Assessing the Effect of Adolescent Employment on Involvement in Criminal Activity
Robert Brame, Shawn D. Bushway, Raymond Paternoster and Robert Apel
Journal of Contemporary Criminal Justice 2004 20: 236
DOI: 10.1177/1043986204266891

The online version of this article can be found at:
http://ccj.sagepub.com/content/20/3/236

Published by:
http://www.sagepublications.com

Additional services and information for Journal of Contemporary Criminal Justice can be found at:

Email Alerts: http://ccj.sagepub.com/cgi/alerts
Subscriptions: http://ccj.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.com/journalsPermissions.nav
Citations: http://ccj.sagepub.com/content/20/3/236.refs.html

>> Version of Record - Aug 1, 2004

What is This?
Assessing the Effect of Adolescent Employment on Involvement in Criminal Activity

ROBERT BRAME  
University of South Carolina

SHAWN D. BUSHWAY  
RAYMOND PATERNOSTER  
University of Maryland

ROBERT APEL  
University of South Carolina

This article considers the problem of estimating the effect of a binary independent variable (employment) on a binary outcome variable (involvement in criminal activity) for a nationally representative sample of adolescents (ages 15-18). The authors' bivariate analysis confirms a common finding from the literature, that adolescent employment is associated with increased risk of involvement in criminal activity. They then turn to the problem of assessing whether this association is sensitive to plausible assumptions about the impact of other variables (both observed and unobserved) on both employment and crime. This assessment reveals that both the sign and magnitude of the maximum likelihood estimate of the employment effect are quite sensitive to these assumptions. Based on this evidence, they conclude that future efforts to understand the adolescent work-crime relationship will benefit from resolving the ambiguities identified by their analysis.

Keywords: propensity scores; sensitivity analysis; observational studies; work and crime

1. INTRODUCTION

Over the course of the past two decades, a number of researchers have carefully documented the relationship between adolescent employment and involvement in a variety of different kinds of problem behaviors—including criminal behavior. Although there are many important differences among studies of the linkages between adolescent employment and criminal behav-
ior, most of them suggest that adolescent employment is associated with increased risk of involvement in criminal behavior (Agnew, 1986, pp. 28-29; Bachman & Schulenberg, 1993, p. 228; Gottfredson, 1985, p. 428; Ploeger, 1997, pp. 664-669; Shannon, 1991, pp. 21, 63; Steinberg, Fegley, & Dornbusch, 1993, p. 176; Wright, Cullen, & Williams, 1997, p. 213; but see Shannon, 1982).

The uniformity of this finding is interesting, because criminological theory provides plausible reasons for predicting that employment will have both crime-reducing and crime-increasing (criminogenic) effects. For example, social control theory (see, e.g., Agnew, 1986, p. 21; Hirschi, 1969) predicts that—at a minimum—involvement in a job will create space and time constraints that reduce the physical opportunity to offend. Control theories also suggest that jobs can facilitate the development of strong attachments to conventional institutions, accelerated maturity, and heightened sense of personal responsibility, which should also lead to greater self-imposed constraints against offending (Thornberry & Christenson, 1984, p. 400). Indeed, plausible explanations for a negative relationship between adolescent employment and crime can be found in most contemporary criminological theories (Bachman & Schulenberg, 1993, pp. 220-221; Thornberry & Christenson, 1984, pp. 400-401).

It is also true, however, that plausible bases can be found within these same theories for expecting a criminogenic effect of adolescent employment (Agnew, 1986, pp. 22-23). For example, social control theorists argue that attachment to school and family will create barriers to involvement in criminal activity. If these attachments do indeed reduce involvement in criminal activity, and adolescent employment interferes with them, it would be quite reasonable to anticipate that adolescent employment leads to increased risk of involvement in criminal activity. In sum, there are a number of sensible reasons to predict that adolescent employment will either increase or decrease the risk of involvement in criminal activity (see, e.g., Ploeger, 1997, p. 661; Shannon, 1991, p. 21; Wright et al., 1997, pp. 203-207).

The idea of constraining employment opportunities for youngsters is a relatively recent development within Western societies. In fact, it was not until the Progressive Era (late 19th and early 20th centuries) that child labor laws became commonplace in the United States (Coleman, 1974; Kett, 1977). Currently, in most states youngsters are not permitted to hold jobs that generate a paycheck until they reach the age of 14 or 15 years. It seems clear, then, that studies identifying potentially harmful effects of adolescent employment within current age constraints raise an important policy issue. Specifically, for our purposes, if employment increases the risk of involvement in criminal behavior, it might well be worthwhile to impose additional restric-
tions on paycheck-generating job opportunities for young people (Agnew, 1986, p. 20; Wright et al., 1997, p. 207). Whether this would be a sensible policy shift, however, depends on the reason the association between adolescent employment and criminal involvement exists. Unfortunately, this is a difficult question to answer (Ploeger, 1997, p. 660; Steinberg et al., 1993, pp. 171-172).

There are a number of reasons for this difficulty. First, there is significant ambiguity about how both the independent and dependent variables should be measured. As for the independent variable—adolescent employment—some researchers interviewed respondents about employment in a “regular paying part-time job” (i.e., Steinberg et al., 1993, p. 173), whereas others interviewed the youngsters’ parents and asked the parents about “the approximate number of hours of paid work in which their children participated during the previous week” (Wright et al., 1997, p. 209), and still others interviewed respondents and asked them whether “they had worked in the community for pay, excluding allowance, during the past year” (Ploeger, 1997, p. 663). The operational definition of work, therefore, varies somewhat from study to study. In similar fashion, the dependent variable—involvement in criminal behavior—has been defined in different ways in different studies. Some studies have examined behaviors such as carrying a weapon, theft, and vandalism (Steinberg et al., 1993, p. 174), whereas other studies have used broader measures such as Agnew’s (1986) 26-item Total Delinquency Scale that “measures the respondent’s involvement in a wide range of status, property, and violent offenses” (p. 26). Still other studies have constructed scales that comprise a mixture of criminal and noncriminal behaviors (Wright et al., 1997, p. 208) as well as measures that invoke only official contact with the police (Bachman & Schulenberg, 1993, p. 227). Yet, despite these considerable measurement differences, the literature suggests that work-crime relationships are quite robust. Thus, measurement issues cannot be the only source of ambiguity.

A second basis for the difficulty in drawing conclusions about the source of the relationship between adolescent employment and criminal activity lies in our inability to say whether the relationship is causal. The critical issue is that studies of the relationship between adolescent employment and criminal behavior are observational rather than experimental in nature. Consequently, it is possible that certain individual characteristics influence variation in both adolescent employment and criminal behavior. If these characteristics are not held constant, then we have the familiar problem of a selection effect due to omitted variables. All contemporary studies of the work-crime relationship among adolescents have devoted significant attention to this issue. The most typical approach is to estimate the causal effect of work on crime while con-
trolling for observed variables that are thought to be potential confounders. In most of the studies of the relationship between adolescent employment and involvement in criminal activity, researchers have found that the imposition of controls for observed confounders reduces the estimated causal effect of employment. The degree to which confounding variables can account for all of the relationship between work and crime, however, remains unclear. Some research suggests that the effect of employment on crime persists (albeit usually a weaker effect) after adjusting for confounding influences, whereas other research suggests that the effect vanishes or weakens dramatically after such adjustments. A key residual concern in such analyses, however, is whether the observed confounders are the only confounders. For example, researchers will often regress a current measure of criminal behavior on prior employment while holding demographic/social background characteristics and/or individual differences in prior criminal behavior constant (Agnew, 1986, pp. 27-28; Bachman & Schulenberg, 1993, p. 233; Gottfredson, 1985, pp. 423-425; Ploeger, 1997, pp. 670-672; Steinberg et al., 1993, pp. 173-175; Wright et al., 1997, pp. 211-213). In sum, most scientists who have carefully studied the issue agree that there is a fairly robust positive bivariate association between adolescent employment and involvement in crime, but there is no consensus about whether this relationship is causal.

A third basis for the difficulty in understanding the relationship between adolescent employment and crime exists even if the association between the two variables is at least partially causal. The concern—initially raised by Thornberry and Christenson (1984) in their longitudinal study of young adults in Philadelphia—is that there may be reciprocal causal effects between employment and criminal involvement. In their Philadelphia analysis, Thornberry and Christenson (1984, p. 404, 409) reported that crime and unemployment both exert positive effects on each other, creating a so-called vicious cycle effect. They obtained this result from a three-wave linear panel model that used lagged measures of unemployment and criminal behavior as instrumental variables for contemporary measures of unemployment and criminal behavior (estimated under the assumption that serial correlation in both measures vanishes). To our knowledge, researchers studying the effects of adolescent employment on crime have not attempted to estimate such models. Although there are many plausible reasons to believe that employment and crime could causally influence each other among adolescents, there are some significant methodological challenges awaiting researchers in this area. For this reason, we believe that this issue is unlikely to be resolved in the near term.

In this article, we use data from the 1997 National Longitudinal Survey of Youth (NLSY) (Center for Human Resource Research, 2002) to examine the
relationship between adolescent employment and involvement in criminal behavior. A key theme of our analysis is the quantification of our uncertainty about the basis for a valid estimate of the effect of adolescent employment on crime. There are several sources of this uncertainty and, after defining them precisely, we will examine the sensitivity of our estimate of the “work effect” to different sets of plausible assumptions that can be made about them. We close with a discussion of what we believe should be high priorities for future research in this area.

2. DATA AND MEASURES

As noted above, the data for this analysis come from the 1997 NLSY. We are primarily interested in obtaining an estimate of the causal effect of “paycheck work” during the 12 months preceding the administration of the survey on self-reported involvement in a number of criminal behaviors measured by the NLSY during the same time period. Our definition of work excludes freelance work and includes only work for which a paycheck was regularly earned. To measure variation on this variable, we interrogated the NLSY event history database for the year preceding the interview for each individual who was 15 years old or older at the time of the interview (the minimum age where work is allowed is 14). This produced a database with $N = 4,211$ individuals. For each of these individuals a code of 1 was assigned if any paycheck work was reported during the preceding year (41.6% unweighted) and a 0 was assigned if no such work was reported during the preceding year (58.4%, unweighted). We were unable to obtain the necessary employment information to appropriately classify 5 individuals; therefore, our sample size at this stage of the analysis is $n = 4,206$.

Our measure of criminal behavior is coded 1 if the respondent self-reported involvement in any of the following six offenses during the year preceding the interview: (a) property destruction, (b) auto theft, (c) theft of items exceeding $50 in value, (d) other property crimes (e.g., buying and/or selling stolen goods), (e) aggravated assault, and (f) selling drugs (both marijuana and so-called hard drugs). This classification rule results in a total of 1,151 offenders (27.37%, unweighted) and 3,055 nonoffenders (72.63%, unweighted) from the total sample size of 4,206.

Finally, we also included measures of age, race or ethnicity, and sex. Another 38 cases were omitted at this stage of the analysis because of missing information on race or ethnicity. This left us with a final working sample size of $n = 4,168$. A descriptive summary of all the variables with both their weighted and unweighted distributions is presented in Table 1. A discussion of the sampling weights for the NLSY is presented in the appendix.
3. ANALYSIS RESULTS

We now turn to our analysis of the work-crime relationship for the NLSY adolescents. The results of our analysis will be presented in three sections. The first section describes a simple assessment of the joint frequency distribution of adolescent employment and criminal involvement. The second section assesses the relationship after conditioning on a propensity score as defined by Rosenbaum and Rubin (1983b, pp. 51-54). In our case, the propensity score is given by the logistic regression-based estimate of the probability of working conditional on the demographic covariates: (a) sex, (b) race or ethnicity, and (c) age at interview date. Finally, in the third section, we follow the methods described by Rosenbaum and Rubin (1983a) to examine the sensitivity of the maximum likelihood estimate of the causal effect of work on crime to different sets of plausible assumptions about potential unobserved confounding influences. It is important to note that the second and third sections are not independent. Specifically, the sensitivity analysis does not trump the maximum likelihood estimate of the work effect conditional on the propensity score; instead, the sensitivity analysis is itself conditioned on the propensity score.

### Table 1

**Marginal Percentages for Analysis Variables (n = 4,168)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Unweighted Analysis (%)</th>
<th>Weighted Analysis (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>51.0</td>
<td>51.4</td>
</tr>
<tr>
<td>Female</td>
<td>49.0</td>
<td>48.6</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>57.7</td>
<td>72.2</td>
</tr>
<tr>
<td>Black</td>
<td>27.4</td>
<td>15.9</td>
</tr>
<tr>
<td>Other</td>
<td>14.9</td>
<td>11.9</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>44.4</td>
<td>43.0</td>
</tr>
<tr>
<td>16</td>
<td>40.5</td>
<td>41.3</td>
</tr>
<tr>
<td>17</td>
<td>14.5</td>
<td>15.1</td>
</tr>
<tr>
<td>18</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Worked in past year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>58.2</td>
<td>54.0</td>
</tr>
<tr>
<td>Yes</td>
<td>41.8</td>
<td>46.0</td>
</tr>
<tr>
<td>Offended in past year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>72.6</td>
<td>71.0</td>
</tr>
<tr>
<td>Yes</td>
<td>27.4</td>
<td>29.0</td>
</tr>
</tbody>
</table>
3.1 Simple Bivariate Analysis

We begin our investigation by assessing the joint frequency distribution of employment and involvement in criminal behavior. Table 2 presents this joint frequency distribution. Our first task is to obtain a maximum likelihood estimate of the unconditional work effect. We adopt the counterfactual notation of Manski (1995, p. 40) and define the variable \( y_{i1} \) as the outcome that person \( i \) would experience conditional on working and \( y_{i0} \) as the outcome that person \( i \) would experience conditional on not working. It is important to note that \( y_{i1} \) is only observed for individuals who work; it is missing data for individuals who do not work. Similarly, \( y_{i0} \) is only observed among the subset of individuals who do not work; it is missing data for individuals who work. Now we let \( z_i = 1 \) denote the event that individual \( i \) actually works, and we let \( z_i = 0 \) denote the event that individual \( i \) does not work. Following Manski (1995, p. 42), and letting \( p(\cdot) \) denote probability, we state the assumption that

\[
p(y_{i1} = 1) = p(y_{i1} = 1 | z = 1) = p(y_{i1} = 1 | z = 0)
\]

and

\[
p(y_{i0} = 1) = p(y_{i0} = 1 | z = 1) = p(y_{i0} = 1 | z = 0),
\]

which is a formal way of saying that the treatment an individual receives does not depend on that individual’s eventual outcome set \( \{y_{i1}, y_{i0}\} \). Following the notation of Rosenbaum and Rubin (1983a, p. 213), this assumption implies that we can estimate the average causal effect of work on crime (which we call \( \theta \)) by

\[
\theta = E(y_{i1}) - E(y_{i0}) = p(y_i = 1) - p(y_i = 0) = \tau_1 - \tau_0 \tag{1}
\]

where \( E(.) \) denotes expectation. Although this assumption is not at all plausible for our analysis of the effect of work on crime, it forms a useful starting point.

### TABLE 2

**Joint Distribution of Adolescent Employment and Involvement in Criminal Activity (Weighted)**

<table>
<thead>
<tr>
<th>Self-Reported Criminal Activity in Year Before Interview</th>
<th>Employment in Year Before Interview</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not Employed</td>
</tr>
<tr>
<td>No self-reported criminal acts</td>
<td>1,634.00</td>
</tr>
<tr>
<td>At least one self-reported criminal act</td>
<td>617.05</td>
</tr>
<tr>
<td>Total</td>
<td>2,251.07</td>
</tr>
</tbody>
</table>

\( p(\text{Offend} | \text{Work Status}) \) = 0.2741 \( p < .05 \).
point. We emphasize that Equation 1 is a central focus of this article and that it represents the difference between two conditional probabilities—the probability of offending given employment and the probability of offending given nonemployment. The main source of variation in this analysis revolves around different assumptions about the exogeneity of adolescent employment. Based on the information in Table 2, our maximum likelihood estimate of \( \theta \) is given by

\[
\hat{\theta} = \hat{\tau}_1 - \hat{\tau}_0 = 0.3081 - 0.2741 = 0.0340,
\]

where \( \hat{\tau}_1 = \frac{590.54}{1916.93} = 0.3081 \) and \( \hat{\tau}_0 = \frac{617.05}{2251.07} = 0.2741 \). This result implies that adolescent employment is associated with about a 10% increased risk of involvement in criminal behavior. This result replicates the sign of the estimated work effect found in much of the scientific literature on the relationship between adolescent employment and crime.3

### 3.2 Analysis Conditioning on Propensity Score

We now turn to the task of estimating the work effect after taking our demographic covariates into account. Here, we relax the assumption of the previous section to allow for conditional independence of treatment and outcomes instead of requiring complete independence of treatment and outcomes. Toward this end, we define a new variable, \( s_i \), that denotes membership in a population stratum that relates observed potential confounding variables to variation in the independent variable, \( z_i \) (i.e., whether individual \( i \) works or not). The key requirement that this stratifying variable will satisfy is that assignment to \( z_i \) is random conditional on the stratum, \( s_i \) (assuming no unobserved confounding).

A key issue here is how to create a stratification variable that has the necessary property of conditional random assignment (again, assuming no unobserved confounding). A stratification variable that meets this requirement is a version of what Rosenbaum and Rubin (1983b, pp. 51-52) call a propensity score. In our case, the estimated propensity score, \( \hat{\lambda}(d; \hat{\eta}) \), for individual \( i \) is equal to the estimated conditional probability of working given a vector of observed demographic characteristics including indicator variables for sex, race or ethnicity, and age. We use the logistic functional form to write the estimated probability of working by

\[
\hat{\lambda}(d; \hat{\eta}) = p(z_i = 1 | (d; \hat{\eta}) = \frac{\exp (d' \hat{\eta})}{1 + \exp (d' \hat{\eta})},
\]
where \( \hat{\eta} \) is an \( 18 \times 1 \) vector of maximum likelihood parameter estimates and \( d \) is a \( 1 \times 18 \) vector comprising a cell with a 1 followed by 17 indicator variables representing all of the demographic characteristics and all possible interactions of those characteristics for each individual \( i \). The log-likelihood function for this problem is given by

\[
\log_e(L; \eta) = \log \left( \sum_{i=1}^{N} \left( z_i \times \frac{\exp(d_i \cdot \hat{\eta})}{1 + \exp(d_i \cdot \hat{\eta})} + (1 - z_i) \times \frac{1}{1 + \exp(d_i \cdot \hat{\eta})} \right) \right)
\]

where \( \log_e(\cdot) \) denotes the logarithm to the base e (i.e., the natural logarithm).

Following the example of Rosenbaum and Rubin (1983a, pp. 214-215, 1983b, pp. 51-52), we created five strata that represent clusters of individuals with similar values of the propensity score, \( \hat{\lambda}(d; \hat{\eta}) \). These strata were created by ranking the propensity score within the sample and then dividing the sample into five approximately equal-sized quintile groups. According to Rosenbaum and Rubin (1983b), “subclassification with five sub-classes is sufficient to remove at least 90% of the bias for many continuous distributions” (p. 52). Table 3 presents the distribution of the propensity score, \( \hat{\lambda}(d; \hat{\eta}) \), and displays the corresponding placement of observations into their respective strata.

As Rosenbaum and Rubin (1983a, p. 213) note, this stratification scheme ideally corresponds to the well-understood case of a randomized block experiment, such that assignment to values of \( z_i \) is random conditional on knowledge of the stratifying variable, \( s_i \). In the terminology of Manski (1995, pp. 41-42), if this condition holds, then the treatment (i.e., whether one works or not) is independent of the outcome (i.e., whether one offends or not) conditional on \( s_i \) (i.e., the propensity score subclass), but if we do not condition on \( s_i \), then the treatment and outcome are not independent and Equation 1 will be a biased estimator of \( \theta \). Formally, we assume that

\[
p(y_1 = 1 | s) = p(y_1 = 1 | s, z = 1) = p(y_1 = 1 | s, z = 0)
\]

and

\[
p(y_0 = 1 | s) = p(y_0 = 1 | s, z = 1) = p(y_0 = 1 | s, z = 0),
\]

and the log-likelihood function is given by

\[
\log_e(L; \beta, \gamma, \phi) = \sum_{i=1}^{N} w_i \times \log \left( \sum_{t} \left[ p(y_{it} | t = z_i, s_i; \beta) \times p(z_i | s_i; \gamma) \times p(s_i; \phi) \right] \right).
\]

where \( \phi_j = p(s_i = j) \), which obeys the constraint that \( \sum_{j=1}^{k} \phi_j = 1.0 \), and, using the logistic functional form, the conditional distribution of \( z_i \) (i.e., whether one
works or not) given \( s_i \) (i.e., indicator for propensity score subclass membership) is given by

\[
p(z_i = 1 \mid s_i) = \frac{\exp(\gamma_j)}{1 + \exp(\gamma_j)},
\]

and, similarly, the conditional distribution of \( y_{it} \) given \( s_i \) is obtained by

\[
p(y_{it} = 1 \mid t = z_i, s_i) = \frac{\exp(\beta_j)}{1 + \exp(\beta_j)}.
\]

We maximize this log-likelihood function using numerical methods (Tanner, 1996, pp. 26-27). Next, our maximum likelihood estimate of \( \gamma \), given estimates of \( \beta \), \( \gamma \), and \( \phi \), can be calculated by

\[
\hat{\gamma}_j = \sum_{j=1}^{k} \phi \left[ \frac{\exp(\beta_j)}{1 + \exp(\beta_j)} \right].
\]

**TABLE 3**

<table>
<thead>
<tr>
<th>Propensity Score</th>
<th>Unweighted Analysis</th>
<th>Weighted Analysis</th>
<th>Stratum ((s_i))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\lambda}(d, \eta) )</td>
<td>Frequency</td>
<td>%</td>
<td>Frequency</td>
</tr>
<tr>
<td>0.174</td>
<td>144</td>
<td>3.5</td>
<td>108.193</td>
</tr>
<tr>
<td>0.202</td>
<td>211</td>
<td>5.1</td>
<td>131.407</td>
</tr>
<tr>
<td>0.226</td>
<td>270</td>
<td>6.5</td>
<td>148.536</td>
</tr>
<tr>
<td>0.253</td>
<td>143</td>
<td>3.4</td>
<td>113.471</td>
</tr>
<tr>
<td>0.299</td>
<td>546</td>
<td>13.1</td>
<td>642.895</td>
</tr>
<tr>
<td>0.327</td>
<td>259</td>
<td>6.2</td>
<td>139.340</td>
</tr>
<tr>
<td>0.353</td>
<td>130</td>
<td>3.1</td>
<td>103.251</td>
</tr>
<tr>
<td>0.394</td>
<td>221</td>
<td>5.3</td>
<td>143.123</td>
</tr>
<tr>
<td>0.406</td>
<td>538</td>
<td>12.9</td>
<td>649.060</td>
</tr>
<tr>
<td>0.406</td>
<td>106</td>
<td>2.5</td>
<td>87.419</td>
</tr>
<tr>
<td>0.446</td>
<td>101</td>
<td>2.4</td>
<td>51.318</td>
</tr>
<tr>
<td>0.484</td>
<td>78</td>
<td>1.9</td>
<td>50.411</td>
</tr>
<tr>
<td>0.538</td>
<td>58</td>
<td>1.4</td>
<td>48.934</td>
</tr>
<tr>
<td>0.550</td>
<td>42</td>
<td>1.0</td>
<td>33.456</td>
</tr>
<tr>
<td>0.566</td>
<td>458</td>
<td>11.0</td>
<td>586.586</td>
</tr>
<tr>
<td>0.615</td>
<td>516</td>
<td>12.4</td>
<td>662.672</td>
</tr>
<tr>
<td>0.715</td>
<td>175</td>
<td>4.2</td>
<td>246.064</td>
</tr>
<tr>
<td>0.778</td>
<td>172</td>
<td>4.1</td>
<td>221.865</td>
</tr>
<tr>
<td>Total</td>
<td>4,168</td>
<td>100.0</td>
<td>4,168</td>
</tr>
</tbody>
</table>
for \( t = (0,1) \) and, as in Equation 1 above, our maximum likelihood estimator for the work effect is simply
\[
\hat{\theta} = \hat{\tau}_1 - \hat{\tau}_0.
\]

Applying this method to the data from Tables 2 and 3, we obtain maximum likelihood estimates of \( \hat{\tau}_1 = 0.3016 \) and \( \hat{\tau}_0 = 0.2732 \). This implies that the maximum likelihood estimate of \( \theta \) is equal to \( \hat{\theta} = 0.0284 \), which represents a 16.5% reduction from our original unconditional estimate of \( \theta \) from the previous section. Nevertheless, even after conditioning on the propensity score, our estimate of the effect of work on crime continues to be positive.

The primary focus of our effort in this section has been to develop an inference about the work effect in the population after conditioning on observed confounders. Next, we turn to an assessment of the sensitivity of our conclusion to different sets of assumptions about unobserved confounding.

3.3 Sensitivity Analysis Net of Propensity Score

In this final analysis section, we address the question of whether the inference developed in Section 3.2 is sensitive to different sets of plausible assumptions that could be adopted about the influence of unobserved confounders. To some extent, the boundaries of any sensitivity analysis will be based on arbitrary cutoff points. This notwithstanding, we believe that if our substantive conclusions depend heavily on small to moderate variations in the sensitivity parameters, we should avoid the development of inferences about population parameters that depend on those parameters taking particular values.

We begin the analysis by defining an unobserved binary covariate called \( u_i \) (Rosenbaum and Rubin 1983a, p.213). It is useful, for our purposes, to view \( u_i \) as a crime-producing trait that is coded 1 if individual \( i \) possesses the trait and 0 otherwise. This crime-producing trait is a fiction that is used exclusively for the purpose of proxying for unobserved characteristics of individuals that produce criminal behavior. The key issue that we address in our sensitivity analysis is whether our conclusions about the maximum likelihood estimate of \( \theta \) depend heavily on assumptions that we make about \( u_i \).

Indeed, the validity of the analyses presented in Sections 3.1 and 3.2 are already contingent on particular assumptions about \( u_i \); namely, those analyses assume that the probability of work is independent of \( u_i \). We now relax this constraint and define several key parameters that govern the joint distribution of criminal activity, work, the propensity score strata defined in Section 3.2, and \( u_i \). Following Rosenbaum and Rubin (1983a, p. 216), the joint distribution of these variables is given by Bayes’ rule,
\[ p(y_t, z_i, u, s) = \prod_{i=1}^{N} [p(y_i \mid u, s_i) \times p(z_i \mid u, s_i) \times p(u_i \mid s_i) \times p(s_i)], \]

where the parameterization for \( s_i \) (the likelihood of membership in each specific subclass, \( j \), is given by

\[ p(s_i = j) = \phi_j \]

for \( j = 1, 2, \ldots, k \), and once again, we obey the constraint that \( \sum_{j=1}^{k} \phi_j = 1.0 \).

Next, the parameterization for \( z_i \) (i.e., the probability that individual \( i \) is not employed) is

\[ 4p(z_i = 0 \mid s_i) = \frac{\pi_j}{1 + \exp(\gamma_j)} + \frac{1 - \pi_j}{1 + \exp(\gamma_j + \alpha_j)}, \]

where \( \pi_j = p(u = 0 \mid s) \) such that conditional on the event \( s_i = j \) we have \( \pi_j + (1 - \pi) = 1.0 \). We then use the law of total probability to average over \( u_i \), and we therefore write the distribution of the outcome \( y_i \) (whether an individual offends) by

\[ p(y_i = 0 \mid z_i = t, s) = p(y_i = 0 \mid s, u_i = 0) \times p(u_i = 0 \mid z_i = t, s) + p(y_i = 0 \mid s, u_i = 1) \times p(u_i = 1 \mid z_i = t, s), \]

which, as Rosenbaum and Rubin (1983a, p. 216) note, can be written more compactly as

\[ p(y_i = 0 \mid z = t, s) = \frac{w_{sz}}{1 + \exp(\beta_n)} + \frac{1 - w_{sz}}{1 + \exp(\beta_n + \delta_n)}, \]

where \( w_{sz} \) is given by

\[ w_{sz} = p(u = 0 \mid z, s = j) = \pi_j \times \left( \frac{\exp(z \alpha_g) \times [1 + \exp(\gamma_j)]^{1 - \pi_j}}{1 + \exp(\gamma_j + \alpha_g)} \right)^{-1}, \]

where \( z \in \{0, 1\} \). Since \( u_i \) is unobserved, it turns out that all of the parameters in these probability mass functions are not identified. However, if we fix the values of \( \pi, \alpha, \) and \( \delta \), the parameter vectors for \( \beta \) and \( \gamma \) are conditionally identified, and so the maximum likelihood estimates of \( \beta \) and \( \gamma \) are conditionally valid. The estimates of \( \phi_j \) for \( j = 1, 2, \ldots, k \) are unconditionally valid because they do not depend on the unobservables.

Within this parameterization, \( \alpha_g \) represents the effect of \( u_i \) (the unobserved crime trait) on whether an individual works during the year preceding the
interview. Specifically, \( \exp(\alpha_j) \) indexes the change in the odds of working conditional on \( u_i \) such that

\[
\exp(\alpha_j) = \frac{\text{odds ratio } (z \mid u,s = j)}{\text{odds ratio } (z \mid u = 0,s = j)} = \frac{p(z = 1 \mid u = 1,s = j) / p(z = 0 \mid u = 0,s = j)}{p(z = 1 \mid u = 0,s = j) / p(z = 0 \mid u = 0,s = j)}.
\]

We impose the constraint that \( \alpha_j \) is fixed for all \( j = 1, 2, \ldots, k \) propensity score subclasses. In other words, our sensitivity analysis will assume that the relationship between the unobserved crime trait and employment during the past year is the same for all subclasses of the propensity score. We will, however, assess the sensitivity of the maximum likelihood estimate of \( \theta \) to different fixed values of \( \exp(\alpha) \). Specifically, we will allow this odds ratio to vary from \( \frac{1}{3} \) to 3.0. By so doing, we will be able to see whether \( \hat{\theta} \) depends heavily on different plausible assumptions that we might make about the relationship between the unobserved crime trait and employment.

In addition, \( \delta_{st} \) represents the effect of the unobserved crime trait \( u_i \) on the outcome variable \( y \) (criminal offending), conditional on \( t \). It also has an odds ratio interpretation. For individuals who worked during the year preceding the interview, the odds ratio is

\[
\exp(\delta_{s1}) = \frac{\text{odds ratio } (y_1 \mid u,z = 1)}{\text{odds ratio } (y_1 \mid u,z = 0)} = \frac{p(y_1 = 1 \mid u = 1,z = 1) / p(y_1 = 0 \mid u = 1,z = 1)}{p(y_1 = 1 \mid u = 0,z = 1) / p(y_1 = 0 \mid u = 0,z = 1)}
\]

and for individuals who did not work during the year before the interview, the odds ratio is

\[
\exp(\delta_{s0}) = \frac{\text{odds ratio } (y_0 \mid u,z = 0)}{\text{odds ratio } (y_0 \mid u,z = 0)} = \frac{p(y_0 = 1 \mid u = 1,z = 0) / p(y_0 = 0 \mid u = 1,z = 0)}{p(y_0 = 1 \mid u = 0,z = 0) / p(y_0 = 0 \mid u = 0,z = 0)}
\]

For purposes of this analysis, we will impose the constraint that the effect of the unobserved crime trait, \( u_i \), on offending activity is the same regardless of whether one worked. We will also assume that the effect of \( u_i \) on offending activity is the same within each propensity score subclass. Our sensitivity analysis will examine the extent to which the maximum likelihood estimate of \( \theta \) depends on different plausible assumptions about the fixed value of \( \delta \). Here, we will impose the constraint that \( \exp(\delta) > 1.0 \) (i.e., that the unobserved crime trait can only increase the odds of offending activity). We assess the impact of varying \( \exp(\delta) \) up to 3.0. At its maximum, then, we examine the effect of allowing the unobserved crime trait to triple the odds of engaging in criminal behavior.
The final parameter of interest in our sensitivity analysis is \( \pi_j \). Recall that this term indexes the conditional probability distribution of the unobserved crime trait given membership in a specific subclass, \( j \), of the propensity score distribution. Here, we impose two constraints on our sensitivity analysis. First, we assume that \( \pi_j \) is the same for all propensity score subclasses. This means that our notation simplifies from \( \pi_j \) to \( \pi \). Second, we assume that no more than 50% of the population possesses the unobserved crime trait. Our sensitivity analysis will allow \( p \) to vary from .9 (which implies that 10% of the population possesses the trait) to .5 (which implies that half of the population possesses the trait). The rationale for this restriction is that involvement in criminal behavior is relatively (but not exceptionally) rare.

With these parameters in hand, our estimate of \( \tau_t \) (i.e., the conditional probability of offending given treatment \( t \) working or not working) is given by

\[
\tau_t = \sum_{j=1}^{J} \phi_j (1 - \pi_j) \times \left[ \frac{\exp(\beta_j + \delta)}{1 + \exp(\beta_j + \delta)} \times \pi_j \times \frac{\exp(\beta_j)}{1 + \exp(\beta_j)} \right]
\]

for \( t = (0,1) \) and, as before, our maximum likelihood estimate of \( \theta \) is given by

\[
\hat{\theta} = \hat{\tau}_1 - \hat{\tau}_0
\]

where \( \tau_t \) denotes the conditional probability of offending given no employment within the previous year and \( \tau_t \) denotes the conditional probability of offending given employment within the previous year. In short, if \( \theta > 0 \), we conclude that work increases the risk of crime, whereas if \( \theta < 0 \), we conclude that work reduces the risk of crime.

Table 4 presents our maximum likelihood estimates of \( \theta \) conditional on several different sets of assumptions about the three different sensitivity parameters described above. The majority of the results in this table (22 out of 24) suggest that adolescent employment is criminogenic—a result that is itself consistent with much of the extant literature. In some instances, our estimates of \( \theta \) actually become substantially larger than what we observed in the previous sections. In other instances, however, our estimates of \( \theta \) are negative. Because all of the sensitivity parameters in Table 4 are plausible, it follows that our analysis is not capable of identifying the sign of the employment effect. With this in mind, we now turn to a more detailed discussion of the results. We have two main observations to make.

First, the analysis indicates that the work effect is highly sensitive to assumptions that we make about the relationship between the unobserved crime trait and the likelihood of employment. For example, if possession of the crime trait decreases the odds of working by a factor of 1.5 (i.e., \( \exp(\alpha = \ldots) \).
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\exp(\alpha) = \frac{\text{odds}</td>
<td>u = 1}{\text{odds}</td>
<td>u = 0} = 1/3$</td>
</tr>
<tr>
<td>$\exp(\delta) = \frac{\text{odds}</td>
<td>u = 1}{\text{odds}</td>
<td>u = 0} = 3.0$</td>
</tr>
<tr>
<td></td>
<td>$\tau_1$</td>
<td>0.306</td>
</tr>
<tr>
<td></td>
<td>$\theta$</td>
<td>0.036</td>
</tr>
<tr>
<td>$\exp(\alpha) = \frac{\text{odds}</td>
<td>u = 1}{\text{odds}</td>
<td>u = 0} = 2/3$</td>
</tr>
<tr>
<td></td>
<td>$\tau_1$</td>
<td>0.303</td>
</tr>
<tr>
<td></td>
<td>$\theta$</td>
<td>0.032</td>
</tr>
<tr>
<td>$\exp(\alpha) = \frac{\text{odds}</td>
<td>u = 1}{\text{odds}</td>
<td>u = 0} = 1.5$</td>
</tr>
<tr>
<td></td>
<td>$\tau_1$</td>
<td>0.300</td>
</tr>
<tr>
<td></td>
<td>$\theta$</td>
<td>0.025</td>
</tr>
<tr>
<td>$\exp(\delta) = \frac{\text{odds}</td>
<td>u = 1}{\text{odds}</td>
<td>u = 0} = 3.0$</td>
</tr>
<tr>
<td></td>
<td>$\tau_1$</td>
<td>0.297</td>
</tr>
<tr>
<td></td>
<td>$\theta$</td>
<td>0.019</td>
</tr>
</tbody>
</table>
\[
\begin{align*}
\exp(\alpha) &= \frac{\text{odds} \mid u = 1}{\text{odds} \mid u = 0} = 3.0 \\
\exp(\delta) &= \frac{\text{odds} \mid u = 1}{\text{odds} \mid u = 0} = 1.5 \\
\exp(\delta) &= \frac{\text{odds} \mid u = 1}{\text{odds} \mid u = 0} = 3.0 \\
\end{align*}
\]

<table>
<thead>
<tr>
<th>( \tau_0 )</th>
<th>( \tau_1 )</th>
<th>( \theta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.277</td>
<td>0.297</td>
<td>0.020</td>
</tr>
<tr>
<td>0.282</td>
<td>0.291</td>
<td>0.010</td>
</tr>
<tr>
<td>0.283</td>
<td>0.290</td>
<td>0.007</td>
</tr>
<tr>
<td>0.284</td>
<td>0.297</td>
<td>0.300</td>
</tr>
<tr>
<td>0.288</td>
<td>0.274</td>
<td>0.273</td>
</tr>
<tr>
<td>0.005</td>
<td>-0.023</td>
<td>-0.027</td>
</tr>
</tbody>
</table>
2/3), then the positive work-crime relationship estimated in Section 3.2 becomes even stronger. On the other hand, if possession of the unobserved crime trait triples the odds of employment (i.e., \( \exp(\alpha = 3.0) \)), then it is plausible to obtain negative estimates of the effect of work on crime.

Second, the conclusions seem to hinge—to a lesser extent—on the strength of the association between the unobserved crime trait and involvement in crime and the prevalence of that trait in the population. For example, if the trait substantially increases the odds of work and the odds of offending, then the effect of work becomes very close to zero or even negative (depending on the prevalence of the trait). In addition, if the trait triples the odds of work and the odds of offending, then the effect of work is negative and about the same magnitude as our original estimates reported in the previous section as long as at least 30% of the population possesses the unobserved crime trait.

In sum, the results of this analysis suggest that a variety of different conclusions about the strength of the association between employment and crime are consistent with the data. Based on what we have seen with this exploratory analysis, the effect of work on crime could be positive, negative, or even substantively unimportant. Our conclusion about which of these processes is most plausible depends, therefore, on assumptions that are not testable without additional data that will convincingly proxy for the unobserved crime trait or that will provide exogenous variation in adolescent employment.

4. DISCUSSION AND CONCLUSIONS

In this article, we have produced three sets of analyses pertaining to the adolescent work-crime relationship. Our first analysis suggests—consistent with most of the scientific literature—that adolescent employment is associated with increased risk of involvement in criminal activity. The estimated effect from this analysis was \( \hat{\theta} = 0.0340 \). Our second analysis indicates that this effect remains positive but is reduced by about 16.5% when we condition on a propensity score formed exclusively from basic demographic characteristics (i.e., age, sex, and race).

Our third and final analysis examines the robustness of this conclusion to variation in the assumptions used to identify the employment effect. This analysis reveals that the maximum likelihood estimate of the effect of employment is highly sensitive to plausible variation in assumptions about the relationship between unobserved sources of criminal propensity and the decision to work. Under some sets of assumptions, our analysis suggests that employment reduces the risk of involvement in criminal activity, whereas other sets of assumptions would lead us to the conclusion that adolescent employment is criminogenic. To make progress, then, it is necessary to obtain additional information about this relationship.
In our view, there are at least two approaches that could prove useful for narrowing the range of uncertainty about the employment effect. The first approach would involve the collection of a more extensive set of covariates that might be used to convincingly proxy for the unobserved crime trait than what was used in this analysis. This would be particularly useful if there is significant variation in the unobserved crime trait across the adolescent population. The clear difficulty here lies in the establishment of a domain of consensus about the appropriate set of covariates to form this type of propensity score. Nevertheless, we believe that researchers should continue their efforts to control for plausible sources of spurious association.

The second approach would involve identifying variables that causally influence the probability of employment but have no direct causal effect on the probability of involvement in criminal activity. For example, if regulations governing minimum work age vary from place to place, it might be possible to argue that such regulations have an influence on the probability distribution of adolescent employment but no direct influence on involvement in criminal activity. A recent paper by Tyler (2003) exploits this type of variation to study the effects of work on crime. We are currently exploring whether this type of place-to-place variation in minimum work age regulations exists in the NLSY. If it does, it may be possible to develop more convincing estimates of the effect of adolescent employment on involvement in criminal behavior. We believe that efforts to identify other variables that may act as convincing instruments for adolescent employment should be a high priority for researchers in this area.

In sum, most researchers agree that there is a positive empirical association between adolescent employment and involvement in criminal behavior. What is not yet well understood is why this association exists. Our analysis of the adolescent employment-crime relationship using data from the 1997 NLSY provides additional evidence in support of this basic consensus but also suggests the need for careful interpretation of that relationship.

NOTES

1. Another related concern is that it is not always clear whether the independent and dependent variables are temporally linked in an optimal way. For example, even though the measure of employment used by Wright, Cullen, and Williams (1997, pp. 208-209) refers to employment during the previous week, some of the questions about criminal activity (i.e., whether the child had ever been in trouble with the police) make no restrictions on the time frame in which the variable is free to vary. We agree with Gottfredson (1985, p. 424) that it is difficult to say what the optimal link would be, but studies clearly vary in their treatment of this issue.
2. The studies by Gottfredson (1985), Shannon (1982), Bachman and Schulenberg (1993), and Ploeger (1997) are the most prominent contributions to this latter group. Gottfredson (1985) found that the positive effect of work on crime vanished after she controlled for (a) drug use and interpersonal aggression (females only), (b) parental attachment (males only), and (c) involvement in extracurricular activities (both males and females). Shannon (1982) found inconsistent (some positive and some negative) work effects both before and after controls for other individual differences were considered. Bachman and Schulenberg’s (1993, p. 232) longitudinal analysis of the Youth in Transition data suggested that most negative correlates of work predate rather than postdate initiation of employment. Finally, Ploeger (1997, p. 671) found that preexisting differences between individuals alone failed to account for the work-crime relationship but imposition of a control for delinquent peer exposure along with these differences explained almost all of the relationship.

3. A likelihood ratio test of the null hypothesis that $\tau_1 = \tau_0$ produces a test statistic of 5.792, which is statistically significant when referred to a $\chi^2$ table with a single degree of freedom. Because this test does not take the sampling weights into account, however, it can only be interpreted as an approximation.

4. Because there were only a very small number of individuals at age 18 years at the time of data collection (weighted frequency was 22.638, or 0.5% of the sample), we combined them with the 17 year olds for purposes of constructing the propensity score.

**APPENDIX**

The National Longitudinal Survey of Youth (NLSY) was administered according to a complex sampling design. Although the details of this design are beyond the scope of this article, it is important to note that all individuals in the target population (adolescents in the United States) did not have an equal probability of being selected for the NLSY. In general, then, different individuals carry different weights in our analysis. The NLSY sample weight is given by

$$w_i = \frac{1}{\text{Pr}(\text{selected for NLSY sample})}.$$  

We apply the following normalizing transformation to the NLSY weight to obtain

$$w_i^* = \frac{w_i}{\sum_{i=1}^{N} w_i} = \frac{w_i}{\bar{w}},$$
which ensures that the sum of the sampling weights (i.e., \( \sum_{i=1}^{N} w_i^* \)) is equal to the size of the sample (\( N \)) (Lee, Forthofer, & Lorimor, 1986, p. 73). Point estimation with these weights is straightforward, and we therefore use the weights above for all of the analyses reported in this article after Tables 1 and 2. Unfortunately, variance estimation for estimators such as those used in this article is a nonstandard problem and not easily implemented. For purposes of this article we will, therefore, devote our attention primarily to the point estimates. Weighted average point estimates for all of our variables are presented in the right-hand panel of Table 1.

REFERENCES


Robert Brame is an associate professor in the Department of Criminology and Criminal Justice at the University of South Carolina. His research interests include juvenile delinquency, domestic violence, and capital punishment.

Shawn D. Bushway is an assistant professor of criminology in the Department of Criminology and Criminal Justice at the University of Maryland and a fellow with the National Consortium on Violence Research. He received his Ph.D. in public policy analysis and political economy in 1996 from the H. John Heinz III School of Public Policy and Management at Carnegie Mellon University. His current research focuses on understanding the process of desistance and disparity in sentencing outcomes.

Raymond Paternoster is a professor in the Department of Criminology and Criminal Justice at the University of Maryland. His research interests include theories of criminal offending, capital punishment, and quantitative methods.

Robert Apel is an assistant professor in the Department of Criminology and Criminal Justice at the University of South Carolina. His current research interests include the relationship between employment and offending, patterns in victimization, and applied econometrics. His recent publications appear in Criminology and Social Forces.