Career discontinuities and the gender wage gap: a quantile approach

Paolo Naticchioni

University of Cassino and Southern Lazio, IZA

Michele Raitano

University of Rome La Sapienza

Giulia Santangelo

University of Rome La Sapienza

June 21, 2013

Abstract

This paper investigates the impact of career discontinuities in working history, i.e. experience gaps, on the workers' career profile, to assess whether and to what extent they affect differently wages of female and male workers and hence contribute to the gender gap. We make use of a unique Italian administrative dataset, the AD-SILC, to precisely identify both the experience gaps in career histories for women and men, and the reasons for such gaps, i.e. maternal leave. We implement a recent quantile methodology that allows controlling for unobserved heterogeneity in a quantile framework. Our results show that in fixed effects experience gaps have a negative impact on male weekly wage dynamics, effects that is decreasing along the wage distribution. Women display an additional wage penalty with respect to men, which is instead more uniformly distributed along the wage distribution. We also point out that the additional penalty for women is only partially due to cumulative weeks in maternal leave and to childbirth events.

Keywords: Gender Wage Gap, Career discontinuities, Quantile Regressions.
JEL Codes: C33, J16, J31

1 Introduction

This paper contributes to the literature concerning the gender wage gap along the wage distribution. Previous literature showed that it is crucial taking into account the whole distribution, since the gender gap is not uniformly distributed because of phenomena such as the glass ceiling, i.e. gender gap higher for top earners, and the sticky floor, i.e. women trapped in low wage dynamics at the bottom of the distribution. For the glass ceiling, one of the pioneering works is Albrecht, Bjorklund and Vroman (2003) that use quantile regressions finding an increasing gender pay gap along the wage distribution in Sweden. Other papers have then extended this finding to most of the OECD countries (Arulamapalan et al., 2007). Although the glass ceiling phenomenon is observed in most OECD countries, the understanding of the reasons behind it represents an open field of research (Booth, 2007), with relatively few papers testing explanations from an empirical point of view (De la Rica et al., 2010, Bertrand and Hallock, 2001, Matano and Naticchioni, 2013, Raitano, 2009, among others). Similar remarks apply for the sticky floor, i.e. solid empirical evidence and not much about explanations.

The main contribution of this paper concerns the analysis of discontinuities in working history on the workers' career profile, in order to assess whether and to what extent they affect differently wages of female and male workers and hence contribute to the gender gap. The main problem in this literature concerns the fact that data usually do not allow a precise identification of workers' careers and the related discontinuities. Our unique database makes it possible. We use of a longitudinal dataset, called Administrative Statistics on Income and Living Conditions (AD-SILC), built by Brodolini Foundation and the Italian Ministry of Treasure, which has been constructed matching longitudinal information coming from administrative archives gathered by INPS (National Institute of Social Security) with survey data collected by ISTAT (National Institute of Statistics). In particular, Administrative SILC contains very detailed micro-data from the 2005 cross-sectional wave of the IT-SILC survey (i.e. the Italian database of the European Union Survey on Income and Living Conditions, EU-SILC), which have been merged with information collected in various administrative archives managed by INPS, which record, since the beginning of their working life, the individuals working histories and various characteristics of private and public employees, self-employed, recipients of unemployment benefits and retired. Furthermore, from our data it is possible to identify maternal leave and its length, from 1975.

The uniqueness of the data arises in fact from the possibility to have, on the one hand, the whole working history of individuals from administrative sources (from 1975), and, on the other hand, variables such as education, occupation, family composition, children from the SILC survey (variables usually not available in administrative data).

We investigate cohorts of females and males entered in the labour market from 1975 to 2000 and aged at the entry year less than 30. In this way it is possible to follow their working career up to the end of 2009 (i.e. in the period 1975-2009).

Using our data it is possible to reconstruct the cumulative individuals' discontinuities in working career, i.e. the experience gap, by calculating for each year the difference between the cumulative potential working experience of an individual, exploiting the information concerning the entry year in the labour market, and the cumulative actual experience. Our dataset allows a precise computation of the actual experience, since it is possible to recover the number of weeks worked in each year from the beginning of the career, from the different INPS administrative archives. Furthermore, it is also possible to identify the amount of weeks in maternal leave.

Another interesting contribution of our research is that we investigate the whole wage distribution, i.e. whether career patterns and job interruptions affect in a heterogeneous way men and women located at different quantiles of earnings distribution, in such a way addressing issues such as the glass ceiling and the sticky floor effects. From a methodological point of view, we enrich the standard quantile regression model of Koenker and Bassett (1978) by means of the quantile regression model with fixed effects proposed by Canay (2011). This method allows exploiting in a quantile setting the longitudinal component of our data set and to get rid of the fixed effects associated to unobserved heterogeneity. This is crucial in our analysis, since career discontinuity can take place because of family duties but also because of heterogeneity in ability levels that are not observed. Our estimates are derived from a pooled sample of females and males.

We regress the log weekly earnings on potential experience (and its squared), the experience gap, and the interaction term between experience gap and the female dummy, which is our main variable of interest. We also include some control variables, time dummies and unobserved fixed effects. This will allow evaluating the differential effect of working discontinuities and career profiles between the two gender groups, and the patterns of this differential along the wage distribution. Our results show that experience gap has a negative and substantial effect on wages for all workers. This effect decreases along the wage distribution, suggesting that unskilled workers are the ones who suffer the most from career discontinuities. Interestingly, women display an additional penalty with respect to men, and this additional penalty is not decreasing along the wage distribution. We also show that experience gaps directly related to weeks in maternal leave and to childbirth explain only a part of the additional penalty that females display with respect to males.

2 Theoretical and empirical framework

The analysis of the gender wage gap, underlying related explanations, has been long since a matter of wide interest in the labour economics literature. In particular, empirical researches focused on investigating whether gender gaps can be accounted for by differences between men and women in their characteristics versus differences in the returns to the same characteristics. First studies on this topic concentrated on the average gap between male and female workers (see Altonji and Blank (1999) for a survey). More recently, works on the gender gap have begun to explore whether and to what extent the observed gender gap might differ across the wages distribution. The extension of the analysis of earning differentials to the overall wage distribution allows to investigate phenomena such as the glass ceiling and the sticky floor. The first refers to the situation in which the gender pay gap is higher for top earners and widens therefore at the top of the wage distribution. The sticky floor, instead, indicates a widening gender wage gap at the top of bottom of the wage distribution induced by the fact that women are trapped in low wage dynamics at the bottom of the distribution.

For the glass ceiling, one of the pioneering works is Albrecht, Bjorklund and Vroman (2003) that using quantile regressions document the existence of a significant glass ceiling effect in Sweden in the 1990, due to a wide gender gap at the top of the Swedish wage distribution, larger than the corresponding gap in the United States. Other papers have then extended this finding to most of the OECD countries. Arulamapalan et al. (2007) use data from the European Community Household Panel in a quantile regression framework to analyze gender pay gaps across the wage distribution for eleven countries and find that the magnitude of the gaps, which can be attributed to differing returns, varies substantially across the different countries and across the wage distributions. Booth (2007) investigates the extent of glass ceilings and sticky floors across a number of different European countries and concludes that almost without exception the gender pay gap was largest towards the top of the wages distribution. He also underlines that this glass ceiling effect is bigger in the private sector than in the public, and that it seems to be induced at least partially by discrimination.

De la Rica et al. (2010) refer to the so-called *mommy track* hypothesis (see Mincer and Polacheck, 1978), and highlight the importance of monopsonistic features, possibly related to women's lower labour mobility due to their attachment to household tasks, in explaining the higher level of the wage gap in the performance-pay (PP) component of total hourly wages, with respect to the gap in non-PP compensation. First, as stressed in the occupational segregation literature, women may select themselves into jobs with scarce or either absent PP, as public sector jobs, because they anticipate that these positions are more compatible with their larger household responsibilities. Secondly, women can be disadvantaged with respect to men since employers expect lower female work attachment.

Olivetti and Petrongolo (2008) in virtue of the negative correlation between the gender wage gaps across countries and gender employment gaps, analyze gender wage gaps correcting for sample selection induced by unemployment. They find that higher median wage gaps on imputed rather than actual wage distributions for several OECD countries and note that this difference is small in countries like the United States, the United Kingdom, and most central and northern EU. Blundell et al. (2007), instead, examine changes in the distribution of wages using bounds based on theoretically motivated restrictions from economic theory to allow for the impact of nonrandom selection into work. They find convincing evidence in the United Kingdom of an increase in inequality within education groups, changes in educational differentials, and increases in the relative wages of women.

The human capital models explain the gender pay gap by gender differences in human capital accumulation. First of all, since women are more likely to have intermittent labour market participation and are more likely to work part-time, male and female workers show differences in accumulated work experience, which is assumed to be correlated with levels of human capital. Secondly, the anticipation of future intermittence may affect current investments in human capital by females. Finally, the greater domestic commitments of women can reduce their effort into work, especially in case of married women with dependent children.

The effects of a more intermittent career by women with respect to men have been investigated by Manning and Swaffield (2008) and Bjerk (2008). The first focus on the gender gap in wage growth in the early years after labour market entry and explore three main hypotheses: human capital, job-shopping and "psychological" theories. They identify labour market intermittence by the gap between wage observations, assuming that a gap of more than one year must imply a period not in employment. They find that women have longer gaps between wage observations than men, implying that women do have weaker labour market attachment than men. They also notice that the effect of labour market intermittence and part-time working on the gender gap is more important between 10 and 20 years after labour market entry, that is when most women have breaks in employment associated with childcare. Bjerk (2008) explores how inequality of opportunity with respect to hiring and promotion may arise between gender groups. His theoretical model makes clear that, if females are more likely than males to take time out of the labour market, they will have fewer opportunities to signal their skill and/or fewer opportunities to successfully complete tasks. This can thwart promotions to the top jobs and produce glass ceilings.

The first works which focused on domestic commitments of women are those of Blau and Kahn (1992) and Waldfogel (1998). The last refers to the gap in pay between women with children and women without children as the "family gap", and analyses for the United States and Britain the controversial effects of maternity leave policies. He underlines that maternity leaves can reduce the family gap, by increasing the likelihood that working mothers return to their employers after childbirth, but that at the same time leaves impose costs on employers, which may be passed on to women in the form of lower wages or employment. In a more recent paper, Blau and Kahn (2011) also highlight the importance of the availability of information about actual work experience for analyzing women's post-school human capital accumulation, residual wage inequality, and the gender pay gap.

Despite the wide interest on the analysis of the gender wage gap and of various factors which can contribute to it, research on the Italian case is limited. Raitano (2009) in a comparison of EU 15 European countries highlight that the smallest hourly wage gaps can be observed in Belgium, Germany, Italy and Ireland. Addabbo and Favaro (2007) apply quantile regression analysis and an adaptation of the procedure suggested by Machado and Mata (2005) using data on the Italian sample of the European Community Household Panel (ECHP). Analyzing groups with diverse educational levels, they find that for low-educated workers, lower levels of education or experience are responsible for the gap, while for highly-educated females, better characteristics than highly-educated men can partially reduce the gap. Addabbo et al. (2007), studying the wage gap along the entire distribution, focus on the relationship between human capital characteristics, in particular different educational levels, and outcomes in differences in pay. They detect some patterns of sticky floor in the sample of low educated females and some glass ceiling pattern among high educated females. Picchio and Mussida (2012) study the gender wage gap by educational attainment in Italy using the ECHP data for the period 1994–2001, exploiting the estimator of probability density functions in the presence of covariates and sample selection proposed by Picchio and Mussida (2011) and microsimulation to decompose the gender wage gap. After controlling for nonrandom sample selection, they find a marked evidence of sticky floor for low-educated women and glass ceiling for highly-educated women. Matano and Naticchioni (2013) make use of a unique matched employer-employee panel database for Italy and show that even after controlling for first level bargaining, sorting and endogeneity, there is a different degree of rent-sharing between men and women that increases along the wage distribution, and hence rent sharing plays a role in explaining the glass ceiling effect.

3 Data

The empirical analyses of this paper are based on a panel dataset on individual working histories, called AD-SILC, recently built for Italy merging longitudinal information provided by administrative archives with a sample survey dataset. AD-SILC has been developed merging the IT-SILC 2005 sample (i.e. the Italian version of EU-SILC 2005) with the administrative records on individual working histories, since their entry in the labour market up to 2009, collected by the INPS.

Administrative archives include detailed information on the universe of individuals (and firms) since the beginning of their working career up to the most recent years and these archives refer to each type of worker (i.e. private and public employees, self-employed). For each employment spell (or period receiving welfare or unemployment benefits) in each year, these archives record the number of related worked weeks, the gross earnings (including employees' contributions), the starting and final date of the spell and (for private employees) the characteristics of the firms. The observation units are then the different contributive periods experienced by an individual during a year (e.g. if during a year an individual works for two firms, he/she is self-employed for some weeks and unemployed receiving a subsidy for other weeks, he/she will be characterized by four records). Therefore these archives offer a comprehensive picture of the working history of the Italian labour force and, by definition, they are not plagued by attrition (if someone disappears from the archives it means that he/she has stopped to work or has left Italy). However, being not relevant to administrative purposes, these archives do not record some information that are instead crucial to analyze determinants of individual working statuses and earnings, e.g. educational attainments, family composition, family background. On the contrary, the IT-SILC 2005 wave collects detailed information about several time invariant individual characteristics. Due to their complementary characteristics, IT-SILC and INPS archives information have then been merged in this new panel dataset using individual fiscal codes (recorded in both datasets) as the matching key.

Hence, starting from a cross section of about 47,000 individuals interviewed in IT-SILC 2005, the AD-SILC panel contains around 1,150,000 observations. Currently AD-SILC is the only panel dataset available for Italy that follows individuals over time and collects detailed information on individual working statuses (e.g. employment, self-employment, receipt of "cassa integrazione" – the benefit paid in case of suspension of the working activity – unemployment, maternity or sickness allowances), individual and household characteristics (e.g. education) and on firms characteristics (e.g. sector and firm size).

To the scope of our analysis – i.e. to study the association between career path and earnings for private employees – AD-SILC has many pros.

- It records the gross wage related to each job relationship during a year (i.e. it records the annual wage in case of continuous employment, while in case of jobs episodes lasting less than a year, the wage is referred to the length of this period); wages are recorded in current Euros (then converted in 2010 prices using the CPI). It records the number of weeks during a year spent working (periods spent in different firms or different contractual arrangements are distinguished), or receiving an unemployment, a "cassa integrazione guadagni" (CIG), a maternity or a sickness allowance. It allows to compute weekly gross wages (dividing gross earnings by the number of related working weeks); in this paper weekly wages are computed considering the longest employment period in a year and dividing the associated earnings to the working weeks. Furthermore, it allows to identify the cohort of entry of each individual in the labour market, as the first working episode of each individual; in order not to consider isolated job relationship during the youth, in this paper we identify as the entry cohort the year related to the first working episode lasting at least 13 weeks; as a consequence, potential experience is defined as the time span from the entry and a specific point of time.

- It allows to exactly measure individual labour market experience (i.e. the cumulated working weeks since the entry in the labour market), distinguishing potential experience from the effective one and also distinguishing the causes of not working periods (i.e. unemployment, inactivity, CIG, maternity, sickness). As a consequence, a cumulated working weeks gap can be computed (e.g. the cumulated weeks since the entry in the labour market spent not working or receiving a CIG benefit).

- It allows to compute the cumulated weeks spent receiving a maternity allowance during the working career, that can be considered a proxy of the care needs of a woman.

- It allows to identify the childbirth, i.e. at least 20 consecutive weeks receiving a maternity allowance (20 weeks is the length of the mandatory maternity allowance in Italy).

- It records several information that can be used in an empirical analysis, such as: the province where the firm is situated; the contractual arrangement of each employment relationship (i.e. full-time versus part-time); the job qualification (i.e. white-collar, blue-collar, apprentice) of employment period; the characteristics of the employer – i.e. firm size and sector of activity (recorded at 3 digits NACE) (since 1987). Finally, all the variables recorded in IT-SILC 2005 are added to the longitudinal information collected in administrative archives.

In this paper we only refer to private employees earnings, because self-employed wages are plagued by huge problems of underreport and truncation, and public employees earnings are only available since 1996 (but the periods spent working as public employees or self-employed are taken into account for computing individual experience). The dependent variable is the log of gross weekly wages from private employment (i.e. including personal

Table 1. Summary Statistics by Gender		
	Mean	
	Male	Female
Log of gross weekly wages	6,028	5,775
Potential experience	11,737	10,798
Experience gap	1,856	1,860
Cumulative matern.leave	0,002	0,193
Childbirth dummy	0,001	0,052
Education		
Primary education	0,497	0,403
Secondary education	0,448	0,530
Tertiary education	0,055	0,066
Occupation		
White collar	0,263	0,492
Blue collar	0,661	0,431
Apprentiship	0,075	0,077
Part time contracts	0,027	0,208
Geographic area		
Centre	0,230	0,243
Islands	0,063	0,030
Northeast	0,277	0,326
Northwest	0,262	0,295
South	0,168	0,105

Data source: AD-SILC

income taxes and employees' social insurance contributions). To reduce the impact of outliers we do not use, for each year, the top 0.5% and the bottom 0.5% of the weekly wage distribution.

The analyses are carried out referring to the cohorts of males and females entered in the labour market since 1975 to 2000 and aged at the entry year less than 30. Their career is observed since the entry up to the end of 2009 (i.e. in the period 1975-2009). For each year we consider individuals aged 15-64. We do not use years prior 1975 because the earnings variable has substantial missing values.

The final sample is then composed by 159,132 longitudinal observations concerning 11,156 individuals entered in the labour market in the period 1975-2000 and then followed up to 35 years.

Table 1 displays the descriptive statistics for the wage variable and the control variables calculated on the two subsamples of male and female workers.

Mean values of the wage variable are lightly lower for female: the logarithm of gross weekly wages is equal to 6,028 and 5,775 respectively for males and females. The difference in values of potential experience between male and female workers is about one year: for male workers the average time span spent from the entry in the labour market is equal to 11,737 years, while for females it is equal to 10,798. The average values for the experience gap, that is the cumulated weeks (calculated as annual percentage) since the entry in the labour market spent not working, estimated in the two subsamples are equal to 1,856 for males and to 1,860 for females workers.

The two groups exhibit obviously notably differences with respect to variables related to work interruptions induced by childcare or childbirth. The average amount of cumulated weeks spent receiving a maternity leave (calculated as annual percentage) is equal to 20% for females, while is nearly null for male workers. Values associated to the childbirth dummy variable indicate that 5,2% of females, and only 0,1% of male workers, spent at least 20 weeks receiving a maternity allowance.[1]

Female workers show higher levels of education with respect to male workers. Among male workers, the percentage of individuals with primary education is equal to 49,7% and the percentages of individuals with secondary and tertiary education are equal, respectively, to 44,8% and 5,5%. Instead, in the subsample of female workers the percentage of individuals with primary education is equal to 40,3%, the percentage of individuals with secondary education is equal to 53%, and the percentage of individuals with tertiary education is equal to 6,6%.

The distribution of workers among different classes of occupation is very different in the two subsamples of male and female workers. The biggest differences are related to the classes of white and blue collars: 66,1 % of male workers is concentrated in the class of blue collars, while only 26,3% of male workers are white collars. In the female subsample, the percentage of white collars is much higher (49,2%) and the percentage of blue collars is lower (43,1%) with respect to the male group. Percentages of apprentices are similar in the two groups (7,5% and 7,7% respectively).

The adoption of part time contracts is considerably higher among female workers, for whom the percentage of this kind of contracts raises the 20%, while it is equal to the 2,7% for male workers. The distribution of workers in the five considered geographical areas is

[1] The number of weeks when a worker received a maternity allowance is evaluated in the time span of two consecutive years.

not very different according to gender.

4 Econometric methodology

In this paper we make use of quantile estimation methods in order to evaluate heterogeneous effects of variables of interest along the whole outcome distribution. We first apply the standard model for quantile regressions introduced by Koenker and Bassett (1978) to perform a quantile analysis on cross sectional data. Secondly, we use the methodology proposed by Canay (2011) for the estimation of quantile regressions for panel data.

This latter methodology allows to explicitly take into account the unobserved individual heterogeneity that can bias the cross sectional estimates, and to control for some time invariant unobserved covariates by the inclusion of individual fixed effects in the model specification. Exploiting the longitudinal component of the panel data, the individual fixed effects can be eliminated by a simple transformation of the data.

The model proposed by Canay relies on the assumption that the unobserved fixed effects are location shift variables (i.e. variables that affect all quantiles in the same way) and supplies a two-step estimator, which is consistent and asymptotically normal when both n and T go to infinity. In the first step the unobserved individual fixed effects are estimated. Since they are assumed to affect all units in the same way regardless of quantiles, i.e. regardless of where units are located along the outcome distribution, the fixed effects can be estimated at the conditional mean. In the second step the predicted fixed effects are deducted from the dependent variable in order to control for unobserved individual heterogeneity. Finally, computed values for the dependent variable, net of fixed effects, are regressed on other regressor variables by traditional quantile regressions.

Consider the following model:

$$Y_{it} = X_{it}'\theta\left(U_{it}\right) + \alpha_i, \qquad t = 1, ..., T, \qquad i = 1, ..., n,$$

where t refers to the time period and i to the observed individual.

$(Y_{it}, X_{it}) \in \mathbb{R} \times \mathbb{R}^k$ are observable variables. Y_{it} is the outcome variable. The vector of covariates, X_{it}, is assumed to include a constant term, i.e. $X_{it} = (1, X_{it}^{s\prime})$ with $X_{it}^{s\prime} \in \mathbb{R}^{k-1}$.

$(U_{it}, \alpha_i) \in \mathbb{R} \times \mathbb{R}$ are unobservable. U_{it} is a random variable aggregating all unobserved factors, except fixed effects α_i. U_{it} represents the disturbance and α_i represents the unobserved individual heterogeneity.

$\tau \in (0,1)$ is the quantile of interest. For a given quantile τ, $\theta(\tau)$ indicates the quantile specific effect. It represents the parameter of interest and can be estimated by quantile regressions. The function $\tau \to X'\theta(\tau)$ is assumed to be strictly increasing in $\tau \in (0,1)$.

If α_i were observable, it would follow that

$$P\left[Y_{it} \leq X_{it}'\theta(\tau) + \alpha_i | X_i, \alpha_i\right] = \tau,$$

under the assumption that $U_{it} \sim U[0,1]$ conditional on $X_i = (X_{i1}', ..., X_{iT}')'$ and α_i.

The presence of the term α_i, related to the unobserved fixed effects, in the model proposed by Canay (2011) constitutes the innovation element with respect to the standard quantile regression model introduced by Koenker and Bassett (1978). In this model this unobserved effect α_i is assumed to have a pure location shift effect, i.e. α_i captures unobserved covariates $Z_i'\beta(\tau)$ that enter the model and are constant over time, and have coefficients constant across τ, with $\beta = \beta(\tau)$ for all τ.

Identification of parameters

If there are at least two time periods available, $T \geq 2$, then, under independence restrictions and existence of moments, $\theta(\tau)$ is identified.

Let $S_t = X_t'\theta(U_t)$ (the dependence on i is omitted for convenience here). $Y_t = S_t + \alpha$ can be defined as a convolution of S_t and α conditional on X, provided α and U_t are independent conditional on X, meaning that the variable of interest can only be observed with some contamination which is modelled as an independent additive error.

It follows that the conditional distributions of S_t and α can be identified from the conditional distribution of Y_t, using a deconvolution argument similar to the method proposed by Neumann (2007) for the estimation of a distribution in a deconvolution model with panel data and an unknown distribution of the additive errors.

Finally, after exploiting the fact that U_t is conditionally $U[0,1]$, together with some regularity conditions, $\theta(\tau)$ can be identified.

For ease of exposition let $T = 2$ and let the lower case $x = (x_1, x_2)$ denote a realization of the random variable $X = (X_1, X_2)$. The identification of the model is based on the following assumptions:

Assumption 1. Denote by $\phi_{S_t|x}$ and $\phi_{\alpha|x}$ the conditional on $X = x$ characteristic functions of the distributions $P_{S_t|x}$ and $P_{\alpha|x}$, respectively.

(a) conditional on $X = x$ the random variables S_1, S_2 and α are independent for all $x \in \overline{\chi}$, where $\overline{\chi}$ denotes the support of $X = (X_1, X_2)$;

(b) $\Gamma \equiv$

$\left\{\omega : \phi_{S_t|x}\left(\omega 2^{-k}\right) \neq 0 \text{ for } t \in \{1,2\} \text{ and } \phi_{\alpha|x}\left(\omega 2^{-k}\right) \neq 0,\ k = 0, 1, ..\right\}$

is a set dense in $\mathbf{R}$ for all $x \in \overline{\chi}$.

Assumption 2. (a) $U_{it} \perp (X_i, \alpha_i)$ and $U_{it} \sim U[0,1]$;

(b) $\Omega_{UU} \equiv E\left[\left(\theta\left(U_{it}\right) - \theta_\mu\right)\left(\theta\left(U_{it}\right) - \theta_\mu\right)'\right]$, where $\theta_\mu = E\left[\theta\left(U_{it}\right)\right]$, is non-singular with finite norm;

(c) letting $X_t = (1, X_t^s)$ for $t = 1, 2$, there exists no $A \subseteq \mathbf{R}^{k-1}$ such that A has probability 1 under the distribution of $X_2^s - X_1^s$ and A is a proper linear subspace of $\mathbf{R}^{k-1}$;

(d) (Y_t, X_t) have finite first moments for $t = \{1, 2\}$.

Assumption 1(a) implies that α_i does not change across quantiles as α_i is independent of (U_1, U_2). Assumption 1(b) excludes characteristic functions that vanish on non-empty open subsets of $\mathbb{R}$ but allows the characteristic function to have countably many zeros. Assumption 1 implies that Y_t is a convolution of S_t and α conditional on $X = x$.

Assumption 2(a) is a standard assumption for quantile regression models except that here it is extended to the panel case and U_{it} is also assumed independent of α_i. Assumption 2(b) implies that $\theta_\mu \in \mathbb{R}^k$ exists and this entails that the location of S_t is well defined. Assumption 2(c) is a standard rank-type condition on the subvector of regressors that excludes the constant term and implies that θ_μ is identified.

Under Assumptions 1 and 2, the location θ_μ is identified and the parameter of interest $\theta(\tau)$ is *point* identified from the distribution of the observed data.

Estimation

Considering that $\theta_\mu = E\left[\theta\left(U_{it}\right)\right]$ and letting $u_{it} = X_{it}'\left[\theta\left(U_{it}\right) - \theta_\mu\right]$, it is possible to write a conditional mean equation for $Y_{it} = X_{it}'\theta\left(U_{it}\right) + \alpha_i$, as follows

$$Y_{it} = X_{it}'\theta_\mu + \alpha_i + u_{it},$$

with $t = 1, ..., T$, $i = 1, ..., n$, $E\left(u_{it} \mid X_i, \alpha_i\right) = 0$.

This equation implies that α_i is also present in the conditional mean of Y_{it}. Therefore, the model can be estimated by two consecutive steps. In the first step a $\sqrt{T}$-consistent estimator of the fixed effect α_i is computed at the conditional mean given a $\sqrt{nT}$-consistent estimator of θ_μ (for instance, by a standard fixed effect estimation). In the second step, it is then possible to define a new dependent variable net of fixed effects, $\widehat{Y}_{it} \equiv Y_{it} - \widehat{\alpha}_i$, and estimate $\theta(\tau)$ by a quantile regression of the random variable $\widehat{Y}_{it}$ on X_{it}, solving the standard minimization problem of quantile approach:

Step 1. Let $\widehat{\theta_\mu}$ be a $\sqrt{nT}$-consistent estimator of θ_μ.
Define $\widehat{\alpha}_i \equiv \mathbf{E}_T \left[Y_{it} - X'_{it}\widehat{\theta_\mu} \right]$.
Step 2. Let $\widehat{Y}_{it} \equiv Y_{it} - \widehat{\alpha}_i$. Define the two-step estimator of $\widehat{\theta}(\tau)$ as:
$\widehat{\theta}(\tau) \equiv \underset{\theta \in \Theta}{\arg\min} \mathbf{E}_{nT} \left[\rho_\tau \left(\widehat{Y}_{it} - X'_{it}\theta \right) \right]$.

The parameter of interest is identified for fixed T and this simple transformation of the data eliminates the fixed effects as $T \to \infty$.

The random variable $\widehat{Y}_{it}$ weakly converges in probability, as $T \to \infty$,[2] to the variable Y^*_{it} and the two-step estimator defined above is showed by Canay to be consistent and asymptotically normal under some regularity conditions.

5 Econometric specification

To evaluate the effect of discontinuities in working history on wages, we estimate a traditional Mincerian wage equation including variables related to the potential experience and the experience gap, and control variables concerning individual characteristics and working features.

The following equation represents hence the estimated wage equation:

$$\ln w_{it} = \alpha_\theta + \beta_\theta \cdot pot\_\exp_{it} + \gamma_\theta \cdot (pot\_\exp_{it})^2 + \delta_\theta \cdot gap_{it} + \eta_\theta \cdot gender_{it} + \lambda_\theta \cdot (gap_{it} \cdot gender_{it}) + \nu_\theta \cdot X_{it} + +d_{it} + t_{it} + \varepsilon_{\theta it},$$

where i indicates the observed individual, t the time period and θ the considered quantile. $\ln w_i$ is the logarithm of the weekly gross wage, in euro, in real term (base year 2010).

[2] In our empirical analysis the variable T assumes an avarage value across individuals equal to 15,50, which is a high value for panel data.

The first variable of interest, $pot\_\exp_i$, is the potential experience, that is the experience that an individual could acquire remaining in the labour market without experiencing any kind of career interruption. We consider also the square of this variable, in order to detect non linearity in potential experience effects. The second variable of interest, gap_i, is a measure of discontinuities in working paths, that is the difference between the potential experience and the effective experience acquired by workers, calculated by the number of work weeks. This term is referred to male workers, since in the specification the interaction term between experience gap and gender (female), $(gap_i \cdot gender_i)$, is included. It instead measures wage differentials, with respect to male, for female workers.

Coefficients β_θ and γ_θ related to potential experience provide the returns to remain without discontinuities in the labour market (and its square). The coefficient δ_θ referred to the experience gap provides the penalty for having discontinuities in experience for men, while the coefficient of the interaction between experience gap and gender, λ_θ, identifies the penalization for having discontinuities for women.

X_i is the vector of covariates. As control variable we include education (primary, secondary, tertiary), occupation (white collar, blue collar, apprenticeship), tenure and tenure square, part time contract. d_{it} is a dummy variable related to geographic areas (five macro-areas in Italy: Northwest, Northeast, Centre, South and Islands), while the term t_{it} is referred to year dummies (from 1975 to 2009). $\varepsilon_{\theta i}$ is the error term. When using the Canay (2011) approach we also take into account individual fixed effects, which include gender and education variables. We estimate the above-mentioned equation for five quantiles of the distribution of log-wages, namely for $\theta = (0.10, 0.25, 0.50, 0.75, 0.90)$.

6 Results

Table 2 shows coefficients and the related t-statistics in italics. All estimates also include year dummies (from 1975 to 2009) and macro area dummies (5 geographic areas). Using standard quantile regressions the returns for potential experience are concave, as expected, and rather high in magnitude. At the 10th percentile wages increase by 3.2% by one additional year of potential experience, and the trend is slightly decreasing up to the median (2.6%) and slightly increasing in the upper tail of the distribution (2.9%). The gender gap is around 20% along the wage distribution (25% at the 10th percentile). As for the experience gap, the coefficients are negative as expected: the longer the gap accumulated, the greater the wage penalty: one additional year of experience gap entails a wage penalty equal to 3.7% at the 10th percentile, 2.5% at the median and 2.7% at the 90th percentile. These estimates represent the trends for male, since in the specification there is also the interaction term between gap and gender (being female). The coefficients for the interaction terms capture the differential wage effect for female with respect to male, and they are positive, rather small (0.5% for each additional year of accumulated experience gap), and constant along the wage distribution. Although the magnitude is rather small, it suggests that females display positive additional penalties with respect to male when experience gaps increase.

Table 2. Experience trends and Wage dynamics. Quantile regression.

	Percentiles				
	10	25	50	75	90
Potential experience	0,032	0,028	0,026	0,028	0,029
	44,81	*62,19*	*71,12*	*63,84*	*48,16*
Potential experience^2	-0,001	0,000	0,000	0,000	0,000
	-25,08	*-32,58*	*-30,69*	*-24,56*	*-16,67*
Experience gap	-0,037	-0,027	-0,025	-0,026	-0,027
	-59,08	*-67,64*	*-71,08*	*-59,10*	*-43,12*
Gender	-0,252	-0,203	-0,192	-0,193	-0,193
	-72,72	*-92,26*	*-103,65*	*-87,43*	*-63,44*
Experience gap*gender	0,003	0,004	0,004	0,005	0,005
	4,25	*7,43*	*9,17*	*8,22*	*6,32*
Secondary education	0,070	0,069	0,076	0,083	0,092
	20,85	*32,63*	*43,27*	*40,73*	*33,12*
Tertiary education	0,135	0,181	0,252	0,314	0,331
	19,49	*41,53*	*70,78*	*77,02*	*59,72*
White collar	0,219	0,200	0,209	0,249	0,294
	59,15	*85,80*	*109,48*	*113,87*	*100,10*
Apprentiship	-0,343	-0,324	-0,249	-0,211	-0,198
	-58,12	*-86,99*	*-79,03*	*-56,39*	*-38,28*
Part time contracts	-0,676	-0,653	-0,576	-0,492	-0,417
	-130,44	*-201,82*	*-214,24*	*-155,46*	*-93,56*

T-statistics in italics.

Notes: All estimates include geographic area and time dummies.

However, cross sectional quantile regressions are biased, since they cannot control for unobserved heterogeneity. In other words, identification of coefficients in cross sectional quantile regressions is based on the comparison between individuals with different unobserved ability, which might affect the participation decisions to the labour market and the career discontinuities. For such a reason, we implement the methodology proposed by Canay (2011) to control for unobserved heterogeneity in a quantile setting. In the baseline fixed effects specification we include potential experience and its square, the experience gap, and the interaction between experience gap and gender. Note that the main effect for gender is included in the individual fixed effect.

Estimates in Table 3 confirm that returns to potential experience are concave and rather high, as expected. The magnitude is consistent with other papers on the Italian case: combining the linear and the quadratic term, after for instance 10 year of experience the average returns amount to 33.7% at the median. Interestingly, the returns to potential experience are decreasing along the wage distribution, i.e. after 10 years the returns are equal to 43.5% at the 10th percentile, 37.2% at the 25th, 31.2% at the 75th, and 27.9% at the 90th. This clearly suggests that the learning by doing mechanism related to experience is more beneficial for unskilled workers. This could also be related to the fact that for unskilled workers the collective bargaining, which in Italy is strongly based on experience profiles, is more effective.

As far as the experience gap is concerned, the coefficients are negative, as expected: the higher the accumulation of discontinuities in the labour market, the greater is the wage penalty, ranging from 2.4% at the 10th percentile to 0.7% at the 90th percentile. It is also worth noting that since we include in the specification the interaction between experience gap and gender, the coefficients for experience gap refer to the male patterns. Interestingly, also in this case the pattern is decreasing along the wage distribution: for one year of cumulated discontinuity the penalty amounts to 2.5% at the 10th percentile, 1.8% at the 25th, 1,3% at the median, 1% at the 75th, and 0.7% at the 90th. The interaction term refers instead to the differential experience gap effect between male and female. This term comes out to be negative and significant, suggesting that when exploiting time variations in experience gap within individuals with the same unobserved ability, females display a greater wage penalty. For any additional year of experience gap, the additional penalty for females is equal to 1.5% at the 10th percentile (4% overall), and around 1.1% from the 25th to the 90 percentile (overall effect equal to 2.9% at the 25th, 2.4% at the median, 2.1% at the 75th, 1.8% at the 90th).

Since experience and experience gap have different magnitude, it is possible to have an idea of the comparison between the two effects using standard deviations. An increase in one standard deviation of the experience variable (8 years) entails a wage increase equal to 27.7% at the median, 36.1% at the 10th percentile, and 22.7% at the 90th percentile. At the same time, an increase in one standard deviation in experience gap (2.8 years for male and 3.4 years for females) would determine a wage penalty for men equal to 7.2% at the 10th percentile, 3.9% at the median, and 2.1% at the 90th percentile, and for female equal to 13.3% at the 10th percentile, 8.4% at the median, and 6.2% at the 90th. It is also possible to compute the ratio between the penalties due to an increase in one standard deviation of experience gap and the returns to experience due to a one standard deviation in experience. Interestingly, the penalties are more important for both males and females at the bottom of the distribution, since the ratio is equal to 20% for males and to 36.8% for females at the 10th percentile, to 14.1% for males and 30.1% for females at the median, and to 9.2% for males and 27.1% for females at the 90th percentile. The figures are quite impressive, especially for females: if a women experiences one standard deviation in experience gap, her returns to experience would reduce by 36.8% at the 10th percentile, almost twice greater than the figure for men (20%).

Table 3. Experience trends and Wage dynamics. Quantile fixed effects.

	Percentiles				
	10	25	50	75	90
Potential experience	0,052	0,043	0,038	0,035	0,030
	47,01	*44,30*	*41,17*	*35,96*	*29,63*
Potential experience^2	-0,001	-0,001	0,000	0,000	0,000
	-25,75	*-26,48*	*-24,22*	*-16,13*	*-6,74*
Experience gap	-0,024	-0,018	-0,013	-0,010	-0,007
	-18,02	*-16,04*	*-11,19*	*-8,45*	*-4,83*
Experience gap*gender	-0,015	-0,011	-0,011	-0,011	-0,011
	-6,48	*-6,65*	*-7,26*	*-7,23*	*-7,04*
White collar	0,106	0,102	0,098	0,100	0,108
	15,04	*16,09*	*15,97*	*16,24*	*17,01*
Apprentiship	-0,260	-0,247	-0,204	-0,171	-0,152
	-25,64	*-36,54*	*-37,90*	*-32,21*	*-27,34*
Part time contracts	-0,577	-0,538	-0,480	-0,426	-0,388
	-68,61	*-64,67*	*-64,32*	*-60,95*	*-52,70*

T-statistics in italics.

Notes: All estimates include geographic area and time dummies.

Our results show that experience gap has a substantial negative impact of wage dynamics, and this impact is stronger for females, even when controlling for unobserved ability. In the following we further investigate this issue, trying to identify the explanations behind this result. The main candidate explanation concerns the familiar duties of women, and more specifically the duties related to children. A possible way to capture this dimension is to introduce in the specification proxies for these duties. An interesting feature of our administrative database is that it includes the number of weeks for maternal leave for each workers. Since the required weeks for maternal leave are concentrated in the first years of the child, while the duties are supposed to continue over time, we introduce in the specification the cumulate variable of weeks in maternal leave (as percentage of the maximum working weeks, 52). Results are showed in Table 4. As expected, the coefficient for cumulative weeks in maternal leave is negative and high in magnitude. It is slightly decreasing along the distribution, since it is equal to 21.3% at the 10th, 13,4% at the 25th, 11% at the median, 9.9% at the 75th, and 9.2% at the 90%. When including this variable, the penalties related to the interaction term between experience gap and gender decrease as expected, suggesting that familiar duties, and in particular child care, lie behind the wage differential between males and females due to the experience gap.

As an additional investigation, in Table 5 we include, apart from the cumulative weeks in maternal leave, a dummy for childbirth, i.e. for having had in that year a maternal leave greater than 20 weeks in the year, to capture not a cumulative effect but rather a shock effect. From Table 5 it comes out that the childbirth dummy has a strong and decreasing negative impact on wages, being equal to 37% at the 10th percentile and to 1.9% at the 90th percentile. Introducing the childbirth dummy strongly decreases the impact of the cumulative weeks in maternal leave, as expected, while it does not affect much the coefficients related to the interaction between experience gap and females. This evidence suggests that familiar duties explain only a part of the wage penalties due to accumulated experience gaps.

Table 4. Experience trends and Wage dynamics, including cumulative weeks in maternal leave. Quantile fixed effects.

	Percentiles				
	10	25	50	75	90
Potential experience	0,054	0,045	0,040	0,037	0,032
	46,21	*44,83*	*42,06*	*37,84*	*31,01*
Potential experience^2	-0,001	-0,001	0,000	0,000	0,000
	-25,35	*-26,84*	*-24,28*	*-17,92*	*-7,86*
Experience gap	-0,028	-0,021	-0,016	-0,012	-0,009
	-20,89	*-18,00*	*-13,25*	*-10,28*	*-6,01*
Experience gap*gender	-0,008	-0,008	-0,008	-0,009	-0,009
	-3,46	*-4,59*	*-5,37*	*-5,66*	*-5,47*
Cumulative matern.leave	-0,213	-0,134	-0,110	-0,099	-0,092
	-16,24	*-16,19*	*-14,43*	*-13,99*	*-11,69*
White collar	0,107	0,102	0,097	0,097	0,106
	14,85	*16,04*	*15,78*	*15,82*	*16,33*
Apprentiship	-0,254	-0,242	-0,200	-0,167	-0,146
	-26,18	*-34,21*	*-36,28*	*-33,90*	*-26,87*
Part time contracts	-0,539	-0,511	-0,460	-0,408	-0,374
	-56,51	*-61,61*	*-63,47*	*-60,89*	*-51,86*

T-statistics in italics.

Notes: All estimates include geographic area and time dummies.

Table 5. Experience trends and Wage dynamics, including cumulative weeks in maternal leave and childbirth dummy. Quantile fixed effects.

	Percentiles				
	10	25	50	75	90
Potential experience	0,055	0,045	0,040	0,037	0,032
	46,97	*45,51*	*42,46*	*37,89*	*30,95*
Potential experience^2	-0,001	-0,001	0,000	0,000	0,000
	-26,26	*-28,40*	*-24,99*	*-17,85*	*-7,83*
Experience gap	-0,028	-0,021	-0,015	-0,012	-0,009
	-20,55	*-17,84*	*-13,02*	*-10,34*	*-5,99*
Experience gap*gender	-0,010	-0,008	-0,009	-0,009	-0,009
	-4,13	*-5,00*	*-5,72*	*-6,05*	*-5,80*
Cumulative matern.leave	-0,100	-0,090	-0,086	-0,087	-0,090
	-8,15	*-9,70*	*-10,37*	*-11,00*	*-10,28*
Childbirth dummy	-0,370	-0,188	-0,081	-0,032	0,019
	-19,88	*-17,44*	*-10,05*	*-4,36*	*1,48*
White collar	0,108	0,103	0,097	0,098	0,106
	15,04	*16,09*	*15,71*	*15,84*	*16,39*
Apprentiship	-0,256	-0,242	-0,201	-0,168	-0,147
	-26,28	*-36,53*	*-36,62*	*-33,43*	*-26,96*
Part time contracts	-0,553	-0,516	-0,462	-0,411	-0,376
	-58,21	*-61,80*	*-63,71*	*-62,60*	*-52,09*

T-statistics in italics.

Notes: All estimates include geographic area and time dummies.

7 Conclusion

The main contribution of this paper concerns the identification of discontinuities in working history on the workers' career profile, to assess to what extent they affect differently wages of female and male workers and in such a way contribute to the gender gap. We make use of a longitudinal dataset, called Administrative SILC, which records, since the beginning of their working life, the individuals working histories and various characteristics of private and public employees, self-employed, recipients of unemployment benefits and retired.

Using our data it is possible to reconstruct the cumulative individuals' discontinuities in working career, the experience gap, and the cumulative actual experience. We estimate the impact of the experience gap on wages of males and females. Furthermore, we investigate the whole earning distributions by means of the quantile fixed effects methodology (Canay, 2011), which allows to control for unobserved heterogeneity. This is crucial in our analysis, since career discontinuities can take place because of family duties but also because of heterogeneity in ability levels that are not observed. Our estimates are derived from a pooled sample of females and males.

Our results show that career discontinuities entail a negative and substantial effect on wages for all workers, and that this effect decreases along the wage distribution. This suggests that unskilled workers are the ones who suffer the most from career discontinuities. Further, we show that women suffer from an additional wage penalty with respect to men, and this additional penalty is only partially related to children duties.

References

[1] Addabbo T. and Favaro D. (2007), "Education and Wage Differentials by Gender in Italy," CHILD Working Paper No. 4, Centre for Household, Income, Labour and Demographic economics, Italy.

[2] Addabbo T., Favaro D. and Magrini S. (2007), "The Distribution of the Gender Wage Gap in Italy: Does Education Matter?," RECent Working Paper No. 4, Center for Economic Research, University of Modena and Reggio Emilia.

[3] Albrecht J., Bjorklund A., and Vroman S. (2003), "Is there a Glass Ceiling in Sweden?," *Journal of Labour Economics*, 21(1), pp.145-177.

[4] Altonji J.G. and Blank R.M. (1999), "Race and Gender in the Labor Market," in O.C. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Volume 3C, Elsevier Science, chapter 48, pp. 3143-3259.

[5] Arulampalam W., Booth A.L., and Bryan M.L. (2007), "Is there a glass ceiling in Europe? Exploring the gender pay gap across the wages distribution," *Industrial and Labor Relations Review*, 60(2), pp.163-186.

[6] Becchetti L., Massari R., and Naticchioni P. (2013), "The Drivers of Happiness Inequality: Suggestions for Promoting Social Cohesion," ECINEQ Working Papers 2010, University of Cassino working no.6/2011, forthcoming in Oxford Economics Papers.

[7] Bertrand M. and Hallock K. (2001), "The Gender Gap in Top Corporate Jobs," *Industrial and Labor Relations Review*, 55, 3-21.

[8] Bjerk D. (2008), "Glass Ceilings or Sticky Floors? Statistical Discrimination in a Dynamic Model of Hiring and Promotion," *Economic Journal*, Royal Economic Society, 118(530), pp. 961-982.

[9] Blau F. and Kahn L. M. (1992), "The Gender Earnings Gap: Learning from International Comparisons," American Economic Review, American Economic Association, 82(2), pp. 533-38.

[10] Blau F. and Kahn L. M. (2011), "The Feasibility and Importance of Adding Measures of Actual Experience to Cross-Sectional Data Collection," IZA Discussion Papers 5873, Institute for the Study of Labor (IZA).

[11] Blundell R., Gosling A., Ichimura H., and Meghir C. (2007), "Changes in the Distribution of Male and Female Wages Accounting for Employment Composition using Bounds," *Econometrica*, 75 (2), 323-363.

[12] Booth A.L. (2007), "The Glass Ceiling in Europe: Why are Women Doing Badly in the Labour Market?," *Swedish Economic Policy Review*, 14, pp. 121-144.

[13] Canay I. A. (2011), "A simple approach to quantile regression for panel data," *Econometrics Journal*, volume 14, pp. 368-386.

[14] De la Rica S., Dolado J., and Sanchez R.V. (2010), "Performance Pay and the Gender Wage Gap: Evidence from Spain," IZA Discussion Papers 5032, Institute for the Study of Labor (IZA).

[15] Fortin N., Lemieux T. and Firpo S. (2011), "Decomposition Methods in Economics," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*,Volume 4A, Elsevier Science, chapter 1, pp.1-102.

[16] Koenker R., and Bassett G. Jr. (1978), "Regression quantiles", Econometrica ,46, pp. 33-50.

[17] Manning A. and Swaffield J. (2008), "The gender gap in early-career wage growth", *Economic Journal*, Royal Economic Society, 118(530), pp. 983-1024.

[18] Machado J.A. and Mata J. (2005), "Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression," *Journal of Applied Econometrics*, 20(4), pp. 445-465.

[19] Matano A. and Naticchioni P. (2013), "Rent Sharing as a Driver of the Glass Ceiling Effect," *Economics Letters*, 118 (1), 55-59.

[20] Mincer J. and Polacheck S. (1978), "An Exchange: The Theory of Human Capital and the Earnings of Women: Women's Earnings Reexamined", *Journal of Human Resources*, University of Wisconsin Press, 13(1), pp. 118-134.

[21] Neumann M. (2007), "Deconvolution from panel data with unknown error distribution," *Journal of Multivariate Analysis,* 98, pp. 1955-68.

[22] Olivetti C. and Petrongolo B. (2008), "Unequal Pay or Unequal Employment? A Cross-Country Analysis of Gender Gaps," *Journal of Labor Economics*, 26 (4), pp. 621–654.

[23] Picchio M. and Mussida C. (2011), "Gender Wage Gap: A Semi-Parametric Approach With Sample Selection Correction," *Labour Economics*, Elsevier, 18(5), pp. 564-578.

[24] Picchio M. and Mussida C. (2012), "The Gender Wage Gap by Education in Italy," Discussion Paper 2012-021, Tilburg University, Center for Economic Research.

[25] Raitano M. (2009), "Differenze di genere nel mercato del lavoro dei paesi della UE15," *Rivista delle Politiche Sociali*, 2.

[26] Waldfogel J. (1998), "Understanding the “Family Gap” in Pay for Women with Children," *Journal of Economic Perspectives*, 12(1), pp. 137-156.