From Juvenile to Adult Offender:

An Investigation into the Determinants of Juvenile Arrests and the Relationship between Juvenile and Adult Arrests.

Guyonne Kalb[*†]

Melbourne Institute of Applied Economic and Social Research

University of Melbourne

and

Jenny Williams[†]

Department of Economics

University of Adelaide

JEL classification codes: C35, J13, K42

Keywords: juvenile criminality, adult criminality, treatment effects model, discrete choice model, EM algorithm

Abstract

This paper investigates the determinants of juvenile delinquency and the relationship between juvenile and adult arrests. An ordered probit model for juvenile arrest is estimated separately for males and females. The results for males and females indicate that juvenile arrests are more likely for non-whites and for those who leave education early. Males and females behave differently, in that males are more likely to be repeat offenders. A treatment effects model is used to estimate juvenile and adult arrest equations jointly for males. This shows that the effect of juvenile arrest on adult arrest is largest for white men and men with fewer years of schooling.

---

[*] Corresponding author: Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Victoria 3010, Australia. Tel.: +61 3 8344 7025, fax: +61 3 8344 5630, e-mail address: g.kalb@unimelb.edu.au.

[†] We would like to thank Tim Fry for helpful discussion and suggestions. The Economics department at the University of Adelaide provided funding for travel and accommodation for us to work jointly on this project. Jenny Williams thanks the Economics department at the University of Illinois at Chicago for hosting her while part of this paper was written. Responsibility for any errors remains with the authors.

I. Introduction

A positive relationship between juvenile and adult arrest is confirmed in several studies of adult crime, indicating that juvenile delinquency is the most common pathway to adult criminality.[1] However, none of these studies investigate whether this correlation is due to unobserved heterogeneity that increases the probability of both juvenile and adult arrest, or because juvenile arrest has a direct impact on the probability of adult criminality. Understanding the mechanism that drives the relationship between juvenile and adult arrest is nonetheless important from a policy perspective. In the U.S. for example, the number of juvenile arrests for drug abuse violations increased by 132 per cent between 1990 and 1999 (compared to 29 per cent for adults).[2] If juvenile arrest has a direct impact on the probability of adult arrest, then the war on drugs may have the unintended long-term consequence of increasing the adult criminal population. The purpose of this paper is to shed some light on this issue, by disentangling the role of juvenile arrest and unobservable characteristics on the probability of adult arrest. To do so, we use a treatment effects model. This model is appropriate because it allows juvenile arrest (the "treatment") to have a direct impact on the probability of adult arrest, controlling for the potential correlation between unobservables which influence selection into the "treatment" of juvenile arrest and the outcome of interest, adult arrest.

A second contribution of this paper is its empirical investigation into the factors that are commonly believed to increase the risk of juvenile delinquency. From a policy perspective, this is of importance –independent of the relationship between juvenile and adult criminality– given that 32 per cent of property arrests are juveniles.[3] Moreover, recent research by Widom (2000)

---

[1] See for example Williams and Sickles (2000); Grogger (1998); Witte and Tauchen (1994); Good, Pirog-Good and Sickles (1986).
[2] See Snyder (2000).
[3] ibid.

and Ritchie (2000) suggests that the major factors contributing to delinquency and subsequent criminality in females and males differ. In particular, physical and sexual abuse is identified as the primary cause of delinquency in girls. This has led to calls for the criminal justice system to define justice for females and males differently, due to the victim status of female offenders. We attempt to shed some light on this issue, by carrying out a comparative analysis of the determinants of delinquency in males and females.

The empirical analysis draws on data from the Delinquency in a Birth Cohort II study. These data were collected for the purpose of examining delinquent and criminal activities of a birth cohort and they contain juvenile and adult arrest records for the population of individuals born in the city of Philadelphia in 1958. The analysis is based on a sample from this population, who were administered a retrospective survey in 1988. The retrospective survey includes detailed information on family composition during childhood and the respondent's history of physical and sexual abuse, in addition to more standard variables such as the respondent's education and marital status. This rich source of individual level background information for the sample, along with the complete official records of juvenile and adult arrests make this data well suited to investigate both the determinants of juvenile delinquency and the relationship between juvenile and adult criminality.

Because the Delinquency in a Birth Cohort II study covers both males and females, we are able to carry out a comparative study of criminality across gender. However, very few adult arrests are observed for women in the sample. Therefore, the comparative analysis is confined to the determinants of juvenile delinquency. Estimation of the model of juvenile delinquency for females is complicated by missing observations on the variable which links criminal justice arrest records and follow-up survey information. As a consequence 30 per cent of women in the sample cannot be uniquely matched to criminal justice records. The issues arising in estimating a model for juvenile arrests from this "scrambled data" are similar in nature to those generated by grouped or censored data. We therefore use an adaptation of the EM algorithm to estimate the model for female juvenile delinquency rather than maximizing the likelihood directly.

To anticipate the results, we find with respect to juvenile arrests, that the effect of race and leaving school before the age of 17 have similar effects for males and females. A major difference between males and females is that a change in characteristics mostly shifts males from no arrests to more than one arrest, whereas females move between no arrests and one arrest. The results for the full model of juvenile and adult arrests show that juvenile arrests affect adult arrests both directly and indirectly through unobservables. The direct effect of juvenile arrests on adult arrests is largest for men with low education levels and for whites.

The rest of the paper is laid out as follows. The next section places this research in context by discussing salient literature on the economics of crime. In section III we describe the data used in the analysis. Section IV discusses the methodology used and the results from estimation are presented in section V. The paper finishes with some concluding remarks in section VI.

II. Background

Following the seminal work of Ehrlich (1973), the economics literature on crime has focused on investigating the impact of criminal justice variables and measures of labour market success on crime.[4] Research that uses individual level data is almost exclusively based on males (see for example, Witte and Tauchen, 1994; Grogger, 1998; Freeman, 1991, 1996; Williams and

[4] The relationship between juvenile delinquency and adult crime is implicitly recognized by Donohue and Levitt (2000), who using state level data, find evidence that the fall in the U.S. crime rates since 1991 can be largely attributed to the legalization of abortion and the resultant decrease in the birth of individuals who would be at risk of criminality.

Sickles, 2000).[5] These studies condition on past criminal involvement or juvenile arrests when examining current period criminal or labour market outcomes and they find that past criminal history is a strong predictor of current criminality. Without exception, these studies treat past criminal experience as exogenous. If, however, unobservable characteristics that increase the probability of juvenile arrests also increase the probability of adult arrests (and decrease the probability of labour market success), then the impact of past criminal experience on current period outcomes will tend to be overstated.

While not modelling juvenile delinquency explicitly, most studies of crime control for a range of background characteristics that may contribute to criminality in both childhood and adulthood.[6] These background characteristics include family structure during the individual's childhood, such as: whether the father, mother, or both parents were present in the childhood home, whether the mother worked; and the number of siblings the individual has. Another class of variables measures the socio-economic status of the respondent's family during childhood, such as: ethnicity, race, occupational status of the household head during high school and household income. The effect of role models and peer effects are captured using variables such as: gang membership, criminal history of family members, drug and alcohol problems of family members, parent's education and religiosity. Whether the respondent is in school and educational attainment are also commonly included to capture the opportunity cost of time. The results from including these covariates are mixed.

---

[5] Jurik (1983) is an exception.

[6] Quantitative studies that focus on the determination of juvenile delinquency are remarkably rare in the economics literature. We are only aware of one paper, Phillips and Votey (1987). The criminology literature provides some salient studies on the determinants of juvenile delinquency such as, Good, Pirog-Good and Sickles (1986) and Laub and Sampson (1988). The former paper focuses on the crime-work nexus, while the latter investigates the role of family background and family process in causing delinquency.

The presence of a father in the childhood home is generally found to have a negative impact on measures of criminal involvement. However, while this effect is found to be significant by Comanor and Phillips (1999), Williams and Sickles (2000) find that the father's presence has an insignificant effect on the probability of adult arrest. Similarly, Case and Katz (1991) find that having both parents present in the family home has a negative but insignificant effect on the probability of participating in crime. Family size, measured by the respondent's number of siblings is found to have a positive but generally insignificant effect on the probability of engaging in crime by both Williams and Sickles (2000) and Grogger (1998).

The literature also reports mixed findings with respect to race, and measures of the respondent's socio-economic status and household income.[7] For example, Comanor and Phillips (1999) find that being black is associated with significantly fewer encounters with the law, while both Williams and Sickles (2000) and Grogger (1998) find a positive but insignificant relationship. In contrast Phillips and Votey (1987) and Witte and Tauchen (1994) find that blacks are significantly more likely to be involved in crime. Both Witte and Tauchen (1994) and Phillips and Votey (1987) find that the socio-economic status of the respondent's family is negatively related to the probability that an individual engages in crime. However, this effect is only found to be significant in the latter study, which focuses on juvenile arrests, whereas Witte and Tauchen (1994) examine adult arrests.

In terms of peer effects and role models, both Comanor and Phillips (1999) and Case and Katz (1991) find that being raised in a household where the adults went to church on a regular basis has a negative effect on the respondent reporting criminal activity. However, the effect is only significant in the former study. Case and Katz (1991) also find that the parent's educational

[7] With respect to race, differences in the sample analyzed and definition of participation in crime also lead to conflicting results between studies based on the same data sets.

attainment has a negative but insignificant impact on criminality. On the other hand, the authors find that variables reflecting negative role models and peer effects, such as a family member with a drug or alcohol problem or a family member in jail, have a significantly positive impact on the probability that the respondent engages in crime. Grogger (1998) and Williams and Sickles (2000) also find that having a family member who has had contact with the criminal justice system has a positive impact on participating in crime, however, the effect is not statistically significant in these studies. Agreement is also found for the effect of gang membership in youth on arrest in adulthood, with both Witte and Tauchen (1994) and Williams and Sickles (2000) reporting a positive and significant effect.

With respect to the role of education, Grogger (1998), Witte and Tauchen (1994) and Phillips and Votey (1987) find that measures such as high school graduation or years of education do not significantly affect the probability of engaging in crime.[8] However, both Witte and Tauchen (1994) and Phillips and Votey (1987) do find that being in school has a significantly negative effect on the probability of arrest. In contrast, using the subset of males from the 1958 Philadelphia Birth Cohort Study who had completed school, Williams and Sickles (2000) find that the number of years of schooling has a significantly negative effect on the probability of adult arrest.

In the following analysis, we draw on this previous research for identifying key variables of interest in determining the probability of juvenile arrest. Of particular interest are variables reflecting family structure in the respondent's childhood household, parent's education, race and the respondent's education. In light of the recent research linking female criminality to physical

[8] After controlling for wages, Grogger (1998) finds that education has no significant effect on the probability of engaging in crime.

and sexual abuse in childhood, (see for example Ritchie, 2000; Widom, 2000) we also include measures of abuse in the model for juvenile delinquency.

III. Data

A. Description of the 1958 Philadelphia Birth Cohort Study

Inclusion in the 1958 Philadelphia Cohort Study (Figlio, Wolfgang and Tracy, 1991) is based on the criteria of being born in 1958 and living in Philadelphia between one's tenth and eighteenth birthdays. The 27,160 members of this universe were identified using the Philadelphia school census, the U.S. Bureau of Census and school records. Once the members of the cohort were identified, data collection by Figlio and his team occurred in two phases.

The first phase involved obtaining the complete official criminal history of the cohort. These data were collected between 1979 and 1984 and cover the criminal careers, as recorded by the police, and juvenile and adult courts, for all 27,160 members of the cohort.[9] The second stage of the Study entailed a retrospective follow-up survey for a sample from the cohort. Figlio, Wolfgang and Tracy (1991) employed a stratified sampling scheme to ensure that they captured the most relevant background and juvenile arrest characteristics of the cohort, and yield a sample size sufficient for analysis. The male population was stratified by race, socio-economic status, arrest history (0, 1, 2-4, 5 or more arrests) and juvenile "status" arrests, which are arrest categories only applicable to individuals less than 18 years of age. The female population was stratified by race, socio-economic status and arrest history (0, 1, 2 or more arrests). From the

[9] The information for juveniles was obtained from the Philadelphia police, Juvenile Aid Division (JAD). Once individuals reach the age of 18, police encounters are recorded on regular police forms (rap sheets) and reported to the FBI. Information about adult arrests was obtained from the Philadelphia Police Department, the Common and Municipal Courts, and the FBI, ensuring arrests both within and outside the boundaries of Philadelphia are included in the data set.

resulting strata, a sample of males and females were randomly selected, with equal draws from all strata. This framework oversamples from the more sparsely populated strata covering repeat and chronic juvenile offenders. Of the 1,992 individuals randomly selected for the survey, 576 men and 201 women were interviewed in 1988. Areas of inquiry covered by the survey, which are of interest to this research, include composition of current and childhood households, history of sexual and physical abuse, personal and social demographic characteristics, parent's education, respondent's education and respondent's marital status.

By combining the information from the retrospective survey and official arrest records, we have information on the respondent's family structure while growing up, such as whether the mother worked in paid employment outside the home, presence of both parents and number of siblings; characteristics of the respondent's parents such as their education; self-reported information on physical and sexual abuse; other important variables, such as the respondent's educational attainment, marital history and number of children; and the official arrest records. The dependent variables for the following analysis are an ordered categorical variable for juvenile arrest (0 if not arrested, 1 if arrested once, 2 if arrested more than once) and an indicator for adult arrest. Sample statistics for these data are provided in appendix A. The unweighted statistics are based on sample averages. The weighted statistics take the stratified nature of the sample into account and reflect statistics for the underlying population.

B. Missing data

The observations on the dependent variables (juvenile and adult arrests) and the independent variables were collected at different points in time (phase 1 and phase 2 of the data collection) and are physically contained in different files. The file containing the dependent variables has observations on all 27,160 persons in the 1958 Philadelphia birth cohort from which people are sampled for the retrospective survey. Most observations from the retrospective survey can be linked to the respondent's arrest records through a common identification number called the link code. However, a group of 59 women who were surveyed about family background and other characteristics has not been provided with this link code. As a result, the independent and dependent variables cannot be matched for these women. The number of observations from the arrest file that are potential matches for the unlinked women can be reduced considerably by the information we have. This information includes variables that occur in both the criminal and retrospective survey files, such as birth month and race, and the total number of individuals surveyed from each of the strata. A more detailed discussion of the available information can be found in appendix B. Rather than dropping the unmatched women from the sample, we address the problem by adopting an estimation technique, which accounts for the scrambled nature of the data. This is discussed in section IV.B.

IV. Methodology

Adult criminality is often a continuation of juvenile delinquency. This may be because involvement in the juvenile criminal justice system has a direct impact on future involvement in crime. Such an effect could be through scarring, whereby juvenile arrests leave delinquents with fewer legitimate alternatives; or through human capital accumulation, whereby exposing a juvenile delinquent to criminal peers increases his criminal human capital and criminal networks. Alternatively, juvenile criminality could be correlated with adult criminality because unobserved individual-specific heterogeneity that increases the probability of juvenile criminality also increases the probability of adult criminality.

In order to empirically disentangle the role of these two possible explanations of the correlation between juvenile and adult arrests, we use a treatment effects model. We are interested in whether juvenile arrest (the "treatment") affects the outcome variable adult arrest directly, after controlling for the potential correlation between selection into juvenile arrests and adult arrests. We also explore whether the effect of juvenile arrest differs across various groups in the population. In particular, we consider whether juvenile arrests have a differential impact on the probability of adult arrest by race and years of education.

Juvenile arrests may affect adult arrests differently for whites and non-whites, for example, if there is greater stigma (or prestige) associated with arrest for one group compared to the other. Freeman (1991) has suggested that arrest amongst black youths is so common as to have eroded any stigma associated with arrests.[10] To investigate this issue, we include an interaction term between race (is non-white) and an indicator for juvenile arrests in the equation for the adult arrest probability. If juvenile arrests increase the probability of adult arrest by stigmatising those arrested, a negative coefficient on the interaction term is consistent with Freeman's conjecture that the stigma of arrest is smaller for the non-white population, for whom the prevalence of arrest is higher.

We also explore the possibility that the impact of juvenile arrest on adult arrest differs by level of education. One reason we might expect to find such an effect is if education provides a pathway out of crime, ameliorating the effect of juvenile arrest on the probability of adult arrest. If this were the case, we would expect the sign of the coefficient on the interaction term to be negative. To allow for this possibility, an interaction term between years of education and an indicator for juvenile arrests is included in the model.

---

[10] The Ethnographic work of Anderson (1999) identifies an alternative "street" culture, in which prestige is associated with arrest.

A. The Treatment Effects Model

Suppose that latent juvenile criminal activity, denoted y_1^*, depends on a vector of observable characteristics x_1 —such as race, family structure, history of physical or sexual abuse, parent's education, whether the respondent's mother participated in paid employment and whether the respondent dropped out of school—, and unobservables ε. Latent adult criminal activity, denoted y_2^*, is assumed to depend upon the observable characteristics x_2 —such as the variables contained in x_1, marital status, the number of children the respondent has and years of schooling—, an indicator for whether the individual is arrested as a juvenile (JA=1 if arrested as a juvenile and JA=0 otherwise), interaction terms between the indicator for juvenile arrest and race and between the indicator for juvenile arrest and years of schooling, and unobservables ν. Assuming ε and ν are potentially correlated standard normal random variables, the system of equations looks as follows:[11]

$$
\begin{aligned}
&y_1^* = \beta_1 x_1 + \varepsilon \\
&y_2^* = \beta_2 x_2 + \alpha_1 \mathrm{JA} + \alpha_2 (\mathrm{JA} * \mathrm{race}) + \alpha_3 (\mathrm{JA} * \text{years of schooling}) + \upsilon \\
&(\varepsilon, \upsilon) \sim \text{Bivariate Normal}(0, 0, \sigma_\varepsilon, \sigma_\upsilon, \rho_{\varepsilon\upsilon})
\end{aligned}
\tag{1}
$$

While criminal activity is not directly observed in the data, we do observe information about arrests. During youth, we observe whether an individual has, no arrests, one arrest, or two or more arrests. We denote this categorical variable y_1, where

[11] Since the underlying continuous variables y_1^* and y_2^* cannot be observed themselves and only categorized versions of these variables are observed, the variance of these variables cannot be estimated. Therefore σ_ε and σ_υ are normalized to one.

$$
\begin{aligned}
y_1 &= 0 \quad \text{if not arrested as a juvenile} \\
&= 1 \quad \text{if arrested once as a juvenile} \\
&= 2 \quad \text{if arrested more than once as a juvenile}
\end{aligned}
$$

We also observe whether an individual is arrested as an adult, which we denote y_2, where

$$
\begin{aligned}
y_2 &= 1 \quad \text{if arrested as an adult} \\
&= 0 \quad \text{otherwise}
\end{aligned}
$$

Let y_1 and y_2 be related to latent juvenile and adult criminal activity, y_1^* and y_2^*, according to the following rules:

$$
\begin{aligned}
y_1 &= 0 \text{ if } y_1^* \le 0 \\
y_1 &= 1 \text{ if } 0 < y_1^* \le c \\
y_1 &= 2 \text{ if } y_1^* > c
\end{aligned}
\qquad \text{and} \qquad
\begin{aligned}
y_2 &= 1 \text{ if } y_2^* > 0 \\
y_2 &= 0 \text{ otherwise}
\end{aligned}
$$

where, c is a threshold parameter to be estimated together with the other parameters.

Then the likelihood for this treatment effects model is the joint probability density function for y_1 and y_2:

$$
\begin{aligned}
f(y_1, y_2) = {} & P(y_1 = 0, y_2 = 0)^{(1-y_{11})(1-y_{12})(1-y_2)} P(y_1 = 0, y_2 = 1)^{(1-y_{11})(1-y_{12})y_2} \bullet \\
& P(y_1 = 1, y_2 = 0)^{y_{11}(1-y_{12})(1-y_2)} P(y_1 = 1, y_2 = 1)^{y_{11}(1-y_{12})y_2} \bullet \\
& P(y_1 = 2, y_2 = 0)^{(1-y_{11})y_{12}(1-y_2)} P(y_1 = 2, y_2 = 1)^{(1-y_{11})y_{12}y_2}
\end{aligned}
\tag{2}
$$

where y_{11} is an indicator for being arrested once as a juvenile (that is y_{11}=1 if y_1=1, and zero otherwise) and y_{12} is an indicator for being arrested more than once as a juvenile (that is y_{12}=1 if y_1=2, and zero otherwise). Since y_1^* and y_2^* are bivariate normal, the joint density for y_1 and y_2 can be expressed as (see Greene (1997), pp.907-909):

$$
\begin{aligned}
& f(y_1, y_2) = \\
& -q_{11}(1 - y_{12})\Phi_2\left(-\beta_1 x_1, q_2\left(\beta_2 x_2 + \alpha_1 JA + \alpha_2 (JA * \text{race}) + \alpha_3 (JA * \text{yrs of school})\right), -q_2\rho\right) + \\
& (y_{11} + y_{12})\Phi_2\left(q_{12}(\beta_1 x_1 - c), q_2\left(\beta_2 x_2 + \alpha_1 JA + \alpha_2 (JA * \text{race}) + \alpha_3 (JA * \text{yrs of school})\right), q_{12} q_2 \rho\right)
\end{aligned}
\tag{3}
$$

Where:

$q_{11} = (1 - 2y_{11})$
$q_{12} = (1 - 2y_{12})$
$q_2 = (1 - 2y_2)$
Φ_2 is the bivariate normal cumulative density function

This model can be estimated by maximum likelihood using data on a random sample of individuals. The data used in this study, however, are generated by a stratified random sample, where stratification is based on the number of juvenile arrests. This variable appears in the model as a dependent variable. Manski and Lerman (1977) and Manski and McFadden (1981) show that a simple weighting of the observations and a correction of the covariance matrix are sufficient to deal with this type of endogenously stratified data. The weights are calculated by dividing the population proportions of the strata by the sample proportions of the strata. The covariance matrix is calculated as HGH, where H is the negative inverse of the hessian of the weighted log-likelihood and G are the summed outer products of the first derivatives of the weighted log-likelihood.

A second issue for estimation, only arising when estimating the model for females, is that the variable linking the follow-up survey data with the juvenile and adult arrest information is missing for 59 out of the 201 observations for women. As a consequence, the dependent variables are not uniquely observed for this subsample of women. We discuss the method for dealing with the "scrambled data" for females in the next section. In addition to this missing information, only six out of the 142 linked women are observed to have had adult arrests in the sample period. Due to this infrequency of observed female adult arrests, we limit examination of female criminality to the juvenile period.

B. Using the EM Algorithm to Overcome the Problem of Scrambled Data

The EM algorithm is often used in situations where data are missing or where variables are censored.[12] The problem in this paper is of a similar nature. The difference between the standard missing or censored variable problem and the problem in this research is that the number of possible values for each variable is finite. The EM algorithm is normally used in a continuous context, but because of this finite number of possible values, a discrete approach is more appropriate[13].

As previously discussed, the dependent variable $Y_i=\begin{pmatrix} Y_{1i} \\ Y_{2i} \end{pmatrix}$(where Y_1 is the number of juvenile arrests and Y_2 is the number of adult arrests) has not been linked to the independent variables, X, for the 59 unmatched women. There is a choice from at most six different possible values for Y_i. That is, $Y_i \in \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}$. The number of possible matches from the arrest file is restricted for some observations because of the limited occurrence of women with adult arrests (Y_{2i}=1) and women with more than one juvenile arrest (Y_{1i}=2), and because the respondent's birth month is contained in both files and has to be matched. Each record in the cohort file can at most be linked once. Thus, selecting a value for one woman may restrict the choice of values for the other 58 women. This means that for the women without link codes, a combined likelihood expression has to be evaluated instead of an expression for each individual, because the matching to values for the dependent variables for one woman cannot be done independently from the matching for the other women. The unit of observation is this group of

[12] EM stands for the two steps involved in this method an Expectation and a Maximization step. Dempster, Laird and Rubin (1977) were the first to recognize the usefulness of the method for incomplete data.

[13] The lemma underlying the derivation of the EM algorithm in the continuous case is also valid for the discrete case (see lemma 1e.6 (i) and (ii) in Rao (1973)).

59 women rather than each woman individually. The information available is observed for the whole group at once instead of for each woman separately. Appendix B reports on the available information.

Instead of integration over a range of possible values for a missing or censored variable, in this discrete version of the EM algorithm, a summation over all possible combinations of values for the variable Y is performed. Explicitly writing down this set of possible combinations is cumbersome, since the constraints on the possible combinations need to be fulfilled for all 59 observations simultaneously. The choice for one observation can affect the possible choices for another observation. Therefore for ease of notation, a set A is defined which is the set of possible combinations for the 59 non-matched observations that fulfil all constraints[14]. We defer the discussion of computational issues till the end of this section.

The contribution to the likelihood function of the 59 unmatched women can be constructed by taking the joint probability density function for the 59 observations without link code and summing over all combinations contained in the set A:

$$
\begin{aligned}
L(\theta|A) &= \sum_{i,k,\ldots,p\in A} \mathrm{pdf}\left(Y_1 = y_i, Y_2 = y_k, \ldots, Y_{59} = y_p | \theta, X_1, \ldots, X_{59}\right) \\
&= \sum_{i,k,\ldots,p\in A} \mathrm{pdf}\left(Y_1 = y_i | \theta, X_1\right) \ldots \mathrm{pdf}\left(Y_{59} = y_p | \theta, X_{59}\right) \qquad (4)
\end{aligned}
$$

Taking logarithms to obtain the log likelihood:

$$
l(\theta|A) = \ln\left\{ \sum_{i,k,\ldots,p\in A} \mathrm{pdf}\left(Y_1 = y_i | \theta, X_1\right) \ldots \mathrm{pdf}\left(Y_{59} = y_p | \theta, X_{59}\right) \right\} \qquad (5)
$$

[14] See appendix B.

where θ is the parameter vector, consisting of β_1, β_2, α_1, α_2, α_3 and $\rho_{\varepsilon\upsilon}$, that has to be estimated; and pdf stands for the probability density function.

Using the EM algorithm results in a rewritten log likelihood expression (see Dempster, Laird and Rubin, 1977):

$$l(\theta| \mathrm{A}) = \mathrm{Q}(\theta, \varphi; \mathrm{A}) - \mathrm{H}(\theta, \varphi; \mathrm{A}) \tag{6}$$

where φ is a vector defined over the same domain as θ and the above is valid for any value of φ. The function H[15] is not relevant when searching for the maximum of $l(\theta| \mathrm{A})$. In the function Q, the contribution $\log(f(\theta,y))$ of an unobserved latent variable y to the log likelihood function is replaced by its expectation over the set of values in which its true value is known to lie:

$$\mathrm{Q}(\theta,\varphi;\mathrm{A}) = \sum_{i,k,\dots,p\in \mathrm{A}} \left\{ \ln\left[f\left(y_i|\theta,X_1\right)\right] + \ln\left[f\left(y_k|\theta,X_2\right)\right] + \dots + \ln\left[f\left(y_p|\theta,X_{59}\right)\right]\right\} \bullet \Pr(Y_1 = y_i,\dots,Y_{59} = y_p|\varphi,\mathrm{A}) \tag{7}$$

Applying the EM algorithm to the problem of the scrambled data sets and assuming a distributional form: $\mathrm{pdf}(y| \varphi,X) = f(y| \varphi,X)$, the expression Q can be calculated. Assuming that $f(y| \varphi,X)$ is known, $\Pr(Y_1, Y_2,\dots,Y_{59} \;| \varphi,\mathrm{A})$ can be constructed:

$$\Pr(Y_1 = y_i,\dots,Y_{59} = y_p|\mathrm{A},\varphi) = \frac{\mathrm{pdf}\left(Y_1 = y_i|X_1,\varphi\right)..\mathrm{pdf}\left(Y_{59} = y_p|X_{59},\varphi\right)}{\sum\limits_{i,k,\dots,p\in \mathrm{A}} \mathrm{pdf}\left(Y_1 = y_i|X_1,\varphi\right)..\mathrm{pdf}\left(Y_{59} = y_p|X_{59},\varphi\right)}$$

$$= \frac{f(y_i|X_1,\varphi)..f(y_p|X_{59},\varphi)}{\sum\limits_{i,k,\dots,p\in \mathrm{A}} f(y_i|X_1,\varphi)..f(y_p|X_{59},\varphi)} \quad \text{for } \{y_i,\dots,y_p\}\in \mathrm{A} \tag{8}$$

$$= 0 \qquad \text{elsewhere}$$

[15] $\mathrm{H}(\theta,\varphi;\mathrm{A}) = \sum\limits_{i,k,\dots,p\in \mathrm{A}} \ln\left(\Pr(Y_1 = y_i,\dots,Y_{59} = y_p|\theta,\mathrm{A})\right) \bullet \Pr(Y_1 = y_i,\dots,Y_{59} = y_p|\varphi,\mathrm{A})$

The Q-function in equation (7) has to be maximized with respect to θ, where φ is given. Dempster, Laird and Rubin (1977) have shown that iteratively maximizing this function using the previous optimal values θ^{q-1} for φ leads to convergence. Only an arbitrary value $\varphi=\theta^{0}$ is needed to start the process and the iterations are finished when $\theta^{q} = \theta^{q-1}$. This procedure leads to an estimated θ that is a stationary point of the log likelihood function $l(\theta|\mathrm{A})$. Since the θ found in this way is not necessarily a maximum, second order conditions should be checked. If we experiment with some different starting values, we may be reasonably certain that the maximum found is global.

As previously noted, writing down the set A of possible combinations would be quite complicated. The requirements need to be checked simultaneously for all 59 observations for each possible match that can be made. However, the exact calculation in the E-step can be replaced by a Monte Carlo implementation of the E-step (Tanner, 1993). Using a simulation technique to approximate the expected value will be convenient, since we can then just check the requirements for each simulated combination and only allow those that fulfil all conditions.

However, before observations can be sampled from the function, all elements of the probability function in (8) would have to be calculated. Therefore, to circumvent having to write down all possible combinations a further step is needed, since the probability of each possible combination occurring still contains the sum over all possible combinations in the denominator. Instead of using the exact probabilities in the simulation, we use an approximation and correct for the approximation by weighting the draws with the ratio of the true probability and the approximated probability. In Bayesian literature much use is made of this method named importance sampling[16], which under certain conditions allows one to draw from a simpler

[16] See Kloek and Van Dijk (1978) and Van Dijk and Kloek (1980).

distribution than the actual one. In order to get good results the approximating distribution has to be sufficiently similar to the original distribution. This means that the approximating distribution should have large probability where the actual distribution has large probability, so that all important combinations will be drawn from the importance function.[17]

After sampling from the simple distribution each drawing is weighted by the ratio of the probability of this drawing calculated with the actual distribution to the probability calculated with the simple distribution. When calculating weighted averages over this sample, the constant terms in both probabilities will cancel out. This means in this case that the computation intensive denominator, consisting of the sum over all possible combinations can be left out of the calculations. Further details on the importance sampling procedure are given in appendix C.

V. Results

We begin by presenting the results for the ordered probit models of juvenile arrest for males and females in subsection A. The results for the full treatment effects model estimated over the sample of males are presented in subsection B.

A. A Comparison of the Determinants of Juvenile Arrest for Boys and Girls

The results for the ordered probit models of juvenile arrest for boys and girls are contained in Table 1. The dependent variable for the model is an ordered categorical variable that takes on a value of zero if the individual was not arrested, one if the individual was arrested once, and two if the individual was arrested more than once. The upper panel of Table 1 contains results for males and the lower panel presents results for females. Parameter estimates and t-statistics can be found in the first two columns of the table. Columns 3, 4 and 5 contain the marginal effect of a change in an explanatory variable on the probability that the individual is

[17] The number of drawings T from the importance function $I(Y_1,Y_2,..., Y_{59} | A, \theta_q)$ can be less when this function is more similar to the actual distribution.

observed to have zero, one, or greater than one juvenile arrest, respectively. The marginal effect of a discrete variable is defined as the change in the probability of zero, one, or greater than one juvenile arrest associated with the variable of interest changing from zero to one, while keeping all other characteristics at their observed values. The change in probability is averaged over all observations. The predicted average probability of being in each juvenile arrest category is contained in the row corresponding to the constant term.

Table 1 about here

The results in Table 1 are informative in terms of the determinants of juvenile delinquency and the extent to which these determinants differ across gender. We find that there is substantial concordance between the two estimated models. In terms of what matters, we find that both boys and girls who are not white are significantly more likely to be arrested than white children. Being non-white decreases the probability of no juvenile arrests by 10.9 percentage points on average for males and 9.4 percentage points for females. The probability of having exactly one juvenile arrest is 3.3 percentage points greater for non-white males and 7.4 percentage points greater for non-white females, compared to their white counterparts. The probability of more than one juvenile arrest is estimated to be 7.6 percentage points higher for non-white males and 2.1 percentage points higher for non-white females compared to whites. These are large effects, given that the average predicted probability of having more than one arrest is only 16.1 per cent for males and 2.2 per cent for females.

Leaving school before the age of 17 is also associated with a greater probability of juvenile arrest for both boys and girls. For males, leaving school before the age of 17 decreases the probability of no juvenile arrests by 25.8 percentage points. The decrease is 29.7 percentage points for females. As with the effect of race, we find that the decrease in the probability of no juvenile arrest associated with leaving school early is accompanied by a large increase in the probability of having *more than one* arrest for males (19.7 percentage points) and a large increase in the probability of having a *single* arrest for females (20.6 percentage points)[18].

The results in Table 1 are also informative about some commonly held beliefs about the causes of juvenile delinquency. We find that growing up without a father in the childhood home has a positive but statistically insignificant effect on the probability of juvenile arrest for both boys and girls. Due to a focus on the absence of a father in the literature, we experimented with disaggregating the measure no father present to no father present due to divorce or separation, and no father present due to death. These variables were neither individually nor jointly significant. We also investigated whether the presence of a stepfather affected the respondent's juvenile delinquency. This variable was also not significant.

We also fail to find any statistical evidence that physical or sexual abuse during childhood is associated with increased probability of juvenile delinquency for girls or boys. However, with just 9 per cent of the sample of 575 males and 7 per cent of the sample of 201 females reporting sexual or physical abuse, we simply might not have the power to identify this effect. The results also fail to find evidence of the popularly held belief that working mothers contribute to higher levels of juvenile delinquency. On the contrary, the model for females finds statistically significant evidence that having a mother who stays at home increases the probability of juvenile arrest. Moreover, the point estimate for males also suggests that having a mother who stays at home is associated with an increased probability of juvenile arrest and this

[18] This is clearly an important result. This finding is consistent with several hypotheses regarding the influence of education on juvenile delinquency. These hypotheses include schooling, which works through the time constraint as in "the Devil makes work for idle hands" hypothesis; through shaping preferences away from criminal activities which may be the peer effect; through household budget constraints, if leaving school before age 17 is because parents cannot afford to keep their children in school; and through having poor earning opportunities because of low levels of human capital.

effect is statistically significant at the 10%-level on the basis of a one sided test. Although not reported, we investigated whether this result was sensitive to whether the respondent's mother was a single mother by including an interaction term between mother not working and no father in the respondent's childhood home. This interaction term was not statistically significant.

While the estimated models of male and female delinquency share many common findings, there are nonetheless distinct differences. In particular, having a father who graduated from high school reduces the probability that a male is arrested as a juvenile, but is not a significant determinant of juvenile arrests for females. For males, the effect of having a father who has graduated from high school is of a similar magnitude as the effect of race, although opposite in sign. Because education is generally correlated with employment and income, we would expect that fathers with more education are more likely to be employed and earning a higher wage than those with lower levels of education. This suggests that the father's education may be a proxy for socio-economic status. An alternative explanation is that fathers with higher levels of education are better able to act as a positive role model for their sons and provide information about legitimate opportunities to their sons. Although not included in the results in Table 1, the mother's education was also included in the model for males and females. We found that the mother's education had no significant effect on the probability of juvenile arrest and it was therefore excluded from the model.

We also find that the probability that a male is arrested as a juvenile is increasing with the number of siblings he has. This result is common in the literature, which is limited to males, and is generally interpreted as reflecting the fact that children from larger families receive less individual attention. It is therefore somewhat surprising that the number of siblings is not significant and does not have the expected sign in the model for female juvenile arrests.

The first row of columns 3, 4 and 5 reports the predicted proportion of the sample with zero, one, or greater than one juvenile arrest. This information reveals that males and females differ in the pattern of delinquent behaviour. Overall, 29 per cent of males are predicted to have at least one juvenile arrest, compared to 14 per cent of females. While this is consistent with the accepted wisdom that delinquency is a larger problem for males than females, it nonetheless indicates that delinquency is a substantial problem for females as well. The higher rate of juvenile arrest experienced by males can be attributed to a greater prevalence of repeat offenders. Males are much more likely to have more than one arrest compared to females (0.16 and 0.02 respectively). The proportion of males and females who have only one juvenile arrest is similar (0.13 for males and 0.12 for females). This greater tendency for males to experience multiple arrests is also evident in the marginal effects. The marginal effect of changing a characteristic mostly shifts males between no arrests and more than one arrest, whereas females mostly move between no arrests and one arrest. This is reflected in the size of the threshold parameter, which is much higher for females than for males, indicating that women are unlikely to be in the group with the highest arrest rate. This raises the question of whether juvenile arrests serve as a greater deterrent for future criminal activity for females than for males. While the infrequency of adult arrests for females in the follow-up sample of the 1958 Philadelphia Birth Cohort Study suggests that this might well be the case, the paucity of data on arrests for women precludes us from addressing this question formally. Although we are unable to carry out a comparative analysis of the effect of juvenile arrests on the probability of adult arrest, we investigate this relationship for males in the following section.

B. A Treatment Effects Model of the Effect of Juvenile Arrests on Adult Arrests

The results for the full treatment effects model estimated for the male sample is contained in Table 2. As with Table 1, the first two columns of Table 2 contain the coefficient estimates and t-statistics, while the remaining columns contain the marginal effects. The upper panel contains the results for juvenile arrests, while the lower panel contains the results for adult arrests. Because the estimates for the juvenile arrest model for males in the previous section are consistent (although inefficient relative to the full treatment effects model), we do not re-examine the results for juvenile arrests here. Rather, we focus the discussion on the results from the adult arrest equation of the full treatment effects model.

A primary contribution of this study is determining the underlying factors accounting for the positive correlation between juvenile and adult arrests. As discussed in section IV, we allow two possible mechanisms for this correlation. First, we allow for a direct impact of juvenile arrest on the probability of adult arrest, which may vary by race and years of schooling. Second, we allow the error terms in the juvenile arrest and adult arrest equations to be correlated (that is, unobservable characteristics that increase the probability of juvenile arrest also increase the probability of adult arrest). In addition, the model for adult arrest includes the control variables that are included in the juvenile arrest equation, except for the age at which schooling was left. The latter variable is excluded because we have information on total years of schooling for adults. We also include whether the respondent is married and the number of children he has, to capture family structure in adulthood.

Table 2 about here

The results in Table 2 show that there is evidence of both mechanisms of correlation between juvenile and adult arrests. The point estimate for the correlation between the unobservables is 0.71 and statistically significant. The point estimate on the indicator for juvenile arrests in the adult arrest equation is also significantly positive. The results in Table 2 further show that the interaction terms of race and years of schooling with juvenile arrests are significant and negative in sign. This provides evidence that the effect of juvenile arrests on adult arrests differs by race and years of schooling. Discussion of the marginal effects of juvenile arrest, race and schooling are postponed until the discussion of Tables 3 and 4, which present the combined effects of race, years of schooling and juvenile arrests on the probability of adult arrest.

In terms of the control variables, we find that growing up in a household without a father, suffering physical or sexual abuse as a child, or having a mother who worked, have no significant effect on the probability of arrest in adulthood. We find that the probability of adult arrest is greater for individuals with more siblings. This suggests that coming from a larger family increases the probability of adult arrest, even after controlling for the impact of family size on the probability of juvenile arrest. In contrast, the level of education of the respondent's father has no independent effect on the probability of the respondent being arrested as an adult, after controlling for its impact on the probability of juvenile arrest. The results indicate that growing up without a mother significantly decreases the probability of adult arrest, despite having no impact on juvenile arrest. In terms of variables capturing family structure in adulthood, we find that marital status has a large significant effect on the probability of adult arrest, with married individuals less likely than their single counterparts to be arrested. Married men have an average arrest probability that is 8.1 percentage points lower. This is consistent with the results of Akerlof (1998). Unlike Akerlof's study however, we do not find that having children negatively affects the probability of adult arrest. The point estimate for the parameter corresponding to the number of children the respondent has at the time of survey is positive and insignificant.

In order to interpret the effect of juvenile arrests, race and years of schooling on adult arrests, we must consider the combined effect of interaction as well as leading terms on these variables. The top panel of Table 3 contains the parameter estimates and t-statistics for the joint effect of race and juvenile arrests, while the bottom panel contains the parameter estimates and t-statistics for the joint effect of years of schooling and juvenile arrests. Table 4 presents the effect of juvenile arrest on the probability of adult arrest by race and years of schooling.

Tables 3 and 4 about here

It can be seen from the magnitude of the parameters for the joint effect of race and juvenile arrest contained in Table 3 that, irrespective of race, individuals with juvenile arrests are more likely to have an adult arrest than those without juvenile arrests. Further, conditional on juvenile arrests whites are less likely than non-whites to have an adult arrest. However, Table 4 shows that the increase in the adult arrest probability as a result of having juvenile arrests is largest for white respondents (17.1 percentage points compared to 12.5 percentage points).

Returning to Table 3, we find that schooling partly counteracts the effect of juvenile arrests on the probability of an adult arrest, ceteris paribus. However, individuals with juvenile arrests are still more likely to have an adult arrest than individuals without juvenile arrests and fewer years of schooling. Examining the effect of juvenile arrests on the probability of adult arrest by education level in Table 4, we see that individuals who were arrested as a juvenile and who completed 10 years of schooling are 31.7 percentage points more likely to be arrested as an adult, compared to an individual with the same level of schooling and no juvenile arrests. However, if the individual had completed two additional years of schooling, the increase in the probability of adult arrest compared to an individual with the same level of schooling and no juvenile arrests is only 14.9 percentage points greater. This finding suggests that legitimate human capital may prevent juvenile offenders from persisting in criminal behaviour.

It is not evident, however, that the crime-preventative effect of education is working through wages, as is typically hypothesized. This would imply that years of schooling would have a significantly negative impact on the probability of arrest. We find that years of schooling, per-se, has no significant effect on the probability of adult arrest. Only the interaction term between schooling and juvenile arrest is significant.

In summary, the results from estimating the treatment effects model provide evidence that unobserved characteristics which increase the probability of juvenile arrest also increase the probability of adult arrest. However, juvenile arrests also have a direct impact on the probability of adult arrest. This latter effect is more important for whites than non-whites and for those with lower levels of education.

VI. Conclusion

While juvenile delinquency has received little attention in the economics literature, understanding its causes and relationship to adult criminality is important in terms of developing crime prevention and sentencing policies. To further an understanding of these issues, this research uses data from the Delinquency in a Birth Cohort II study to investigate a number of popular hypotheses regarding the determinants of juvenile delinquency, as well as the underlying factors that account for the positive correlation between juvenile and adult arrests. Importantly, we also examine whether the determinants of criminality differ by gender. Due to the paucity of information in the data on adult arrests for women, the comparative analysis focuses on juvenile delinquency.

The results for the comparative analysis of juvenile delinquency indicate that some determinants of delinquency are common across gender, while others are gender specific. In terms of common determinants, we find that the probability of juvenile arrests for both males and females is greater for non-whites, who leave school before the age of seventeen. Contrary to recent research, we fail to find evidence that sexual or physical abuse is a significant determinant of either female or male delinquency, or that coming from a single parent household is associated with a greater likelihood of delinquency. We also find important differences both in the determinants and patterns of male and female delinquency. In particular, father's education and number of siblings are significant determinants of male –but not female– delinquency, with juvenile arrest less likely for boys whose father has graduated from high school and who have fewer siblings. In terms of the pattern of delinquency, we find that while a similar proportion of males and females experienced a single juvenile arrest, males were about twice as likely as females to have more than one arrest.

The results from estimating the joint model of juvenile and adult arrests suggest that while there is a significantly positive correlation in the unobservables associated with juvenile and adult arrests, juvenile arrest also has a direct impact on the probability of adult arrest. Juvenile arrest is found to significantly increase the probability of adult arrest, with the magnitude of this increase varying by race and years of schooling. We find that the increase in the probability of adult arrest associated with juvenile arrest is greater for whites than non-whites. This may be evidence of a reduction in stigma associated with arrest for non-whites resulting from the high prevalence of contact with the criminal justice system experienced by this group. We also find that the impact of juvenile arrest on the probability of adult arrest is smaller for individuals with a greater number of years of schooling. For example, a juvenile delinquent with 10 years of education is 30 percentage points more likely to have an adult arrest than a similarly educated non-delinquent, whereas a juvenile delinquent with 12 years of schooling is 15 percentage points more likely to have an adult arrest than an otherwise identical non-delinquent. The results fall short, however, of implying that the crime preventative effect of education is working through wages as is commonly hypothesized, because education per-se is not found to have a significant effect on the probability of adult arrest.

An issue not addressed in this paper is whether education can in fact be treated as exogenous to the crime decision. If characteristics that lead people to stay in school also result in them being less likely to participate in crime, then the results will overstate the impact of education on criminal outcomes. This remains an issue for future research.

An important result from this research is that juvenile arrest does have a direct impact, raising the probability of adult arrest for males. This suggests that adult criminals are not simply "bad apples", in that the correlation between juvenile delinquency and adult arrest is not solely due to correlation in unobserved heterogeneity. From a policy perspective, this could be taken as evidence in support for recent moves to divert juveniles away from formal criminal justice sanctions through youth courts. It also provides evidence that criminal justice interventions in youth have the potential to influence adult outcomes. The policy question is then how to identify interventions that prevent youth from engaging in crime. While this research has found that race and education are the key to this question, further research is needed to understand the mechanisms by which these factors impact criminal choice.

REFERENCES

Akerlof, G.A., "Men Without Children," ***The Economic Journal*** 108 (1998), 287-309.

Anderson, E., ***Code of the Street: Decency, Violence, and the Moral Life of the Inner City***, (New York: W.W. Norton, 1999).

Case, A., and L. Katz, "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths," Working Paper, National Bureau of Economic Research, Cambridge, MA (1991).

Comanor, W.S., and L. Phillips, "The Impact of Income and Family Structure on Delinquency," Unpublished manuscript, University of California, Santa Barbara (1999).

Dempster, A.P., N. Laird, and D.B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," ***Journal of the Royal Statistical Society B*** 39 (1977), 1-38.

Donohue, J.J., and S.D. Levitt, "The Impact of Legalized Abortion on Crime," Working Paper, National Bureau of Economic Research, Cambridge, MA (2000).

Ehrlich, I., "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation," ***Journal of Political Economy*** 131 (1973), 521-565.

Figlio, M., P.E. Tracy, and M.E. Wolfgang, "Delinquency in A Birth Cohort II: Philadelphia 1958-1986: Computer File," Philadelphia, PA: Sellin Center for Studies in Criminology and Criminal Law and National Analysts, Division of Booz-Allen and Hamilton, Inc.: Producers, 1990. Ann Arbor, MI: ICPSR (Second Release 1991).

Freeman, R.B., "Crime and the Economic Status of Disadvantaged Young Men," Mimeo, Conference on Urban Labor Markets and Labor Mobility, Virginia: Arlie House (1991).

Freeman, R.B., "Why Do So Many Young American Men Commit Crimes and What Might We Do About It?" ***Journal of Economic Perspectives*** 10 (1996), 25-42.

Good, D., M. Pirog-Good, and R.C. Sickles, "An Analysis of Youth Crime and Employment Patterns," ***Journal of Quantitative Criminality*** 2 (1986), 219-235.

Greene, W.H., ***Econometric Analysis***, third ed. (New Jersey: Prentice Hall, 1997).

Grogger, J., "Market Wages and Youth Crime," ***Journal of Labor Economics*** 16 (1998), 765-791.

Jurik, N.C., "The Economics of Female Recidivism," ***Criminology*** 21 (1983), 603-622.

Kloek, T., and H.K. van Dijk, "Bayesian Estimates of Equation System Parameters: An Application of Integration by Monte Carlo," ***Econometrica*** 46 (1978), 1-19.

Laub, J.H., and R.J. Sampson, "Unraveling Families and Delinquency: A Reanalysis of the Glueck's Data," ***Criminology*** 26 (1988), 355-379.

Manski, C.F., and S.R. Lerman, "The Estimation of Choice Probabilities from Choice Based Samples," ***Econometrica*** 45 (1977), 1977-1988.

Manski, C.F., and D. McFadden, "Alternative Estimators and Sample Designs for Discrete Choice Analysis," in C.F. Manski and D. McFadden (Eds.), ***Structural Analysis of Discrete Data with Econometric Applications***. (Cambridge, Massachusetts: MIT press, 1981).

Phillips, L., and H.L. Votey, "Rational Choice Models of Crime by Youth," ***Review of Black Political Economy*** 16 (1987), 129-187.

Rao, C.R., ***Linear Statistical Inference and Its Applications***. (New York: John Wiley & Sons, 1973).

Ritchie, B.E., "Exploring the Link Between Violence Against Women and women's involvement in Illegal Activity," in ***Research on Women and Girls in the Criminal Justice System:***

Plenary Papers of the 1999 Conference on Criminal Justice Research and Evaluation – Enhancing Policy and Practice Through Research, Vol 3, NCJ 180973 (2000).

Snyder, H. N., “Juvenile Arrests 1999,” U.S. Department of Justice, (2000)

Tanner, M.A., ***Tools for Statistical Inference; Methods for the Exploration of Posterior Distributions and Likelihood Functions***, Springer series in statistics, (New York: Springer-Verlag, 1993).

Van Dijk, H.K., and T. Kloek, “Further Experience in Bayesian Analysis using Monte Carlo Integration,” ***Journal of Econometrics*** 14 (1980), 307-328.

Widom, C.S., “Childhood Victimization and the Derailment of Girls and Women to the Criminal Justice System,” In ***Research on Women and Girls in the Criminal Justice System: Plenary Papers of the 1999 Conference on Criminal Justice Research and Evaluation – Enhancing Policy and Practice Through Research***, Vol 3, NCJ 180973 (2000).

Williams, J., and R.C. Sickles, “An Analysis of the Crime as Work Model: Evidence from the 1958 Philadelphia Birth Cohort Study,” Working Paper, University of Adelaide (2000).

Witte, A.D., and H. Tauchen, “Work and Crime: An Exploration Using Panel Data,” ***Public Finance*** 49 (1994), 155-167.

APPENDIX A— SUMMARY STATISTICS

TABLE A.1.— SUMMARY STATISTICS OF VARIABLES IN THE SAMPLE USED

	Men (N=575)		Women (N=201)	
	weighted	**unweighted**	**weighted**	**unweighted**
Indicator for juvenile arrests (=1)†	0.1290	0.2383	0.1075	0.3028
Indicator for juvenile arrests (>1)†	0.1606	0.3826	0.0205	0.3099
Indicator for adult arrests†	0.1186	0.2104	0.0028	0.0423
E(juvenile arrests (=1))			0.1193	0.3117
E(juvenile arrests (>1))			0.0215	0.3650
E(adult arrests)			0.0031	0.0409
Race is non-white	0.5274	0.4609	0.5263	0.4478
No father in childhood home	0.1659	0.1896	0.1309	0.2090
Stepfather	0.0844	0.0904	0.0789	0.0945
No father (divorced)	0.1047	0.1113	0.0758	0.1244
No father (deceased)	0.0237	0.0383	0.0454	0.0448
No father (other)	0.0374	0.0400	0.0100	0.0398
Stay at home mother	0.3359	0.3704	0.2394	0.3383
No father*stay at home mother	0.0593	0.0661	0.0477	0.0597
Physically abused in childhood	0.0769	0.0713	0.0144	0.0348
Sexually harassed in childhood	0.0143	0.0139	0.0604	0.0597
No mother in childhood home	0.0301	0.0261	0.0294	0.0299
No mother (divorced)	0.0244	0.0157	0.0268	0.0199
No mother (deceased)	0.0026	0.0070	0.0001	0.0050
No mother (other)	0.0030	0.0035	0.0017	0.0050
Don’t know mother’s education	0.0570	0.0765	0.0310	0.0448
Mother’s education <hs grad	0.2627	0.3113	0.3061	0.3433
Mother’s education is hs grad	0.6194	0.5600	0.5713	0.5522
Mother’s education > hs grad	0.0308	0.0261	0.0623	0.0299
Don’t know father’s education	0.0849	0.1148	0.0917	0.0945
Father’s education <hs grad	0.3034	0.3183	0.2916	0.3035
Father’s education is hs grad	0.3866	0.3391	0.4081	0.3433
Father’s education > hs grad	0.0592	0.0383	0.0777	0.0498
Number of siblings	3.0500	3.2243	3.2571	3.2736
left school <16 years old	0.0392	0.0574	0.0326	0.0896
left school at 16 years old	0.0891	0.1635	0.0470	0.1642
left school at 17 years old	0.2794	0.2730	0.3199	0.3085
Years of schooling	12.6074	11.8609		
Number of children	1.0316	1.2591	1.5585	1.7960
Married	0.3595	0.3948	0.4870	0.4030
Juv.off.*race†	0.1937	0.2748	0.1132	0.4507
Juv.off.*(years of schooling)	3.3632	7.1235		

† For women, the statistics for these variables are only calculated over the 142 complete observations.

APPENDIX B— INFORMATION USED FOR MATCHING

The following information is used for matching:

- the unmatched records have a cohort identification number over 24780
- the unmatched records are white women
- birth month (occurs in both files)
- we know the number of white females in the birth cohort who have no arrests, one arrest and more than one arrest, and the number of women in these categories who were surveyed and hence that amongst the unlinked women there are 10 without arrests, 20 with one arrest and 30 with more than one arrest.
- assuming that people do not report more arrests than they have actually had (but rather underreport their arrests), the number of self-reported arrests in the survey identifies in some cases the strata the women are in. We found that six women are in strata 5 (one arrest) and 14 in strata 6 (more than one arrest)
- there are only a restricted number of people with adult arrests in each subgroup, which further limits the number of possible values to choose from.
- the number of possible different outcomes for the dependent variables can only range from one to six (three different outcomes for juvenile arrests and two outcomes for adult arrests).

Of the 60 women without a link code one woman was dropped. Since we do not know to which strata the woman belongs, the above information remains unchanged. The cohort file, from which we can choose a match for the 59 surveyed women with a missing link code, contains 2046 records for white women with no juvenile arrests, 33 of whom have had an adult arrests; 282 records for white women with a single juvenile arrest, 27 of whom had adult arrests; and 52 records for white women with more than one juvenile arrest, eight of whom had adult arrests.

APPENDIX C— IMPORTANCE SAMPLING

The importance function considered here is constructed from a sequence of probabilities. The procedure for generating data for Y goes as follows:

Choose one unmatched record containing a vector X with characteristics from the main data set to start from and calculate probabilities of observing each of the attainable values for Y in the cohort file (CF) given X and a value for $\theta = \theta^q$:

$$\Pr\left(Y_1 = y_i \middle| X_1, \theta^q\right) = \frac{f\left(y_i \middle| X_1, \theta^q\right)}{\sum_{k \in A} f\left(y_k \middle| X_1, \theta^q\right)}, \text{ for } y_i \in CF \tag{C.1}$$

Draw a value for Y_1 from this discrete distribution. Then go to the next record on X and repeat this procedure after removing the value $y_{1,t}{*}= y_{t_1}$ that is drawn in the previous step from the cohort file. t_1 indicates the t^{th} simulated y value belonging to X_1. There are at most six different values to choose from.

The probability distribution function then looks like:

$$\Pr\left(Y_2 = y_i \middle| X_2, \theta^q\right) = \frac{f\left(y_i \middle| X_2, \theta^q\right)}{\sum_{\substack{k \in A \\ k \neq t_1}} f\left(y_k \middle| X_2, \theta^q\right)}, \text{ for } y_i \in CF \setminus \{y_{t_1}\} \tag{C.2}$$

where A \ B means the set A after deletion of the elements in set B.

From this distribution function y_{t_2} will be drawn. In the next step y_{t_1} and y_{t_2} will be excluded from the set of possible values for Y. After each draw the set of attainable values for Y contains fewer value. Following this procedure, a value for Y will be drawn for all records on X in the matching group in a sequential way. For each series of draws of y-values the procedure will start from a different record in the main data set. This is done since the probability of certain combinations X and Y occurring is dependent on the order of observations X_i (i=1,...,59) for which y-values are drawn. The last observation on X potentially has fewer values to choose from since several possible values have already been combined with other observations on X. Starting from the same observation on X over and over again could disadvantage certain combinations of values y and x. By alternating the starting point it is expected that all combinations that would occur with high probability according to the actual probability distribution, will also be drawn from this much simpler importance function.

Although an attempt is made to get as close as possible to drawing from the actual probability distribution function, we know that this will never exactly be the case. To correct for drawing from an incorrect distribution, the simulated contributions to the log likelihood have to be weighted. The appropriate weights can be found by dividing the value of the actual probability density function by the approximated value at the simulated data point[19]:

$$w_t = \frac{f\left(Y_1 = y_{t_1} \middle| X_1, \theta^q\right) \ldots f\left(Y_{59} = y_{t_{59}} \middle| X_{59}, \theta^q\right)}{\dfrac{f\left(Y_1 = y_{t_1} \middle| X_1, \theta^q\right)}{\sum\limits_{k} f\left(Y_1 = y_k \middle| X_1, \theta^q\right)} \cdot \dfrac{f\left(Y_2 = y_{t_2} \middle| X_2, \theta^q\right)}{\sum\limits_{\substack{k \\ k \neq t_1}} f\left(Y_2 = y_k \middle| X_2, \theta^q\right)} \cdots \dfrac{f\left(Y_{59} = y_{t_{59}} \middle| X_{59}, \theta^q\right)}{\sum\limits_{\substack{k \\ k \neq t_1, t_2, \ldots, t_{58}}} f\left(Y_{59} = y_k \middle| X_{59}, \theta^q\right)}} \qquad (C.3)$$

We only accept those series of draws that fulfil all requirements as described in Appendix B, that is $\{y_{t_1}, \ldots, y_{t_{59}}\} \in A$.

Using the weights w_t, $Q(\theta,\theta_q;A)$ is now approximated by:

[19] Note that only the numerator of both density functions appears in (C.3). The denominators are constants and can be left out of the formula.

$$\frac{1}{\sum_{t=1}^{T} w_t} \sum_{t=1}^{T} w_t \left\{ \log\left[f\left(y_{t_1} | \theta, X_1 \right) \right] + \ldots + \log\left[f\left(y_{t_{59}} | \theta, X_{59} \right) \right] \right\} \tag{C.4}$$

and the information matrix can be approximated by using (see Tanner, 1993):

$$\begin{aligned}
\left. \frac{\partial^2 l\left(\theta | A, X_1, \ldots, X_{59}\right)}{\partial \theta^2} \right|_{\theta^q} = & \frac{1}{\sum_{t=1}^{T} w_{tj}} \sum_{t=1}^{T} w_{tj} \left. \frac{\partial^2 \left\{ \log\left[f\left(y_{t_1} | \theta, X_1 \right) \right] + \ldots + \log\left[f\left(y_{t_{59}} | \theta, X_{59} \right) \right] \right\}}{\partial \theta^2} \right|_{\theta^q} + \\
& \frac{1}{\sum_{t=1}^{T} w_t} \sum_{t=1}^{T} w_t \left(\left. \frac{\partial \left\{ \log\left[f\left(y_{t_1} | \theta, X_1 \right) \right] + \ldots + \log\left[f\left(y_{t_{59}} | \theta, X_{59} \right) \right] \right\}}{\partial \theta} \right|_{\theta^q} \right)^2 - \\
& \left(\frac{1}{\sum_{t=1}^{T} w_t} \sum_{t=1}^{T} w_t \left. \frac{\partial \left\{ \log[\left[f\left(y_{t_1} | \theta, X_1 \right) \right] + \ldots + \log\left[f\left(y_{t_{59}} | \theta, X_{59} \right) \right] \right\}}{\partial \theta} \right|_{\theta^q} \right)^2
\end{aligned} \tag{C.5}$$

TABLE 1. — ORDERED PROBIT OF JUVENILE ARRESTS FOR MEN AND WOMEN

	Men				
	Estimated model		Marginal effects[a]		
	Parameter	t-value	JA=0	JA=1	JA=2
Constant	-1.1386	-6.39	0.7087	0.1302	0.1612
Race is non-white	0.3496	2.32	-0.1091	0.0333	0.0759
No father in childhood home	0.1697	0.85	-0.0542	0.0151	0.0391
Stay at home mother	0.2242	1.43	-0.0708	0.0199	0.0509
Physically or sexually abused	-0.1617	-0.63	0.0482	-0.0148	-0.0333
No mother in childhood home	0.2806	0.61	-0.0920	0.0233	0.0687
Father's education is ≥ hs grad	-0.3041	-1.93	0.0946	-0.0291	-0.0656
Number of siblings	0.0677	1.91	-0.0190	0.0065	0.0125
Left school < 17 years old	0.7527	3.38	-0.2582	0.0617	0.1965
Left school at 17 years old	0.3142	1.96	-0.0990	0.0311	0.0678
Threshold parameter	0.4903	10.03			
	Women				
	Estimated model		**Marginal effects[a]**		
	Parameter	**t-value**	**JA=0**	**JA=1**	**JA=2**
Constant	-1.5790	-5.19	0.8640	0.1181	0.0220
Race is non-white	0.4889	2.73	-0.0941	0.0736	0.0205
No father in childhood home	0.3284	0.87	-0.0720	0.0546	0.0175
Stay at home mother	0.4767	2.17	-0.1040	0.0781	0.0259
Physically or sexually abused	-0.3160	-0.90	0.0538	-0.0427	-0.0111
No mother in childhood home	-0.5347	-1.22	0.0801	-0.0646	-0.0155
Father's education is ≥ hs grad	-0.0951	-0.41	0.0186	-0.0144	-0.0042
Number of siblings	-0.0383	-0.73	0.0082	-0.0062	-0.0021
Left school < 17 years old	1.0905	2.15	-0.2968	0.2062	0.0906
Left school at 17 years old	0.2492	1.17	-0.0471	0.0384	0.0088
Threshold parameter	1.0799	6.88			

Note a: In the row for the constant, the predicted proportions of the sample with 0, 1 and 2 or more juvenile arrests are presented.

TABLE 2. — PROBABILITY OF JUVENILE AND ADULT ARRESTS FOR MEN

	Estimated model		Marginal effects[a]		
Juvenile arrests	**parameter**	**t-value**	**JA=0**	**JA=1**	**JA=2**
Constant	-1.1189	-6.31	0.7083	0.1306	0.1612
Race is non-white	0.3440	2.29	-0.1076	0.0329	0.0748
No father in childhood home	0.1700	0.86	-0.0544	0.0152	0.0392
Stay at home mother	0.2201	1.43	-0.0696	0.0197	0.0500
Physically or sexually abused	-0.1365	-0.55	0.0410	-0.0125	-0.0284
No mother in childhood home	0.3024	0.66	-0.0997	0.0250	0.0747
Father's education is $\geq$ hs grad	-0.3035	-1.94	0.0947	-0.0291	-0.0656
Number of siblings	0.0688	1.98	-0.0194	0.0067	0.0127
Left school < 17 years old	0.7492	3.72	-0.2589	0.0603	0.1986
Left school at 17 years old	0.2475	1.63	-0.0775	0.0247	0.0528
Adult arrests			**AA=0**	**AA=1**	
Constant	-2.0788	-2.51	0.8864	0.1136	
Race is non-white	1.4030	5.67	-0.1069	0.1069	
No father in childhood home	0.1900	0.75	-0.0280	0.0280	
Stay at home mother	0.2816	1.28	-0.0425	0.0425	
Physically or sexually abused	0.1031	0.36	-0.0222	0.0222	
No mother in childhood home	-1.0199	-2.02	0.0997	-0.0997	
Father's education is $\geq$ hs grad	0.1791	0.68	-0.0409	0.0409	
Number of siblings	0.1106	2.26	-0.0122	0.0122	
Years of schooling	-0.0362	-0.81	0.0567	-0.0567	
Number of children	0.0360	0.61	-0.0059	0.0059	
Married	-0.5650	-3.77	0.0812	-0.0812	
Juvenile arrests	2.7285	2.29	-0.1439	0.1439	
Juv.arr.*race	-1.0368	-4.07			
Juv.arr.*years of schooling	-0.1955	-2.94			
Rho	0.7076	3.66			
Threshold parameter	0.4907	9.81			

Note a: In the row for the constant of the juvenile equation, the predicted proportions of the sample with 0, 1 and 2 or more juvenile arrests are presented. In the row for the constant of the adult equation, the predicted proportions of the sample with and without adult arrests are presented. In the other rows the value of the relevant variable is changed from 0 to 1 (except in the row for years of schooling where the value is changed from 10 to 12 years of schooling), after which the resulting change in probability is calculated.

TABLE 3— THE INTERACTION EFFECTS IN THE MODEL FOR MEN (DERIVED FROM THE COEFFICIENTS IN TABLE 2)

	Joint effects	
Race	**parameter**	**t-value**
White, no juvenile arrest[a]		
Non-white, no juvenile arrest	1.4030	5.67
White, juvenile arrest	2.7285	2.30
Non-white, juvenile arrest	3.0947	2.55
Years of schooling		
No schooling, no juvenile arrest[a]		
1 year schooling, no juvenile arrest	-0.0362	-0.81
No schooling, juvenile arrest	2.7285	2.30
1 year schooling, juvenile arrest	2.4968	2.17

Note a: This is the person of comparison.

TABLE 4— THE EFFECT OF JUVENILE ARRESTS ON THE PROBABILITY OF ADULT ARREST BY RACE AND YEARS OF SCHOOLING

	Effect of juvenile arrests[b]	
	AA=0	**AA=1**
On average in whole sample	-0.1439	0.1439
By race		
White	-0.1710	0.1710
Non-white	-0.1247	0.1247
By years of schooling		
People without schooling	-0.8539	0.8539
People with 1 year of schooling	-0.8513	0.8513
People with 10 years of schooling	-0.3167	0.3167
People with 12 years of schooling	-0.1489	0.1489

Note a: These are the effects of juvenile arrests on the probability of adult arrests, measured by the difference in the probability of adult arrests conditional on having juvenile arrests and the probability of adult arrests conditional on having no juvenile arrests.