Justice Quarterly

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rjqy20

For Whom Do Sanctions Deter and Label?
Robert G. Morris & Alex R. Piquero
Published online: 24 Nov 2011.

To cite this article: Robert G. Morris & Alex R. Piquero (2011): For Whom Do Sanctions Deter and Label?, Justice Quarterly, DOI:10.1080/07418825.2011.633543
To link to this article: http://dx.doi.org/10.1080/07418825.2011.633543

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
For Whom Do Sanctions Deter and Label?

Robert G. Morris and Alex R. Piquero

Deterrence and labeling theories make opposing predictions regarding the effect of sanctions on subsequent crime. Deterrence anticipates that sanctions deter, while labeling anticipates that sanctions amplify future crime. The knowledge base with respect to this question is vast, and while a handful of studies provide evidence of a deterrent effect, the majority of studies indicate a null effect. Our study examines whether an arrest leads to an increase in subsequent crime, but extends the knowledge base by considering whether an arrest has the same effect across offender trajectories and by employing techniques that deal with sample selection bias. Thus, we assess for whom sanctions deter or exacerbate subsequent offending. Results indicate that for greater risk youth, arrest amplifies subsequent delinquency, net of other effects, but not among lower risk youth. Thus, experiencing an arrest aggravates subsequent delinquency among some but not all persons. Implications and directions for future research are identified.

Keywords deterrence; labeling; arrest; trajectories; propensity score matching

Robert G. Morris is an Assistant Professor of Criminology in the School of Economic, Political and Policy Sciences at the University of Texas at Dallas. His research interests include topics surrounding technology and crime, recidivism, corrections, white-collar crime, and quantitative methods. His work has published in journals such as the Journal of Quantitative Criminology, Crime and Delinquency, Criminal Justice and Behavior, and the Journal of Criminal Justice. He is a recent recipient of the prestigious University of Texas System Regents’ Outstanding Teaching Award. Alex R. Piquero is Ashbel Smith Professor of Criminology in the School of Economic, Political and Policy Sciences at the University of Texas at Dallas. He is also Adjunct Professor Key Centre for Ethics, Law, Justice, and Governance, Griffith University and Co-Editor of the Journal of Quantitative Criminology. His research interests include criminal careers, criminological theory, and quantitative research methods. He has received several research, teaching, and mentoring awards, and has been elected Fellow of both the American Society of Criminology and the Academy of Criminal Justice Sciences. Correspondence to: Robert G. Morris, School of Economic, Political and Policy Sciences at the University of Texas at Dallas, 800 West Campbell Road, GR 31, Richardson, TX 75080-3021, USA. E-mail: morris@utdallas.edu
Introduction

There may be no two theories that make clearly opposing predictions about the direction of the effect of arrest on subsequent criminal activity than deterrence and labeling theory (Ward & Tittle, 1993). The line-in-the-sand between the theories concerns the sign of the arrest effect. For deterrence theory, an arrest is believed to produce a specific deterrence effect such that an arrest reduces subsequent offending in general and may also indirectly inhibit future crime by increasing an individual’s perception of the certainty and severity of punishment. On the other hand, labeling theory hypothesizes a deviance amplification effect, where an arrest/sanction experience is expected to increase subsequent offending via several mechanisms such as: signaling to others that the offender has committed a criminal offense and is somehow a dangerous person, or that the arrest/sanction somehow leads to the development of a deviant self-image that results in the individual enveloping an offender persona and profile. Both of these labeling mechanisms would serve to shutter prosocial opportunities and open antisocial ones. Given the striking predictions made by the two theories, it is not surprising that a healthy knowledge base has emerged.

Recently, criminologists have vocalized the importance of not only bettering our understanding of whether sanctions mitigate or aggravate subsequent behavior, but have also started to ask when and especially for whom do sanctions deter or label (Piquero, Paternoster, Pogarsky, & Loughran, 2011). Such questions are of scientific and practical import as the field unpacks the nuances of official sanctions on future behavior. In this paper, we extend this line of work in two important ways. First, we examine the effect of arrest on subsequent criminal activity across distinct offending trajectories to assess whether an arrest experience operates in the same manner across divergent offender groups. Second, we employ a counterfactual approach that allows us to consider the true causal effect of arrest on subsequent criminal activity after matching individuals on key characteristics that helps to buttress sample selection bias concerns. Such methods have not been utilized in prior research using nationally representative data. Before we present the results of our analyses, we provide a brief overview of extant research findings regarding the effect of arrest on subsequent criminal activity and then indicate how our study extends this line of work.

The Effect of Arrest on Subsequent Crime

As noted above, the effect of arrest on subsequent crime can be attributed to deterrence or labeling theory. According to specific deterrence, arrest and the imposition of sanctions should deter the future criminality of apprehended individuals by serving to increase the perceived risk of arrest, while according to general deterrence the influence of arrests/sanctions should serve as a
Deterrence and Labeling

There is an extensive empirical literature on the effect of arrest/sanctions on subsequent offending from both deterrence (Blumstein, Cohen, & Nagin, 1978; Nagin, 1998; Pratt, Cullen, Blevins, Daigle, & Madensen, 2006) and labeling (Matsueda, 1992; Smith & Paternoster, 1990) perspectives. For present purposes, it would suffice to provide a few key highlights. Huizinga and Henry (2008) undertook a comprehensive review of longitudinal research on the effect of arrest and justice system sanctions on subsequent behavior. The range of studies is wide, with some investigating subsequent offending of individuals who were arrested relative to those who were not arrested, while others compared various forms of sanctioning following arrest. There are also a few experiments focusing on the effects of arrest—most notably the domestic violence experiment studies (Sherman, 1992). Although there is some evidence that sanctions provide a deterrent effect, this effect is variable emerging in some but not all studies and varying according to sample composition and individual characteristics (Piquero et al., 2011). The most consistent finding to emerge from this large literature is one of a non-significant (deterrent) effect with a few studies indicating a harmful (crime-exacerbating) effect of arrest (e.g. Farrington, 1977).

Before outlining the varied findings within this literature, it is important to review one particularly noteworthy study by Smith and Paternoster (1990), who examined whether referral of a case to juvenile court (relative to diverting it) affected future delinquent/criminal activity among a sample of Florida juvenile offenders. A key focus of their investigation was their observation that a potentially higher rate of subsequent crime among those referred to court may not reflect a labeling-hypothesized deviance amplification effect per se, but instead reflect a powerful selection artifact. That is, those youth referred to juvenile court may have more of the factors that are related to future offending compared to those individuals who are diverted. In short, any differences in

1. The above deterrence hypothesis is consistent with the theory's classic view. Recent empirical analyses assessing deterrence have identified a resetting effect, which attempts to explain the counter-intuitive (at least to classic depictions of deterrence) finding that an arrest may increase subsequent offending. This account invokes a judgment and decision-making bias known as the "gambler's fallacy," in which punished offenders reset their sanction certainty estimate, perhaps believing they would have to be very unlucky to be caught again (Pogarsky & Piquero, 2003, p. 95).
subsequent offending may be due to preexisting differences in criminal propensity. Their analyses revealed three key results. First, preliminary findings showed that youth recommended by intake personnel for formal juvenile court processing were significantly more delinquent during the follow-up period which at first glance supports the labeling perspective. Second, a bivariate probit model predicting referral to juvenile court and recidivism showed that a recommendation for referral to juvenile court was not significantly associated with recidivism, indicating that a selection artifact was likely operating to create the incorrect view that juvenile court referral is deviance-generating. Third, an instrumental variables regression replicated the bivariate probit findings pointing to a null effect for referral on subsequent re-offending. In sum, there was strong evidence of selection bias such that a previously significant labeling effect was likely due to a selection artifact. If causal inferences are to be made from official sanctions the selection effects established by Smith and Paternoster must be confronted, yet earlier studies tended to not address this issue directly.

Regarding non-significant findings, Huizinga, Elliott, and Dunford (1986) used data from the National Youth Survey (NYS) to examine the effect of arrest on subsequent delinquency. To do so, arrestees and controls were matched on several important individual characteristics in the two years prior to arrest. Results failed to identify significant differences between the arrestees and their controls in delinquency/offending outcomes. Use of a delinquency typology also showed that most arrestees, among the most delinquent respondents, either increased or maintained the same level of delinquency compared to their non-arrested matched controls. It is important to note, however, that the delinquency typologies developed by Huizinga et al. were based on observed frequencies alone and did not objectively account for any preexisting developmental profiles, as is now common in the developmental literature via trajectory analysis and is the case for the study at hand.

A series of investigations on the effect of arrest on subsequent offending have also been carried out with the Denver Youth Survey (DYS), a longitudinal study of inner-city, high-risk youth in Denver. Huizinga and Esbensen (1992) found that arrest did not decrease offending, whereas a follow-up study by Huizinga, Esbensen, and Weiher (1996) that matched each adolescent who had experienced their first arrest with a similar control individual who had never been arrested, found that subsequent delinquency among most arrestees was either equal to or greater than their matched control after experiencing first arrest. Also using the DYS data, Matsueda, Kreager, and Huizinga (2006) found that as the experiential risk of apprehension increased so did the perceived chance of apprehension which, in turn, was associated with a decrease in future offending. Cross-national analyses using the DYS along with a sample of adolescents from Bremen Germany (that included matched-controls) found that in general, arrest had little effect on subsequent offending and that in some instances arrest increased future offending (Huizinga, Schumann, Ehret, & Elliott, 2003).

A companion longitudinal study, the Rochester Youth Development Study (RYDS) comprising a large sample of youth residing in Rochester, NY, has also
investigated the effect of arrest on subsequent offending. Bernburg and Krohn (2003) tested Sampson and Laub's theory of cumulative disadvantage, which claims that justice interventions influence later offending by knifing off prosocial opportunities such as educational attainment and employment prospects. Results showed that official intervention (police or juvenile justice) as a juvenile reduced the likelihood of educational attainment and increased periods of nonemployment in early adulthood, that official intervention as a youth-increased crime in early adulthood and that this effect was partly mediated by life chances such as educational achievement and employment, and that official intervention had stronger adult crime amplification effects among the disadvantaged (i.e. Blacks, impoverished males).

In an extension of their previous study with the RYDS data, Bernburg, Krohn, and Rivera (2006) examined whether deviant social groups mediated the link between juvenile justice intervention and subsequent delinquency. These authors found that youths with juvenile justice intervention experience were more likely to be members of a gang and to report having delinquent peers during the successive period compared to youth without such experience, that juvenile justice intervention increased subsequent participation in serious delinquency via involvement in deviant social groups (street gangs and delinquent peers), and that deviant social groups were strong correlates of subsequent offending.

Several other studies have used a range of samples to investigate the arrest effect. Paternoster and Piquero (1995) found that South Carolina high school students who were recently sanctioned for alcohol or marijuana use in the 10th grade (compared to students who were not sanctioned) reported greater alcohol and marijuana use one year later. Using data from the Edinburgh Study, a longitudinal study of youth in Edinburgh Scotland, along with propensity matched control and intervention (for whom there was a police decision to charge) groups, McAra and McVie (2007) found no significant difference in subsequent serious delinquency between the two groups after the police decision to charge. Smith and Gartin (1989) used data from the Racine Birth Cohort and found that arrest had a deterrent effect on the number of future police contacts—however with increasing previous police contacts, the deterrent effect waned over time. Thus, while an arrest appeared to reduce subsequent offending—especially among novice offenders, more experienced offenders tended to offend after contact with the police.

2. Smith and Gartin's study, along with the original Minneapolis Domestic Violence Experiment Study (Sherman & Berk, 1984) were among the few studies to report a decrease in subsequent delinquency following official arrest.

3. Several other studies examine the extent to which various forms of sanctioning following arrest (different kinds of police and court alternatives) differentially impact subsequent criminal offending. As the variation in the type of sanctioning is beyond the scope of our presentation, we forego a detailed review of these studies. Huizinga and Henry (2008, p. 236) summarize this body of evidence, finding that like the arrest evidence shows, sanctions in comparison to no sanctions result in either no change or an increase in subsequent delinquency and/or that increased severity of sanctions results in either no change or an increase in subsequent offending. Only a small minority of studies provides evidence of a decrease in subsequent delinquency with increasing sanctioning.
Current Study

The amalgam of prior results may be due to some combination of mixed thinking about arrest/sanctions, sample compositions, and research methods that have not considered potential selection bias. And further, while prior literature has tested components of the differential deterrability hypothesis (Andenaes, 1974; Piquero et al., 2011; Pogarsky, 2002), this literature has not examined the effect of arrest on subsequent offending across trajectory groups using propensity score matching.

In this paper, we examine the deterrence and labeling hypotheses that sanctions deter subsequent crime or that they amplify future crime, respectively. Two specific questions are addressed: (1) does arrest increase or decrease subsequent offending? and (2) does any such effect vary across distinct groups of offenders? In so doing, we extend prior research in two important ways. First, we examine the effect of arrest on next-period offending across distinct offending trajectories. Studying the effect of arrest on subsequent offending across trajectories is important on both theoretical and policy grounds. As Huizinga and Henry (2008, p. 223, emphasis added) note:

> The effect of arrest and sanctions may also vary by type of offender. Not all individuals commit delinquencies/crime for the same reason. Offenders vary by age, by gender, by stage in a delinquent career, and level of seriousness and frequency of offending. Thus, the effect of an arrest might vary between first time arrestees who are low level offenders and first time arrestees who are "seasoned" serious offenders; or between individuals first arrested at age 10 and those first arrested at age 17, and so on.

In this regard, our effort follows one of their explicit directions for future research: to examine whether the effects of arrest may vary by characteristics such as stage of delinquent or criminal career (p. 249). Second, we follow Smith and Paternoster’s (1990, p. 1129) suggestion that researchers increase attention to possible selection bias issues when studying the labeling hypothesis of deviance amplification, which has not been carefully considered. In this regard, our approach utilizes a blending of group-based trajectory analysis and propensity score matching.

Causal inferences specific to the arrest effect are made possible by capitalizing on group-based analysis and counterfactual analysis in tandem. By doing so in a transparent manner (Haviland, Nagin, Rosenbaum, & Tremblay, 2008), we eliminate potential selection bias within the arrest effect by: (a) accounting for underlying differences between delinquency risk profiles that manifest prior to (b) matching youths of equivalent propensities for being arrested for the first time over a clearly defined treatment period. By matching youth within specific risk-profiles, we improve the comparability of youths, specific to our focal outcome, subsequent delinquency. This approach builds upon the
studies reviewed above by reducing selection bias concerns surrounding the arrest effect, based on data from a large-scale longitudinal study.

Methods

Data

Data were culled from waves 1 through 6 of the NYS (Elliott, Huizinga, & Ageton, 1985; Elliott, Huizinga, & Menard, 1989). Respondents were ages 11 through 17 during 1976 (wave 1) and were interviewed annually for five consecutive years (waves 1 through 5) and thereafter in three-year intervals through wave 9. The study sample was based on a national probability sample of households in the USA selected for participation via a multistage cluster sampling design (for additional detail on the design and reliability of NYS data, see Elliott et al., 1989). The demographics of wave 1 participants represented national averages for 1976. The measures employed by NYS were extensive and involved many social, psychological, and behavioral self-reports in addition to contact/experience with the criminal justice system. The measures specific to this study are outlined below. Descriptive statistics are presented in Table 1.

The NYS is an amalgamation of different aged-youth at the first and subsequent waves, which may pose some concern about studying development in offending over time in general and then the extent to which an arrest experience connotes the same thing for different aged members of the NYS. We note here that the use of the NYS by wave was solely to form the trajectories (discussed below) while the analysis concerning our primary question, the effect of arrest on subsequent offending across trajectories, focuses on wave 5, at which point the age-range of persons first arrested is between 15 and 21 (mean = 17.6, SD = 1.83), with most cases clustered in the 16-18 age range and only five first-arrest cases at 21.4

Measures

The primary focus of this study surrounds measures of self-reported delinquency and self-reported official contact with police (i.e. arrest). At each wave, respondents were asked to report the frequency for which they had been involved in eleven different delinquent behaviors during the year preced-

4. We recognize that there may be some concern about the use of wave/age, but this has not been resolved. In two separate analyses of life-course offending patterns in the MacArthur Foundation’s Pathways to Desistance Study, a longitudinal study of 1,354 serious adolescent offenders in Philadelphia and Phoenix who at the baseline interview were between ages 14 and 18, substantive conclusions about offending trajectories and the covariates that distinguish between them are not sensitive to the use of wave or age (on the use of wave, see Mulvey et al. (2010); on the use of age, see Monahan, Steinberg, Cauffman, and Mulvey (2009)).
Table 1 Descriptive statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Low</th>
<th>Moderate</th>
<th>Chronic</th>
<th>Combined</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arrested, wave 5</td>
<td>.03 .18</td>
<td>.09 .29</td>
<td>.38 .49</td>
<td>.07 .25</td>
<td>65.1***</td>
</tr>
<tr>
<td>Age, wave 1</td>
<td>13.53 1.93</td>
<td>13.72 1.82</td>
<td>14.30 1.64</td>
<td>13.63 1.89</td>
<td>6.1**</td>
</tr>
<tr>
<td>Female</td>
<td>.59 .49</td>
<td>.32 .47</td>
<td>.21 .41</td>
<td>.48 .50</td>
<td>63.1**</td>
</tr>
<tr>
<td>White</td>
<td>.80 .40</td>
<td>.80 .40</td>
<td>.82 .39</td>
<td>.80 .40</td>
<td>.4</td>
</tr>
<tr>
<td>SES, wave 1</td>
<td>42.73 16.58</td>
<td>43.62 16.04</td>
<td>44.36 15.49</td>
<td>43.10 16.35</td>
<td>1.8</td>
</tr>
<tr>
<td>Public assistance, wave 1</td>
<td>.16 .36</td>
<td>.17 .38</td>
<td>.27 .45</td>
<td>.17 .37</td>
<td>5.6**</td>
</tr>
<tr>
<td>GPA, wave 1</td>
<td>3.93 .75</td>
<td>3.64 .80</td>
<td>3.38 .78</td>
<td>3.81 .79</td>
<td>32.3***</td>
</tr>
<tr>
<td>Social disorganization, waves 1 and 5</td>
<td>9.37 1.94</td>
<td>9.75 2.14</td>
<td>9.83 2.25</td>
<td>9.52 2.03</td>
<td>4.4*</td>
</tr>
<tr>
<td>Peer influence, mean of waves 1-4</td>
<td>11.83 3.18</td>
<td>15.75 4.10</td>
<td>22.11 4.66</td>
<td>13.67 4.52</td>
<td>418.3***</td>
</tr>
</tbody>
</table>

**Offending**

| Wave | Offending | Low          | Moderate     | Chronic      | Combined     | F            |
|------|-----------|--------------|--------------|--------------|--------------|
| Wave 1| .12 .39   | 1.06 1.29    | 3.59 2.47    | .61 1.28     | 572.1***     |
| Wave 2| .04 .20   | 1.13 1.01    | 4.19 1.75    | .61 1.22     | 1412.1***    |
| Wave 3| .04 .20   | 1.02 1.04    | 4.03 2.24    | .57 1.24     | 1005.4***    |
| Wave 4| .05 .23   | 1.00 1.25    | 3.10 2.55    | .52 1.20     | 484.5***     |
| Wave 5| .18 .67   | .87 1.37     | 2.74 2.79    | .54 1.29     | 200.3***     |
| Wave 6| .21 .83   | .96 1.72     | 2.60 2.80    | .58 1.47     | 135.3***     |

n = 881 453 83 1,417

Note. F = ANOVA test comparison between low, moderate, and chronic groups.
*p < .05; **p < .01; ***p < .001.
ing each wave’s interview. The specific forms of delinquency include: auto theft, theft of goods worth $50 or less, theft of goods worth more than $50, attacking someone with the intent of hurting or killing them, participating in group (gang) fights, selling marijuana, joyriding, burglary, knowingly purchasing stolen goods, and pressuring someone sexually. The prevalence of each delinquency indicator (i.e. one or more occurrence = 1, no occurrences = 0) was summed, resulting in an overall count of the number of different delinquent acts (i.e. a variety score). Delinquency counts from waves 1 through 4 were used to estimate the trajectory model—discussed below. Delinquency counts at waves 5 and 6 were used as outcome measures for which the treatment effect of arrest is modeled. It is important to reiterate that wave 6 data was collected three years after wave 5; however, at wave 6, respondents were asked to report delinquency retrospectively for the complete time period between waves 5 and 6. That is, they were asked to report on behaviors that occurred during 1983, 1982, and 1981 (wave 5 data was collected in 1980). All retrospective information was incorporated into the wave 6 delinquency measure, thus it reflects delinquency reported to have occurred for the three years prior to wave 6 interviews. In the end, our study addresses the effect of arrest on delinquency occurring in close proximity to having been arrested (i.e. that which occurred during the same year as the arrest) as well as delinquency reported in the three-year period following the arrest.

**Treatment variable**

Equally paramount to the present study was whether a respondent was ever arrested by law enforcement. During wave 5, respondents were asked: “Have you ever been arrested by the police for anything other than a minor traffic offense?” Follow-up questions were asked about the timing and reason/s for the most recent three arrests. Specifically, respondents were asked “How many times?” they had been arrested as well as the specific date (month and year) of the arrest. A binary variable was created where a value of 1 indicated having been arrested for the first time during the 12-month period preceding the wave 5 interview (the treatment window) and 0 was indicative of never having been arrested. At wave 5, a total of 117 of 1,494 (7.8%) respondents reported having ever been arrested. Of those, 93.1% (n = 109) reported having been first arrested during the treatment window (the remaining eight reporting first arrest prior to the treatment window and were thus excluded). Ninety-four of the 109 (86.2%) wave 5 first time arrestees had complete data on the delinquency measures for each previous wave and thus represent our treatment group. Representing our control group, a total of 1,349 respondents

5. Regarding missing data, 21.8% (n = 376) of the original 1,725 NYS respondents did not have complete data across waves 1 through 4 on the outcome measures and were thus excluded from the analysis. Between cases with complete vs. incomplete data over this period, significant differences were found on mean levels of delinquency for females, but not for males. Several studies have demonstrated that attrition in the NYS minimally affects substantive findings (e.g. Bosick, 2009; Brame & Paternoster, 2009; Elliott et al., 1989; Jang, 1999; though see Brame & Piquero, 2003).
reported having never been arrested through wave 5 and also had complete
data on the delinquency measures for waves 1 through 4.

It is important to note that this analysis is not addressing the first arrest in
general, but instead focuses only on those whose first arrest occurred during a
specified time period and those who were not arrested during the same speci-
fied time period, and never prior to that time. By first-time arrest, we mean
that up until wave 5, a person has never been arrested, but during wave 5 they
were arrested. Among the respondents reporting first arrest at wave 5, no sig-
nificant difference in delinquency reported across waves 1 through 4 was found
between cases with complete vs. incomplete data. To re-iterate, our study
explores the impact of first arrest only. We leave the exploration of the effect
(s) of subsequent arrests on subsequent criminality to future studies.6

Covariates

Several variables were drawn from the NYS to serve as covariates in estimating
the propensity for having been arrested. It is important to reiterate that each
covariate is either time invariant or reflects information from a time period
prior to the treatment period. These include age, sex (female = 1), race
(white = 1, else = 0), socioeconomic status (SES; Duncan (1961) socioeconomic
index reported by the parent at wave 1), the prevalence of public assistance
(reported by the parent at wave 1), grade point average (GPA) on a five-point
scale, an index of social disorganization, an index of negative peer influence,
and whether the respondent lived in a rural vs. urban/suburban environment

6. Further, it is unlikely that the complexity ends with the idea that the first sanction is always the
most important. The matter of timing, as highlighted by the life-course perspective, is an im-
portant concern, one that we do not explicitly address in this study. However, as discussed below, we
use important information on the timing of an arrest to ascertain differences in the arrest effect
for different groups of individuals. The life-course perspective has gained widespread attention in
criminology, and one theory in particular, Sampson and Laub’s (1993) age-graded informal social
control theory, has been the subject of much attention. In explicating a perspective that is friendly
to both continuity and change, Sampson and Laub not only consider the importance of the strength
and quality of life events (and bonds) in relation to persistence/desistance from crime, but also
attend to the timing of such events/bonds as well. In our study, we focus on a life event (first
arrest) that occurs in late adolescence/emerging adulthood. Some life events, such as criminal jus-
tice system contact, “cut off opportunities” for prosocial prospects later in life (Sampson & Laub,
1997, p. 150). As the theorists note: “The theory specifically suggests a “snowball” effect—that
adolescent delinquency and its negative consequences (e.g., arrest, official labeling, incarceration)
increasingly “mortgage” one’s future, especially later life chances molded by schooling and
employment” (p. 147). They (p. 148) go onto add that: “Arrest, conviction, and imprisonment are
clearly stigmatizing, and those so tarnished face structural impediments to establishing strong
social ties to conventional lines of adult activity …” In sum, the timing of an event, i.e. getting
married or arrested at age 14 compared to age 28, is likely to be critical in understanding whether
and how a life event alters a behavioral trajectory (cf. Theobald & Farrington, 2011). In our study,
the life event of interest is the first experience with an arrest. Of course, this may be different for
a second- or third-time arrest and while very interesting, we must reserve this for future inquiry
with a more “offender-based” sample, given the low power associated with multiple arrests in our data.
(rural = 1, else = 0). Also included was the averaged count of delinquency reported for waves 1 through 4.\(^7\)

Peer influence was operationalized via nine items asking respondents to report how many of their close friends had purposely damaged property, used marijuana, stolen something worth less than $5, stolen something more than $50, hit someone without any reason, used alcohol, broken into a car or house, sold drugs, and suggested that the respondent do something illegal. The responses were recorded on a five-point scale where higher values reflected increased exposure to delinquent peers. At waves 1 through 4, respectively, each set of indicators was summed and the mean of the summed scores across waves was used as an indicator of peer influence from waves 1 through 4.

Information on social disorganization (neighborhood problems) was reported by the parents during wave 1. Respondents were asked to report on the presence of six neighborhood problems including vandalism, the presence of winos or junkies, abandoned homes, burglaries and thefts, rundown buildings, and assaults/muggings. Responses were recorded on a three-point scale where 1 indicated that the indicator was "not a problem," 2 indicated that it was "somewhat of a problem," and 3 indicated that it was a "big problem." Scores on each indicator were summed to create the social disorganization index where higher scores reflect increased social disorganization. Missing data on SES (less than 10% of cases) were imputed with the sample mean.

Analytic Procedures

The specific goal of this study is to determine whether the effect of being arrested aggravates or mitigates subsequent delinquency within and between distinct offending trajectories from late adolescence into emerging adulthood. In other words, we ask whether having been arrested for the first time serves as a turning point with respect to subsequent criminal activity.\(^8\) Focusing on the effect of a(n) (first-time) arrest is important for many reasons, not the least of which is that early (as opposed to later) contact with the justice system is likely to have the most significant impact on youth because it can set the stage for adverse consequences in school, home, and the community (Bernburg et al., 2006, p. 72).

Experiments are ideally suited to test such an effect; however, a true experimental research design that is generalizable to a large collective is often unrealistic. Fortunately, the experimental design can be closely

---

\(^7\) Models were also run with separate delinquency indicators for waves 1 through 4, as well as for exposure to delinquent peers (discussed below). Further, models were run with varying measures of parental management, reported by parents at wave 1. Results were consistent with those presented here.

\(^8\) The mean age of respondents at wave 5 was 17.67 years.
approximated with longitudinal observational data through counterfactual designs, such as propensity score matching (Rosenbaum, 1984). This can be done by specifically and carefully defining both treatment and outcome measures, assuring that the treatment precedes the outcome, and allowing for treatment to occur during a specified time period. Recently, scholars have used such approaches targeting life course turning points that potentially impact delinquent/criminal behavior occurring after the turning point has taken place, for example the effect of joining a gang (e.g. Haviland et al., 2008; see also Barnes, Beaver, & Miller, 2010; DeLisi, Barnes, Beaver, & Gibson, 2009). Herefore, none have addressed first arrest as a treatment predicting further offending while removing potential bias by distinguishing between risk profiles in a systematic way as undertaken in the current study.

Analogous to Haviland et al.’s (2008) study of the effect of joining a gang, our focus is on youth who had no experience with arrest prior to wave 5 of the NYS. However, a portion of these youth were arrested at some point between waves 4 and 5. Utilizing propensity score matching in tandem with group-based trajectory modeling (i.e. finite mixture modeling), we matched youths who reported having been arrested for the first time at wave 5 to controls who were not arrested at wave 5. This was done within each specific trajectory profile, respectively. We then relied on data capturing their involvement with delinquency for several years after wave 5. To ensure an unbiased treatment effect, we utilized several pre-treatment covariates in order to balance treated youths with prospective controls, so they may be appropriately compared. As in experiments, groups are comparable only on the attributes that are actually measured. Here, potential bias due to unmeasured covariates is considered, as best as possible, via sensitivity analysis (Guo & Fraser, 2010; Rosenbaum, 2002).

As noted by Haviland et al. (2008), using this sort of research design for approximating an experimental design has several advantages. First, observational datasets, such as the NYS, are very rich in the measures of respondent characteristics. Second, behavioral patterns that represent the final outcome are observed pre-treatment. This information can be used to classify respondents into differing trajectory profiles that demarcate very different behavioral patterns, thus accounting for underlying group heterogeneity. By pairing un-arrested youths with arrested youths from within the same trajectory profile, based on correctly time-ordered observed covariates, we come closer to comparing apples to apples, so to speak. In doing so, the assessment of a treatment effect becomes increasingly (though never perfectly) immune to potential bias resulting from both underlying group heterogeneity as well as in the treatment effect itself. More specifically, youths who demonstrated a high level of delinquency for an extended period pre-treatment (waves 1 through 4) are compared to others with a similar behavioral profile. In the end, the patterns of delinquency are classified into a small number of finite groups and propensity scores are used to estimate the probability of being arrested, which
ultimately results in a propensity score that is conditioned by observed covariates. Again, the use of group-based modeling, as done herein, results in findings that apply specifically to each trajectory profile. As life course theories of development (Elder, 1985, 1998; Thornberry, Krohn, Lizotte, Smith, & Tobin, 2003) surmise that both the direction and degree of a treatment effect (or the effect of a covariate in predicting an outcome) may be trajectory-specific, we focus on the effect of arrest on subsequent delinquency across different trajectories.9

Results

Trajectory Analysis

Trajectory analysis, also referred to as finite mixture modeling (Land, McCall, & Nagin, 1996; Nagin, 2005), growth mixture modeling (Muthén, 2004), and latent class growth modeling, is designed to examine behavioral changes over age or time and has been used routinely to assess longitudinal trends in delinquency/crime (Piquero, 2008). Although there are subtle and important mathematical differences across the various approaches, they all share a common goal of approximating the distributions of developmental trajectories specific to a population from a finite number of groups, or latent classes, the exact forms of which are not known. Here, trajectory groups are comprised of respondents with complete data on delinquency from waves 1 through 4 and who reported never having been arrested prior to the treatment period. Results presented below include: (1) form/shape of each delinquency trajec-

9. Haviland et al. (2008) highlight the importance of correct time ordering in an analytical approach such as this. We would be mistaken not to reiterate such here. Delinquency at wave 6 (and to at least some extent wave 5) and beyond are behaviors that occurred post initial treatment status (i.e. never having been arrested). Further, the covariates relied upon here are either (a) time invariant or (b) represent phenomena occurring or already in place prior to treatment. Thus, causal inferences resulting from the design carry more weight than non-counterfactual procedures commonly used for the analysis of observational, or non-experimental, data. Related is the issue of persistence and/or desistence of the treatment. Haviland et al. stress the importance of considering the transient nature of their treatment effect, gang membership. Gang members may or may not still be gang members at any given point after status initiation. In other words, youth could opt-out of treatment even after the treatment had been administered. A key difference between our study and theirs is that being arrested is an event that cannot be taken away from a subject, thus issues surrounding the modification of treatment states can be downplayed. For certain, experiences with arrest vary. For example, an arrest may not result in a conviction or adjudication (e.g. no charges are filed by a district attorney, the arrest is officially expunged, etc.); however, memory of the experience, along with the stigma of being arrested, whether in the form of psychological or sociological distress (i.e. official or subjective labeling), may have a lasting effect and is certainly not subject to self-selection to the alternative state (i.e. un-arrested) (see Cameron, 1964). Being arrested, particularly for the first time, is something that most individuals will not forget, particularly in the short-term. In the end, like Haviland et al., we have presented a transparent approximation of an experimental design to address our research goal in which covariates of the treatment precede the treatment period, the outcomes occur post-treatment period, and treatment is static and not time-varying or subject to self-selection out of treatment.
tory profile, (2) a proportional breakdown of each group specific to the NYS data, and (3) the posterior probability of membership to each specific trajectory profile. The specific procedure used here is consistent with that of finite mixture modeling (Nagin, 2005).

The trajectory model was specified under the assumption that each group’s functional form of delinquency fell under the Poisson distribution, but where observed variances exceeded means, thus the negative binomial link-function was utilized. Further, we specified that delinquency trajectories followed a quadratic function over time. As in any latent class analysis, varying numbers of groups were considered. A balance between Bayesian information criterion (BIC) and model parsimony were used to define the cutoff for additional groups. The 3-group model fit the data best as the 4-group model had only a modest improvement in BIC and simply split one group from the 3-group model, resulting in a fourth group that was sparsely represented.

Table 2 presents the estimates from the trajectory analysis of self-reported delinquency for the 3-group model. We refer to group 1 as the “low-risk” respondents, group 2 as the “medium-risk” respondents, and group 3 as the “chronic” delinquents. About 62% of the 1,424 respondents were estimated to fall into the low-risk group, 32% to the medium-risk group, and the remaining 6% were classified as chronic delinquents, with respondents being assigned to their respective group based on the highest posterior probability of group assignment. Also presented in Table 1 are expected rates of delinquency over time, specific to each class (also see Figure 1). As shown, expected delinquency for the low- and medium-risk groups was fairly stable, though somewhat declining for the latter. For the chronic group however, expected delinquency is clearly non-linear but is markedly higher than the two lower risk groups during any given year. Observed proportions and counts of group members who reported arrest at wave 5 suggest that arrest prevalence

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group 1 (Low)</th>
<th>Group 2 (Moderate)</th>
<th>Group 3 (Chronic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average posterior probability</td>
<td>.92</td>
<td>.90</td>
<td>.91</td>
</tr>
<tr>
<td>Proportion of sample in group</td>
<td>.62</td>
<td>.32</td>
<td>.06</td>
</tr>
<tr>
<td>Average expected rate of offending at wave 1</td>
<td>.12</td>
<td>.99</td>
<td>3.41</td>
</tr>
<tr>
<td>Average expected rate of offending at wave 2</td>
<td>.04</td>
<td>1.01</td>
<td>4.01</td>
</tr>
<tr>
<td>Average expected rate of offending at wave 3</td>
<td>.04</td>
<td>.97</td>
<td>3.83</td>
</tr>
<tr>
<td>Average expected rate of offending at wave 4</td>
<td>.07</td>
<td>.87</td>
<td>2.99</td>
</tr>
<tr>
<td>% Arrested at wave 5</td>
<td>3.1 (27 of 881)</td>
<td>8.6 (39 of 453)</td>
<td>33.7 (28 of 83)</td>
</tr>
</tbody>
</table>

Note. Low = low offending; moderate = moderate offending; chronic = chronic offending.
increases with risk for delinquency. In fact, over one-third of respondents assigned to the chronic group (through wave 4) were arrested for the first time at wave 5.

**Between-group Propensity Score Matching**

Initial official contact with the criminal justice system, defined here as being arrested for the first time (at the wave 5 interview), can be considered a potentially important turning-point that may evince negative consequences from the labeling perspective or preventive effects from the deterrence perspective. Experiencing this event may have a sizable impact on how others perceive the individual, the subjective perception of others’ opinions/attitudes toward the self, the sensitivity of the criminal justice response to the individual if adjudicated delinquent or convicted of a criminal act, or may impact an individual’s life more generally, perhaps ultimately modifying the probability of subsequent conforming or antisocial behavior.\(^{10}\)

---

10. Of course, attempting to infer causation on such an event based on observational data alone must be tempered. As noted by Haviland et al. (2008), individual characteristics of a specific individual that are present prior to experiencing an event (such as arrest for the first time) may impact the treatment effect, either in magnitude or direction. This is one reason why it was important to parse out the pre-existing differences before the fact via trajectory analysis. As discussed above, treated individuals are those who experienced arrest for the first time at wave 5 of data collection and possible controls are those who were not arrested at wave 5 and had never been so previously.
Figure 2 presents mean covariate differences between treated individuals and non-treated individuals. Prior to matching, it can be seen that for some covariates, mean differences exist between those arrested and otherwise. Among metric covariates, the treated group had higher means for self-reported delinquency (waves 1 through 4 as a whole), social disorganization, exposure to delinquent peers, and was somewhat older at wave 1. Ultimately, the box-plots help to illustrate systematic differences that may exist between those having been arrested and those not, both before matching and before mixture modeling occurs. Indeed, covariates may confound one another if underlying group heterogeneity remains unaccounted for. Further, if respondents had been randomly assigned to arrest vs. non-arrest, such as in an experiment, we would expect that mean differences across each covariate would be null, with some minimal expectation of error. This is not the case for six covariates (delinquency, exposure to delinquent peers, age, sex, race, and rural residence). Without proper adjustment, any attempt to assess the effect of arrest on subsequent delinquency would likely be subject to selection bias. The use of propensity score matching on trajectory group stratified data helps to remove such bias in the treatment effect.

Propensity score matching was developed to provide a means of achieving balance among covariates between treatment and control groups. As discussed in Rosenbaum and Rubin (1983), covariate imbalance can be remedied by matching on a unidimensional propensity score. This is possible since the propensity score is a function of the covariates themselves. Further, once balance is appropriately achieved and sensitivity to unobserved covariates has been critiqued, then a propensity score can be used to estimate an unbiased treatment effect. While it is important to keep in mind that matching on a propensity score cannot directly control for imbalance due to unobserved covariates (which can only be done through randomization), the process does remove imbalance and allows for the comparison of mean differences on an outcome between treatment and control groups, provided that the observed covariates represent a time prior to the treatment.11

Here, we use propensity scores to match first time arrestees with comparable non-arrestees, where exposure to arrest was limited to an equivalent time period (defined above), and do so within each trajectory group, respectively. Moreover, we matched each arrestee with a non-arrestee whose estimated propensity, or probability, for being arrested was in very close proximity to the arrested match (caliper = .05). Matching was done without replacement and was carried out using the psmatch2 program in Stata version 11. Figure 3 presents the estimated propensity scores for those arrested along with their potential controls, both as a combined sample and between trajectory groups. As shown, the difference in propensity for the arrest between treated and

11. The instrumental variables approach can also handle selection based on unobservables, assuming that one can find a true instrument.
Figure 2  Box plots of six metric covariates prior to matching for first-time arrestees at wave 5 and for potential controls who were not arrested.

Note. The p values for the dichotomous covariates (not shown) female, white, public assistance, and rural were .001, .023, .010, and .018 with the arrested group being more male, non-white, receiving more public assistance, and less rural. The above box plots are based on all trajectory groups combined.
control-group members is rather substantive, prior to trajectory group classification. The overlap in propensity scores between treatment and control groups for each specific trajectory group suggests that acceptable matches were likely.

The Effect of Arrest on Subsequent Delinquency

Matching within each trajectory group limits the possibility for covariates that could act as confounders to other covariates in estimating arrest propensity. As shown above, finite mixture modeling helped to eliminate much potential treatment bias. The estimates presented in Table 3 represent the effect of

12. Secondary Propensity Score Matching (PSM) models were estimated between arrested youth and smaller random samples of non-arrested youth, within each respective class, to bring outcome values to closer proportions. Doing so improved the model fit/classification accuracy of the logit models used to estimate propensity scores and modestly improved the number of matches within each class; however, the substantive findings resulting from the PSM went unchanged.

13. The statistics used to demonstrate balance between the matched groups for each trajectory profile and the entire samples are presented in Appendix A. The matching procedure for each group, as well as for the sample in whole, was successful in balancing mean scores for each covariate and for each grouping of respondents and was equivalent across outcomes based on both reductions in standardized bias and on hypothesis tests.
being arrested between trajectory classes post-matching. The arrest effect across each trajectory group, as well as for a matched combined sample, is graphically illustrated with box plots in Figure 4. For the wave 5 delinquency outcome, the results show that, without the use of finite mixture modeling (i.e. the combined group), respondents who were arrested for the first time at wave 5 had much higher rates of delinquency at wave 5. This effect, however, is somewhat confounded by the trajectory group classifications. Specifically, for wave 5 delinquency, the effect is particularly strong for chronic delinquents. To this point it is important to note that causal inferences made from the wave 5 outcome should be avoided, largely because we have no way of knowing whether some or all of the delinquency reported by the respondent occurred prior to arrest, or otherwise. To remedy this issue, we shift to an outcome representing a time after treatment, wave 6 delinquency.

Using the wave 6 delinquency outcome it can be seen that the arrest effect is positive for each trajectory group. Yet, larger differences emerge in the effect for the treated group for the medium and, to a greater extent, the chronic class (discussed in more detail below). To corroborate the evidence from the propensity matching approach, we also carried out an assessment of the arrest effect via an inverse probability of treatment weighted (IPTW) regression analysis, the results (predicted rate) of which are presented in the right-hand side of Table 3. IPTW “consistently estimates the causal effect of a time-dependent treatment when all relevant confounding factors have been measured” (Guo & Fraser, 2010, p. 331). Ultimately, the IPTW approach is a regression analysis where the outcome is regressed on a single predictor, which is the covariate weighted [inverse] propensity, or probability, of having received the treatment (i.e. the conditioned treatment effect). Based on the IPTW results, the expected rates of offending for those arrested and for matched controls were 1.05 and .10, 1.92 and .59, and 3.14 and 1.42 for the low-risk, medium-risk, and chronic offending groups, respectively. These findings indicate that the experience of an arrest may exacerbate future delinquency and that there may be differences in the rate of delinquency that we might expect may differ between groups as a result.

The findings initially presented above lend themselves to multiple avenues of interpretation and are based on varied analyses used to determine whether the arrest effect differs across groups of offenders. First, and at the most basic level, the observed median post-arrest offending data at W6 (see Figure 4) clearly shows that median offending rates are indeed aggravated by arrest for the medium-risk and chronic delinquency groups, but generally not for the low-risk group. In fact, median delinquency is zero for both treatment and control group among the low-risk group. This finding suggests that the

---

13. The statistics used to demonstrate balance between the matched groups for each trajectory profile and the entire samples are presented in Appendix A. The matching procedure for each group, as well as for the sample in whole, was successful in balancing mean scores for each covariate and for each grouping of respondents and was equivalent across outcomes based on both reductions in standardized bias and on hypothesis tests.
Table 3  The effect of arrest at wave 5 on offending at waves 5 and 6

<table>
<thead>
<tr>
<th>Offending difference</th>
<th>t</th>
<th>Offending difference</th>
<th>t</th>
<th>Sensitivity</th>
<th>Predicted offending (IPTW)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Untreated</td>
</tr>
<tr>
<td>Wave 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>1.27</td>
<td>10.25*</td>
<td>1.15</td>
<td>2.59*</td>
<td>2.35</td>
</tr>
<tr>
<td>Moderate</td>
<td>1.38</td>
<td>6.25*</td>
<td>1.08</td>
<td>2.69*</td>
<td>1.35</td>
</tr>
<tr>
<td>Chronic</td>
<td>2.28</td>
<td>3.54*</td>
<td>2.14</td>
<td>2.44*</td>
<td>1.80</td>
</tr>
<tr>
<td>Combined</td>
<td>2.13</td>
<td>16.83*</td>
<td>1.36</td>
<td>4.49*</td>
<td>2.90</td>
</tr>
<tr>
<td>Wave 6</td>
<td>.91</td>
<td>5.94*</td>
<td>.74</td>
<td>1.57</td>
<td>1.25</td>
</tr>
<tr>
<td>Low</td>
<td>1.27</td>
<td>4.54*</td>
<td>.87</td>
<td>2.61*</td>
<td>1.55</td>
</tr>
<tr>
<td>Moderate</td>
<td>1.98</td>
<td>3.03*</td>
<td>.95</td>
<td>1.03</td>
<td>1.35</td>
</tr>
<tr>
<td>Chronic</td>
<td>1.89</td>
<td>12.61*</td>
<td>1.14</td>
<td>3.31*</td>
<td>1.60</td>
</tr>
</tbody>
</table>

Note. Offending difference refers to the difference in mean delinquency between those arrested and those not arrested. Gamma refers to Rosenbaum sensitivity analysis; the numeric refers to the necessary strength (odds ratio) of an un-modeled covariate to potentially bias matching results.

*p < .05 (two-tailed).
arrest effect, while positive, is attenuated for the low-risk group and particularly amplified for the chronic group.

Figure 4 Offending at waves 5 and 6 for respondents who were arrested for the first time during wave 5 and their matched controls.
Second, we can compare the substantive differences in the treatment effect between each trajectory group and explore $t$-test findings on the arrest effect across matched groups (shown in Table 3). This particular finding suggests that the arrest effect is positive and statistically significant only for the medium-risk group; however, the substantive results suggest an upward trend in the arrest effect on future delinquency from low risk to chronic delinquents. We strongly caution readers in drawing conclusions on this finding solely based on the $t$-test statistics because the process of matching, and further still by dividing the sample into trajectory classes, will result in a reduction to statistical power (i.e., sample sizes are reduced). Nevertheless, the totality of the results implies that differences on the arrest effect suggest that the arrest effect is stronger for higher risk profiles and effectively null for the lowest risk group.

Third, the IPTW findings (i.e., predicted rates of delinquency, also presented in Table 3) suggest that the arrest effect is positive across each class, but that the arrest effect is strongest for the chronic group. In fact, based on these findings we would expect only a single type of delinquency reported for arrestees of the low-risk group compared to matched controls (1.05 vs. .10). For the medium-risk group, we expect an additional 1.33 reports of delinquency from those arrested compared to matched controls (1.92 vs. .59), and for the chronic group an extra 1.73 reports of delinquency from those arrested compared to matched controls (3.14 vs. 1.41). Equivalently, these findings suggest that compared to the low-risk group, we can expect nearly three times more delinquency reports from the chronic group post-arrest ($\text{pr}$(delinquency of chronic-risk group)/$\text{pr}$(expected delinquency of low-risk group) = $[3.14/1.05] = 2.990$) and nearly twice the amount of delinquency from the medium-risk group ($[1.92/1.05] = 1.823$), as a result of being arrested. Further, we would expect over one and a half times more delinquency from the chronic group, compared to the medium-risk group ($[3.14/1.92] = 1.635$), post-arrest.

These findings suggest that the arrest effect on delinquency is substantively amplified for the chronic group, compared to the lower risk groups. It is important to note that since we are dealing with expected counts of delinquency, we should consider substantive differences in the expected/observed rates of delinquency post-treatment over proportional within-group effect size where the expected outcome is small, or not substantively large, such as the case for the low-risk group where the treatment group has an expected delinquency rate of only one.

Fourth, and returning to observed mean differences between treated youth and their matched controls, supplemental $t$-tests (one-tailed) indicated significant differences in W6 offending among first-time arrestees between the low- and medium-risk groups ($t = 1.78; p = .039$), between the low risk and chronic groups ($t = 3.67; p < .001$), and between the medium-risk and chronic groups ($t = 2.58; p = .006$). Thus, among those arrested in each class, delinquency post-arrest is more pronounced as the risk profile increases—particularly among the chronic group.
Based on a preponderance of evidence outlined above (i.e. observed median differences, observed mean differences, predicted rates, differences between predicted rates, and differences between observed delinquency between arrestees only), the effect of arrest appears to be strongest for chronic offenders and attenuated among the medium-risk profile. For the least involved youth, the arrest effect can be argued as null.

Before we conclude the presentation of our results, it is important to highlight a component of the study that deals with the sensitivity of the findings to unmeasured covariates. Shown in Table 3, \( I' \) represents the required magnitude, in the form of odds ratios, of an unmeasured covariate for it to potentially bias the results of the matching procedure. Hidden bias may arise in the matching procedure if there are unobserved variables that would have significantly affected assignment to the treatment group. The sensitivity analysis utilized here (Rosenbaum, 2002) provides test statistics that assess the strength requirement of an unobserved indicator/s for it to have been capable of interference with the treatment effect (see also Becker & Caliendo, 2007). Note that here, \( I' \) varies between the outcome and between trajectory classes and ranges between 1.25 and 2.35. The results of the sensitivity assessment suggest that, overall, we can have a reasonable amount of confidence that the arrest effect is not due to bias stemming from unmeasured covariates. Further detail on the utility, explanation, and execution of sensitivity analysis in propensity score matching is presented in Becker and Caliendo (2007).

**Discussion**

The over-arching goal of this paper was to examine whether the experience of an arrest produced an increase in subsequent crime as hypothesized by labeling theory, or whether the experience of an arrest produced a decrease in subsequent crime as hypothesized by deterrence theory. Our study contributes to the knowledge base on the effects of arrest in two ways by examining: (1) whether the effect varied across distinct offender trajectories and (2) whether the effect varied as a result of attention to sample selection bias issues. Neither of these two aspects of our study had been considered jointly in previous research. Our analysis of data from the NYS indicated that arrest substantively amplified subsequent delinquency among the most chronic trajectory group and to a much lesser extent for the medium-risk group, while the arrest effect had virtually no crime-exacerbating effect among low-risk youth.

Returning to Huizinga et al.’s (1986) analysis of the NYS data recall that they failed to find a general arrest effect on subsequent delinquency. However, among those defined as the most delinquent in their typology scheme, the majority of those arrested reported either no change or an increase in delinquency, compared to those not arrested. While we also relied on the NYS data, the methodology and findings presented in this study are different in several important ways. Huizinga et al.’s finding on the arrest effect for
the most seriously involved youth is similar to our findings on the crime exacerbating effect of arrest among youth in the chronic developmental profile; however, we addressed the arrest effect question in a more sophisticated manner, using analytical techniques that were not available when Huizinga et al. completed their analysis. These techniques were developed specifically to account for underlying group heterogeneity (here between developmental behavioral profiles) and also helped to account for selection bias in being arrested. In other words, we were able to systemically and objectively classify youth into risk profiles based on longitudinal data reported annually over the five-year period prior to first arrest. Huizinga et al. simply defined their most chronic group (serious career offenders) as youth who reported two or more years of continuous involvement in delinquency. Additionally, we matched youth using a more comprehensive array of covariates than did Huizinga et al. Taken together, our strategy provided for more confidence in making comparisons between arrested youth and matched controls and better addressed selection bias issues in being arrested. In the end, we found that such heterogeneity was important for understanding the effect of arrest on subsequent offending across the different trajectories. Individuals on the most criminogenic trajectories respond to arrest with considerably more delinquency compared to their less serious counterparts. The findings of the present study of the arrest effect are clearer, more distinguishable, and more robust than those of prior works.

That arrest does not appear to deter but may actually stimulate subsequent delinquency among some persons (particularly those most heavily involved in delinquency) may indeed represent a deviance amplification effect as posited by labeling theorists. But the finding may also reflect other processes and explanations. One of these deals with what criminologists refer to as the "resetting" effect, insomuch as punished individuals reset their estimation of the certainty of being caught or apprehended by authorities (Piquero & Paternoster, 1998; Piquero & Pogarsky, 2002; Pogarsky & Piquero, 2003). In this regard, the effect of an arrest among persons experiencing such an event may be considered a "hazard of the trade," or a cost of doing crime. Since most (high-rate) offenders offend a great deal and are rarely arrested (much less punished) (Blumstein, Cohen, Piquero, & Visher, 2010; Decker, Wright, & Logie, 1993), this positive arrest effect could have little to do with labeling per se. In the same vein, our finding that the chronic group is more delinquent after experiencing arrest is consistent with Pogarsky's (2002) findings, which reveal that the most engaged offenders often fail to even consider the cer-

14. Indeed, the resetting effect has been supported in several studies and across varying sampling frameworks. For example, Pogarsky and Piquero (2003) found that the gambler’s fallacy was predicted by increases in risky behaviors among college students, but only for those who were below the median for self-reported risky behaviors. They also found that sanction-certainty was equivalent between those who had and had not experienced punishment. If resetting is driving our findings, then it is those who demonstrate increased antisocial behavior who reset, and perhaps not youth who generally abstain.
tainty of being caught for misdeeds. Alternatively, one could argue that the effect of arrest fades and that the forces driving delinquency are not perceptions of the certainty of sanction, as with the resetting effect, rather delinquency may be attenuated for a period immediately after arrest, but the memory of arrest, and perhaps its side-effects, decay over time. Equally interesting is the finding that youth who are at a low risk of delinquency, but who offend and are arrested, appear to be relatively immune to the negative consequences of arrest that are more pronounced for riskier developmental profiles (i.e. we would expect them to engage in little to no delinquency at all, which might also be expected had they not been arrested).

Finally, it is also possible that the findings specific to more delinquent youth, particularly the chronic group, could be due in part to state dependence (Nagin & Paternoster, 1991). Regarding persistence over time, state dependence implies that the positive correlation between past, present, and future offending is dependent on the initiation of past offenses as well as those that follow, even after controlling for persistent individual differences. In other words, prior involvement “reduces what may have been reasonably effective inhibitions against future crime” (p. 166). More specifically, if we consider arrest as a possible byproduct of offending, then the findings for the chronic group can be explained via state dependence. Similar to the life-event findings presented by Sampson and Laub (1993), experiencing arrest for the first time during or following a period of virtually undetected persistency may have adverse ramifications on later life circumstances (e.g. arrest may weaken social bonds later in life that may otherwise stave off a history of childhood misbehavior or conversely, enhance prospects for prosocial opportunities later in life) among high-volume offenders, which would support Sampson and Laub’s (1997) labeling-friendly cumulative disadvantage perspective. This may also explain our lack of findings for an arrest effect among the low-risk group.15 There may be a number of other reasons why the more engaged offending group findings emerge as they do in our study and subsequent research should further unpack the mechanisms underlying the arrest-exacerbation effect among more serious offending groups.

These findings and interpretations must be tempered by some limitations related to the data and measures. For example, relying on an outcome that represents a three-year window of opportunity for post-treatment delinquency (i.e. wave 6) is both a strength and weakness. This outcome is limited by the reality that data for wave 6 was collected three years after wave 5, thus, retrospective information regarding delinquency occurring during the years immediately following wave 5 could be subject to memory decay. Further, our treatment measure was based on self-reported arrest, rather than on official

15. Still, one could argue that continuity in crime, in and of itself, should serve to make a young individual increasingly susceptible to negative life consequences that may follow arrest (loss of job, damaged social relationships). However, it is worthwhile to consider that this effect may or may not apply equally to first arrest occurring later in life, possibly after bonds may have been established even in lieu of a chronic history. Such questions are left to future studies.
records—the latter is unfortunately not publically available for the NYS. Additionally, our findings must be viewed cautiously because the treatment group was relatively small. Thus, similarly designed studies using data from larger, high-risk populations should be carried out to corroborate our findings.

Given the cross-section of ages in the NYS and potential age heterogeneity concerns (see Lauritsen, 1998), it would be good for future studies to focus on the effect of arrest on subsequent offending across different ages. Although there are too few cases available for such an analysis in the current study because of our methodological focus, the age issue is an important focus of future investigations.

Another limitation has to do with the range of covariates used to estimate propensity scores, which were restricted to those available in the NYS. For example, it would be worthwhile to consider individual-level measures tapping stable personality characteristics unavailable in the NYS (such as negative emotionality) as they may play an important role in further explaining the arrest effect. Also, since the NYS only contains information on the arrest event, future studies should consider obtaining data from respondents’ perceptions of the sanctioning event, i.e. how they emotionally reacted to it and so forth. As Sherman (1993) notes, the effect of a sanction on subsequent behavior will be contingent on several factors including the person’s interpretation of the sanction, how bonded they are to the sanctioner, and their emotional reaction (see also Piquero, Langton, & Gomez-Smith, 2004). These sorts of mediating mechanisms serve as an important direction for future research especially given the lack of attention given to them in prior studies (likely due to data limitations). Our data also did not permit us to assess post-arrest punishment (i.e. incarceration, length of incarceration, etc.) to determine whether additional punishment incurred labeling/deterrent effects. An additional point of future inquiry could examine the timing (and subsequent effect) of a person’s actual first arrest as well as potentially mediating mechanisms underlying the arrest effect.

As our investigation was primarily focused on linking arrest to future criminal behavior in general and then across trajectory groups in particular, we did not have the requisite data that would permit a detailed investigation of the potential mechanisms linking arrest to subsequent offending. As a reviewer noted, this relationship may be due to a resetting effect, negative reactions from the environment that cut off future opportunities, or because the offender comes to accept a criminal identity. Consideration of deterrence, life course, and labeling theory, respectively in this regard, would be an appropriate next step in unpacking the arrest/subsequent offending linkage. On this score, it would be especially interesting to consider how these theoretical perspectives may be important among some offenders and not others and also potentially different at different phases of the life course. For example, the life course mechanism may be more relevant early or later in the life course, while the resetting effect may be more pronounced later in the life course and later in offender careers after they have accumulated a healthy stock of offending experience and its consequences. Lastly, because the NYS is a cross-
section of respondents, issues regarding change over time in delinquency must be carefully considered (Brame & Piquero, 2003; Lauritsen, 1998).

These limitations notwithstanding, our results suggest that an important turning point specific to crime occurring during late adolescence (ages 15-21) in the realization of an arrest, aggravates subsequent delinquency among some, but not all, persons. Being arrested, in and of itself, increases the likelihood of additional delinquency for youth, but this effect results in considerably more delinquency among youth who are more involved in delinquency prior to arrest, compared to those who have a history of little or no involvement with delinquency. Indeed, youth who demonstrate a lower risk for delinquency appear to be relatively immune to the negative impact of arrest (i.e. the effect is arguably null). Expansion of our study to other topics as outlined above, including whether our findings are further moderated by race and neighborhood status——where an arrest may be viewed as a badge of honor (Anderson, 1999)——await further empirical attention.

References


Rosenbaum, P. R. (1984). The consequences of adjustment for a concomitant variable that has been affected by the treatment. *Journal of the Royal Statistical Society, Series A*, 147, 656-666.


## Appendix A. Predictor Imbalance Before and After Matching (Wave 5 Outcome)

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Low</th>
<th></th>
<th></th>
<th></th>
<th>Moderate</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>( t )</td>
<td>% Bias</td>
<td>Mean of</td>
<td>Mean of</td>
<td></td>
<td>( t )</td>
<td>% Bias</td>
</tr>
<tr>
<td></td>
<td>(matched)</td>
<td></td>
<td>Possible</td>
<td>Matched</td>
<td>Controls</td>
<td></td>
<td>Possible</td>
<td>Matched</td>
</tr>
<tr>
<td>Age, Wave</td>
<td>1.49</td>
<td>39.4</td>
<td>13.52</td>
<td>12.81</td>
<td>0.88</td>
<td>-18.8</td>
<td>13.46</td>
<td>13.73</td>
</tr>
<tr>
<td>Female</td>
<td>0.00</td>
<td>0.0</td>
<td>0.37</td>
<td>0.37</td>
<td>1.07</td>
<td>-23.7</td>
<td>0.18</td>
<td>0.33</td>
</tr>
<tr>
<td>White</td>
<td>0.00</td>
<td>6.3</td>
<td>0.37</td>
<td>0.37</td>
<td>0.32</td>
<td>-6.6</td>
<td>0.85</td>
<td>0.79</td>
</tr>
<tr>
<td>SES, Wave 1</td>
<td>0.23</td>
<td>-3.9</td>
<td>42.04</td>
<td>40.93</td>
<td>0.02</td>
<td>0.5</td>
<td>44.85</td>
<td>43.50</td>
</tr>
<tr>
<td>Public Assist.</td>
<td>0.95</td>
<td>-9.4</td>
<td>0.22</td>
<td>0.25</td>
<td>0.28</td>
<td>0.16</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>GPA</td>
<td>0.19</td>
<td>-5.4</td>
<td>3.67</td>
<td>3.70</td>
<td>0.97</td>
<td>22.2</td>
<td>3.61</td>
<td>3.64</td>
</tr>
<tr>
<td>Peer Influence ((X, \text{W1-4}))</td>
<td>0.33</td>
<td>9.3</td>
<td>13.64</td>
<td>13.30</td>
<td>1.03</td>
<td>-25.1</td>
<td>16.80</td>
<td>15.51</td>
</tr>
<tr>
<td>Rural environment</td>
<td>0.64</td>
<td>16.1</td>
<td>0.26</td>
<td>0.18</td>
<td>0.55</td>
<td>-12.3</td>
<td>0.18</td>
<td>0.26</td>
</tr>
<tr>
<td>Mean Offending, Waves 1-4</td>
<td>0.58</td>
<td>-18.7</td>
<td>0.15</td>
<td>0.10</td>
<td>2.3</td>
<td>1.15</td>
<td>1.04</td>
<td>1.14</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Covariate</th>
<th>Chronic</th>
<th></th>
<th>Mean of Possible Controls</th>
<th>Mean of Matched Controls</th>
<th>Combined</th>
<th></th>
<th>Mean of Possible Controls</th>
<th>Mean of Matched Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$</td>
<td>t</td>
<td>$</td>
<td>% Bias (matched)</td>
<td>Mean of Treated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, Wave</td>
<td>0.81</td>
<td>24.6</td>
<td>14.07</td>
<td>14.33</td>
<td>1.06</td>
<td>-16.1</td>
<td>13.55</td>
<td>13.62</td>
</tr>
<tr>
<td>Female</td>
<td>0.00</td>
<td>0.0</td>
<td>0.18</td>
<td>0.21</td>
<td>0.91</td>
<td>12.4</td>
<td>0.25</td>
<td>0.49</td>
</tr>
<tr>
<td>White</td>
<td>0.40</td>
<td>11.6</td>
<td>0.86</td>
<td>0.84</td>
<td>0.78</td>
<td>11.9</td>
<td>0.84</td>
<td>0.79</td>
</tr>
<tr>
<td>SES, Wave 1</td>
<td>0.10</td>
<td>-2.9</td>
<td>45.10</td>
<td>42.70</td>
<td>0.71</td>
<td>-11.2</td>
<td>44.33</td>
<td>42.97</td>
</tr>
<tr>
<td>Public Assist.</td>
<td>0.34</td>
<td>-10.2</td>
<td>0.23</td>
<td>0.19</td>
<td>0.27</td>
<td>-22.3</td>
<td>0.28</td>
<td>0.16</td>
</tr>
<tr>
<td>GPA</td>
<td>0.34</td>
<td>-11.6</td>
<td>3.36</td>
<td>3.36</td>
<td>0.67</td>
<td>10.5</td>
<td>3.59</td>
<td>3.81</td>
</tr>
<tr>
<td>Social Disorganization (x, W1-4)</td>
<td>0.10</td>
<td>3.1</td>
<td>10.25</td>
<td>9.62</td>
<td>0.75</td>
<td>11.2</td>
<td>9.93</td>
<td>9.53</td>
</tr>
<tr>
<td>Peer Influence (x, W1-4)</td>
<td>0.37</td>
<td>12.7</td>
<td>22.74</td>
<td>22.16</td>
<td>0.28</td>
<td>4.7</td>
<td>17.07</td>
<td>13.30</td>
</tr>
<tr>
<td>Rural environment</td>
<td>0.59</td>
<td>-13.6</td>
<td>0.05</td>
<td>0.24</td>
<td>0.20</td>
<td>-2.7</td>
<td>0.17</td>
<td>0.31</td>
</tr>
<tr>
<td>Mean Offending, Waves 1-4</td>
<td>0.64</td>
<td>21.1</td>
<td>4.23</td>
<td>3.59</td>
<td>0.71</td>
<td>12.3</td>
<td>1.48</td>
<td>0.49</td>
</tr>
</tbody>
</table>