

An Empirical Analysis of Sexual Violence

Hugo M. Mialon and Sue H. Mialon¹

Abstract

We perform an empirical analysis of the determinants of sexual violence against women and reporting behavior by victims of sexual violence in the United States during the era of the women's movement. We find that neighborhood income inequality and divorce increase the probability of sexual victimization. We then use these variables as part of a set of instruments to correct for victim selection by offenders in estimating the probability of reporting victimization to the police. This allows us to identify factors that influence the probability that a woman in the general population would report if she were sexually victimized. These factors can potentially be controlled by policy-makers to make women in general more credible in their threats to report, which would deter potential sex offenders. (JEL K42 C34 J10)

¹ Department of Economics, Emory University, Atlanta, GA 30322 (e-mail: hmialon@emory.edu and eco_sjh@shsu.edu). We thank Joel Waldfogel for a conversation that stimulated the project, and Dan Hamermesh, Preston McAfee, and Max Stinchcombe for valuable comments.

1. Introduction

Following Becker (1968), Stigler (1970), and Ehrlich (1973), a large economics literature has addressed the impact of sanctions on criminal behavior. However, sanctions, no matter how large, cannot reduce the incidence of crime if victims do not report crimes to the police. The problem is especially severe in the case of rape. Rape victims may suffer from shame and social stigma. And if they report to the police, their cases may proceed to trial, where intimate details of their personal lives may be further publicized. To victims, the pain and humiliation associated with reporting to the police may outweigh any benefits from punishing the offenders. For this reason, rape is the violent crime least reported by its victims (Hart and Rennison, Bureau of Justice Statistics, 2003).

Understanding police notification behavior in rape cases is therefore essential to designing policies to reduce the incidence of rape. The National Crime Victimization Survey (NCVS), ongoing since 1973, has offered researchers the opportunity to study victim reporting empirically. The survey, which is representative of the general population, asks respondents whether or not they have experienced a criminal victimization, and if so, whether or not they informed the police about it. This data has given rise to a large literature in the fields of sociology and criminology on the determinants of police notification (e.g., Block, 1974; Hindelang and Gottfredson, 1976; Skogan, 1977, 1984; Greenberg *et al.*, 1979; Greenberg *et al.*, 1982; Braithwaite and Biles, 1980; Laub, 1981, 1997; Ruback, 1984; Gottfredson and Gottfredson, 1988; Bachman, 1993, 1998).

However, all these studies focus only on the subset of victims to estimate the probability of police notification as a function of various victim and incident characteristics. Therefore, they can only identify the factors that influence whether actual rape victims report to the police. From a policy perspective, it is more interesting to know the factors that influence

whether potential victims in general would report to the police if they were victimized. These are the factors that can be controlled to make threats to report to the police more credible in general, which would deter potential sex offenders who may be counting on their victims to remain silent in order to escape apprehension.

We hypothesize that the factors that influence whether women in general would report to the police if they are sexually victimized are not the same as those that influence whether women who have already been sexually victimized would report. Victims are a selected subset of the population. They have been selected by their attackers to be victimized. Attackers do not choose victims at random.² They may in part select victims on the basis of characteristics that affect the likelihood that victims report to the police. Because existing studies do not take this into account, their estimates of the determinants of reporting by potential victims may be statistically biased.

In this paper, we estimate the determinants of rape victim silence correcting for selection bias. First, we estimate potential victims' probability of being sexually victimized. Second, we estimate victims' probability of remaining silent controlling for their probability of being sexually victimized. This requires identifying variables, which affect the probability of being victimized, but not the probability of reporting the victimization to the police. One of our choices is inspired by the sexual competition theory of rape. The "resource deprivation hypothesis" (Thornhill and Palmer, 2000) predicts that in an environment where females select males based on their ability to provide resources that would facilitate the survival of their potential offspring, males who are strongly out-performed by competing males with

² Sexual offenders choose victims on the basis of their age, for example. Felson and Krohn (1990) find that 92 percent of rape victims from 1973-1982 were 40 or younger. Perkins et al. (1996) and Perkins and Klaus (1996) analyze data from the NCVS of 1993, finding that victimization rates were highest in the age range 17-24 and next highest in the age range 12-16. Greenfield (1997) reviewed data from more than 24 data sets, again finding great overrepresentation of adolescent and young adult victims. The analysis in our paper shows that other characteristics, such as marital and employment status, also influence the probability of being sexually victimized. This suggests that attackers do not choose victims at random.

respect to resource acquisition are generally more likely to rape.

The hypothesis is consistent with data on rapists in the penal system, which indicate that rape is disproportionately committed by males with relatively lower socioeconomic status (Thornhill and Thornhill, 1983). More recently, a study of adolescent male sexual behavior by Figueredo *et al* (1999) found that the majority of adolescents who had committed sexual offenses had a history of failed romantic and sexual relationships, and repeated rejection and sexual frustration. Moreover, many of them exhibited social and psychological deficiencies, including learning disabilities. The authors speculate that these deficiencies may have contributed to their competitive disadvantage in the mainstream sexual marketplace and led them to adopt coercive sexual strategies to overcome the disadvantage.

If competitively disadvantaged males are more prone to rape, then income inequality in a woman's neighborhood could correlate with her probability of being sexually victimized. Greater income inequality correlates with greater competitive disadvantage of certain males relative to others. In an unequal environment, the males with fewer ability and resources would be easily rejected by females in favor of their stronger competition, which may increase their sexual frustration, and therefore their propensity for sexual aggression. Thus income inequality among men in a woman's neighborhood, because it implies a greater competitive disadvantage of certain men relative to others, could increase the woman's probability of being sexually victimized. But it does not affect in an obvious way her decision to report the victimization to the police. Therefore, neighborhood income inequality is potentially an identifying variable.

However, data on neighborhood characteristics are difficult to obtain because the public-use NCVS is not area-identified. There are two ways around this problem. First, one can try to obtain restricted, temporary access to the more recent area-identified data from

the NCVS.³ Second, one can use the public-use files for the first ten years of the survey, 1973-1983, which have been supplemented with information from the U.S. Census regarding neighborhood demographic and income characteristics, including the Gini index that measures income inequality. In this paper, we use the 1973-1983 data because it is publicly available so that our results can be easily replicated. However, it should be clear that the analysis conducted in this paper can be applied to the more recent area-identified data, when it becomes accessible.

Using the 1973-1983 data, we find that neighborhood income inequality increases the probability of sexual victimization, and we use it as part of a set of instruments to correct for victim selection and generate better estimates of the determinants of rape victim silence. We find that the victim's age, family income, and marital status, and her neighborhood's urbanization and female labor force participation rates, each has a statistically significant impact on her reporting behavior. Several of these results contrast sharply with those of previous studies. For example, Bachman (1993, 1998) finds that the victim's age and marital status have no effect on her reporting behavior. Because the previous studies did not correct for victim selection, their results were not fully informative and possibly misleading.

Our paper provides the first estimates of rape victim reporting that are corrected for victim selection. Our paper is also among the first to identify an empirical relationship between income inequality and rape. There is a large literature on the effects of unemployment on crime (e.g., Britt, 1994, and Raphael and Winter-Ember, 2001). Raphael and Winter-Ember find that the relationships between unemployment and violent crime in general, and between male unemployment and rape in particular, are statistically weak. There is also a literature

³ The U.S. Bureau of Justice Statistics does not presently have the resources to add the neighborhood characteristic data to the more recent public-use files. Moreover, at present, researchers cannot even gain restricted, temporary access to the area-identified files because of unresolved data documentation problems. They may eventually be available on a restricted basis when the problems are resolved (private correspondence with Michael Rand, Chief, Victimization Statistics, U.S. Bureau of Justice Statistics).

on the effect of income inequality on crime (e.g., Fajnzylber *et al*, 2002, Doyle *et al*, 1999, Chiu and Madden, 1998). Chiu and Madden demonstrate the effect theoretically. Fajnzylber *et al* find a positive correlation between crime and inequality empirically, both within and between countries. They also find that the correlation reflects causation from inequality to crime. In contrast, Chiu and Madden, find no effect at all. However, the inequality-crime studies have not focused on rape in particular.

The remainder of the paper is organized as follows. The second section describes the data. The third section presents the estimation results with and without correction for victim selection. The final section summarizes the results and discusses policy implications.

2. Data

Data are from the National Crime Victimization Survey (NCVS) for the period 1973-1983, which are supplemented with neighborhood demographic information from the U.S. Census (NCJ-143959). The NCVS is a survey of about 80,000 persons (in 43,000 households), and it is representative of persons age 12 and older residing in U.S. households (Garafalo, 1990). Respondents are asked whether they have experienced a criminal victimization during the previous six months. For each victimization reported, they are asked detailed questions about the offender, the offense, and whether they reported the incident to the police.

We focus only on the subset of women older than 12. The survey variables employed to study rape victim silence are summarized in Table 1. *Silence* is the main dependent variable, and it is coded to 1 if a woman responded that she had suffered a rape or attempted rape but did not inform the police about it, and to 0 if she did inform the police about it. *Occurred in private* is coded to 1 if the sexual attack occurred in a private place, such as the victim's own home or a hotel room, and to 0 if it occurred in a public place, such as a school, commercial

Table 1. Means (SE) of Variables Describing Respondents and Incidents

Variable	All	Rape Victims	Non-Victims
Silence		0.444 (0.497)	
Occurred in private		0.334 (0.472)	
Attempted		0.541 (0.499)	
Medical attention		0.354 (0.478)	
Offender alone		0.811 (0.392)	
Offender age above 18		0.800 (0.401)	
Offender black		0.374 (0.484)	
Offender relative		0.065 (0.246)	
Only victim		0.782 (0.413)	
Age	35.193 (16.622)	24.515 (11.353)	35.609 (16.621)
Married	0.491 (0.500)	0.173 (0.378)	0.506 (0.500)
Black	0.120 (0.325)	0.177 (0.382)	0.118 (0.322)
Education	14.419 (6.177)	13.965 (5.982)	14.495 (6.184)
Family income	9.724 (3.270)	8.128 (3.612)	9.784 (3.250)
Employed	0.521 (0.500)	0.461 (0.499)	0.521 (0.500)
Rural location	0.189 (0.392)	0.146 (0.354)	0.191 (0.393)
Neighborhood flp	0.420 (0.080)	0.429 (0.083)	0.420 (0.080)
Neighborhood Gini	0.326 (0.062)	0.332 (0.062)	0.325 (0.062)
Neighborhood income	10.107 (3.294)	9.709 (3.164)	10.133 (3.310)
Neighborhood poverty	0.101 (0.089)	0.110 (0.085)	0.101 (0.089)
Neighborhood ue	0.047 (0.026)	0.049 (0.027)	0.047 (0.026)
Neighborhood divorce	0.076 (0.054)	0.090 (0.061)	0.075 (0.053)
Total obs.	117,478	571	111,002

Note: "All" includes victims of crimes other than rape

"Non-Victims" are not victims of rape or any crime at all

building, office, or park. *Attempted* is coded to 1 if the sexual attack was attempted, and to 0 if it was completed. *Medical attention* is coded to 1 if the victim was injured by the sexual attack to the extent that she needed immediate medical attention, and to 0 otherwise.

Offender alone is coded to 1 if there was a single offender, and to 0 if there were multiple offenders. *Offender age above 18* is coded to 1 if the victim thought that the (youngest) offender was older than age 18, and to 0 otherwise. *Offender black* is coded to 1 if (at least one) offender was black, and to 0 otherwise. *Offender relative* is coded to 1 if the victim was attacked by a relative, which includes a spouse, ex-spouse, or family member, and to 0 otherwise. *Only victim* is coded to 1 if the victim was the only person victimized by the attack, and to 0 if others were victimized with her. *Age, Married, Black, Education, Family income, Employed* are variables indicating the victim's age, marital status, race, years of education, and employment status, respectively.

In the empirical model, we group victims' ages into category dummies, *Age12-17*, equal to 1 if the victim's age is between 12 and 17, *Age 18-29*, equal to 1 if the age is between 18 and 29, and *Age30-above*, the excluded category. These dummies capture the age of consent and distinguish adolescents from young women, and young women from mature women. Adolescents and young women may be more likely to be raped than mature women, since the probability of rape may depend in part on the victim's attractiveness or fertility. Girls and young women may also suffer greater psychological pain from rape, which may affect the likelihood that they report it to the police.

Five socioeconomic characteristics of each woman's *Neighborhood* are also recorded (neighborhoods are about the size of census tracts, that is, approximately 4,000 persons).⁴ These

⁴ The neighborhood characteristics are 1970 Census "neighborhoods" matched on a household-by-household basis to the NCVS data. The neighborhoods were not designed with reference to maps and therefore rarely correspond with tracts or other units. They were formed by computer using geographic keys associated with 1970 Census housing records. The resulting neighborhoods are contiguous and relatively compact.

are the female labor force participation rate (*flp*), the Gini measure of income inequality (*Gini*), the proportion of families living below the poverty level (*poverty*), the average income (*income*), the unemployment rate (*ue*), and the divorce rate (*divorce*). Values of *Gini* closer to 0 indicate less income inequality, while values closer to 1 indicate more inequality. *Gini* is expected to increase victimization

The summary statistics for each of these variables are presented in three columns. The first is for the entire sample, including rape or attempted rape victims, victims of crimes other than rape, and non-victims. The second column is for the subset of rape or attempted rape victims. Notice that of the respondents who suffered a rape or attempted rape, 44.4 percent did not report the crime to the police. The third column is for the subset of non-victims. The interesting comparison is between rape or attempted rape victims and non-victims because rape may be correlated with other violent crimes. Rape or attempted rape victims are much younger on average (24.5 years) than non-victims (35.6 years). Moreover, a smaller proportion of rape or attempted rape victims are married and employed than non-victims. Average income inequality in the neighborhoods of rape or attempted rape victims is greater than that in neighborhoods of non-victims. The t-statistic testing the null hypothesis of equality of the mean *Gini* indices across the two groups is 2.691, which passes the 1-tailed test at the 5 percent significance level. The summary evidence suggests a positive correlation between income inequality and rape.

3. Model

We want to estimate the effects of individual and social characteristics of respondents on the probability of reporting sexual victimization to the police, correcting for victim selection by offenders. To do this, we employ the probit model with sample selection (van de Ven and

van Praag, 1981), which adapts the standard Heckman selection model (Gronau 1974, Lewis 1974, Heckman 1976) to the case of a binary dependent variable.

Applied to the present context, the standard probit model assumes a latent relationship between an ‘index’ measuring reporting behavior by rape victim i at time t and a vector of explanatory variables X_{it} :

$$Silence_{it}^* = X_{it}\beta + u_{1it} \text{ where } u_{1it} \sim N(0, 1). \quad (1)$$

That this equation expresses a latent relationship implies that we do not observe the actual values of $Silence_{it}^*$. Instead we observe its sign, which will determine the value of the observed binary variable $Silence_{it}$ in our probit equation. In the standard probit model, the probability that $Silence_{it} = 1$ is given by:

$$\Pr(Silence_{it} = 1) = \Pr(Silence_{it}^* > 0) = \Phi(X_{it}\beta). \quad (2)$$

where Φ denotes the normal distribution function. The problem is that $Silence_{it}$ is only observed if i suffered a rape, and the subset of rape victims may be a selected subset of the population: they have been selected by offenders to be victimized, perhaps in part because they are less likely to report. To control for this, we introduce a selection equation:

$$Rape_{it} = Z_{it}\gamma + u_{2it} \text{ where } u_{2it} \sim N(0, 1). \quad (3)$$

where $Rape_{it}$ takes on the value 1 if i is observed in the sample of rape or attempted rape victims, and 0 otherwise. The event is determined by a vector of explanatory variables Z_{it} . The probability that $Rape_{it} = 1$ is:

$$\Pr(Rape_{it} = 1) = \Phi(Z_{it}\gamma). \quad (4)$$

The correlation between u_1 and u_2 is denoted by ρ . The joint probability that a woman is

raped and chooses to be silent is:

$$\Pr(\text{Silence}_{it} = 1, \text{Rape}_{it} = 1) = \Phi_2(X_{it}\beta, Z_{it}\gamma, \rho). \quad (5)$$

where Φ_2 is the cumulative bivariate normal distribution function. The log likelihood is:

$$\begin{aligned} L = & \sum_{\substack{\text{Rape}_{it}=1 \\ \text{Silence}_{it}=1}} \ln \{ \Phi_2(X_{it}\beta, Z_{it}\gamma, \rho) \} \\ & + \sum_{\substack{\text{Rape}_{it}=1 \\ \text{Silence}_{it}=0}} \ln \{ \Phi_2(-X_{it}\beta, Z_{it}\gamma, -\rho) \} + \sum_{\text{Rape}_{it}=0} \ln \{ 1 - \Phi(Z_{it}\gamma) \}. \end{aligned} \quad (6)$$

The Heckman-probit procedure is carried out in three stages. First, the standard probit model (2) is estimated. Second, the selection probit model (4) is estimated. Third, the full probit model with sample selection (5) is estimated. If $\rho \neq 0$, then the sum of the log likelihoods of (2) and (3) will not equal the log likelihood of (5), presented in (6). In this case, just estimating the standard probit model (2) yields biased results, whereas the Heckman-probit procedure provides consistent, asymptotically efficient estimates for all the parameters.

The set Z_{it} of explanatory variables in the selection equation is different from the set X_{it} of explanatory variables in the silence equation, although the two sets have variables in common. First, as we argued previously, *Neighborhood Gini*, the index of income inequality, should increase the probability of being sexually victimized and thus it is included in Z_{it} . However, it is harder to imagine how neighborhood income inequality could affect an individual rape victim's decision to report the crime to the police, and therefore *Gini* is not included in X_{it} .⁵ For similar reasons, most of the other neighborhood characteristics are

⁵ Of course, it is possible to devise stories why neighborhood inequality might affect the propensity to report, but only with a stretch of the imagination. For example, perhaps income inequality is positively correlated with gang formation, and gangs exact revenge for any contact with the police on the part of victims. No instrument is perfectly clean, but a neighborhood characteristic such as income inequality is probably one of the cleanest available for our study.

also included in Z_{it} but not in X_{it} . Lower *Neighborhood income* and greater *Neighborhood poverty* and *ue* could increase victimization, since poor economic conditions could lead to increased frustration, which could in turn lead to more violent crime. *Neighborhood divorce* could also increase sexual victimization. If men are positively selected for marriage and more marriages dissolve, then more of the most attractive men are on the marriage market, so less attractive and previously unmarried men face harder competition, which may in turn increase their frustration and lead them to sexual aggression.

The only neighborhood characteristic that we include in both X_{it} and Z_{it} is *Neighborhood flp*. The labor force participation of women is directly related to women's independence and control of resources, which should reduce women's probability of being victimized. Moreover, the greater the neighborhood female labor force participation rate, the more resources (such as shelters and counseling services for battered women), might be available to rape victims in that neighborhood, which might encourage them to denounce their attackers. Also, during the period 1973-1983, for which we have data, the women's movement was in full swing. One important way the movement manifested itself is through an increase in female labor force participation. The extent of female labor force participation in a neighborhood may therefore be a proxy for the strength of the women's movement in that neighborhood. The stronger is the women's movement, the less social stigma might be attached to women speaking out against sexual violence, and the more resources might be available to them to cope with the victimization. Leaders of the movement, such as Bella Abzug, Betty Friedan, and Gloria Steinem, insisted that violence against women become a political issue; made the public better understand that rape, domestic violence, and sexual harassment are crimes; and obtained public funding for shelters for battered women. In these ways, the women's movement, and the concomitant increase in female labor force participation, may

have encouraged women to speak out against sexual violence.

Rural location is also included in X_{it} as well as in Z_{it} . Because it may be harder to remain anonymous in rural areas, and because these areas may be more traditional in their views of the role of women, victims in rural areas may report crimes to the police at a lesser rate than victims in urban areas. The personal demographic characteristics of respondents are also included in both X_{it} and Z_{it} . A woman's age, for example, may influence both the probability that she is sexually victimized and the probability that she reports the victimization to the police. Sexual predators may have a preference for younger women, since youth correlates with attractiveness and fertility. On the other hand, the consequences of rape, such as pregnancy or social stigma, are potentially more severe for younger women, so that they may suffer greater pain than other women if they are sexually victimized, which may influence the probability that they denounce their attacker. Lastly, offender and incident characteristics, such as whether the rape was attempted or completed, can only be included in X_{it} , because they are only observed for victims.

4. Results

The coefficient estimates and marginal effects of the Heckman-probit model are presented in the first two columns of Table 2 (White-corrected standard errors are in parenthesis). The Wald test rejects the hypothesis of independent equations ($\rho = 0$), indicating that rape and attempted rape victims are in fact selected. Consider first the results for the selection equation in the bottom half of the table. The estimated effect of neighborhood income inequality on the probability of suffering a rape or attempted rape is qualitatively positive and statistically significant (see the first column). Interpreting the marginal effect (in the second column), we find that a 10 percent increase in income inequality in a woman's neighborhood

Table 2. Coefficient Estimates (robust SE) of the Determinants of Rape Victim Silence

Variable	Heckman-Probit	Mfx	Probit	Mfx
Silence				
Occurred in private	-0.0451 (0.0479)	-0.0538	-0.2460 (0.1339)*	-0.0974
Attempted	-0.0246 (0.0326)*	-0.0306	-0.1851 (0.1287)	-0.0736
Medical attention	-0.2097 (0.0410)***	-0.3427	-0.8695 (0.1366)***	-0.3312
Single offender	0.0537 (0.0665)	0.0598	0.2192 (0.2215)	0.0865
Offender age above 18	-0.0392 (0.0677)	-0.0468	-0.1561 (0.2290)	-0.0622
Offender black	-0.0941 (0.0374)**	-0.1124	-0.2477 (0.1430)*	-0.0982
Offender relative	0.0435 (0.0792)	0.0519	0.0973 (0.2318)	0.0388
Only victim	-0.0380 (0.0431)	-0.0478	-0.1412 (0.1533)	-0.0562
Age 12-17	-0.6163 (0.0744)***	-0.0152	0.0129 (0.2080)	0.0051
Age 18-29	-0.3280 (0.0544)***	-0.0556	0.1933 (0.1445)	0.0767
Married	0.2723 (0.0695)***	0.0512	-0.3796 (0.1590)**	-0.1483
Black	0.0009 (0.0637)	0.0465	0.0786 (0.1868)	0.0313
Education	0.0049 (0.0041)	0.0066	0.0188 (0.0112)*	0.0075
Family income	0.0235 (0.0065)***	0.0049	-0.0114 (0.0179)	-0.0046
Employed	0.0633 (0.0511)	0.0388	-0.1173 (0.1290)	-0.0467
Rural location	0.1581 (0.0707)**	0.1527	0.4341 (0.1796)**	0.1710
Neighborhood flp	-0.6520 (0.3312)**	-0.5173	-0.5044 (0.7495)	-0.2009
Rape (selection equation)				
Age 12-17	0.6717 (0.0590)***	0.0144		
Age 18-29	0.4143 (0.0385)***	0.0044		
Married	-0.3466 (0.0433)***	-0.0033		
Black	0.0442 (0.0482)	0.0004		
Education	0.0005 (0.0029)	0.0000		
Family income	-0.0304 (0.0056)***	-0.0003		
Employed	-0.1053 (0.0348)***	-0.0009		
Rural location	-0.0132 (0.0512)	-0.0001		
Neighborhood flp	0.2400 (0.2511)	0.0021		
Neighborhood Gini	0.4535 (0.2142)**	0.0040		
Neighborhood income	0.0013 (0.0048)	0.0000		
Neighborhood poverty	-0.4956 (0.2282)**	-0.0044		
Neighborhood ue	0.0643 (0.6138)	0.0006		
Neighborhood divorce	0.7498 (0.4346)*	0.0066		
Total obs.	103,301		502	
Censored obs.	102,799			
Wald $\chi^2(17)$	241.56			
Wald $\chi^2(1)$ test of $\rho = 0$	4.90			
Pseudo R^2			0.111	

Notes: (1) White-corrected standard errors are in parenthesis, (2) *, **, *** significant at 10, 5, 1 % level, respectively, (3) each equation includes a constant.

increases her probability of being sexually victimized by 0.0004, which represents 8.3 percent of the probability of being sexually victimized. In contrast, neighborhood unemployment and average income do not have a statistically significant effect on the victimization probability, and the poverty rate significantly reduces it. Rape is not necessarily more prevalent in poorer neighborhoods because most males in such neighborhoods are roughly equally poor, so that they are not really at a competitive disadvantage with respect to one another in attracting women. Rather, it is income inequality that correlates with competitive disadvantage, which in turn increases the probability of being sexually victimized.

Like neighborhood income inequality, neighborhood divorce increases the probability of sexual victimization. A higher divorce rate may mean that more men who had previously been selected by women are once again in the mating market, which may imply a greater competitive disadvantage for men who had not previously been selected, and thus a greater probability of sexually victimization for women. A 10 percent increase in the divorce rate increases the probability of sexual victimization by 0.00066 or 13.8 percent. The data also reveal that younger women are significantly (in a substantive and statistical sense) more likely to be sexually victimized than older women. On the other hand, marriage, employment, and family income each reduce the sexual victimization probability, while race and education do not play a significant role.

The results of the probit equation for victim silence are presented in the upper half of Table 2. The first two columns present the estimates and marginal effects with the correction for victim selection by offenders. These are to be compared to the coefficient estimates and marginal effects of the standard probit model without selection correction, presented in the last two columns. The first two columns and the last two columns have several results in common. For example, whether or not we correct for selection, women in rural areas are

more likely to be silent if they are sexually victimized than women in urban areas. Looking at the marginal effect, the silence probability in rural areas is larger by 0.15. Moreover, in both models, the race and education of women do not have a statistically significant effect on their probability of silence given victimization. The insignificance of education with respect to reporting behavior (and the probability of being sexually victimized) suggests that sex education may have been lacking in the U.S. during the 1970s, or that sex educators may have failed to address the problem of rape.

However, the first two columns in Table 2 also exhibit remarkable differences from the last two columns. For example, victim age and family income and neighborhood female labor force participation do not have a statistically significant effect on rape victim silence in the standard probit model, but do have one when the model is corrected for victim selection by offenders. Intuitively, victims were selected by offenders on the basis of certain characteristics, for example, an air of shyness (which may be correlated with observables such as age and employment status), which influence their probability of remaining silent, and which render them less likely to be influenced by other factors that were included in the silence equation, such as neighborhood female labor force participation. Therefore, if we only look at the subset of victims, it appears that neighborhood female labor force participation does not significantly affect the probability of silence. However, when we control for the characteristics that determine the selection process, we find that it does significantly affect the probability of silence for a woman in the *general* population (at the 10 percent level). In particular, any woman in the general population is less likely to be silent if she is victimized the greater is the female labor force participation rate in her neighborhood. Looking at the marginal effect (in the second column), a 10 percent increase in the female labor force participation rate in a woman's neighborhood reduces the probability that she would be

silent if she were sexually victimized by 0.0517 or 8 percent.

Similarly, although younger victims are (statistically) no more or less likely to be silent than older victims, younger women in the general population are (statistically) less likely to be silent if they are victimized than older women in the general population. And while family income does not affect the probability that a victim is silent, it increases the probability that any woman in the general population is silent if she is victimized.⁶ Perhaps the most striking result in Table 2 is that marriage significantly reduces the probability of silence for victims, but significantly increases the probability of silence given victimization for any woman in the general population (the coefficient estimate on *Married* switches sign between the first and third columns, and is statistically significant in both cases). In general, a married woman might prefer to remain silent if she is sexually victimized for fear that her husband would not understand what happened.

However, the sample of rape or attempted rape victims includes cases of marital rape, accounting for 3.5 percent of these cases. Marital rape cases may cloud the results on the effects of marriage on victim silence. Moreover, the victim sample also includes cases of rape by family members, accounting for 3 percent of cases. Cases of rape by relatives may cloud the results in general because our instruments may not be very strong predictors of this type of rape. For example, income inequality might have a relatively smaller effect on the probability of being victimized by a relative than by a non-relative.

To improve the quality of the instruments, we therefore re-estimate the model focusing only on victims who were attacked by non-relatives, and using non-victims as the control group. Thus we estimate the probability of being sexually victimized by a non-relative, then the probability of remaining silent, then the full model with selection by non-relative

⁶ This may reflect the conservative nature of upper-class families in the U.S. during the 1970s. Wealthier, more prestigious families may have wanted to conceal the fact that one of their members had been sexually assaulted for fear of compromising the family's good name.

Table 3. Coefficient Estimates (robust SE) of the Determinants of Rape Victim Silence
Non-relative Rape Selected

Variable	Heckman-Probit	Mfx	Probit	Mfx
Silence				
Occurred in private	-0.0367 (0.0404)	-0.0444	-0.2601 (0.1389)*	-0.1028
Attempted	-0.0554 (0.0296)*	-0.0731	-0.3103 (0.1398)**	-0.1231
Medical attention	-0.2274 (0.0312)***	-0.3929	-0.9391 (0.1462)***	-0.3547
Single offender	0.0447 (0.0709)	0.0511	0.1994 (0.2241)	0.0788
Offender age above 18	-0.0377 (0.0715)	-0.0456	-0.1622 (0.2295)	-0.0646
Offender black	-0.0866 (0.0352)**	-0.1047	-0.1968 (0.1433)	-0.0781
Only victim	0.0064 (0.0380)	0.0077	-0.0430 (0.1590)	-0.0171
Age 12-17	-0.6135 (0.0749)***	-0.0093	0.0670 (0.2166)	0.0267
Age 18-29	-0.3234 (0.0542)***	-0.0505	0.2133 (0.1522)	0.0846
Married	0.2953 (0.0568)***	0.0391	-0.2738 (0.1677)	-0.1077
Black	0.0334 (0.0658)	0.0610	0.1340 (0.1934)	0.0534
Education	0.0052 (0.0043)	0.0081	0.0208 (0.0117)*	0.0083
Family income	0.0214 (0.0065)***	0.0072	-0.0226 (0.0186)	-0.0090
Employed	0.0696 (0.0495)	0.0338	-0.1099 (0.1343)	-0.0437
Rural location	0.1811 (0.0730)**	0.1559	0.3647 (0.1961)*	0.1443
Neighborhood flp	-0.8806 (0.3165)***	-0.6989	-0.9865 (0.8002)	-0.3928
Non-relative Rape (selection equation)				
Age 12-17	0.6736 (0.0603)***	0.0135		
Age 18-29	0.4028 (0.0394)***	0.0039		
Married	-0.3595 (0.0451)***	-0.0032		
Black	0.0224 (0.0496)	0.0002		
Education	0.0016 (0.0030)	0.0000		
Family income	-0.0299 (0.0059)***	-0.0003		
Employed	-0.1070 (0.0359)***	-0.0009		
Rural location	-0.0371 (0.0525)	-0.0003		
Neighborhood flp	0.3313 (0.2428)	0.0027		
Neighborhood Gini	0.5212 (0.2148)**	0.0042		
Neighborhood income	0.0026 (0.0052)	0.0000		
Neighborhood poverty	-0.4526 (0.2357)**	-0.0037		
Neighborhood ue	0.2484 (0.5245)	0.0020		
Neighborhood divorce	0.7636 (0.3439)**	0.0062		
Total obs.	103,267		468	
Censored obs.	102,799			
Wald $\chi^2(16)$	237.83			
Wald $\chi^2(1)$ test of $\rho = 0$	16.14			
Pseudo R^2			0.112	

Notes: (1) White-corrected standard errors are in parenthesis, (2) *, **, *** significant at 10, 5, 1 % level, respectively, (3) each equation includes a constant.

offenders. The results are presented in Table 3. The Wald test rejects the hypothesis of no selection ($\rho = 0$) with even greater certainty than it did when we allowed all rapes to be selected. The effect of income inequality on the probability of being sexually victimized by a non-relative is greater than its effects on the probability of being sexually victimized in general. A 10 percent increase in income inequality in a woman's neighborhood increases the probability that she is sexually victimized by a non-relative by 0.00052, which represents 11.5 percent of the probability of being sexually victimized by a non-relative.⁷ Moreover, the effect of divorce on the probability of sexual victimization by a non-relative is of equal magnitude but more statistically significant than its effect on the probability of victimization in general (it is now significant at the 5 percent level).

By focusing on victimization by non-relative offenders, the neighborhood instruments are stronger, which improves the estimates of the determinants of rape victim silence. With this improvement, we find that neighborhood female labor force participation reduces the probability of silence, and the effect is now statistically significant at the 1 percent level (whereas it was significant at the 5 percent level before). Looking at the marginal effect, a 10 percent increase in the female labor force participation rate in a woman's neighborhood reduces the probability that she would be silent if she were sexually victimized by 0.0699 or 10.9 percent. Moreover, the silence probability is smaller if the woman is younger or if the offender is black, but greater if the woman's family is wealthier. And being married increases the probability that a woman would be silent if she were sexually victimized by 0.04. None of these factors would seem to have any effect on the silence probability if we did not correct for victim selection by offenders. Comparing the first two columns to the last two columns, it is clear that the selection correction has a substantial effect on the results.

⁷ The Gini measure of income inequality in the U.S. rose from 0.38 in 1968 to 0.47 in 2001, a rise of almost 10 percentage points. Therefore, the result suggests that if the U.S. returned to the level of inequality it had in the late sixties, rape by non-relatives would be reduced by 11.5 percent.

Ignoring victim selection by offenders masks the importance of many factors that turn out to be instrumental in determining whether women speak out against sexual violence.

5. Conclusion

We have found that neighborhood income inequality and divorce statistically increase rape victimization, suggesting that socioeconomic policies that promote economic equality and marital stability could reduce sexual victimization. We have used these factors as instruments to correct for victim selection by offenders in estimating the determinants of women's reporting behavior if they are sexually victimized. Because previous studies did not correct for victim selection, they were only informative about the factors that influence whether women who have already been sexually victimized report to the police. Such factors are necessarily few, because victims are selected by offenders on the basis of characteristics that make them less likely to report to the police. In contrast, because we corrected for victim selection, our results are informative about the factors that influence the probability that any woman in the general population would report to the police if she were sexually victimized. We find a host of previously undiscovered factors that influence reporting behavior in the general population.

For example, married women are less likely to report to the police if they are sexually victimized, suggesting that policies to encourage women to speak out against sexual violence should be geared toward their husbands too; older women are less likely to report to the police, suggesting that the policies should not be geared exclusively toward younger women; and women in neighborhoods with low rates of female labor force participation are less likely to report to the police, suggesting that the policies should be geared toward women in these areas. In addition, the analysis confirmed that women in rural areas are less likely to report,

and that education has little effect on women's reporting behavior, suggesting that sex education should be expanded to include education about rape and its social consequences. By encouraging women in general to speak out if they are sexually victimized, such policies could reduce the probability of sexual victimization for each of them. These policies could potentially make all women in the general population more credible in their threats to report to the police if they are sexually victimized. And this might go a long way toward deterring sex offenders, as they may be counting on the silence of their victims in order to avoid apprehension and the severe penalties that society reserves for their crime.

REFERENCES

- Becker, G.S. (1968): "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, Vol. 76, No. 2, 169-217
- Bachman, R. (1993): "Predicting the Reporting of Rape Victimization," *Criminal Justice and Behavior*, Vol. 20, No. 3, 254-270
- Bachman, R. (1998): "The factors related to rape reporting behavior and arrest: New evidence from the National Crime Victimization Survey," *Criminal Justice and Behavior*, Vol. 25, No. 1, 8-29
- Britt, C.L. (1994): "Crime and Unemployment among Youths in the United States, 1958-1990: A Time Series Analysis," *American Journal of Economics and Sociology*, Vol. 53, No. 1, 99-109
- Block, R. (1974): "Why notify the police: The victim's decision to notify the police of an assault," *Criminology*, Vol. 11, 555-569
- Braithwaite, J. and Biles, D. (1980): "Empirical verification and Black's The Behavior of Law," *American Sociological Review*, Vol. 45, 334-338
- Chiu, W.H. and Madden, P. (1998): "Burglary and Income Inequality," *Journal of Public Economics*, Vol. 69, No. 1, 123-141
- Doyle, J.M., Ahmed, E., and Horn, R.N. (1999): "The Effects of Labor Markets and Income Inequality on Crime: Evidence from Panel Data," *Southern Economic Journal*, Vol. 65, No. 4, 717-138
- Ehrlich, I. (1973): "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation," *Journal of Political Economy*, Vol. 81, No. 3, 521-65
- Fajnzylber, P., Lederman, D., and Loayza, N. (2002): "Income Inequality and Violent Crime," *Journal of Law and Economics*, Vol. 45, No. 1, 1-40

- Felson, R. and Krohn, M. (1990): ‘Motives for Rape,’ *Journal of Research in Crime and Delinquency*, Vol. 27, 222-242.
- Figueredo, A. J., Sales, B. D., Becker, J. V., Russell, K., and Kaplan, M. (2000): “A Brunswikian Evolutionary-developmental Theory of Adolescent Sex Offending,” *Behavioral Sciences and the Law*, Vol. 18, 309-329
- Garofalo, J. (1990): “The national crime survey, 1973-1986: Strengths and limitations of a very large data set,” in Doris L. MacKensie, Phyllis J. Baunach, and Roy R. Roberg (eds.), *Measuring Crime: Large-Scale, Long-Range Efforts*, Albany: State University of New York Press
- Gottfredson, M.R. and Gottfredson, D.M. (1988): *Decision Making in Criminal Justice*, Second Edition, New York: Plenum
- Greenberg, M.S., Wilson, C.E., Ruback, R.B., and Mills, M.K. (1979): “Social and emotional determinants of victim crime reporting,” *Social Psychology Quarterly*, Vol. 42: 364-372
- Greenberg, M.S., Ruback, R.B., and Wescott, D.R. (1982): “Decision making by crime victims: A multimethod approach,” *Law and Society Review*, Vol. 17, 47-84
- Greenfield, L. (1997): *Sex Offenses and Sex Offenders: An Analysis of Data on Rape and Sexual Assault*, Bureau of Justice Statistics, US Department of Justice.
- Gronau, R. (1974): “Wage Comparisons: A Selectivity Bias,” *Journal of Political Economy*, Vol. 82, 1119-1155
- Hart, T. and Remison, C. (2003): *Reporting Crime to the Police, 1992-2000*, Washington, D.C.: Bureau of Justice Statistics, U.S. Department of Justice.
- Hindeland, M. and Gottfredson, M. (1976): “The criminal justice system: Public attitudes and involvement,” In William McDonald (ed.), *Criminal Justice and the Victim*, Berverly Hill, Calif.: Sage
- Heckman, J.J. (1976): “The common structure of statistical models with truncation, sample selection, and limited dependent variables and a simple estimator for such models,” *The Annals of Economic and Social Measurement*, Vol. 5, 475-492
- Laub, J. (1981): “Ecological Considerations in Victim Reporting to the Police,” *Journal of Criminal Justice*, Vol. 9, No. 1, 419-430
- Laub, J. (1997): “Patterns of criminal victimization in the United States,” In Robert C. Davis, Arthur J. Lurigio, and Wesley G. Skogan, (eds.) *Victims of Crime*, Thousand Oaks, Calif.: Sage
- Lewis, H. (1974): “Comments on Selectivity Biases in Wage Comparisons,” *Journal of Political Economy*, Vol. 82, 1119-1155
- Perkins, C. and Klaus, P. (1996): *Criminal Victimization 1994*. National Criminal Victimization Survey. Bulletin, Bureau of Justice Statistics, US Department of Justice.

- Perkins, C., Klaus, P., Bastian, L., and Cohen, R. (1996): *Criminal Victimization in the United States, 1993*. National Criminal Victimization Survey Report, Bureau of Justice Statistics, US Department of Justice.
- Raphael, S. and Winter-Ember, R. (2001): "Identifying the Effect of Unemployment on Crime," *Journal of Law and Economics*, Vol. 44, No. 1, 259-283
- Ruback, R.B. (1984): "Social influence and crime-victim decision making," *Journal of Social Issues*, Vol. 40, 51-76
- Stigler, G. (1970): "The Optimum Enforcement of Laws," *Journal of Political Economy*, Vol. 78, No. 3, 526-536
- Skogan, W.G. (1977): "Dimensions of the dark figure of unreported crime," *Crime and Delinquency*, January, 41-50
- Skogan, W.G. (1984): "Reporting crimes to the police: The status of world research," *Journal of Research in Crime and Delinquency*, Vol. 21, 113-1376
- Thornhill, R. and Palmer, C.T. (2000): *A Natural History of Rape: Biological Bases of Sexual Coercion*, Cambridge: The MIT Press.
- Thornhill, R. and Thornhill, N.W (1983): "Human Rape: An Evolutionary Analysis," *Ethology and Sociobiology*, Vol. 4, 137-173
- Van de Ven, W. P. M. M and van Praag, B. M. S. (1981): "The demand for deductibles in private health insurance: A probit model with sample selection," *Journal of Econometrics*, Vol. 17, 229-252