

# Effects of Family Background on Crime Participation and Criminal Earnings: An Empirical Analysis of Siblings

Liliana E. Pezzin

Medical College of Wisconsin - Health Policy Institute and  
Department of Medicine

## RESUMO

*Este estudo usa dados do National Longitudinal Survey of Youth para medir a extensão pela qual interações sociais de família explicam a variância na probabilidade de participação em crime e na intensidade e sucesso em atividades criminais. A estimação é baseada em um modelo de equações múltiplas cujas perturbações são interligadas por uma variável inobservável comum. A virtude do método proposto é usar dados referentes a irmãos - que compartilham as mesmas características maternas e paternas em relação a fatores de família que possivelmente influenciam a sua própria decisão de engajar em atividades ilegais - para estimar o efeito do background familiar na decisão de participar em crime. Os resultados empíricos indicam um alto nível de correlação entre fatores inobserváveis medindo efeitos de família e a propensão de irmãos em participar de atividades criminosas (0.44 a 0.55). Efeitos de família explicam 25% da variância em renda criminal. Finalmente, os resultados sugerem que estimativas que ignoram o background familiar induzem vieses significantes no efeito de variáveis, tais como raça e educação na propensão de jovens a participar de crimes patrimoniais.*

## PALAVRAS-CHAVE

*efeitos de família, crime, variáveis latentes, dados longitudinais.*

## ABSTRACT

*This study exploits the sibling structure of the National Longitudinal Survey of Youth data to measure the degree to which family background explains the variance in the propensity to engage in criminal activities and in the intensity and success of crime participation as measured by the level of criminal earnings. A multiple-equation model whose reduced form disturbances are connected by a common unobservable variable having a variance-components structure is developed and estimated. Estimation results indicate a high level of association (net of observable measures of family background) between the unobserved factors affecting siblings' propensity to engage in criminal activities in a family, with estimated intra-family correlations ranging from 0.44 to 0.55. Sharing a common family background explains around 25% of the variance of the unconditional criminal income. The results suggest that ignoring family background effects leads to a significant upward bias in the effects of race and education on the propensity to engage in income-generating crime.*

## KEY WORDS

*family effects, crime, latent variable, variance components models.*

JEL Classification  
K42, C33

Understanding the causes of youth criminality is a major objective of social research. The most troubling features of this problem are its prevalence and incidence among adolescents, the disparity in the distribution of criminals by race and background, and the persistence of the negative consequences of a criminal record throughout the youth's adult life.

A long tradition of social thought has maintained that an individual's environment influences his propensity for crime. Recently, research in criminology has begun to refocus its attention on the role of family in explaining delinquency. (HIRSCHI, 1983; LOEBER & STOUTHAMER-LOEBER, 1986; FARRINGTON, 1987; LAUB & SAMPSON, 1988; GOTTFREDSON AND HIRSCHI, 1990). In explaining the origins of delinquency, criminologists have argued that informal social controls derived from the family (for example, parental supervision, monitoring and parent-child attachment) mediate the effects of individual and structural background variables and are the most powerful predictors of juvenile delinquency. (LOEBER & STOUTHAMER-LOEBER, 1986).

While the recent resurgence of interest in the economics of crime has sparked a large and growing literature (see FREEMAN, 1999 for a review), relatively little attention has been paid to the family. Early studies that discuss the role of the family concentrate on measuring differences in market-valued characteristics, such as education, that result from differential parental investment decisions. (BECKER, 1974; BECKER & TOMES, 1976; PEZZIN, 1994). Separate veins of research have focused on the deterrent effect of legal sanctions, both at the aggregate and individual level analyses (BECKER, 1968; EHRLICH, 1973; WITTE, 1980; SCHMIDT & WITTE, 1989; LEVITT 1998a and 1998b) and the linkage between crime and labor market opportunities, particularly the relationship between unemployment and crime. (SJOQUIST, 1973; BLOCK & HEINEKE, 1975; GROGGER, 1992; FREEMAN & RODGERS, 1999; GOULD, WEINBERG & MUSTARD, 2002; MUSTARD, 2003).

More recent work has used economic models with local interactions (SCHEINKMAN & WOODFORD, 1994) to examine social feedback effects on crime. Sah (1991) developed a dynamic theoretical model in which

individual's perceptions concerning their probabilities of punishment are influenced by information generated within their local economy. These subjective probabilities, in turn, influence individual's choices regarding criminal involvement. The resulting relationships are then used to explain how criminality might evolve over time and why crime rates might differ among different societal groups even when they face similar economic fundamentals. That model, however, was not designed to explain the extent to which membership in the same family unit might interfere with the transmission of criminal choices across individuals, time and space. Glaeser, Sacerdote and Scheinkman (1996) develop and estimate a model where social interactions create enough covariance across individuals to explain the high cross-city and cross-precinct variance in crime rates in the United States. A key finding of the paper is that criminal's decisions in a metropolis are highly dependent suggesting a large degree of social interaction in criminal behavior. The authors also interpret the finding of higher levels of social interactions among female headed households as evidence that the average social interactions among criminals are higher when there are not intact family units. Using data collected from a single prison in Brazil, Mendonça, Loureiro and Sachsida (2002) report that variables capturing social interaction, such as good family environment, had negative effects on violent crime rates among this inmate population. As in Glaeser *et al.*, the model does not address the issue of heterogeneity in crime participation within families, however, nor does it allow for the possibility that family effects may not be completely captured through observable characteristics. Because families affect not only the initial distribution of endowments of their members but also their perceptions, expectations and opportunities, the aggregate-level social interaction models may not capture the effect of family heterogeneity on youth criminality. If the potential effects of family background, broadly conceived, are not completely reflected in observable characteristics that individuals have or choose to acquire, ignoring unobservable or unmeasured family background effects may lead to overlooking important aspects of the creation and perpetuation of economic (and other) differences among individuals and families across generations that lead to crime.

Kinometrics is a methodology that allows one to use samples with genetically linked relatives or kin to study the roles of family background and environment. There are two major concerns in kinometrics. The first is to eliminate potential biases by controlling for **unmeasured or unobserved** variables when estimating the relationship between measured variables, for example, controlling for "ability" and "motivation" when estimating the returns to schooling. The second concern is to measure the combined and separate effects of unobserved family and environment variables.

The prospect of controlling for common family influences by modelling the similarities of siblings has helped to motivate a number of economic studies of the social stratification process. Kinometrics has been applied in a variety of areas including educational and earnings attainment, occupational success and cognitive skills. (CHAMBERLAIN & GRILICHES, 1975 and 1977; CORCORAN, *et. al.*, 1976; TAUBMAN, 1976; BEHRMAN & TAUBMAN, 1976 and 1977; OLNECK, 1977; GRILICHES, 1979; KEARL & POPE, 1986; SOLON *et. al.*, 1990; LAM & SCHOENI, 1993). Heterogeneity in households as opposed to individual characteristics has been shown to be important in obscuring the causal effects of parental behavior on outcomes and attainment of children.

Drawing on previous literature on kinometrics, this study uses sibling data to examine the effects of family background on youths' decisions to participate in income-generating crime and on the quality of their "occupational match". The paper's main goal is to measure the degree to which family background, whose source may be genetic, environmental or behavioral, explains the variance in the propensity to engage in criminal activities and in the intensity and success of crime participation as measured by the level of criminal earnings. A second goal of the study is to measure the bias that results from ignoring unobserved family heterogeneity when estimating such relationships. I exploit the family panel structure of the National Longitudinal Survey of Youth data, a nationally representative U.S. survey, to test for the existence of such effects and to measure their relative strength. Controlling for observable individual and family characteristics, the estimated familial resemblance may be interpreted, then, as the upper bound of unobserved

shared genetic, behavioral and environmental factors in the decision to choose crime as a career.

A multiple-equation model whose reduced form disturbances are connected by a common unobservable variable having a variance-components structure is developed and estimated. Unobserved family heterogeneity is modelled as a random effect, possibly correlated with the observed variables, that allows for variation in both the siblings' propensity to participate in income-generating crime and their realized criminal income. The specification biases that arise from the simultaneous presence of sample selectivity and unobserved heterogeneity are addressed by extending the standard sample selection model to the unobserved effects framework.

Estimation results indicate that, despite the importance of many individual-specific characteristics in influencing siblings' probability of crime participation, a substantial fraction of its variance (44 to 55 percent) is due to family heterogeneity. The results also suggest that failure to condition on these unobservable family-specific characteristics leads to a significant upward bias in the effects of race and education on the probability of crime participation.

## *I. MODEL SPECIFICATION AND ESTIMATION*

The analysis proceeds in two stages. First, a panel data-latent variable model of crime participation is developed and used to estimate the effects of family background and other variables on the decision to engage in income-generating crime. Second, a conditional variance components model of the determinants of illegal income is estimated for the sub-sample of active criminals.

### *I.1 Family Effects and the Decision to Participate in Crime*

The underlying crime participation model presented here is a simple variant of standard models of crime (BECKER, 1968) and occupational choice under uncertainty. Throughout the analysis it is postulated that the discrete choice made by the individual is not between crime and legal employment, but between having some or none of his income generated by illegal activi-

ties. The implicit assumption is that there is an important qualitative difference between legitimate and illegitimate income-generating activities, in terms of both the level of risk involved and the nature of the time allocation.

Let  $Yc_{ij}$  be the monetarized returns from criminal activity and  $C_{ij}$  be the cost of punishment associated with a positive level of criminal activity  $\theta_{ij}$  where  $\theta_{ij}$  is the proportion of time devoted to crime by individual  $i$  of family  $j$  ( $0 < \theta_{ij} \leq 1$ ). Let  $Yl_{ij}$  represent the legal earnings and  $P$  the probability of punishment. Let us also introduce the variables  $\alpha_j^c$  and  $\alpha_j^l$  to represent all other factors determining the desirability of the (full or part-time) criminal and legal occupations. These family-specific variables include non-pecuniary characteristics of the two occupations that cannot be captured by the measured values of  $Yc_{ij}$ ,  $C_{ij}$  and  $Yl_{ij}$  as well as differences in tastes across individuals of different families that may arise as a result of asymmetric family preferences, endowments or attributes.

In its simplest form, suppose  $Vc_{ij}$  represents the expected value of engaging in criminal activities to individual  $i$  of family  $j$  as perceived at the outset, while  $Vl_{ij}$  is the value of an alternative, exclusively legal occupation such that

$$Vc_{ij} = [(1 - P) Yc_{ij}(\theta_{ij}) + P C_{ij}(\theta_{ij})] + \alpha_j^c \quad (1)$$

$$Vl_{ij} = Yl_{ij} + \alpha_j^l \quad (2)$$

Note that the expected utility of action  $\theta$  is given by a weighted average of the expectations of criminal rewards and criminal costs, conditional on activity level  $\theta_{ij}$ , with the probability of punishment serving as the weight.

The model posits that the individual will choose to participate in crime if at the time of the decision  $t$  the inequality  $Vc_{ij} > Vl_{ij}$  holds, that is, if the expected value of the criminal option  $Vc_{ij}$  exceeds the expected value of the alternative legal occupation  $Vl_{ij}$ . The probability of crime participation at time  $t$  is

$$P(\text{crime})_{ij} = \text{Prob}(\tilde{V}_{c_{ij}} - \tilde{V}_{l_{ij}} - \alpha_j > 0) \quad (3)$$

where  $\tilde{V}_{c_{ij}} = [(1-P)Y_{c_{ij}}(\theta_{ij}) + P C_{ij}(\theta_{ij})]$ ,  $\tilde{V}_{l_{ij}} = Y_{l_{ij}}$  and  $\alpha_j = \alpha_j^l - \alpha_j^c$ .

In this context, there exists a latent variable  $I_{ij}^*$  which can be interpreted as the unobserved measure of the difference in expected utility between participating in crime, full or part-time, and choosing an exclusively legal occupation such that

$$I_{ij}^* = \tilde{V}_{c_{ij}} - \tilde{V}_{l_{ij}} - \alpha_j \quad (4)$$

Unfortunately, none of the values in equation (4) are directly observed. To implement the model, it is assumed that the heterogeneity variable  $\alpha_j$  is distributed in the population of families according to a distribution function  $F(\alpha_j)$ . Criminal and legal rewards  $Y_{c_{ij}}$  and  $Y_{l_{ij}}$  are assumed to depend upon a vector of observable individual attributes  $Z_{ij}$ , including the youth's age, his/her accumulated human and criminal capital stock as well as other relevant exogenous variables capturing the labor and criminal market opportunities.  $C_{ij}$  is a function of family background and status variables  $W_{ij}$  that affect the individual's general valuation of criminal costs. These variables are assumed to capture both objective (expected probability of punishment,  $P$ ) and subjective (anxiety and other non-monetary costs) aspects of criminal costs. Finally, the stochastic components  $\varepsilon_{ij}^c$  and  $\varepsilon_{ij}^l$  are included to capture the contribution of white-noise errors, varying over individuals and families, that may influence the value of criminal and (exclusively) legal occupations.

$$I_{ij}^* = Z_{ij}(\delta_c - \delta_l) + W_{ij}\gamma - \alpha_j - \varepsilon_{ij} \quad (5)$$

The individual's propensity to engage in criminal activities is then given by where the coefficient vectors  $(\delta_c - \delta_l)$  and  $\gamma$  measure the reduced-form change in the utility difference  $V_{c_{ij}} - V_{l_{ij}}$  due to a unit change in  $Z_{ij}$  and  $W_{ij}$ , respectively, and  $\varepsilon_{ij} = \varepsilon_{ij}^l - \varepsilon_{ij}^c$ .

To simplify the notation below, let  $Z_{ij}(\delta_c - \delta_l) + W_{ij}\gamma = X_{ij}\beta$  where  $X_{ij} = [Z_{ij} W_{ij}]$  and  $\beta' = [(\delta_c - \delta_l) \gamma]$ . Defining  $I_{ij}$  as the observed indicator of whether or not the  $i^{\text{th}}$  individual of family  $j$  participated in crime,

$$\begin{aligned}
 I_{ij} &= 1 \text{ if } I_{ij}^* \geq 0 & i=1KN; \quad j=1KJ \\
 I_{ij} &= 0 \text{ if } I_{ij}^* < 0
 \end{aligned} \tag{6}$$

and assuming that  $\alpha_j / X \approx N(0, \sigma_\alpha^2)$  and  $\varepsilon_{ij} / X \approx N(0, \sigma_\varepsilon^2)$  are mutually independent, the probability that individual  $i$  of family  $j$  will participate in crime is given by

$$\text{Prob}(I_{ij} = 1) = \text{Prob}(I_{ij}^* > 0) = \text{Prob}(X_{ij}\beta - \alpha_j > \varepsilon_{ij}) \tag{7}$$

In this expression,  $\varepsilon_{ij}$  measures the individual-specific component not shared by siblings in the same family, while  $\alpha_j$  captures the unmeasured permanent component common to the siblings in the  $j^{\text{th}}$  family. This parameter can be affected in complex ways by the family's attributes and endowments as well as by parent's behavior regarding investment in their children. Under this interpretation,  $\alpha$  is likely to be correlated with some of the regressors  $X$ , particularly the education level ( $x$ ) of all the siblings in the household. To allow for such correlation, I adopt the approach suggested by Chamberlain (1984), positing that

$$\alpha_j = \tau_1 x_{1j} + \dots + \tau_N x_{Nj} + \eta_j \quad j = 1 \dots J \tag{8}$$

where  $\eta_j/x_j \approx N(0, \sigma_\eta^2)$  is assumed to be uncorrelated with  $x$  and  $\varepsilon_{ij}$ .

For purposes of estimation, let  $\rho = \sigma_\eta^2 / (\sigma_\eta^2 + \sigma_\varepsilon^2) = \sigma_\eta^2 / \sigma^2$ , where  $\sigma^2 = \sigma_\eta^2 + \sigma_\varepsilon^2$  is normalized to one, and define  $\tilde{\eta}_j = \eta_j / \sigma_\eta$ . Given these assumptions, and imposing the cross-sibling symmetry constraint  $\tau_1 = \dots = \tau_N = \tau$ , the likelihood function for the observed sample of  $I_{ij}$  is given by

$$L = \prod_{i=1}^N \int_{-\infty}^{\infty} \prod_{j=1}^J \Phi[(X_{ij}\beta - \tau x_{kj} - \rho^{1/2} \tilde{\eta}_j) / (1-\rho)^{1/2}] [2I_{ij}-1] f(\tilde{\eta}_j) d\tilde{\eta}_j \tag{9}$$

Heterogeneity is controlled for by integrating the likelihood for each observation with respect to the density of  $\tilde{\eta}$ , using a Gauss-Hermite quadrature procedure. This specification yields a multivariate probit model of the "random effects" variety in the sense that it allows us to estimate the effects of



(observable) family-invariant regressors, such as the race/ethnicity indicators, through the use of a marginal maximum likelihood estimation procedure. However, in contrast to the conventional random effects specification, the unobserved component is allowed to be correlated with the observable explanatory variables through the statistical dependence specified in equation (8). (PEZZIN, 2003).

The significance of unobserved family-effects in determining crime participation hinges on the estimated size of  $\rho$ , the correlation between the total disturbance  $(\eta_j + \varepsilon_{ij})$  for the same household. If the estimated value is close to zero then little family heterogeneity is implied, and all variation in the decision to participate in illegal activities can be ascribed to observed variables capturing net earnings differences or individual-specific heterogeneity. If, on the other hand,  $\rho$  is significantly positive, then family heterogeneity is implied meaning that siblings will differ in their propensity to participate in crime even in the absence of variation in their individual characteristics.

### *1.2 Family Effects and Illegal Earnings*

The second objective of the paper is to measure the degree to which family background explains the variance in the intensity and success of crime participation as measured by the level of criminal earnings. For this analysis, I consider the determinants of crime income for the sub-sample of "active" criminals, i.e., individuals who chose to participate in crime during the survey reference year. The effect of unobserved family background is tested by estimating a variation of the standard human capital specification for individual income where the dependent variable is the natural logarithm of the individual's criminal income, obtained from information on total income and the individual's reported percentage of total income generated from illegal activities. This variable captures both the decision to participate in crime and the intensity of criminal behavior.

The model can be written as

$$\begin{aligned}
 \ln(Y_{Cij}) &= Z_{ij}\delta_c + \pi_j + v_{ij} \\
 \text{observed iff } I_{ij} &= 1 \\
 I_{ij}^* &= X_{ij}\beta - \tau_{xj} - \eta_j - \varepsilon_{ij} \\
 I_{ij} &= 1 \text{ if } I_{ij}^* \geq 0 \\
 I_{ij} &= 0 \text{ if } I_{ij}^* < 0
 \end{aligned} \tag{10}$$

where  $I_{ij}^*$  is the latent index of the probit crime participation equation which determines whether individual  $i$  of family  $j$  participated in crime during the survey reference period. In addition to previously defined notation,  $\pi_j$  is a family-specific effect, assumed to capture unmeasured productivity differences that are constant over individuals of the same family, and  $v_{ij}$  is an independent and identically distributed error term.

As it is well known, estimates derived from self-selected samples may be biased due to correlations between the independent variables and the stochastic disturbance induced by the sample selection rule. The utilization of panel data compounds the problem by confronting the estimation with the simultaneous presence of sample selectivity and unobserved heterogeneity, both of which give rise to specification bias.

In this study, consistent estimates of the criminal earnings equation are obtained by extending Heckman's (1979) sample selection model to the unobserved effects framework. The procedure builds on recent work by Verbeek and Nijman (1992) and Wooldridge (1992) who proposed methods for testing and correcting for selectivity bias in panel data models. As in these studies, it is assumed that both the unobservable effect and the idiosyncratic errors in the selection process are normally distributed. However, rather than obtaining the selection correction from variable additions to a standard probit, here the modified inverse Mill's ratio term is derived from the correlated panel probit model. (PEZZIN, 2003). This feature is important because it provides the expectation of the error terms  $\eta_j + \varepsilon_{ij}$  conditional on the entire response indicator vector  $I_j = (I_{1j} \dots I_{Nj})$  by accounting directly

for the sibling structure of the data used in the crime participation (selection) function.

To simplify notation, define  $\varepsilon_{ij}^* = \eta_j + \varepsilon_{ij}$  as the total disturbance from the crime participation equation. In order to allow for correlation between the family-specific effect  $\pi_j$  and the explanatory variables, assume that  $\pi_j$  depends on  $Z_j$  through the family values of criminal and human capital variables  $z_j$  such that

$$E(\pi_j | Z_j, \varepsilon_{ij}^*) = L(\pi_j | Z_j, \varepsilon_{ij}^*) = \psi' z_j \quad (11)$$

where  $L(\bullet | \bullet)$  denotes the linear projection of  $\pi_j$  onto  $z_j$ ,  $z_j$  includes the education level and criminal experience of all siblings in the household, and

$$E(\varepsilon_{ij}^* | z_j) = 0 \quad .$$

After substituting in Equation (11), the model of Equation (10) can be estimated via a two-step procedure assuming that  $\varepsilon_{ij}^*$  and  $v_{ij}$  are distributed according to a zero mean bivariate normal distribution. In the first step, consistent estimates of the structural selection equation are obtained by maximizing the likelihood function corresponding to the correlated panel probit as described in Section I.1. In the second step, the estimates of  $\Theta = (\delta_c, \psi, b)$  are obtained by applying the pooled ordinary least squares estimator to the expectation of the criminal earnings conditional on  $(Z_{ij}, z_j, I_{ij} = 1)$ ,

$$E(Y_{c_{ij}} | Z_{ij}, z_j, I_{ij} = 1) = Z_{ij} \delta_c + z_j \psi + b \Lambda_{ij} + \omega_{ij} \quad (12)$$

where

$$b \Lambda_{ij} = E(v_{ij} | I_j),$$

$$b = \sigma_{v \varepsilon^*} = \text{cov}(v_{ij}, \varepsilon_{ij}^*),$$

$$\Lambda_{ij} = E(\eta_j + \varepsilon_{ij} | I_j) - \frac{\sigma_\eta^2}{N \sigma_\eta^2 + \sigma_\varepsilon^2} \sum_{s=1}^N E(\eta_j + \varepsilon_{sj} | I_j),$$

$$E(\eta_j + \varepsilon_{ij} | I_j) = \frac{\phi[r_{ij} w_{ij}]}{\Phi[r_{ij} w_{ij}]}$$

and  $\Phi$  and  $\phi$  represent the standard normal distribution and density functions. The terms  $r_{ij} = (2I_{ij} - 1)$  and  $w_{ij} = [ (X_{ij}\beta - \tau_{x_{kj}} + \tilde{\psi}_{z_{kj}} - \sqrt{\rho} \tilde{\eta}_j) / \sqrt{(1-\rho)}$  are evaluated at the consistent parameter estimates obtained from the correlated panel probit equation, and reduce to the standard Heckman selectivity correction term in the absence of family heterogeneity ( $\sigma_{\eta}^2 = 0$ ). Note that the remaining error term in equation (12)  $\omega_{ij} \approx N(0, \sigma_{\omega})$  no longer has an error components structure and is, by construction, independent of  $I_j$ .

By following this procedure, the coefficient on the conditional mean of the selection term provides an estimate of the correlation of the reduced-form participation residual with the sample selection disturbance accounting for unobserved family effects. This approach permits both family-specific and youth-specific unmeasured effects to be incorporated within a marginal productivity framework. (PEZZIN, 2003).

## II. DATA, VARIABLES AND SAMPLE STATISTICS

The data source for this analysis is the Youth Cohort of the National Longitudinal Surveys (NLSY, Center for Human Resource Research, Ohio). The NLSY is based on a nationally representative sample of over 12,000 individuals between the ages 14 and 22 years old. A broad range of questions is covered in the interviews, providing a comprehensive account of the youth's educational and work experience, family background, and sources of earned and supplemental income. In addition to the standard questions, the 1996 wave of the NLSY collected self-reported information regarding respondents' involvement with crime, including contacts with the police, courts and correctional institutions. For the purpose of this study, I focus on a relatively homogeneous class of crimes, primarily property offenses, such as shoplifting, theft, robbery, burglary, drug-dealing and manufacturing, possession or sale of stolen goods and other property crimes. These are the behaviors used as indicators of criminality throughout the analysis.

Because of its sample design, the NLSY provides a unique opportunity to study family effects in crime participation patterns. The survey was designed in a stratified random fashion from an underlying household sampling frame, implying that the same household contributed information on several youths (generally siblings) within the panel. Another unique feature of

the NLSY ideal to the study of family effects on criminal involvement is the fact that its sample is not drawn from an offender population. The sample includes true innocents, experimenters with crime, and individuals with fairly lengthy records who persist in crime as well as those who appear to have experimented with crime but have now "retired". Thus the entire range of possible outcomes from participation in crime is represented among this group.

The sample selection strategy was designed to maximize the use of data for biological siblings with valid records on a set of relevant questions regarding their own education, crime participation and percentage of earnings generated from illegal activities. From the original data, it was possible to identify 5,085 individuals meeting these criteria. Since the present analysis is designed for pairs of siblings, eighty-four households for which only one sibling record was left after the appropriate deletions were dropped from the sample. While the assignment for families with only two individuals was unambiguous, for families with three or more siblings, two siblings were randomly chosen to be part of the final data pairs. Although this procedure entailed sacrificing "good" observations concerning some criminals, it minimized potentially important inter-relations between the sample selection criteria and the values of observed variables and assured the randomness of the selected pairwise sample.

In summary, 2,035 siblings-pairs, or 4,070 individuals, were identified as the final working sample for this study, out of which 1,331 individuals (32.7%) reported having had participated in property crime at some point in their lives, with 721 (17.7%) of them being active criminals at the time of the survey. Although high, these percentages are comparable to estimates reported in other surveys on crime participation. Wolfgang, Figlio and Sellin (1972) estimated that one-fourth to one-half of all young men are active in crime prior to their eighteenth birthday. It is also interesting to note that the siblings' sample proportions of participants in crime are slightly higher than the estimates for the entire NLSY sample (29.3% and 15.2%, respectively). A glossary of the basic variables along with the summary statistics for the sub-samples of non-criminals, criminals and active criminals is provided in Table 1.<sup>1</sup>

---

1 For the purposes of this study, "criminals" are individuals who admitted having participated in crime at some point in their lives. "Active criminals" are individuals who reported having some or all of their total income generated from illegal activities during the one-year survey reference period.

**TABLE 1 - VARIABLE DEFINITIONS AND SUMMARY STATISTICS  
PROPORTIONS/MEANS (STANDARD DEVIATIONS)**

Variable	Definition	Means (Standard Deviation)		
		Non-Criminals	Criminals	Active Criminals
<b>Respondent's Demographic Characteristics</b>				
Age	Age in years	17.62 (2.01)	17.25(2.02)	17.00 (1.97)
Education	Highest grade completed	10.88(1.90)	9.99 (1.81)	9.81(1.78)
Male	=1 if respondent is male; 0 otherwise.	41%	44%	47%
Black	=1 if respondent is Black; 0 otherwise.	40%	44%	47%
School Attendance	=1 if respondent is currently attending school; 0 otherwise.	78%	67%	68%
<b>Family Characteristics</b>				
Mother's Education	Highest grade completed by respondent's mother.	12.04 (2.22)	10.82 (2.94)	10.73(2.93)
Mother Works	=1 if respondent's mother works in the market; 0 otherwise.	56%	55%	60%
Number of Siblings	Number of siblings of index child.	3.60 (2.51)	4.06 (2.57)	4.07 (2.60)
<b>Crime-Related Variables</b>				
Drugs	=1 if respondent reports using illegal drugs; 0 otherwise.	38%	67%	76%
Number of Arrests	Number of arrests		0.29 (0.92)	0.26(0.89)
On Probation	=1 if respondent is currently on probation; 0 otherwise.		24%	21%
Crime Experience	Criminal market experience (age at last crime - age at first crime).		1.01(1.69)	0.54 (1.35)
<b>Neighborhood Variables</b>				
Crime In Neighborhood	Crime rate known to the police per 100,000 population in county of residence	5092 (3672)	5278(2847)	5269 (2850)
Unemployment	Unemployment rate for labor market of county of residence	2.60 (0.77)	2.62 (0.80)	2.65 (0.83)
Residence in metropolitan area	=1 if respondents current residence is in a central area of a SMSA; 0 otherwise.	40%	44%	43%
<b>Sample Size:</b>		2739	1331	721

Note: For the purpose of this study, "criminals" are individuals who admitted having participated in income-generating crime at some point in their lives. "Active criminals" is the sub-group of "criminals" who reported having some or all of their total income generated from illegal activities during the survey reference year. The total sample size is 4070 individuals, the sum of the "criminals" and "non-criminals" columns.

### III. RESULTS

#### III.1 Family Effects and the Decision to Participate in Crime

Table 2 presents the maximum likelihood estimates of the pooled and correlated panel probit specifications for the crime participation probability, which is measured by two indicators: participation in crime at some point in the respondent's life (**participation ever**) and participation in crime during the survey reference year (**participation now**). The common vector of independent regressors includes the set of observable family background variables (mother's education, mother's labor force participation, number of siblings of the respondent youth) as well as a wide array of individual human capital and personal characteristics variables expected to influence rewards and costs of criminal involvement. A set of environmental and locational variables (crime in the neighborhood, unemployment in the neighborhood, and residence in a standard metropolitan statistical area, SM-SA) was also included in all models to capture both legal and illegal labor market opportunities. In addition, variables related to criminal experience and competing uses of time, such as the number of police contacts, illegal drug use and school attendance, were incorporated in the model. Due to the possible endogeneity of these variables, a second specification that excludes the corresponding regressors is presented in the study.

Columns (1) and (2) of Table 2 display the results obtained when treating each sibling as an independent observation and fitting them with a conventional pooled probit equation. These estimates, which differ according to the inclusion of the possibly endogenous variables discussed above, provide the benchmark against which to compare the panel estimates. Table 3 provides the implied crime participation probability changes from the two specifications, which was obtained by simulating the estimated models at the individual level.<sup>2</sup>

---

2 The simulation proceeded as follows. For binary regressors, the predicted probability of crime participation was calculated assuming first a "zero" then a "one" value for the relevant regressor, holding other variables fixed at their individual levels. The marginal effects for the continuous variables were obtained by multiplying the relevant coefficient by the standard normal density evaluated at the individual level. These predictions were then averaged over the sample.

**TABLE 2 - MAXIMUM LIKELIHOOD ESTIMATES OF THE CRIME PARTICIPATION MODELS (ASYMPTOTIC STANDARD ERRORS IN PARENTHESES)**

Variables	Model Specification			
	Participation Ever			Participation Now
	Pooled Probit(1)	Pooled Probit(2)	Panel Probit(3)	Panel Probit(4)
Intercept	-0.127(0.312)	0.971**(0.229)	-0.694**(0.256)	-0.373**(0.258)
<b>Demographic Characteristics of Respondent</b>				
Age	-0.034(0.022)	0.058**(0.017)	0.049**(0.017)	-0.035**(0.019)
Education	0.075**(0.022)	0.122**(0.019)	-0.104**(0.019)	-0.092**(0.022)
Male	0.629**(0.048)	0.759**(0.049)	0.776**(0.047)	0.530**(0.052)
Black	0.195**(0.054)	0.093*(0.048)	0.044(0.051)	0.107**(0.056)
School Attendance	0.242**(0.067)			
<b>Family Characteristics</b>				
Mother's Education	0.012(0.009)	0.012(0.008)	0.013(0.008)	0.005(0.009)
Mother Works	-0.030(0.049)	0.027(0.044)	0.037(0.046)	0.120**(0.052)
Number of Siblings	0.006(0.010)	0.019**(0.009)	0.015*(0.009)	0.015(0.011)
<b>Crime-Related Variables</b>				
Drugs	0.702**(0.420)			
Number of Arrests	0.062 (0.420)			
<b>Neighborhood Variables</b>				
Crime In Neighborhood	-0.047(0.073)	0.011(0.064)	0.011(0.067)	0.028(0.082)
Unemployment	-0.004(0.031)	0.022(0.027)	0.023(0.031)	0.039(0.033)
SMSA	0.039(0.0550)	0.116**(0.049)	0.111**(0.052)	0.033(0.057)
<b>Auxiliary Parameters</b>				
Sibling's education, $\tau$			-0.033**(0.011)	-0.026**(0.013)
Correlation coefficient, $\bar{n}$			0.546**(0.035)	0.445**(0.054)
Log Likelihood	1852.8	2360.5	2327.9	1791.8

Note: The dependent variable in the "participation ever" specification is a binary variable taking the value of one if the respondent reported having participated in crime at some point in his life; zero otherwise. The dependent variable in the "participation now" specification is a binary variable taking the value of one if the respondent reported having participated in crime during the survey reference year (1996); zero otherwise. The sample size is 4070 individuals. The number of observations at one is 1331 for the "participation ever" specifications and 721 for the "participation now" specification. Significance at the  $p < 0.05$  and  $0.05 = p < 0.10$  levels are indicated by \*\* and \*, respectively.



As a first step in the analysis, a formal test of the appropriateness of the random-effects model is performed. The test of the null hypothesis that the unobserved family effect is present amounts to a test of whether the overidentifying restrictions implicit in the pooled model are correct, and can be carried out by estimating the unrestricted (panel probit) and restricted (pooled probit) reduced forms of the crime participation functions. (GREENE, 1993). The likelihood ratio test statistic revealed strong evidence against the conventional pooled model in both crime participation specifications ( $\chi^2_{(2)} = 65.2$  and  $\chi^2_{(2)} = 16.4$ , respectively).

The results suggest an overwhelming and highly significant level of association (net of observable measures of family background) between the unobserved factors affecting siblings propensity to engage in criminal activities in a family. The estimated levels of intra-family correlation  $\rho$  – which tell us what proportion of the variance in status is attributable to unobserved family background – range from 0.44 to 0.55. This finding suggests that even if all variation in observable individual characteristics were eliminated, most of the observed inequality in the propensity to participate in crime would remain.

Another way to assess the importance of unobserved family heterogeneity is to compare the correlated panel probit estimates to those of the pooled probit specification. A comparison of the results from columns 2 and 3 of Table 2 provides an indication of the magnitude of the bias induced by ignoring family effects.

Perhaps the most striking difference that results from accounting for unobserved family heterogeneity is the decline in the race/ethnicity coefficient (black) from 0.093 and significant (column 2) to 0.044 and insignificant (column 3). These coefficients imply that blacks are estimated to be more than twice as likely than non-blacks to participate in crime when family effects are not taken into account (0.031 and 0.014 in Table 3), an indication that the ethnicity variable in the conventional probit incorrectly picks up some common environmental background effect, and acts a proxy indicator for a range of more fundamental unobservables. While race is insignificant in determining participation ever, it is relevant in determining current par-

ticipation even after controlling for family heterogeneity, as indicated by the panel probit coefficient on column 4. Being black increases the probability of current participation by 3.3 percentage points. Notice also that the race coefficient, although consistently positive and significant, falls substantially in the pooled specifications (columns 1 and 2) when possibly endogenous variables related to school attendance, drug use and number of arrests are excluded from the model.

The effect of family investments in human capital is measured by the education level variables. The estimation reveals that education is, by far, the most significant observable variable consistently influencing the likelihood of criminal involvement in all specifications. Each additional year of schooling decreases the probability of crime participation by 2.1 to 4 percentage points (Table 3). Failure to account for correlated family heterogeneity leads to a substantial overestimation of the educational effect: The "true" effect of the youth's own education in the correlated panel probit model, given by the excess of his/her own estimated schooling coefficient over that of his/her sibling ( $\beta - \tau$ ) is -0.071, a predicted effect about 42% lower than that indicated by the corresponding pooled probit. For the participation now specification, the overall value of education is -0.066 (Table 2). The effect of the sibling's education on the youth's crime participation is small but statistically significant in both specifications. Each additional year of schooling by the sibling reduces the probability that an individual has engaged in illegal activities ever or in the reference year by 1.1 and 0.8 percentage points, respectively.

Several other important differences across estimates obtained by using the two procedures emerge from the results in Tables 2 and 3. For example, the estimated effects of observed family background variables on the probability of engaging in crime are substantially altered. Most notably, the coefficient measuring the effects of family size (number of siblings) falls by nearly one-fourth, from 0.019 to 0.015. Evaluating the partial derivative at the individual level, these coefficients imply that an additional sibling significantly increases the probability of crime participation by 4 percentage points in the pooled probit whereas the magnitude of this effect in the panel probit model is an insignificant 0.0027. Heterogeneity bias also appears to mask a posi-

tive effect of the variable measuring the labor market status of the respondent's mother on the decision to participate in crime during the reference year (comparison result not shown here).

TABLE 3 - CHANGES IN THE PREDICTED PROBABILITY OF CRIME PARTICIPATION FOR SELECTED VARIABLES

Effect of a unit change in	Model Specification		
	Participation Ever		Participation Now
	Pooled Probit	Panel Probit	Panel Probit
Race	0.031*	0.014	0.033**
Education			
Own Education ( $\beta - \tau$ )	-0.040**	-0.023**	-0.021**
Sibling's Education ( $\tau$ )	-0.011**	-0.008**	
Age	0.012**	0.011**	-0.010**
Mother's Labor Market Status	0.009	0.012	0.025**
Number of Siblings	0.004**	0.003	0.005

Note: The probability changes are based on estimates reported in columns (2), (3) and (4) of Table 2. For the binary regressors, the probability change was calculated by evaluating the predicted probability of crime participation assuming first a "one" then a "zero" value for the relevant regressor, holding other variables fixed at their individual levels. The marginal effects for the continuous variables were obtained by multiplying the relevant coefficient by the standard normal density evaluated at the individual level. These individual probabilities were then averaged over the sample. Significance at the  $p < 0.05$  and  $0.05 = p < 0.10$  levels are indicated by \*\* and \*, respectively.

The remaining variables generally have expected signs and their magnitudes are not significantly affected by the inclusion of random family effects in the model. The age coefficients in columns (2) through (4) of Table 2 indicate that although older individuals are more likely to have participated in crime at some point in their lives, they are less likely to be active criminals at a given time, a result consistent with empirical findings of a short-lived life-cycle in criminal involvement. (WOLFGANG, FIGLIO & SELLIN, 1972; VIS-CUSI, 1986; PEZZIN, 1992). The estimates also indicate that being male and residing in a metropolitan area significantly increases the likelihood of involvement in illegal activities. Finally, the estimated parameters did not

support the notion that criminal opportunities and consequently crime participation decisions are affected by unemployment and crime levels after controlling for other factors.

### *III.2 Family Effects and Illegal Earnings*

Table 4 presents results from four variations of the estimated crime income equations for the sub-sample of active criminals. The common set of criminal earnings predictors includes a vector of personal characteristics, measures of general and criminal human capital as well as locational and environmental variables assumed to capture the returns to productivity. To account for duration effects, measures of criminal experience and its square are also included.

The specifications vary according to the correction for unobserved family background effects and sample self-selection. The first column, labeled "Least Squares without Selection", displays coefficient estimates for the simplest of the specifications, ignoring both unobserved family effects and self-selection. The specification labeled "Family Effects without Selection" allows for the presence of unobserved family background effects but ignores the potential impact of self-selection bias.

The third column, labeled "Family Effects with Selection", exhibits parameter estimates for the most general specification. It contains the entire set of variables which were included in the preceding models as well as the sibling's education and crime experience variables, and the estimated value for the conditional expectation ( $\hat{\Lambda}$ ). Incorporating this covariate allows both family-specific and sibling-specific unmeasured effects to be accounted for in the estimation of the criminal earnings. A more restrictive specification excluding the possibly endogenous measures of criminal experience from both stages of the estimation leads to the specification presented in column (4) of Table 4.<sup>3</sup>

---

3 If the disturbance in equation (11) reflects, in part, the unobservable preferences of the individual, including his attitude toward illegal activities, it is possible and actually likely that a variable such as crime experience would also be correlated with the disturbance term and that, as a result, this variable, its square and its family average would not be legitimate exogenous regressors. A variant of the Durbin-Wu-Hausman-White specification test was used to verify this hypothesis. The test generated an asymptotic normal test statistic of 1.42, implying that the null hypothesis of exogeneity of crime experience and its square could not be rejected at any level of significance. The sample used here, however, may not have been large enough to uncover any statistically significant differences in the estimates.

TABLE 4 - DETERMINANTS OF CRIMINAL INCOME (STANDARD ERRORS IN PARENTHESES)

Variables	Model Specification			
	Least Squares without Selection (Pooled OLS)	Family Effects without Selection (GLS: $\sigma_{v\epsilon^2} = 0$ )	Family Effects with Selection (Pooled GLS: $\pi_j = \psi_{z_j}$ )	Family Effects with Selection (Pooled GLS: $\pi_j = \psi_{z_j}$ )
Intercept	2.231**(0.442)	2.210**(0.441)	1.560(0.543)	1.551**(0.564)
<b>Demographic Characteristics</b>				
Age	0.126**(0.036)	0.141**(0.036)	0.134**(0.039)	0.135**(0.041)
Education	0.150**(0.039)	0.128**(0.040)	0.095**(0.046)	0.098**(0.048)
Male	0.346**(0.104)	0.326**(0.103)	0.412*(0.129)	0.426*(0.136)
<b>Crime-Related Variables</b>				
Crime Experience	0.128(0.081)	0.125(0.080)	0.293**(0.091)	a
Crime Experience Square	-0.030**(0.016)	-0.029**(0.016)	-0.035**(0.015)	a
On Probation	0.122(0.088)	0.112(0.087)	-0.211**(0.101)	0.224**(0.106)
<b>NeighborhoodVariables</b>				
Crime In Neighborhood	-0.061(0.177)	-0.058(0.179)	-0.040(0.191)	-0.053(0.120)
Unemployment	-0.088(0.056)	-0.082(0.057)	-0.108*(0.060)	-0.095(0.063)
Residence in SMSAAuxiliary Parameters	0.080(0.103)	0.086(0.105)	0.060(0.112)	0.058(0.118)
Sibling's Education	a	a	0.074**(0.030)	0.072**(0.031)
Sibling's Crime Experience	a	a	-0.099(0.078)	a
Selectivity correction	a	a	-0.310**(0.165)	-0.285**(0.167)
Correlation coefficient	a	0.243**(0.131)	a	a
Adjusted R <sup>2</sup>	0.160	0.161	0.197	0.185

Note: The dependent variable is the log of the individual's criminal income. Significance at the  $p < 0.05$  and  $0.05 = p < 0.10$  levels are indicated by \*\* and \*, respectively. Reported standard errors in the last two columns are corrected for heteroskedasticity (White, 1980) and pre-estimation error. (AMEMYIA, 1978).

<sup>a</sup> Not applicable in this specification.

Virtually all the individual characteristics that are measured have the expected effects on criminal income; these individual characteristics, however, do not explain a large percentage of the variance in log income. Age, education and gender (male) have significant positive effects on criminal income for all

specifications. Particularly with respect to education, the results suggest that the skills and talents proxied by years of schooling, although reducing the likelihood of an individual to participate in crime, as discussed above, do indeed enhance criminal rewards of participants.

Overall, the generalized least square (GLS) specification of column (2) yields parameter estimates quite comparable to those obtained from the ordinary pooled model (column 1). Except for the age variable, failure to account for unobserved family effects leads to an overstatement of the impact of all coefficients in the non-selectivity corrected models. The results also suggest a significant unobserved family background effect for the crime income of siblings. Sharing a common family background explains around 25% of the variance of the logarithm of the unconditional criminal income, as indicated by the estimated value of the correlation coefficient ( $\rho$ ) in this model. In terms of both the size of the estimated family background effect for sibling's criminal incomes and the importance of this effect relative to observed individual characteristics, our results reveal patterns consistent with those reported in studies of legal earnings and wealth. (CORCORAN, *et al.*, 1976; KEARL & POPE, 1986).

While the relative magnitudes and significance of the coefficients are essentially the same for the first two specifications, there are significant differences when the selectivity correction term is introduced in the model (column 3). In particular, the criminal experience and the probation status effects become statistically significant. The estimated effect of the return to "occupation-specific human capital" as measured by criminal tenure is positive and rather large relative to the effects of other variables. Based on the estimated coefficients for crime experience and its square, the results indicate a pronounced concave life-cycle pattern to crime income with the peak occurring around age 18.5.<sup>4</sup> As is well known, the proposition that involvement in crime diminishes with age is a prominent empirical regularity in criminology. (WOLFGANG, 1972). The finding of a concave life-cycle pattern in income suggests that diminishing marginal productivity, which results in

---

4 This result is calculated based on estimates shown in column (3) of Table 4 and assumes a criminal career initiated at age 13, the reported median age of first crime for the criminal subsample.

reduced criminal earnings prospects, may play an important role in determining the timing of the decision to terminate a criminal career.

The auxiliary parameters  $\psi$  which are included to account for the possible correlation between the family-specific effects and the independent regressors, indicate that, in contrast to the positive and significant impact of the sibling's education, the sibling's criminal experience is negatively associated with the youth's illegal earnings although this effect is not statistically significant.

Finally, the parameters associated with the selectivity bias correction terms in columns (3) and (4) are both significant and indicative of the direction of the selectivity bias. Since the estimated conditional expectation means and their respective coefficients are negative, the empirical results point to the existence of a positive selection bias. This implies that the criminal earnings distribution actually observed for participants is higher than the distribution that would be observed for the average individual in the sample had he/she chosen the criminal option, a result consistent with the notion of sibling-specific productivity influencing crime participation decisions.

### CONCLUSION

The influence of family background on economic status has persistently interested social scientists and others concerned with social policy. While previous studies have concentrated on the effects of family background on (legal) earnings and occupational achievement, this paper has considered whether unobserved family effects explain part of the variance in the propensity to engage in income-generating crime and in the intensity and success of criminal participation as measured by criminal earnings. The research design was based on a sample of siblings pairs, drawn from a nationally representative survey of youths, which permitted the decomposition of the cross-siblings variance into "between" and "within" family components.

Unobserved family heterogeneity was shown to be an important and robust determinant of the variance in the decision to participate in crime and the

distribution of criminal rewards among individuals. A substantial 44% to 55% of the total variance in crime participation was attributable to differential family effects, even after controlling for a number of observable individual and family characteristics. Unobservable family effects also accounted for about one-fourth of the variance in criminal income. This finding suggests that even if all variation in observable and unobservable individual characteristics were eliminated, most of the observed inequality in the propensity to participate in crime would remain.

The results also highlighted the importance of controlling for family heterogeneity when estimating the effects of other regressors in the model. In particular, the analysis revealed a significant upward bias in the effect of race on the propensity to engage in criminal activities when family background effects were not properly accounted for. Even ignoring the potentially important role of black underreporting, the probability of (ever) participating in crime differed significantly by race in the pooled model whereas this difference was small and not statistically significant in the panel probit specification. Thus, neglecting family heterogeneity provides a very misleading picture of racial differences in the decision to participate in crime. The estimates also suggested a substantial impact of education on crime participation, with each additional year of schooling reducing the probability of crime participation by 2 to 4 percentage points. Accounting for unobserved family effects in the crime participation equation reduced the estimated marginal effect of education by about one half.

While the results emphasize the importance of unmeasured family factors in the decision to participate in criminal activities, the direction and magnitude of the observed individual characteristics offers some hope of success to programs designed to reduce youth crime. In particular, policy interventions directed toward mediating or supplementing family's investments in children, such as compensatory training programs, may play an important role in the set of choices that result in youth criminality. Ultimately, however, whether inequality attributable to family background can be effectively altered by policy intervention depends on a better understanding of the complex process by which families' endowments, constraints and resource allocation decisions affect the behavior of their members.



## REFERENCES

- AMEMYIA, T. The estimation of a simultaneous-equations generalized probit model. *Econometrica* 46, p. 1193-1205, 1978.
- BECKER, G. Crime and punishment: an economic approach. In: LANDES, W.; BECKER, G. (eds.), *Essays in the economics of crime and punishment*. 1968.
- BECKER, G.; TOMES, N. Child endowments and the quantity and quality of children. *Journal of Political Economy*, v. 84, n. 4, p. S143-S162, 1976.
- BECKER, Gary S. A theory of social interactions. *Journal of Political Economy*, v. 82, n. 6, p. 1063-93, 1974.
- BEHRMAN, J.; TAUBMAN, P. Intergenerational transmission of income and wealth. *American Economic Review* 66, p. 436-440, 1976.
- BEHRMAN, J.; TAUBMAN, P.; WALES, T. Controlling for and measuring the effects of genetics and family environment in equations for schooling and labor market success. In: TAUBMAN, P. (ed.), *Kinometrics*. New-York: North-Holland, 1977.
- BEHRMAN, J.; POLLAK, R.; TAUBMAN, P. Parental preferences and provision for progeny. *Journal of Political Economy*, v. 90, n. 1, p. 52-73, 1986.
- BLOCK, M.; HEINEKE, J. A labor theoretic analysis of the criminal choice. *American Economic Review* 65, p. 314-325, 1975.
- CHAMBERLAIN, G. Panel data. In: GRILICHES, Z.; INTRILIGATOR, M. (eds.), *Handbook of econometrics*, vol. II, 1984.
- CHAMBERLAIN, G.; GRILICHES, Z. Unobservables with a variance component structure: ability, schooling, and the economic success of brothers. *International Economic Review*, v. 16, n. 2, p. 442-449, 1975.
- \_\_\_\_\_. More on brothers. In: TAUBMAN, P. (ed.), *Kinometrics*. 1977.
- CORCORAN, M.; JENCKS, C.; OLNECK, M. The effects of family background on earnings. *American Economic Review* 66, p. 430-435, 1976.
- EHRlich, I. Participation in illegitimate activities: a theoretical and empirical investigation. *Journal of Political Economy* 81, p. 521-567, 1973.
- FARRINGTON, D. Early precursors of frequent offending. In: WILSON, J.Q.; LOURY, G. (eds.), *From children to citizens*. Vol. 3. Springer-Verlag, 1987.

- FREEMAN, R. B. The economics of crime. *In*: ASHENFELTER; CARD, D. (eds.), *Handbook of labor economics*. Vol. 3. Amsterdam, Holland: Elsevier Science, 1999.
- FREEMAN, R. B.; RODGERS, W. M. Area economic conditions and the labor market outcomes of young men in the 1990s expansion. *National Bureau of Economic Research Working Paper 7073*, Cambridge University Press, 1999.
- GLAESER, E. L.; SACERDOTE, B.; SCHEINKMAN, J. A. Crime and social interactions. *Quarterly Journal of Economics* 111, p. 507-548, 1996.
- GOTTFREDSON, M.; HIRSCHI, T. *A general theory of crime*. Stanford University Press, 1990.
- GOULD, E. D.; WEINBERG, B. A.; MUSTARD, D. Crime rates and local labor market opportunities in the United States. *Review of Economics and Statistics*, v. 84, n. 1, p. 45-61, 2002.
- GREENE, W. *Econometric analysis*. New York, NY: Macmillan Publishing Company, 1993.
- GRILICHES, Z. Sibling models and data in economics: beginnings of a survey. *Journal of Political Economy*, v. 87, n. 5, p. S37-S64, 1979.
- GROGGER, J. Arrests, persistent youth joblessness and black/white employment differentials. *Review of Economics and Statistics*, v. 74, n. 1, p. 100-106, 1992.
- HECKMAN, J. Sample selection as a specification error. *Econometrica* 47, p. 153-161, 1979.
- HECKMAN, J.; WILLIS, R. Estimation of a stochastic model of reproduction. *In*: TERLECKYJ, N. (ed.), *Household production and consumption*. Columbia University Press, 1975.
- HINDELANG, M.; HIRSCHI, M.; WEIS, J. *Measuring delinquency*. Beverly Hills: Sage Publications, 1981.
- HIRSCHI, T. Crime and the family. *In*: WILSON, J. Q. (ed.), *Crime and public policy*. Institute for Contemporary Studies, 1983.
- KEARL; POPE. Unobservable family and individual contributions to the distributions of income and wealth. *Journal of Labor Economics*, v. 4, n. 3, p. S48-S79, 1986.
- LAM, D.; SCHOENI, R. Effects of family background on earnings and returns to schooling: evidence from Brazil. *Journal of Political Economy*, v. 101, n. 41, p. 710-740, 1993.
- LAUB, J.; SAMPSON, R. Unraveling families and delinquency. *Criminology* 26, p. 355-380, 1988.

- LEVITT, S. D. Why do increased arrest rates appear to reduce crime: deterrence, incapacitation, or measurement error? *Economic Inquiry*, v. 36, n. 3, p. 353-72, 1998a.
- \_\_\_\_\_. Juvenile crime and punishment. *Journal of Political Economy*, v. 106, n. 6, p. 1156-85, 1998b.
- LOEBER, R.; STOUTHAMER-LOEBER, M. Family factors as correlates and predictors of juvenile conduct problems and delinquency. In: TONRY, M.; MORRIS, N. (eds.), *Crime and justice*. Chicago: University of Chicago Press, 1986.
- MENDONÇA, M. J. C.; LOUREIRO, P. R. A.; SACHSIDA, A. Interação social e crimes violentos: uma análise empírica a partir dos dados do Presídio de Papuda. *Estudos Econômicos*, v. 32, n. 4, p. 621-41, out./dez. 2002.
- MUSTARD, David B. Re-examining criminal behavior: the importance of omitted variable bias. *Review of Economics and Statistics*. v. 85, n. 1, p. 205-211, 2003.
- OLNECK. On the use of sibling data to estimate the effects of family background, cognitive skills and schooling. In: TAUBMAN, P. (ed.), *Kinometrics*. New York: North-Holland, 1977.
- PEZZIN, L. E. Incentivos de mercado e comportamento criminoso: uma análise econômica dinâmica. *Estudos Econômicos*, v. 24, n. 3, p. 373-404, jul./set. 1994.
- \_\_\_\_\_. Estimating panel selection models. *Working Paper*, Agency for Healthcare Research and Quality. Rockville, MD, 2003.
- SAH, R. Social osmosis and patterns of crime. *Journal of Political Economy* 99, p. 1272-1295, 1991.
- SCHEINKMAN, J. A.; WOODFORD, M. Self-organized criticality and economic fluctuations. *American Economic Review* LXXXIV, p. 417-421, 1994.
- SCHMIDT, P.; WITTE, A. D. Predicting criminal recidivism using split population models. *Journal of Econometrics* 40, p. 141-159, 1989.
- SJOQUIST, D. Property crime and economic behavior: some empirical evidence. *American Economic Review* 63, p. 439-446, 1973.
- SOLON, G.; CORCORAN, M.; GORDON, R.; LAREN, D. A longitudinal analysis of sibling correlations in economic status. *Journal of Human Resources*, v. 26, n. 3, p. 509-534, 1991.
- TAUBMAN, P. The determinants of earnings: genetics, family and other environments: a study of white male twins. *American Economic Review*, v. 66, n. 5, p. 858-870, 1976.

- VERBEEK, M.; NIJMAN, T. Testing for selectivity bias in panel data models. *International Economic Review*, v. 33, n. 2, p. 681-703, 1992.
- VISCUSI, K. Market incentives for criminal behavior. In: FREEMAN, R. (ed.), *Inner city black youth employment*. NBER Series. Chicago: University of Chicago Press, 1986.
- WILSON, W. J. *The truly disadvantaged*. Chicago: University of Chicago Press, 1987.
- WITTE, A. D. Estimating the economic model of crime with individual data. *Quarterly Journal of Economics* 94, p. 57-84, 1980.
- WOOLDRIDGE, J. Selection corrections for panel data models under conditional independence assumptions. 1992. Unpublished manuscript.
- WOLFGANG, M.; FIGLIO, R.; SELLIN, T. *Delinquency in a birth cohort*. Chicago: University of Chicago Press, 1972.

---

I am grateful to Shelly Lundberg, Robert Pollak, Raaj Sah, Barbara Schone, Jeffrey Wooldridge, and two anonymous referees for many helpful comments and suggestions. The views expressed in the paper are those of the author. No official endorsement by either the Medical College of Wisconsin or the Health Policy Institute is intended or should be inferred. lpezzin@mcw.edu

(Recebido em maio de 2003. Aceito para publicação em janeiro de 2004)