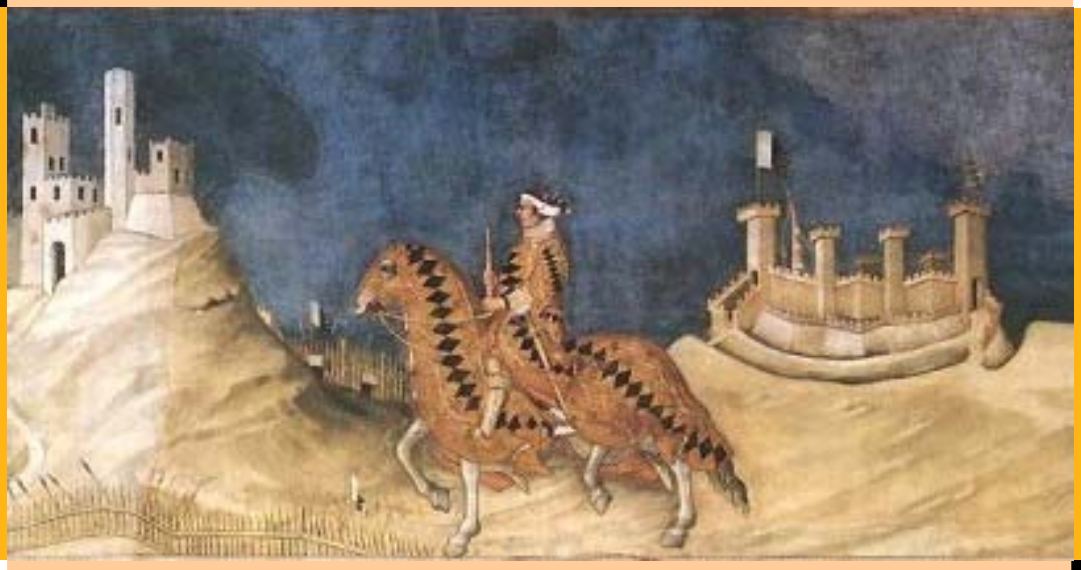
UNIVERSITÀ DEGLI STUDI DI SIENA

QUADERNI DEL DIPARTIMENTO
DI ECONOMIA POLITICA

Dario Sciulli

Conviction, Partial Adverse Selection and Labour Market Discrimination

n. 594 – Maggio 2010



Abstract - This paper analyses data from the 6th sweep of National Child Development Study to investigate the labour market perspective of convicted individuals. Decomposition analysis makes it clear that convicted workers are actually discriminated against both in terms of employment and wage with respect to non-convicted. Adopting a simple theoretical model accounting for partial adverse selection problem in the hiring process, I show that discrimination is not only explained in terms of economic stigma but also may derive from the inefficiency of the police/justice system in detecting crime and punishing offenders. In fact, while firms may apply economic stigma to recover the expected extra-costs from hiring convicted workers, firms rationality may impose to charge on convicted workers also unobservable expected extra-costs deriving from offenders non-convicted hired. The resulting over-stigma is increasing with the probability of offending and with the level of expected extra-costs, while it is decreasing with the probability of convicting offenders.

JEL codes: J71, K14, C21

Dario Sciulli, DMQTE, University of Chieti-Pescara and CEEAplA, d.sciulli@unich.it

Introduction[1]

The interest of economists on the relation between crime and labour market performance has strongly increased in the last decades, as the number of individuals involved in crime has risen (Freeman, 1999). As Holzer (2007) reviewed, most of the empirical evidence shows that criminal record or being sanctioned by the justice system produces negative effects in terms of future employment and earnings. More recently, contradictory results have emerged among studies that stresses the role of pre-existing heterogeneity in sorting individuals both into criminal activities and poor labour market performance. On the one side, Waldfagel (1994) finds negative effects increasing with the number of convictions and Grogger (1995) finds that they are moderate in magnitude and short-living. On the other side, Nagin and Waldfagel (1995) finds that conviction, favouring the exit from low paying apprentice scheme in pursuit of more profitable activities, increases wages in the short-term but reduces chances for advancement in the future. Cho and LaLonde (2005) and Kling (2006) find slightly positive short-term effects and negligible medium-term effects, respectively for specific groups of incarcerated women and among men who have spent time at work during incarceration. Poor labour market performance of individuals with criminal records may be explained in terms of both sides of the labour market. From a supply side point of view, according to the timing (childhood, adolescence or adulthood) of criminal records, crime may be associated with lower educational attainments and/or skills depreciation (Myers, 1983). The resulting lower productivity affects wage and possibly the range of employment opportunities. From a demand side point of view, stigmatization of offenders is believed to be one of the major source of their poor labour market performances. Finn and Fontaine (1985), show that employers are reluctant to hire offenders above all in case of violence against persons. Other economists measured the amount of stigma (for example Lott 1990, Waldfagel 1994, and Grogger 1995), while Rasmusen (1996) have tried to explain it from an economic point of view.

This paper contributes to this stream of literature adopting a discrimination approach to analyze the structure, the magnitude and the sources of the offenders' disadvantage on the labour market in terms of employment and wage. Discrimination may be seen as the valuation in the labour market of workers' characteristics unrelated to worker productivity and it is a cause of labour market failure and a source of income inequality. Apart from social prejudice, discrimination may occur because some groups of

[1] This paper was largely written during my research visit period (August-September 2009) at the Centre for Criminal Justice Economics and Psychology (University of York, UK), that I wish to thank for its support and hospitality. I am also indebted to Prof. Roger Bowles for useful discussions and to the participants at the V SIDE-ISLE Conference, Florence December 2009, for their suggestions. The usual disclaimer applies.

workers are suspected to possess unobservable characteristics negatively affecting productivity[2] or implying extra-costs for employers. From an empirical point of view, discrimination corresponds to a returns' penalty, given some characteristics, worsening labour market outcomes. There are various reasons to focus on discrimination against offenders. First, while labour market discrimination has been studied with respect to many groups (gender, race, age, religion and so on), it has received scarce attention with respect to offenders. Second, discrimination analysis may provide an actual measure of the disadvantage on the labour market of offenders, once the characteristics effect (including productivity differential) is controlled for. Finally, by adopting a simple theoretical framework modelling the adverse selection in the hiring process and accounting for the expected extra-costs from hiring offenders, it is possible to explain the poor labour market performances of offenders from an economic perspective rather than a prejudice perspective. Specifically, I show that the total amount of empirical discrimination against offenders, i.e. their actual disadvantage, is not composed only by stigma but also by what I call "over-stigma" deriving from the inefficiency of the police/justice system in detecting crime and punishing offenders (henceforth inefficient police/justice system). On the one side, employers may suspect that hiring offenders involves future extra-costs because of their higher probability of re-offending[3] (Entorf, 2009). This is *per se* a direct source of economic stigma involving penalties for offenders on the labour market. On the other side, as Rasmusen (1996) implicitly argued, an economy populated by offenders and non-offenders, and with an inefficient police/justice system, is characterized by an asymmetric information problem in the hiring process. In fact, since conviction is the only signal for employers to recognize offenders, and given the inefficiency of the police/justice system, only convicted offenders are recognizable as offenders while non-convicted offenders, and their expected extra-costs, remain unrecognizable. In this case, if firms recover expected extra-costs from unobservable offenders and act rationally, they should load these extra-costs on convicted individuals determining another source of their poor labour market performance. Obviously, this has implication in terms of wage inequality. Moreover, the existence of over-stigma has implications in terms of offending decisions in the context of dynamic offending frameworks (for an application see Bowles and Florackis, 2006), since it should be taken into account by individuals as a source of future loss in case of criminal activities.

Discrimination analysis is carried out by applying the standard Blinder-Oaxaca decomposition technique and its extensions. It allows to decompose

[2] Bushway (1998 and 2004) and Holzer, Raphael and Stoll (2006) apply discrimination thesis to crime and labour market literature to explain the poor labour market performance of ethnic groups (black people) in presence of limited access to criminal archives to identify criminals.

[3] Re-offence may determine both direct and indirect extra-costs. Direct extra-costs may arise in the form of theft, fraud, injury to co-workers and so on. Indirect extra-costs may arise in the form of new search and training costs for employers in case re-offender is dismissed.

the raw employment and wage differential determined by recognizable criminal activities between endowment effect (productivity differential) and coefficient effect that represents the valuation of workers' characteristics unrelated to worker productivity (stigma plus over-stigma). Specific econometric techniques is applied to correct estimates for selection bias and endogeneity problems, possibly determined by the most likely association between some labour market outcomes and criminal activities. Moreover, even thought cross-sectional data does not allow to correct estimates for pre-existing heterogeneity, the presence and the relevance of unobservable heterogeneity is tested and presented in the appendix.

The analysis is based on cross-sectional data from the 6th sweep (1999-2000) of the National Child Development Study (NCDS). The 6th sweep collects a large number of idiosyncratic, familiar and job related information on over 11000 individuals of the original cohort, and it is the only one including questions about conviction records in the time span since the last survey (1991), allowing us to identify convicted and non-convicted individuals.

Empirical results show that convicted individuals are less likely to be employed than non-convicted, while display higher observed wage. However wage decomposition makes it clear that discrimination is working and convicted workers are actually moderately penalized in terms of wage, while observed higher wages derives from segmentation of convicted with poor characteristics into self-employment, unemployment and inactivity. Finally, by combination of theoretical predictions and evidence of employment and wage discrimination suggests that firms expect extra-costs from hiring offenders justifying economic stigma. Moreover, in case of inefficiency of the police/justice system, the labour market performances of convicted individuals may worsen since discrimination against raises to include over-stigma determined by the loading of unobservable extra-costs.

This paper is organized as follows. Section 2 presents the simple theoretical model, while section 3 presents the econometric specification. Section 4 describes the data. Section 5 discusses the results. Finally, Section 6 concludes.

2. Inefficiency of the police/justice system and discrimination

This section presents a simple framework to model the effect on firms' labour costs of imperfect identification of offenders and discusses possible implications in terms of wage and employment discrimination. The identification of offenders is an important issue for firms, since hiring offenders may determine extra-costs due to their higher probability of re-offending. I assume that, since firms are aware of the risk of incurring in extra-costs, they may apply a discount rate on employment and wage (respectively g and d) to offenders to recover extra-costs. This is economic stigma and descends from a moral hazard problem. Without loss of generality for the

model predictions I make two assumptions to simplify the framework. First, I do not account for moral hazard assuming that extra-costs certainly arise. Second, the expected lower productivity of offenders is not modelled, since its effect on labour market outcomes concerns the endowment effect and not the coefficient effect measuring discrimination. I also assume that the only instrument that firms have to identify offenders is the observation of a criminal conviction, as similarly argued by Nagin and Waldfagel (1999). However, since the probability of conviction for offenders (π) is less than one, firms only have weak instrument for identification. This implies a partial adverse selection problem in the hiring process deriving from the inefficiency of the police/justice system, and it is the source of over-stigma.

2.1 Wage

Let us indicate by x the optimal wage paid to workers according to their productivity given their observable characteristics (we refer also to x as the fair wage). The optimal wage policy for firms provides a discount for the extra-costs of the wage paid to offenders, so the labour cost of employing an offender is x-d. So, d is the measure of the economic stigma to recover the future extra-costs deriving from hiring offenders. Let us assume that the proportion of the offending population is γ and that the economy is characterized by a fully efficient police/justice system, i.e. the probability of conviction is 1. In this case all offenders may be indentified and firms may discount the wage to allow for the future losses from re-offending. The standardized optimal labour cost paid by firms in an efficient system (ES) is:

$$wES = (1-\gamma)x + \gamma(x-d) \tag{1}$$

However if the probability of conviction of offenders is less than 1 then only a part of offenders are observable ($\gamma\pi$), while the remaining ($\gamma(1-\pi)$) are unobservable and firms consider them as non-offenders (partial adverse selection). The standardized optimal labour cost paid in an inefficient system (IS) is now:

$$wIS = (1-\gamma)x + \gamma\pi(x-d) + \gamma(1-\pi)x \tag{2}$$

Since $wIS > wES$ and if firms pay their own employees the fair wage, the inefficiency of the police/justice system results in a profit loss for firms (cost of the inefficiency). The cost of the inefficiency in terms of wage is:

$$CIw = wIS - wES = d\gamma(1-\pi) \tag{3}$$

i.e. CIw is increasing in the future extra-costs, and in the proportion of the population offending, while it is decreasing in the probability of convicting offenders. It represents the part of discrimination due to unobservable

offenders. High values of γ and low values of π are associated with high CI, since both increase the probability of facing unobservable extra-costs.

2.2 Employment

An identical approach is used to model the cost of inefficiency in terms of employment. The standardized workforce in an efficient system is:

$$eES = (1-\gamma) + \gamma(1-g) \tag{4}$$

Where g represents the discount on employment to recover the future extra-costs of offenders (economic stigma). It implies that the level of employment is lower with respect to a labour market without offenders. As before, g may be thought as the coefficients effect in the context of employment differential, while the discount due to worse job related characteristics of convicted individuals represents the endowments effect. The standardized workforce in an inefficient system is:

$$eIS = (1-\gamma) + \gamma\pi(1-g) + \gamma(1-\pi) \tag{5}$$

Since $eIS > eES$, if firms hire workers according to their observable characteristics, then the inefficiency of the system determines a surplus of employment with respect to the optimal level (cost of inefficiency). The cost of inefficiency in terms of employment is:

$$CIe = eIS - eES = g\gamma(1-\pi) \tag{6}$$

Conclusions equivalent to the previous ones may be drawn about the relationship between the cost of inefficiency and explanatory variables.

2.3 The loading of unobservable extra-costs

The model predicts that because of partial adverse selection, firms may find themselves in sub-optimal solutions both in terms of wage and employment. Rational behaviour of firms provides for loading the cost of inefficiency on employees, reducing their employment and/or wage through a reduction of characteristics' returns (discrimination due to unobservable extra-costs). This reduction is added to that applied to convicted individuals to recover observable extra-costs. Which group of workers (convicted or non-convicted) will be charged this further penalty? Here we argue that if firms act rationally it is likely that the cost deriving from partial adverse selection is charged on convicted individuals, determining a worsening of their labour market perspective (over-stigma). From an employment point of view, for firms it should be more efficient to discriminate against convicted individuals. On the one side, they are likely to be less productive while discriminating against non-convicted individuals may imply a relatively greater loss in production

level. On the other side, convicted individuals may involve greater extra-costs per capita. From a wage point of view, discrimination against convicted individuals is more likely since, for example, they should be more prone to accept lower wages to work, or because of their lower bargaining power. Finally, the decision of applying both employment and wage discrimination may be dictated by rational constraints with respect to the hiring decisions of firms and, above all, job-search behaviour of employees. For instance, if firms decide to apply only wage discrimination, and the actual wage offered to convicted workers is lower than their reservation wages, then firms would not recover the cost of inefficiency, since no convicted would be employed. To avoid this risk, firms can choose to recover the cost of inefficiency sharing it between wage and employment discrimination.

2.4 Structure of wage differential

From the above, it is possible to infer the structure of the employment and wage differential between convicted and non-convicted people. Let us consider just the structure of the wage differential, since the structure of the employment differential is very close. The first term, h, represents the endowment effect, i.e. the wage differential due to differences in characteristics, and include the differences in productivity level[4]. The second term, d, represents the wage differential due to the reduction of returns to characteristics of convicted individual to recover the observable future extra-costs. The third term, $d(1-\pi)/\pi$, represents the wage differential due to the reduction of returns to characteristics of convicted individual to recover the unobservable future extra-costs. It corresponds to the quota of the cost of inefficiency paid by each convicted individual hired. The sum of the second and the third term represents the wage discrimination, i.e. the wage differential determined by returns' penalty. The sum of the first two terms may be seen as the disadvantage that convicted individuals expect to face in the labour market, while the third term is the so-called over-stigma and strictly depends on the inefficiency of the police/justice system. It represents an unexpected disadvantage in the labour market for convicted individuals.
The figure below describes the individual raw differential structure for wage:

h	***d***	***d[(1-π)/π]***
Endowment effect	Coefficient effect	
Including productivity differential	Stigma	Over-stigma
Expected differential		Unexpected differential
Raw differential		

According to the previous consideration, if empirical evidence confirms the existence of discrimination against convicted individuals, it would be a support for three main points. First, extra-costs exist. Second, firms are

[4] The first term also includes wage differential not directly related with productivity. For example, specific local market effect or firm size effect and so on.

informed of their existence and operate to absorb them. Third, and more importantly, the inefficiency of the police/justice system is a further source of poor labour market performance for convicted individuals (over-stigma), that is added to the wage differential due to their suspected lower productivity and to the future extra-costs that are suspected to be involved (stigma effect).

3. Econometric analysis

This study aims at investigating the existence of discrimination against workers who have conviction records in terms of employment and wage. Discrimination analysis requires estimating employment and wage equations both for convicted workers (C) and non-convicted workers (NC), and applying the decomposition technique due to Oaxaca (1973) and Blinder (1973). A generalization for non-linear outcomes (Fairlie, 2005) is applied to investigate employment discrimination.

Since we are interested in studying discrimination in the sphere of the relationship between conviction and the probability of being hired by employers, the self-employment condition is not investigated and the self-employed are included in the not employed group along with the unemployed and inactive.

A static employment equation for each group k (k = C, NC) is estimated. For each group it reads:

$$e_{ik} = 1\left[\beta_k X_{ik} + \varepsilon_{ik} \geq 0\right] \tag{7}$$

where e is a binary variable indicating whether the individual i is employed in the year 2000, X is a vector of individual, familiar and job related characteristics, β is a set of parameters to be estimated and ε is the error term. If the distribution is symmetric then:

$$\Pr\left(e_{ik} = 1 \mid X_{ik}\right) = \Pr\left(e_{ik} \geq 0 \mid X_{ik}\right) = \Pr\left(\varepsilon_{ik} \leq -\beta_k X_{ik} \mid X_{ik}\right) = F\left(\beta_k X_{ik}\right) \tag{8}$$

if we assume that the error term has a standard normal distribution we obtain a probit model, hence the probability of employment reads:

$$\Pr\left(e_{ik} = 1 \mid X_{ik}\right) = \Phi\left(\beta_k X_{ik}\right) \tag{9}$$

The probit model is usually estimated by maximum likelihood estimation and the log likelihood for a sample of independent observations may be written as:

$$\log L(\beta) = \sum_{i=1}^{N} e_i \log F(\beta X) + \sum_{i=1}^{N} (1 - e_i) \log\left[1 - F(\beta X)\right] \tag{10}$$

The decomposition for the non-linear equation, following Fairlie (2005) may be written as:

$$\bar{e}^{NC} - \bar{e}^{C} = \left[\sum_{i=1}^{N^{NC}} \frac{F\left(X_i^{NC} \hat{\beta}^{NC}\right)}{N^{NC}} - \sum_{i=1}^{N^{C}} \frac{F\left(X_i^{C} \hat{\beta}^{NC}\right)}{N^{C}} \right] + \left[\sum_{i=1}^{N^{C}} \frac{F\left(X_i^{C} \hat{\beta}^{NC}\right)}{N^{C}} - \sum_{i=1}^{N^{C}} \frac{F\left(X_i^{C} \hat{\beta}^{C}\right)}{N^{C}} \right] \quad (11)$$

To uncover the existence of wage discrimination against convicted workers we estimate a wage equation for each group *k*. A standard wage equation reads:

$$w_{ik} = \widetilde{\alpha}_k \widetilde{Z}_{ik} + \xi_k S_{ik} + u_{ik} \quad (12)$$

where *w* is the log of the hourly wage, *Z* is a vector of individual, familiar and job related characteristics, *S* is the educational variable, α is a vector of unknown parameters to be estimated, ξ is the educational level parameter and *u* is the spherical disturbance. Generally, since the probability of observing wage depends on the selection into employment rather than on the other possible economic status due to specific individual characteristics, wage equations are usually corrected by introducing a selection term to obtain consistent estimates. This appears particularly relevant in our case, since convicted workers are more likely to experience selection into inactivity, unemployment or self-employment as a result both of stigma and/or discouragement or self-selection. To estimate a selection-corrected wage equation a Heckman selection model (1979) is applied. Therefore, in the first step we run the previous employment equation to estimate the probability of selecting into employment and to calculate the Inverse Mill's ratio (IMR), while in the second step a wage equation including the selection term is estimated. The IMR is given by:

$$\lambda_{ik}\left(\beta_k X_{ik}\right) = \frac{\phi\left(\beta_k X_{ik}\right)}{\Phi\left(\beta_k X_{ik}\right)} \quad (13)$$

where φ and Φ denote, respectively, the probability density and cumulative distribution functions of the standard normal distribution. In the second step wage equations including the selection term are estimated. Therefore, the Heckman selection model we estimate is:

$$E\left[w_{ik} \mid \widetilde{Z}_{ik}, S_{ik}, e_{ik} = 1\right] = \widetilde{\alpha}_k \widetilde{Z}_{ik} + \xi_k S_{ik} + E\left[u_{ik} \mid \varepsilon_{ik} > -\beta_k X_{ik}\right] \quad (14)$$

and the selection corrected wage equation, or the conditional wage equation, reads:

$$w_{ik} = \widetilde{\alpha}_k \widetilde{Z}_{ik} + \xi_k S_{ik} + \delta_k \hat{\lambda}_{ik} + v_{ik} \quad (15)$$

where the error terms are assumed to follow a bivariate normal distribution with means zero and variances 1 for the selection equation and σ_u for the wage equation, and ρ is the correlation coefficient and lambda the inverse mills ratio.

Another typical problem arising in cross-sectional data models is the endogeneity of covariates that occur when they are correlated with the disturbances due to, for example, omitted variables or measurement error. If a covariate is endogenous, then the orthogonality assumption is violated and ordinary least squares estimates are not consistent. Since endogeneity tests reveal that the educational variable is endogenous, at least in the wage equation of non-convicted people, we apply a two stage least squares regression model (2SLS). The educational level is instrumented by a set of variables available from the NCDS1, NCDS2 and NCDS3, that should be correlated with the instrumented variable and not correlated with the disturbances. Specifically, we use a variable indicating the presence of problems (death of one of parents, unemployment, financial problems, separation/divorce, family conflicts, domestic tension) in the family of origin of the individual when he/she was 7 years old, two variables indicating the attention of fathers and mothers for the education of their children when he/she was 11 years old, variables indicating the parents' expectations about the maximum education level that the children would have reached at age 16, a dummy indicating if the individual was suspended at least once from school. According to the 2SLS method, in the first stage the endogenous variable is estimated both on the exogenous variables of the wage equation and on the vector of instrumental variables (Q_{ik}):

$$S_{ik} = \widetilde{\alpha}_k \widetilde{Z}_{ik} + \delta_k \hat{\lambda}_{ik} + \vartheta_k Q_{ik} + \eta_{ik} \tag{16}$$

In the second stage the original selection corrected wage equation is estimated by including the instrumented educational level

$$w_{ik} = \widetilde{\alpha}_k \widetilde{Z}_{ik} + \xi_k \hat{S}_{ik} + \delta_k \hat{\lambda}_{ik} + \upsilon_{ik} \tag{17}$$

where the spherical disturbance υ_{ik}, is assumed to be uncorrelated with η_{ik}.
The presence of wage discrimination between the two groups is investigated by applying the Oaxaca (1973)-Blinder (1973) decomposition revisited by Oaxaca and Ransom (1994). Based on the conditional wage estimation, that decomposition reads:

$$\overline{w}^{NC} - \overline{w}^{C} = \alpha^{*}\left(\overline{Z}^{NC} - \overline{Z}^{C}\right) + \left[\overline{Z}^{NC}\left(\hat{\alpha}^{NC} - \alpha^{*}\right) + \overline{Z}^{C}\left(\alpha^{*} - \hat{\alpha}^{C}\right)\right] \tag{18}$$

Where Z is the full set of covariates, including the educational variable, and α is the full set of estimated parameters. Finally:

$$\alpha^{*} = \Omega\hat{\alpha}^{NC} + (I - \Omega)\hat{\alpha}^{C} \tag{19}$$

is an approximation of the unobserved non discriminatory earnings structure, with I indicates an identity matrix and

$$\Omega = \left(X^{NC'}X^{NC} + X^{C'}X^{C}\right)^{-1} X^{NC'}X^{NC} \tag{20}$$

is the Oaxaca and Ransom (1994) weighting matrix[5].

4. Data

Econometric analysis is based on the information gathered by 6th sweep of the National Child Development Study (NCDS) in the 1999-2000. The NCDS is a continuing longitudinal study that seeks to follow the lives of all those living in England, Scotland and Wales who were born in one particular week in 1958. The main aim of the study is to improve the understanding of the factors affecting human development over the whole lifespan. The NCDS has its origin in the Perinatal Mortality Survey (PMS) that collected information on a cohort of about 17000 children. Successively, the PMS became the NCDS that has gathered information on the same individuals at different points in time (1965, 1969, 1974, 1981, 1991, 1999-2000 and 2004-2005). The 6th NCDS sweep took place in 1999-2000, when cohort members were aged 41-42 years, providing a large set of information over 11000 of the original cohort individuals. Specifically, the dataset covers topics such as household, housing, relationships, children, family, social relationship and support, income, employment, lifelong learning, health, citizenship and values and, finally a self-completion part that includes information about drug use and contact with the police and crime. Retrospective information from the 1st-3rd and 5th sweeps is also used in the context of endogeneity problems and to construct some variables used in the 6th sweep analysis. The question "Been found guilty by a court since the reference date?" is used here to identify individuals with and without conviction records in the time span between 1991 and 1999. Information about economic status allows us to distinguish employed individuals from those who are self-employed, unemployed and out of the labour force. In this context, since self-employed are not subjected to economic stigma, they are simply considered as not employed. Finally information on net pay, the period covered and usual and overtime hours worked per week, allows us to construct the hourly wage variable[6]. The

[5] Oaxaca (1973) and Blinder (1973) proposed that the weighting matrix is a null matrix or an identity matrix. Reimers (1983) proposed that it is 0.5*I, while Cotton (1988) proposed that it corresponds to *s**I where *s* denotes the relative sample-size of the majority group.

[6] Given that about 15% missing information on working hours, the imputation technique was implemented to avoid the loss of much information. The auxiliary regression used to impute missing data is available upon request.

resultant hourly wage variable was subjected to top and bottom coding at 1% to reduce bias from outliers, and for the same reason we excluded from our sample individuals (25 observations) declaring to work less than 7 hours per week or more than 84 hours per week. Because of missing information the sample used to estimate employment equation is composed of 10330 individuals, of whom 441 have conviction histories, while the sample used to estimate the wage equation is composed of 7151 individuals, of whom 260 have conviction records. Finally, a further restriction is made (5320 individuals, of whom 191 have conviction records) with regard to the wage estimation by 2SLS because of missing data in the instrumental variables.

Table 1 reports descriptive statistics both for the employment equation sample and for the wage equation sample.

Table 1 Descriptive Statistics

	Employment equation				Wage equation			
	NON-CONVICTED		CONVICTED		NON-CONVICTED		CONVICTED	
	Mean	Std.dev.	Mean	Std.dev.	Mean	Std.dev.	Mean	Std.dev.
Male	0.469	0.499	0.821	0.384	0.478	0.500	0.842	0.365
Educational level	2.571	1.491	2.331	1.478	2.636	1.480	2.515	1.482
Ethnic	0.014	0.118	0.025	0.156	0.014	0.116	0.031	0.173
Disability	0.018	0.134	0.023	0.149	0.013	0.115	0.023	0.150
Poor health	0.031	0.172	0.063	0.244	0.011	0.102	0.019	0.138
Drug use	0.320	0.467	0.531	0.500	0.300	0.458	0.462	0.499
Married	0.814	0.389	0.689	0.463	0.828	0.377	0.719	0.450
Children 0-6	0.196	0.397	0.179	0.384	-	-	-	-
Permanent contract	-	-	-	-	0.952	0.214	0.958	0.202
Private sector	-	-	-	-	0.598	0.490	0.765	0.425
Manager/Professional	-	-	-	-	0.318	0.466	0.404	0.492
Non manual worker	-	-	-	-	0.173	0.378	0.092	0.290
Junior non manual	-	-	-	-	0.200	0.400	0.046	0.210
Skilled manual worker	-	-	-	-	0.146	0.353	0.292	0.456
Semi-skilled	-	-	-	-	0.067	0.250	0.085	0.279
Unskilled/Personal/Agricultural	-	-	-	-	0.095	0.293	0.077	0.267
Tenure	-	-	-	-	9.133	7.882	9.142	8.865
Number of unemployment spells	0.084	0.358	0.222	0.607	0.078	0.336	0.177	0.489
Firm size 0-9	-	-	-	-	0.167	0.373	0.135	0.342
Firm size 10-24	-	-	-	-	0.164	0.370	0.150	0.358
Firm size 25-99	-	-	-	-	0.257	0.437	0.296	0.457
Firm size 100-499	-	-	-	-	0.227	0.419	0.265	0.442
Firm size over 500	-	-	-	-	0.185	0.388	0.154	0.361
Part-time job	-	-	-	-	0.234	0.423	0.065	0.248
Vocational	0.303	0.460	0.340	0.474	0.319	0.466	0.365	0.482
Union membership	-	-	-	-	0.374	0.484	0.362	0.481
Computer use	0.658	0.475	0.571	0.495	0.709	0.454	0.638	0.481
Ability learn new skills	-	-	-	-	0.549	0.498	0.619	0.487
Partner employed	0.694	0.461	0.531	0.500	-	-	-	-
North-east	0.062	0.242	0.068	0.252	0.065	0.247	0.073	0.261
Yorkshire-The Humber	0.091	0.288	0.100	0.300	0.093	0.290	0.081	0.273
East midlands	0.074	0.262	0.077	0.267	0.078	0.268	0.100	0.301
East-anglia	0.041	0.199	0.041	0.198	0.041	0.198	0.035	0.183
South-east	0.297	0.457	0.297	0.457	0.288	0.453	0.281	0.450
South-west	0.097	0.296	0.082	0.274	0.092	0.289	0.088	0.285
West-midlands	0.088	0.283	0.093	0.291	0.092	0.290	0.092	0.290
North-west	0.106	0.307	0.084	0.278	0.109	0.312	0.096	0.295
Wales	0.053	0.225	0.052	0.223	0.050	0.218	0.038	0.193
Scotland	0.090	0.286	0.107	0.309	0.091	0.288	0.115	0.320
Observations	9889		441		6891		260	

Source: my elaboration on sixth sweep of NCDS data

Variables used to control for observable heterogeneity include individual, familiar and job related characteristics. Specifically, the variability in the employment equation is controlled for gender, educational level (six levels ranging from no qualification to NVQ4/5), ethnic, disability status, poor health conditions, drug use (present or past), marriage status, presence of children aged 0-6 (including an interaction dummy to control for the different effect of children on fathers and mothers), the number of previous spells of unemployment to proxy the cumulated work experience, a dummy for individuals who received vocational training, ability in computer use, a dummy identifying the partner's employment status, finally regional dummies are introduced to control for country-local specific effects (the South-east, including London, is the base-category). Wage equations variability is controlled for the same variables excluding the presence of children and the employment status of partner. Moreover, a set of job related characteristics are included to control for some specific source of individual wage heterogeneity. Specifically we include a dummy for individuals employed with a permanent contract, employed in private sector, dummies to control for profession (according to the socio-economic group definition), tenure, firm size, part-time employment, union membership and, finally, a dummy to control for ability in learning new skills (taking the value one if the individual has good ability) introduced to proxy the individual ability heterogeneity that is usually unobservable in many datasets. Table 2 shows the mean values of the employment rate and wage between individuals with and without convictions, as preliminary information about the effect of convictions on labour market outcomes.

Table 2 Mean employment and wage

	NON-CONVICTED		CONVICTED			
	Mean	s.e.	Mean	s.e.	t-test	P-value
Employment	0.734	0.004	0.637	0.023	4.497	0.000
Wage	1.791	0.007	1.850	0.038	-1.537	0.124

Source: my elaboration on sixth sweep of NCDS data

The observed employment rate is higher for individuals without convictions than among individuals with conviction records[7], 73.4% versus 63.7%[8], and the difference is statistically significant according to the t-test. Against expectations, the observed wages are higher among individuals with conviction records than among individuals without convictions (1.85 versus 1.79, meaning 6.35 pounds per hour versus 6 pounds per hour) even though the difference is not significant in statistical sense (p-value is about 12%). However, as discrimination analysis below seems to indicate, such advantage

[7] Convicted individuals are more likely to be self-employed (17.4% vs. 12.2%), unemployed (4.8% vs. 2%) and inactive (14.1% vs. 12.4%) than non-convicted, confirming the greater difficulties of integration into labour market due to conviction records.
[8] The full sample shows a greater differential: 72.7% versus 61.1%.

may simply indicate a stronger selection of convicted with "good" characteristics into employment and relegation into self-employment or inactivity of convicted with "bad" characteristics, justifying a push-up effect for the observed wage of employed convicted individuals.

5. Results

Empirical results are based on cross-sectional data models. Tables 3 and 4 present the results of the employment analysis, while tables 5, 6 and 7 present the results of the wage analysis. Since we are interested in discrimination analysis, employment and wage equations are estimated separately for each group: convicted and non-convicted individuals. We first comment on the employment analysis, and then we focus on the wage analysis. Employment equations are estimated by applying probit models and controlling for individual, familiar, job-related and local-specific characteristics to account for observable heterogeneity.

Table 3 Employment equation estimates

	NON-CONVICTED				CONVICTED			
	b	s.e.	P-value	% Effect	b	s.e.	P-value	% Effect
Male	-0.015	0.032	0.643	-1.5%	0.134	0.193	0.488	14.3%
Educational level	0.029	0.010	0.004	2.9%	0.097	0.048	0.046	10.2%
Ethnic	-0.096	0.115	0.403	-9.2%	0.399	0.482	0.407	49.0%
Disability	-0.257	0.102	0.012	-22.6%	0.150	0.443	0.734	16.2%
Poor health	-1.238	0.082	0.000	-71.0%	-0.962	0.298	0.001	-61.8%
Drug use	-0.182	0.031	0.000	-16.7%	-0.341	0.135	0.012	-28.9%
Married	-0.068	0.047	0.145	-6.6%	0.095	0.182	0.602	9.9%
Children 0-6	-0.543	0.049	0.000	-41.9%	-1.175	0.458	0.010	-69.1%
Male*Children 0-6	0.543	0.069	0.000	72.2%	0.828	0.490	0.091	128.9%
Number of unemployment spells	-0.095	0.038	0.013	-9.1%	-0.112	0.111	0.315	-10.6%
Vocational	0.110	0.032	0.000	11.7%	0.126	0.144	0.382	13.4%
Computer use	0.430	0.031	0.000	53.8%	0.328	0.139	0.018	38.8%
Partner employed	0.299	0.039	0.000	34.9%	0.394	0.164	0.016	48.3%
North-east	0.184	0.064	0.004	20.2%	0.090	0.280	0.747	9.5%
Yorkshire-The Humber	0.142	0.054	0.008	15.3%	-0.117	0.239	0.625	-11.1%
East midlands	0.153	0.059	0.009	16.6%	0.615	0.297	0.038	84.9%
East-anglia	0.086	0.074	0.246	9.0%	0.020	0.359	0.955	2.1%
South-west	-0.076	0.051	0.133	-7.3%	0.025	0.259	0.922	2.6%
West-midlands	0.160	0.055	0.003	17.4%	0.023	0.247	0.925	2.3%
North-west	0.131	0.051	0.010	14.0%	0.140	0.268	0.602	15.0%
Wales	-0.036	0.065	0.579	-3.5%	-0.463	0.310	0.136	-37.0%
Scotland	0.115	0.054	0.033	12.2%	0.049	0.233	0.833	5.0%
Intercept	0.218	0.051	0.000	24.3%	-0.146	0.258	0.570	-13.6%
Observations	9889				441			
LR Chi2 (22)	901.11				84.65			
Pseudo R2	0.0787				0.1465			

Source: my elaboration on sixth sweep of NCDS data

Emerging results are quite standard. As expected, educational level positively affects employment probabilities of both groups; interestingly the return is higher for convicted individuals even though the estimate is less significant in terms of p-value then the non-convicted case. Disability reduces the probability of employment only for non-convicted individuals. On the contrary

poor health status significantly affects the employment probabilities of both convicted and non-convicted individuals; however non-convicted individuals seem to suffer a greater penalty. Drug consumption (present or past) also reduces employment probabilities both of convicted and non-convicted individuals. Interestingly the negative effect is stronger for convicted individuals, possibly indicating that the association between conviction and drug use produces a worsening of the stigma against convicted individuals. Consistently with major studies regarding employment opportunities and child care, having children aged 0-6 strongly reduces employment probabilities for both groups, even though the effect is stronger among convicted individuals. On the one side, this suggests a weakness of the child-care system. On the other side, as expected, the presumed weakness of the child-care system appears to affect females rather than males. In fact, the estimation of the interaction dummy introduced to control that duality is indicative that males are not disadvantaged by having children aged 0-6. The number of spells of unemployment cumulated since the entry in the labour market is introduced to capture the effect of an interrupted career pattern that tends to reduce the accumulation of working experience. Have an increased number of unemployment spells reduces the probability of employment of non-convicted people. The estimate is not significant for convicted people, possibly because having previous conviction periods include the possibility of career pattern interruption, hence the specific effect of our control variable may be ineffective from the firms' point of view. Vocational training appears to increase employment probabilities, meaning that specialization in working abilities is appreciated by employers. The covariate is not significant for convicted individuals. Similarly to the previous control variable, acquaintance with a personal computer increases employment probabilities in a labour market characterized by increasing use of new technologies. An employed partner employed also increases the probability of employment. It means that the substitution effect is deactivated while, as specific literature suggests, it is likely that a coupling effect is working, i.e. it is more likely to observe partnership between individuals with similar characteristics (for example education) favouring the employment probabilities of both or conversely. Finally regional dummies are introduced to control for local labour market effects. The South-East (that includes London) is the reference category. According to our estimates among non-convicted individuals, employment probabilities are higher in Northern regions, including Scotland, and in Midlands regions rather than in the reference region, while the other regions do not present significant estimates. Among convicted individuals only East Midlands present a significant, and strongly positive, effect with respect to the South-East, while other regions do not show significant differences.

Table 4 presents evidence from the Fairlie (2005) decomposition, which allows decomposition analysis for non linear outcomes. Employment rate among non-convicted is about 10% greater than among convicted individuals (73.4% against 63.7%). The Fairlie decomposition clarifies that only 4.75% of

that differential (corresponding to 48.9% of the total differential) may be explained in terms of endowments, while the remaining 4.96% (51.1% of the total differential) is unexplained, i.e. attributable to differences in returns (and unobservables). Evidence of discrimination against convicted individuals in terms of employment, is indicative that firms take into account the expected extra-costs of hiring convicted workers. Moreover, according to our model predictions, because of partial adverse selection, part of that discrimination actually is an over-stigma against convicted individuals to recover expected extra-costs of unobservable offenders.

Table 4 Fairlie Decomposition

Observed		Differential	Explained		Unexplained	
Non-convicted	Convicted		Absolute	%	Absolute	%
0,7343	0,6372	0,0971	0,0475	48,91%	0,0496	51,09%

Source: my elaboration on sixth sweep of NCDS data

Table 5 reports wage equations estimation by adopting various liner models. Specifically, in turn, an OLS model, a Heckman (1979) model and a 2SLS model with selectivity correction are used to estimate the effect of observable variables on the dependent variable, i.e. the log of hourly wage. Specifically, the first two columns regard the OLS estimates; third and fourth columns regard Heckman model estimates, which correct OLS estimates for sample selection, and finally the fifth column presents estimation results using a 2SLS model with selectivity correction, that allows me to correct the estimation bias deriving from endogeneity of the educational variable[9].

[9] Table 6 also reports Sargan and Bausmann overidentification tests about the adequacy of used instruments.

Table 5 Wage equations estimates

	OLS						OLS corrected (Heckman model)						2SLS corrected		
	NON-CONVICTED			CONVICTED			NON-CONVICTED			CONVICTED			NON-CONVICTED		
	b	s.e.	P-value	b	s.e.	P-value	b	s.e.	P-value	b	s.e.	P-value	b	s.e.	P-value
Male	0.209	0.015	**0.000**	0.139	0.111	0.213	0.210	0.015	**0.000**	0.146	0.105	0.165	0.179	0.018	**0.000**
Educational level	0.067	0.005	**0.000**	0.068	0.027	**0.011**	0.067	0.005	**0.000**	0.073	0.027	**0.008**	0.156	0.014	**0.000**
Ethnic	0.037	0.050	0.457	0.052	0.210	0.805	0.035	0.050	0.487	0.072	0.202	0.724	0.056	0.076	0.462
Disability	-0.016	0.050	0.747	-0.517	0.245	**0.036**	-0.021	0.051	0.673	-0.508	0.230	**0.027**	0.015	0.059	0.796
Poor health	-0.023	0.056	0.688	-0.481	0.261	**0.067**	-0.049	0.075	0.512	-0.547	0.294	**0.062**	-0.040	0.092	0.658
Drug use	0.026	0.013	**0.043**	0.028	0.072	0.698	0.023	0.014	0.113	0.008	0.083	0.919	0.014	0.018	0.439
Married	0.037	0.016	**0.017**	0.066	0.080	0.412	0.039	0.016	**0.015**	0.076	0.079	0.337	0.042	0.020	**0.032**
Permanent contract	0.038	0.028	0.173	-0.100	0.183	0.586	0.038	0.028	0.172	-0.098	0.170	0.564	0.078	0.033	**0.017**
Private sector	-0.023	0.014	**0.093**	0.233	0.098	**0.019**	-0.024	0.014	**0.090**	0.233	0.091	**0.011**	0.032	0.018	**0.073**
Non manual	-0.173	0.018	**0.000**	-0.059	0.130	0.652	-0.173	0.018	**0.000**	-0.055	0.121	0.650	-0.150	0.022	**0.000**
Junior non manual	-0.312	0.019	**0.000**	-0.303	0.169	**0.075**	-0.312	0.019	**0.000**	-0.299	0.158	**0.058**	-0.226	0.025	**0.000**
Skilled manual	-0.311	0.020	**0.000**	-0.285	0.099	**0.004**	-0.311	0.020	**0.000**	-0.286	0.092	**0.002**	-0.201	0.029	**0.000**
Semi-skilled manual	-0.399	0.027	**0.000**	-0.254	0.144	**0.079**	-0.399	0.027	**0.000**	-0.253	0.133	**0.057**	-0.245	0.039	**0.000**
Unskilled/Personal/Agricultural	-0.590	0.025	**0.000**	-0.422	0.159	**0.009**	-0.590	0.025	**0.000**	-0.418	0.148	**0.005**	-0.444	0.036	**0.000**
Tenure	0.002	0.001	**0.006**	0.005	0.004	0.309	0.002	0.001	**0.006**	0.004	0.004	0.302	0.003	0.001	**0.001**
Number of unemployment spells	-0.077	0.018	**0.000**	-0.062	0.076	0.419	-0.078	0.018	**0.000**	-0.071	0.075	0.342	-0.065	0.022	**0.003**
Firm size 1-9	-0.207	0.021	**0.000**	-0.003	0.137	0.982	-0.207	0.020	**0.000**	-0.004	0.128	0.976	-0.173	0.025	**0.000**
Firm size 10-24	-0.087	0.020	**0.000**	0.046	0.131	0.723	-0.087	0.020	**0.000**	0.047	0.122	0.699	-0.067	0.024	**0.005**
Firm size 25-99	-0.077	0.018	**0.000**	0.084	0.111	0.449	-0.076	0.018	**0.000**	0.084	0.104	0.415	-0.054	0.021	**0.010**
Firm size 100-499	-0.055	0.018	**0.002**	0.029	0.111	0.792	-0.055	0.018	**0.002**	0.027	0.104	0.792	-0.050	0.022	**0.021**
Part-time	-0.105	0.017	**0.000**	-0.032	0.157	0.841	-0.106	0.017	**0.000**	-0.037	0.146	0.803	-0.126	0.021	**0.000**
Vocation	-0.047	0.013	**0.000**	-0.197	0.074	**0.009**	-0.045	0.013	**0.000**	-0.189	0.072	**0.009**	-0.018	0.016	0.246
Union	0.031	0.014	**0.027**	0.109	0.084	0.192	0.031	0.014	**0.027**	0.110	0.078	0.158	0.004	0.017	0.830
Computer ability	0.129	0.014	**0.000**	0.111	0.082	0.177	0.137	0.020	**0.000**	0.124	0.083	0.137	0.107	0.024	**0.000**
Ability learn new skills	0.037	0.012	**0.002**	0.068	0.074	0.359	0.037	0.012	**0.002**	0.069	0.068	0.314	0.001	0.015	0.942
North-east	-0.197	0.025	**0.000**	-0.109	0.142	0.444	-0.194	0.026	**0.000**	-0.104	0.134	0.435	-0.192	0.031	**0.000**
Yorkshire-The Humber	-0.139	0.022	**0.000**	-0.137	0.138	0.320	-0.136	0.022	**0.000**	-0.149	0.132	0.260	-0.113	0.026	**0.000**
East midlands	-0.104	0.023	**0.000**	-0.096	0.127	0.452	-0.102	0.024	**0.000**	-0.071	0.134	0.596	-0.084	0.029	**0.004**
East-anglia	-0.074	0.030	**0.015**	-0.444	0.195	**0.024**	-0.072	0.030	**0.017**	-0.442	0.183	**0.015**	-0.068	0.036	**0.060**
South-west	-0.138	0.022	**0.000**	-0.229	0.136	**0.093**	-0.139	0.022	**0.000**	-0.230	0.127	**0.070**	-0.131	0.026	**0.000**
West-midlands	-0.106	0.022	**0.000**	-0.200	0.128	0.120	-0.103	0.022	**0.000**	-0.195	0.120	0.105	-0.116	0.026	**0.000**
North-west	-0.167	0.021	**0.000**	-0.306	0.129	**0.019**	-0.165	0.021	**0.000**	-0.299	0.122	**0.014**	-0.144	0.026	**0.000**
Wales	-0.206	0.028	**0.000**	-0.227	0.195	0.246	-0.206	0.028	**0.000**	-0.249	0.190	0.191	-0.185	0.033	**0.000**
Scotland	-0.088	0.022	**0.000**	-0.160	0.122	0.192	-0.085	0.022	**0.000**	-0.163	0.114	0.154	-0.097	0.026	**0.000**
Lambda	-	-	-	-	-	-	0.036	0.067	0.594	0.098	0.245	0.689	0.070	0.083	0.399
Intercept	1.752	0.043	**0.000**	1.585	0.256	**0.000**	1.729	0.061	**0.000**	1.508	0.307	**0.000**	1.376	0.093	**0.000**
Rho	-	-	-	-	-	-	0.075			0.197			-	-	-
Observations		6891			260			9519			420			5129	
Censored obs	-	-	-	-	-	-		2628			150		-	-	-
Uncensored obs	-	-	-	-	-	-		6891			260		-	-	-
R2		0.393			0.335		-	-	-	-	-	-		0.351	
Adjusted R2		0.390			0.234		-	-	-	-	-	-		0.347	
F statistics		130.4			3.3		-	-	-	-	-	-		86.4	
Wald Chi2	-	-	-	-	-	-		4653.6			163.0		-	-	-

Source: my elaboration on sixth sweep of NCDS data

The 2SLS is only used to estimate the wage equation of non-convicted individuals, since according to the Wu-Hausman F test and the Durbin-Wu-Hausman Chi2 test education is not endogenous in the wage equation of convicted individuals (table 6).

Table 6 Tests

	NON-CONVICTED			CONVICTED		
	Tests of endogeneity					
		Statistics	P-value		Statistics	P-value
Wu-Hausman F test	F(1,5092)	55.423	0.000	F(1,154)	0.357	0.551
Durbin-Wu-Hausman Chi2 test	Chi2(1)	55.225	0.000	Chi2(1)	0.441	0.507
	Tests of overidentifying					
		Statistics	P-value		Statistics	P-value
Sargan N*R2 test	Chi2(5)	7.022	0.219	Chi2(5)	8.503	0.131
Basmann test	Chi2(5)	6.975	0.223	Chi2(5)	6.989	0.222

Source: my elaboration on sixth sweep of NCDS data

However, while all estimation results are presented, for brevity, we only comment on the results in the last two columns. On the one side, we control for the same variables used in the employment equation, except dummies regarding the presence of children aged 0-6 and the presence of employed partners, since they are expected not to affect wage equations. On the other side, we control for numerous job-related characteristics extremely relevant to determine wage variability across individuals. Evidence of gender duality against women is only found for non-convicted individuals, while among convicted individuals there are no significant differences in terms of gender. As expected, higher educational level increases wages both for convicted and non-convicted individuals. However, contrarily to the employment equation, education presents a stronger positive effect on wage of non-convicted rather than on wage of convicted individuals. Similarly, while disability does not affect the employment probabilities of convicted individuals, it strongly reduces their wages. This strong disadvantage may be partly explained by selection of disabled people into poor job positions. Poor health status also negatively affects wages of convicted individuals, while it does not produce significant effects on wages of non-convicted individuals. To be married also increases the wage of non-convicted workers. Job related characteristics largely explain wage variability, above all among the non-convicted. To be permanently employed increases the wages of non-convicted individuals, being employed in the private sector appears to increase the wage of both groups of individuals. Professional level, defined according to socio-economic group, importantly contributes to wage variability. The reference category is represented by manager and professional; with respect to this group a non-manual profession slightly decreases wages, while a junior non-manual as well as a skilled or semi-skilled manual profession reduces the wage. Unskilled workers and workers employed in personal services or agriculture stay at the bottom of the wage distribution, without relevant difference between convicted and non-convicted. Tenure and number of unemployment spells only significantly affect the wage of non-convicted and both show the expected sign, positive and negative, respectively. Moreover, wages increase as firm size also increases. Estimation results are significant for non-convicted. Similarly part-time reduces the hourly wage of the non-convicted,

consistently with most of the literature that identifies part-time jobs as a source of segregation in the labour market. Unexpectedly, vocational training reduces the wage of convicted individuals. While union membership does not affect significantly the wage of individuals, good computer skills, as expected, increases the wage of non-convicted. The variable controlling for the ability in learning new skills, introduced to proxy the ability of workers as a source of endogeneity and/or selection effect, is not significant either for convicted or for non-convicted individuals. Wages diverge across regions. Employment probabilities are higher in northern regions, while wages are lower. South-East (including London) is the region where jobs are better paid, while convicted individuals appear to be particularly disadvantaged in East-Anglia. Finally, the inverse Mills' ratio (lambda) introduced to correct for selectivity is not significant. This implies that even though selection into non employment of convicted individuals exists, it seems not to be affecting the wage equation of either group.

Table 7 gives the decomposition of predicted wage among endowment, coefficients (returns) and interaction effects, or between explained and unexplained components. Differential due to selection effect is also accounted for. While in the table 7 we present all results from discrimination analysis, for brevity and since results are consistent across econometric models used, we only comment on the results derived from the estimation reported in columns 4 (IV model for non-convicted) and 5 (Heckman model for convicted individuals) of the table 5.

Table 7 Decomposition analysis for wage equations

	OLS estimates				Heckman estimates				IV (non-convicted)-Heckman estimates			
Mean prediction wNC	1.791				1.776				1.762			
Mean prediction wC	1.850				1.799				1.799			
Raw differential (R)	-0.059				-0.024				-0.037			
Due to endowments (E)	-0.077				-0.074				-0.073			
Due to coefficients (C)	0.031				0.065				0.041			
Due to interaction (CE)	-0.013				-0.015				-0.005			
Weighting matrix	0	1	0.5	0.964	0	1	0.5	0.964	0	1	0.5	0.964
Unexplained (U){C+(1-D)CE}	0.018	0.031	0.025	0.031	0.051	0.065	0.058	0.065	0.035	0.041	0.038	0.041
Explained (V) {E+D*CE}	-0.077	-0.090	-0.084	-0.090	-0.074	-0.089	-0.081	-0.088	-0.073	-0.078	-0.076	-0.078
% unexplained {U/R}	-30.7	-52.9	-41.8	-52.1	-215.3	-277.6	-246.4	-275.3	-94.5	-108.8	-101.6	-108.1
% explained (V/R)	130.7	152.9	141.8	152.1	315.3	377.6	346.4	375.3	194.5	208.8	201.6	208.1
Differential due to selection					-0.035				-0.021			

Source: my elaboration on sixth sweep of NCDS data

According to the estimated models the raw differential is 0.037 in favour of convicted individuals, confirming evidence from observed values. Differently from the OLS estimates, measuring a raw differential equal to 0.059 in favour of convicted individuals, models controlling for selectivity reveal that part of the differential favourable to convicted is due to selection into employment. In other words, part of the measured advantage (equal 0.022) is explained by the selection of convicted individuals (with better characteristics) into employment. Alternatively, this may be seen as a consequence of selection

into self-employment or inactivity of convicted people with poor characteristics. Moreover, while the raw differential, net of selection effect, is 0.037 in favour of convicted, according to observable variables the differential in favour of convicted should be higher and equal to 0.073. This means that what is observed as an advantage for convicted individuals actually is a disadvantage. In fact, the coefficients (or returns) effect is favourable to non-convicted individuals (0.041), indicating that for the same level of characteristics, the wage of a convicted worker is lower. Translating into non-logarithm terms, the discrimination corresponds to 25 cents per worked hour, which is almost 500 pounds per year. Wage decomposition results are consistent across weighting matrix used, and differences appear to be quite negligible. Resuming, results from wage discrimination analysis confirm that convicted individuals may be disadvantaged in terms of wage with respect to non-convicted for reasons other than the penalty arising from expected lower productivity. The existence of wage discrimination, may be seen as a consequence of firms' needs to absorb expected extra-costs. Otherwise, the returns effect should be null, i.e. the only wage differential should be attributable to endowments or selection effects. This finding suggests that in determining wages, firms take into account the expected extra-costs. Above all, if the probability of conviction is less than one and firms act rationally, wage discrimination against convicted individuals is greater that the "fair amount" since it also includes the recovery of expected extra-costs of unobservable offenders. This implies that the disadvantage on the labour market for convicted individuals is larger than what is possible to expect from conviction. Finally, the existence of wage discrimination combined with employment discrimination is possibly suggesting that the cost of inefficiency level is not fully recoverable by applying only wage discrimination without violating rational constraints.

Conclusions

This paper focuses on the poorer performance of convicted people in terms of labour market outcomes. On the basis of the 6th sweep of the NCDS data, we found that the observed employment rate is lower among convicted individuals, while the observed wage is higher. Nevertheless, the latter evidence is largely explained by exclusion from employment of convicted individuals with worse characteristics, who are more likely to be self-employed or inactive. Analysis based on decomposition techniques applied to static employment and wage equations, shows that actually convicted individuals are discriminated against on the labour market both in terms of employment and wage. On the basis of a simple model and under the assumption of firms' rationality, we suggest that discrimination against convicted individuals derives from two sources. The first is an "expected" discrimination, which derives from reasons directly imputable to convicted individuals. The second is an "unexpected" discrimination, which derives from

the inefficiency of the police/justice system that implies partial adverse selection in the hiring process. Discrimination proves two main facts. First, firms expect extra-costs from hiring offenders, since they are more likely to offend in the future, and the reduction in the characteristics' return is applied to reduce employment and wage level to endogenize the loss deriving from the extra-costs (economic stigma). Second, because of the inefficiency of the police/justice system, some offenders are unobservable and a partial adverse selection problem arises in the hiring process. The deriving cost of inefficiency of the police/justice system is loaded on convicted individuals, if firms act rationally, worsening their labour market outcomes (over-stigma). The over-stigma is increasing in the extra-costs and in the proportion of the population offending and is decreasing in the probability of convicting offenders, and its magnitude may be seen as a measure of inefficiency of police/justice system.

The main implication of our results is that, with partial adverse selection, convicted individuals face greater disadvantage in the labour market with respect to the expected stigma deriving from worse job-related characteristics affecting productivity and from the existence of extra-costs connected to the risk of recidivism. This may have implications also in terms of dynamic offending decision framework.

While the paper brings to light new evidence and proposes new lines of investigation about the performance of convicted individuals in the labour market, much remains to be done. For example, since panel data are not available for the purpose of our paper, unobservable heterogeneity is not accounted for. Moreover, in the context of identification of discrimination sources, would be a relevant issue to investigate and the study of the relationship between the magnitude of over-stigma and conviction rates of offenders could be useful to support our findings.

Appendix

In the context of my analysis, a selection effect (different from selection bias and endogeneity) may occur if unobserved factors are correlated with both the treatment (conviction) and the outcome (wage and, overall, employment). Since cross-sectional data do not allow to model unobservable heterogeneity a possible source of estimation bias remains uncontrolled, for even though a large number of explanatory variables is used to reduce the unobservable variability. In any case, different methods are available to correct cross-sectional data estimates in case of selection effect or, at least, to test the presence and the relevance of unobservable factors selecting individuals both into treatment and outcome equations. Possibly, applying a bivariate probit model is one of the best solution to solve the problem of employment equation. However, since I need to estimate separately an employment equation for each group, the bivariate probit model cannot be used to correct estimates. On the contrary, it remains a valid instrument to test the presence

and the relevance of unobservable factors potentially affecting employment and conviction probabilities. Since this approach assumes that unobserved factors manifest themselves in the correlation of the errors of both equations, I am only interested in the magnitude and the significance of the correlation term (Rho). Employment probability is explained in terms of the variables presented above, while conviction probability is explained in terms of numerous variables including childhood and adolescence background, and past trouble with the police[10].

Table A1 Biprobit estimates: correlation between error terms

	Likelihood-ratio test of Rho=0	
Rho	Chi2(1)	Prob > Chi2
-0.066	3.642	0.056

Source: my elaboration on sixth sweep of NCDS data

Estimation of the correlation term indicates that, while the presence of unobserved factors affecting both events cannot be excluded, it should be rather small and significant only at 10%, possibly reducing suspects of relevant estimation bias affecting the robustness of the results.

[10] The list of explanatory variables used and the estimation results are available upon request.

References

Blinder, A. S. 1973. Wage Discrimination: Reduced Form and Structural Estimates. 8 *The Journal of Human Resources* 436–455.

Bowles R., Florackis C. 2006. Impatience, Reputation and Offending. Centre for Criminal Justice Economics and Psychology, University of York.

Bushway S.D. 1998. The Impact of an Arrest on the Job Stability of Young White American Men. 35 *Journal of Research in Crime and Delinquency* 454-479.

Bushway S.D. 2004. Labour Market Effects of Permitting Employer Access to Criminal History Records. 20 *Journal of Contemporary Criminal Justice* 276-291.

Cho R., LaLonde R. 2005. The Impact of Incarceration in State Prison on the Employment Prospects of Women. Discussion Paper No. 1792, IZA Institute for the Study of Labor.

Cotton, J. 1988. On the Decomposition of Wage Differentials. 70 *The Review of Economics and Statistics* 236–243.

Entorf H. 2009. Crime and labour market: evidence from a survey of inmates. Discussion paper No. 3976, IZA Institute for the Study of Labor.

Fairlie, R. W. 2005. An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models. 30 *Journal of Economic and Social Measurement* 305–316.

Finn R.H., Fontaine P.A. 1985. The Association between Selected Characteristics and Perceived Employability of Offenders. 12 *Criminal Justice and Behavior* 353-365.

Freeman R.B. 1999. The Economics of Crime. In Handbook of Labor Economics (Eds. O. Ashenfelter and D. Card) vol. 3 chapter 52. Elsevier: Amsterdam.

Grogger J. 1995. The effect of arrests on the employment and earnings of the young men. 110 *The Quarterly Journal of Economics* 51-71.

Heckman. 1979. Sample Selection Bias as a Specification Error. 47 *Econometrica* 153–161.

Holzer H.J. 2007. Collateral Costs: The Effects of Incarceration on the Employment and Earnings of Young Workers. Discussion paper No. 3118, IZA Institute for the Study of Labor.

Holzer, H.J., Raphael S., Stoll M. 2006. Perceived Criminality, Criminal Background Checks and the Racial Hiring Practices of Employers. 49 *Journal of Law and Economics* 451-80.

Kling J.R. 2006. Incarceration length, employment and earnings. 96 *American Economic Review* 863-976.

Lott, J. 1990. The Effect of Conviction on the Legitimate Income of Criminals. 34 *Economic Letters* 381-385.

Myers S.L. 1983. Estimating the Economic Model of Crime: Employment versus Punishment Effects. 98 *The Quarterly Journal of Economics* 157-166.

Nagin D., Waldfogel J. 1995. The Effects of Criminality and Conviction on the Labor Market Status of Young British Offenders. 15 *International Review of Law and Economics* 109-126.

Nagin D., Waldfogel J. 1999. The Effect of Conviction on Income through the Life Cycle. 18 *International Review of Law and Economics* 25-40.

Oaxaca, R. 1973. Male-Female Wage Differentials in Urban Labor Markets. 14 *International Economic Review* 693–709.

Oaxaca, R. L., and M. R. Ransom. 1994. On Discrimination and the Decomposition of Wage Differentials. 61 *Journal of Econometrics* 5–21.

Rasmusen, E. 1996. Stigma and Self-Fulfilling Expectations of Criminality. 39 *Journal of Law and Economics* 519-544.

Reimers, C. W. 1983. Labor Market Discrimination Against Hispanic and Black Men. 65 *The Review of Economics and Statistics* 570–579.

Waldfogel, J. 1994. Does Conviction Have a Persistent Effect on Income and Employment?. 14 *International Review of Law and Economics* 103-119.