years	.079	2.465	.026	.654
2 $<$ tenure $<$ 3 years	.069	2.022	.022	.556
3 $<$ tenure $<$ 4 years	.070	2.231	.032	1.138
4 $<$ tenure $<$ 5 years	.067	1.748	.024	.590
tenure $\geq$ 5 years	.054	1.699	.016	.499

Note: The injury probability is the estimated probability of injury within a year. The injury index is the estimated yearly number of working days lost due to injury.

Table 2: Composition of the sample by industry.

Sector	Men		Women		Men-Women	
	1994	2002	1994	2002	1994	2002
CA: Extraction of fuel minerals	0.08	0.05	0.01	0.01	0.07***[0.02]	0.04***[0.01]
CB:Extraction of non-fuel minerals	0.41	0.37	0.07	0.06	0.34***[0.04]	0.32***[0.03]
DA:Food industries	3.74	3.15	4.53	4.01	-0.79***[0.13]	-0.86***[0.11]
DB:Textile industries	3.01	2.50	15.67	9.95	-12.66***[0.16]	-7.45***[0.13]
DC:Hide and leather industries	1.47	1.30	3.54	2.51	-2.07***[0.09]	-1.21***[0.08]
DD:Wood industry	1.75	1.54	0.69	0.63	1.06***[0.08]	0.90***[0.06]
DE:Paper, printing and publishing	2.52	2.14	2.03	1.70	0.49***[0.10]	0.44***[0.08]
DF:Coke manufacturing and refineries	0.27	0.26	0.12	0.08	0.15***[0.03]	0.18***[0.03]
DG:Chemical product manufacturing	2.72	2.15	1.95	1.70	0.77***[0.10]	0.46***[0.08]
DH:Rubber and plastics	2.14	2.11	1.78	1.59	0.35***[0.09]	0.52***[0.08]
DI:Processing of non-metallic minerals	3.10	2.56	1.40	1.17	1.69***[0.10]	1.40***[0.08]
DJ:Metal and metallic products	10.88	10.63	4.70	4.60	6.19***[0.19]	6.03***[0.16]
DK:Manufacturing and repair of machinery	4.88	5.16	2.08	2.34	2.81***[0.13]	2.82***[0.12]
DL:Manufacturing of electrical machinery	5.24	5.09	5.21	5.32	0.04[0.15]	-0.23*[0.13]
DM:Vehicle manufacturing	3.63	2.61	1.23	1.00	2.41***[0.11]	1.61***[0.08]
DN:Other manufacturing industries	2.42	2.20	2.59	2.16	-0.17*[0.10]	0.03[0.09]
E:Electrical energy, gas and water	2.06	1.36	0.60	0.48	1.46***[0.08]	0.88***[0.06]
F:Construction	15.01	15.88	1.81	1.90	13.20***[0.20]	13.99***[0.18]
G:Commerce	12.31	13.31	20.09	21.94	-7.78***[0.23]	-8.63***[0.21]
H:Hotels and restaurants	4.18	5.17	8.63	10.54	-4.46***[0.15]	-5.37***[0.15]
I:Transport and communications	7.85	7.93	2.35	2.92	5.50***[0.16]	5.01***[0.14]
J:Financial intermediation	8.27	11.28	15.90	21.16	-7.63***[0.20]	-9.89***[0.20]
K:Business services	1.04	1.23	2.01	2.22	-0.97***[0.08]	-0.99***[0.07]

Note: The table reports the proportion of workers in each industry and the difference between men and women. Standard errors in square brackets (200 replications). *** 1 % significant. ** 5% significant, * 1% significant.

Table 3: Descriptive statistics on worker and job characteristics.

	Men		Women		Men - Women	
	1994	2002	1994	2002	1994	2002
Avg. age	36.54	36.75	33.16	34.94	3.39***[0.07]	1.82***[0.06]
Avg. months of tenure	62.36	68.63	55.46	60.92	6.90***[0.30]	7.71***[0.42]
% apprentice	4.64%	5.58%	4.58%	6.79%	0.06 [0.10]	-1.21***[0.10]
% blue-collar	67.03 %	67.62%	49.07%	46.10%	17.96***[0.31]	21.52***[0.28]
% white-collar	26.74 %	25.47%	46.17%	46.84%	-19.43***[3.02]	-21.37***[2.68]
% manager	1.60%	1.33%	0.18%	0.26%	1.14***[0.07]	1.07***[0.06]
% part-time	2.10%	5.08%	17.88%	27.74%	-15.78***[0.16]	-22.66***[0.19]
% jobs in the North	57.12%	56.41%	65.86%	63.43%	-8.74***[0.32]	-7.02***[0.29]
% jobs in the Center	19.01%	19.14%	19.82%	20.33%	-0.81***[0.26]	-1.19***[0.23]
% jobs in the South and Islands	23.87%	24.45%	14.32%	16.24%	9.55***[0.26]	8.21***[0.24]
% firm size $0-9$	26.62%	28.24%	33.46%	31.54%	-6.84***[0.30]	-3.30***[0.27]
% firm size $10-19$	12.88%	13.48%	14.60%	13.19%	-1.72*** [0.22]	0.29 [0.20]
% firm size $20-199$	28.42%	29.81%	28.43%	27.41%	-0.00 [0.29]	2.40***[0.26]
% firm size $200-999$	13.14%	12.93%	10.72%	12.48%	2.42***[0.22]	0.45** [0.19]
% firm size ≥ 1000	17.91%	15.23%	11.77%	15.16%	6.14***[2.39]	0.06 [0.21]
n.obs	72,891	84,408	34,267	45,005		

Note: Standard errors in square brackets (200 replications). *** 1 % significant. ** 5% significant, * 1% significant.

Table 4: Weekly wages in 1994 and 2002.

	Men			Women		
	1994	2002	2002'	1994	2002	2002'
Mean log deflated weekly earnings	5.771 [.002]	5.753 [.002]	5.800 [.002]	5.570 [.002]	5.595 [.002]	5.593 [.002]
5th percentile	5.053 [.006]	5.059 [.006]	5.150 [.005]	4.883 [.007]	4.911 [.007]	4.941 [.008]
10th percentile	5.280 [.003]	5.278 [.003]	5.336 [.002]	5.150 [.006]	5.173 [.004]	5.195 [.006]
25th percentile	5.507 [.001]	5.487 [.001]	5.529 [.001]	5.394 [.002]	5.400 [.002]	5.409 [.002]
50th percentile	5.727 [.002]	5.690 [.001]	5.736 [.002]	5.564 [.002]	5.578 [.001]	5.580 [.002]
75th percentile	6.016 [.002]	5.983 [.003]	6.025 [.002]	5.768 [.002]	5.790 [.003]	5.780 [.003]
90th percentile	6.336 [.003]	6.349 [.004]	6.380 [.004]	6.039 [.005]	6.093 [.006]	6.058 [.006]
95th percentile	6.572 [.005]	6.609 [.005]	6.640 [.007]	6.215 [.007]	6.306 [.006]	6.267 [.007]
Gini index	0.281 [.002]	0.287 [.002]	0.283 [.002]	0.216 [.002]	0.249 [.005]	0.233 [.004]
Theil index	0.174 [.004]	0.182 [.005]	0.175 [.004]	0.098 [.004]	0.188 [.020]	0.158[.015]

Note: The columns denoted by 2002' report counterfactual values for 2002, assuming the same distribution of observable characteristics as in 1994. Bootstrap standard errors in square brackets (200 replications).

Table 5: The probability of injury and the number of working days lost due to injury in 1994 and 2002.

	Men			Women		
	1994	2002	2002'	1994	2002	2002'
Injury probability	0.087	0.062	0.058	0.028	0.025	0.021
	[0.003]	[0.002]	[0.001]	[0.003]	[0.002]	[0.002]
Injury index	2.810	1.865	1.751	0.932	0.700	0.540
	[0.237]	[0.078]	[0.075]	[0.167]	[0.086]	[0.054]

Note: The table reports the mean probability of injury and the mean number of days lost due to injury within a year. The columns denoted by 2002' report counterfactual values for 2002, assuming the same distribution of observable characteristics as in 1994. Bootstrap standard errors in square brackets (200 replications).

Table 6: The gender wage and injury gap.

	Men-Women 1994	Men-Women 2002	Men-Women 2002'
Mean log deflated	0.201	0.158	0.207
weekly earnings	[0.003]	[0.003]	[0.003]
5th percentile	0.170	0.148	0.209
	[0.009]	[0.009]	[0.008]
10th percentile	0.130	0.105	0.141
	[0.006]	[0.004]	[0.005]
25th percentile	0.113	0.087	0.120
	[0.002]	[0.002]	[0.002]
50th percentile	0.163	0.112	0.156
	[0.002]	[0.002]	[0.002]
75th percentile	0.248	0.193	0.245
	[0.003]	[0.003]	[0.003]
90th percentile	0.297	0.256	0.322
	[0.006]	[0.006]	[0.007]
95th percentile	0.357	0.303	0.373
	[0.008]	[0.007]	[0.009]
Injury probability	0.059	0.037	0.037
	[0.004]	[0.002]	[0.002]
Injury index	1.878	1.165	1.211
	[0.305]	[0.139]	[0.106]

Note: The injury probability is the estimated probability of injury within a year. The injury index is the estimated yearly number of days lost due to injury. The column denoted by 2002' reports counterfactual values for 2002, assuming the same distribution of observable characteristics as in 1994. Bootstrap standard errors in square brackets (200 replications).

Table 7: The average injury probability by wage quartile.

	(a) Men			(b) Women			(c) Men-Women		
Wage quartile	1994	2002	2002'	1994	2002	2002'	1994	2002	2002'
1st quartile	0.150	0.094	0.088	0.039	0.040	0.035	0.111	0.053	0.053
	[0.009]	[0.004]	[0.003]	[0.006]	[0.007]	[0.008]	[0.011]	[0.009]	[0.009]
2nd quartile	0.096	0.072	0.068	0.025	0.025	0.020	0.072	0.048	0.048
	[0.004]	[0.003]	[0.003]	[0.005]	[0.003]	[0.002]	[0.005]	[0.004]	[0.003]
3rd quartile	0.078	0.057	0.054	0.030	0.024	0.020	0.048	0.033	0.034
	[0.004]	[0.002]	[0.002]	[0.004]	[0.002]	[0.002]	[0.006]	[0.003]	[0.003]
4th quartile	0.023	0.024	0.023	0.021	0.012	0.011	0.003	0.012	0.012
	[0.003]	[0.002]	[0.003]	[0.007]	[0.002]	[0.001]	[0.008]	[0.003]	[0.003]

Note: The injury probability is the estimated probability of injury within a year. The columns denoted by 2002' report counterfactual values for 2002, assuming the same distribution of observable characteristics as in 1994. Bootstrap standard errors in square brackets (200 replications).

Table 8: The average injury index by wage quartile.

	(a) Men			(b) Women			(c) Men-Women		
Wage quartile	1994	2002	2002'	1994	2002	2002'	1994	2002	2002'
1st quartile	6.262	3.613	3.359	1.051	1.222	0.909	5.221	2.391	2.450
	[0.809]	[0.251]	[0.235]	[0.210]	[0.265]	[0.192]	[0.871]	[0.467]	[0.360]
2nd quartile	2.283	1.934	1.778	0.529	0.549	0.470	1.755	1.386	1.308
	[0.181]	[0.155]	[0.139]	[0.143]	[0.067]	[0.069]	[0.218]	[0.170]	[0.158]
3rd quartile	2.153	1.360	1.346	1.226	0.475	0.400	0.928	0.885	0.947
	[0.429]	[0.129]	[0.139]	[0.408]	[0.065]	[0.043]	[0.603]	[0.124]	[0.137]
4th quartile	0.550	0.551	0.519	0.922	0.554	0.381	-0.372	-0.003	0.138
	[0.101]	[0.061]	[0.065]	[0.487]	[0.200]	[0.114]	[0.552]	[0.226]	[0.135]

Note: The injury index is the estimated yearly number of days lost due to injury for each worker. The columns denoted by 2002' report counterfactual values for 2002, assuming the same distribution of observable characteristics as in 1994. Bootstrapped standard errors in square brackets (200 replications).

Table 9: Statistics on the joint distribution of log wages and injuries.

	(a) Men			(b) Women			(c) Men-Women		
	1994	2002	2002'	1994	2002	2002'	1994	2002	2002'
Concentration index	-0.447***	-0.353***	-0.355***	0.008	-0.185*	-0.201**	-0.456***	-0.168	0.154*
(Inj. Index, log wages)	[0.049]	[0.022]	[0.024]	[0.124]	[0.109]	[0.081]	[0.132]	[0.113]	[0.088]

Note: The table reports the correlation coefficient and the concentration index of the injury index and log weekly earnings. Bootstrap standard errors in square brackets (200 replications).
*** 1 % significant. ** 5% significant, * 1% significant.