Journal of Youth and Adolescence (2022) 51:724–745
https://doi.org/10.1007/s10964-022-01576-7

EMPIRICAL RESEARCH

Does Early Adolescent Arrest Alter the Developmental Course of Offending into Young Adulthood?

Bianca E. Bersani[1] · Wade C. Jacobsen[1] · Elaine Eggleston Doherty[2]

Received: 14 January 2022 / Accepted: 17 January 2022 / Published online: 5 February 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Adolescent involvement in risky behavior is ubiquitous and normative. Equally pervasive is the rapid decline in risky behavior during the transition to adulthood. Yet, for many, risky behavior results in arrest. Whereas prior research finds that arrest is associated with an increased risk of experiencing a host of detrimental outcomes, less understood is the impact of an arrest on the developmental course of offending compared to what it would have looked like if no arrest had occurred—the counterfactual. This study examines the developmental implications of an arrest early in the life course. The sample ($N =$ 1293) was 37% female, 42% non-white, with a mean age of 13.00 years (SD = 0.82, range = 12–14) at baseline and followed annually for 15 years. Analyses combine propensity score matching and multilevel modeling techniques to estimate the impact of early arrest (i.e., 14 or younger) on the development of offending from adolescence into adulthood. The results indicate that early arrest alters the developmental course of offending in two primary ways. First, early arrest heightens involvement, frequency, and severity of offending throughout adolescence and into early young adulthood even after controlling for subsequent arrests. The detrimental influence of early arrest on the developmental course of offending is found regardless of gender or race/ethnicity. Second, even among youth with an early arrest, offending wanes over time with self-reported offending among all youth nearly absent by the mid- to late-twenties. The findings advance understanding of the developmental implications of early arrest beyond typical and expected offending.

Keywords Early arrest · Developmental course of offending · Race and gender · Propensity score modeling · Counterfactual

Introduction

Involvement in risky behavior during adolescence and the transition to adulthood is a regular feature of the early life course (Moffitt, 1993:692; Steinberg, 2004:57) and yet, for many this behavior results in a formal response. Nearly 700,000 youth were arrested in 2019, and 34% of these arrests were of youth 14-years of age or younger (Puzzanchera & Hockenberry, 2021). Prior research finds that arrest during adolescence, particularly at young ages (Ge et al., 2003), may heighten involvement in offending in part by diminishing a youth's engagement in conventional activities like education (Kirk & Sampson, 2013) and employment (Lopes et al., 2012), weakening their social networks (Jacobsen et al., 2021), and increasing their surveillance (Liberman et al., 2014; Wiley, 2015). However, little is known about how early arrest alters the developmental course of offending (Huizinga & Henry, 2008). That is, does early arrest alter the curvature of the offending trajectory such as accelerating delinquent involvement during adolescence or stalling the decline in offending with age. Additionally, while research repeatedly shows that arrest is disproportionately and differentially experienced by persons in various racial, ethnic and gender groups (Crutchfield et al., 2012; Rengifo & Pater, 2017), few studies have tested for demographic differences in the consequences of early arrest or its developmental implications. This study contributes to the literature by examining if and how early arrest, defined as age 14 or younger, alters the developmental course of offending from early adolescence into

Supplementary information The online version contains supplementary material available at https://doi.org/10.1007/s10964-022-01576-7.

✉ Bianca E. Bersani
bbersani@umd.edu

[1] Department of Criminology and Criminal Justice, University of Maryland—College Park, College Park, MD, USA

[2] School of Public Health, University of Maryland—College Park, College Park, MD, USA

adulthood beyond normative involvement occurring from early to late adolescence (see Gottfredson & Hirschi, 2020:55; Moffitt, 1994:43). Among a sample of youth followed annually through young adulthood, matching methods are used to minimize concerns with selection into early arrest, and subsequent arrests are observed to control for later legal system involvement. Analyses are conducted on subsamples stratified by race/ethnicity and gender to assess demographic variation in the impact of early arrest.

The Normative Nature of Offending

Adolescence is a harbinger of dramatic, rapid, and distinctive change in the life course. During this time, youth experiment with identities, expand their social ties, and lay the foundation for future pursuits. A normative manifestation of this change and experimentation involves engaging in offending (Moffitt, 1994; Rutter et al., 1998). For instance, 95% of youth in the Oregon Youth Study reported at least one offending behavior from early adolescence into young adulthood (Wiesner & Capaldi, 2003). Whereas engagement in delinquency is so normal as to be expected, this behavior is typically temporary and limited to adolescence, with desistance occurring swiftly during the transition to adulthood even among high-risk youth (see Bersani & Doherty, 2018).

Taken together, the evidence that both involvement in offending during the early life course and a decline in offending beginning in young adulthood are normative and pervasive reveals an age-crime curve. Observed in the aggregate, the onset and intensification of criminal behavior occurs in early adolescence, peaking during the late teens and early twenties, and precipitously declining soon thereafter; a trajectory that is robust across time, place, and people (Gottfredson & Hirschi, 1990), at least when comparing trends in Western societies (Steffensmeier et al., 2017; but see Duell et al., 2018). Observed at the individual level, research finds evidence of heterogeneity in offending trajectories across the life course (see Piquero, 2008), but the evidence also suggests that behavioral trajectories of offending share a similar unimodal shape or distribution over age (Erosheva et al., 2014).

Early Arrest's Potential to Alter the Development of Offending

Research is replete with examinations of how salient life events can serve as turning points to alter an individual's expected age-crime curve (Nguyen & Loughran, 2018), with the bulk of this research focusing on the impact of events that typically occur in adulthood to examine pathways *out* of crime (Laub et al., 2019). Yet, notions of turning points, trajectories, and a focus on the implications of salient events on the developmental course of offending may be particularly fruitful when examining adverse events in childhood that may lead *into* a pathway of crime (Laub & Sampson, 2020:161; Widom et al., 2018). Although salient life events with the potential to function as negative turning points encompass a diverse range of experiences, formal responses such as arrest that risk public labeling and stigma may be particularly potent (Moffitt, 1993; Sampson & Laub, 1997). Whereas most youth engage in deviance, only a subset will experience a formal response to that deviance. Ample evidence finds that early criminal legal system contact is associated with the perpetuation of delinquency and arrest in later adolescence and adulthood (e.g., Bernburg & Krohn, 2003; Mowen et al., 2018), even after accounting for one's propensity to be arrested (e.g., Ward et al., 2014; Wiley & Esbensen, 2016).

Some scholars contend that these iatrogenic effects of early arrest increase the risk of subsequent delinquency by setting in motion a process of cumulative disadvantage whereby formal responses to deviant behavior serve to disintegrate conventional ties and opportunities (Sampson & Laub, 1997), weaken normative peer relationships (Jacobsen et al., 2021), and increase involvement with antisocial peers (Bernburg et al., 2006). It is possible that this process of cumulative disadvantage may affect not only the prevalence of offending but also the frequency, type (severity), and duration of offending. For instance, while de-escalation is characteristic of desistance (Le Blanc & Loeber, 1998), the experience of formal system contact may escalate the seriousness of offending overtime, even counteracting the desistance effect of aging (Liu et al., 2011). A key component of this process of cumulative disadvantage is an accumulation of subsequent arrests. An early arrest may alter the developmental course of offending in part because it pushes youth toward greater legal system involvement (more arrests). This is because an initial arrest facilitates persistence in offending and also augments the level of surveillance and policing that youth experience regardless of any change in offending (Goffman, 2009; Liberman et al., 2014).

Prior research on the developmental impacts of arrest

A substantial body of research utilizing longitudinal data and rigorous quantitative methods reveals a robust relationship between adolescent arrest and later offending (Bernburg, 2019). This evidence shows that the effect of adolescent arrest increases the short-term risk of offending (i.e., one year later) (Augustyn et al., 2019) as well as long-term risk extending into late adolescence (e.g., Wiley & Esbensen, 2016) and adulthood (e.g., Liberman et al., 2014). The adverse effects of the legal system may be most acute for one's development when contact occurs early in

the life course, prior to age 15 (Ge et al., 2003; Wiley, 2015).

Minimizing selection effects is a critical feature of this line of research because associations between early arrest and negative outcomes may be due to preexisting differences in the likelihood of offending between youth who experience an early arrest and those who do not (e.g., Morris & Piquero, 2013; Ward et al., 2014). For example, research has identified numerous risk factors for arrest including individual characteristics such as impulsivity (e.g., Beaver et al., 2009), early antisocial behavior, and academic failure (e.g., Yun et al., 2014); family factors including poor parental monitoring, family structure and low socioeconomic status (e.g., Kirk, 2009); and peer characteristics such as associations with delinquent peers and gang membership (e.g., Tapia, 2011). Moreover, contextual factors such as social control deficits at school and neighborhood disorder also increase the risk of arrest in adolescence (Kirk, 2009). Collectively, the evidence suggests that the consequences of juvenile arrest are not simply due to differences across youth in the propensity to be arrested (e.g., Wiley & Esbensen, 2016; Liberman et al., 2014).

Whereas scholarship examining the offending implications of police contact is methodologically rigorous and uses longitudinal data, certain features limit the extent to which it can shed light on the developmental consequences of early arrest. First, the bulk of prior research uses survey data with static or point-in-time analytic designs, observing the effect of arrest at one point in time (time 1) on outcomes at a later point in time (time 2). Second, most datasets lack consistent and annualized measures of offending over a significant portion of the life course spanning both adolescence and adulthood, limiting the ability to examine incremental variation in the developmental course of offending across multiple phases of the life course. This body of research suggests that juvenile arrest may increase offending at a later point, but the research does not speak to the impact of early arrest on the developmental unfolding of offending over the life course. Finally, studies that rely on offender-based samples or administrative data contain information on "treated" cases only; those with justice-system contact. Examining developmental trajectories among justice-involved samples allows for the observation of patterns of officially recorded crime but provides little insight into patterns of offending absent early arrest. For example, a longitudinal study with annual arrest rates spanning 20 years examined within-individual cumulative arrest trajectories and found that early arrest was associated with the accumulation of more arrests at a faster pace (Natsuaki et al., 2008). Yet, because the sample was comprised entirely of system-involved individuals, it is unclear what the effect of an early arrest was beyond typical involvement in crime observed over the life course. Moreover, the use of administrative data reveals patterns of legal system involvement or criminal justice careers (see Bushway & Tahamont, 2016:375) but is unable to address patterns of offending for behaviors that occur outside the purview of the legal system.

Recent research has begun to grapple with whether and how early arrest reshapes offending trajectories. Focusing on adolescent development over a four-year period, arrest was found to increase within- and between-individual levels of delinquency with detriments accumulating over time (Mowen et al., 2018). Other studies have employed trajectory analysis to distinguish youth based on their self-reported offending and then tested for the effect of arrest within these groups. Thus far, this research has produced mixed evidence, with one study suggesting that arrest is most detrimental for those in a high-risk offending trajectory group (Morris & Piquero, 2013) and another study suggesting that the effect of arrest was most pronounced for those in a low violent-offending trajectory group (Ward et al., 2014). Trajectory analysis studies have observed the development of offending, but their purpose was to define arrest trajectories within predetermined groups rather than to test whether and how early arrest alters offending trajectories.

Early arrest, race/ethnicity, and gender

While research examining the outcomes of adolescent arrest has advanced rapidly in recent years, few studies have tested for demographic differences in the consequences of early arrest or its developmental implications. Theories offer competing hypotheses about who will bear the greatest burden of arrest. In terms of race/ethnicity, one perspective suggests that criminal legal system involvement is more harmful for people of color due to processes of labeling and cumulative disadvantage whereby involvement in the legal system exacerbates disadvantages and inequalities that are present prior to legal system involvement (Bernburg & Krohn, 2003; Lopes et al., 2012). An alternative perspective argues that legal system contact is more detrimental for Whites (Chiricos et al., 2007) because the sheer prevalence of legal system presence in communities of color makes it a normalized life experience (Pettit & Western, 2004) or part of a regular routine (Jones, 2014). The logic follows that arrest may not create a meaningful social reaction as more people within a racial/ethnic minority experience it (Hirschfield, 2008). A recent study finds that an early arrest poses a significant detriment to Black youth whose odds of an adult arrest were eleven times greater than their non-early arrest Black peers (McGlynn-Wright et al., 2020). In contrast, detriments of early arrest were not found for White youth.

Theories also offer competing hypotheses about the effect of criminal legal system contact when considering

gender. While males have long accounted for most of the justice-involved youth population, young females have an increasing presence in the legal system (Puzzanchera & Hockenberry, 2021) which may be the result of heightened policing and sanctioning of girls, particularly Black girls (Stevens et al., 2011). Whereas some research finds that women are afforded leniency in legal system contact (e.g., Stolzenberg & D'Alessio, 2004) and may be informally diverted from the system, other research suggests that women may receive more severe treatment if they are viewed as doubly deviant, violating both crime and conventional gender norms (e.g., Chesney-Lind & Eliason, 2006).

Current Study

Ample evidence finds that involvement in risky behavior is a routine facet of normal adolescent development. At the same time, arrest can be detrimental to later life course outcomes. What remains unknown are the developmental implications of early arrest—if and how the experience of arrest reshapes the offending trajectory by accelerating criminal behavior in adolescence or disrupting desistance processes in young adulthood. The goal of the current research is threefold. First, it situates the effects of arrest in a developmental context by examining whether (e.g., prevalence), in what ways (e.g., frequency, variety), and how (e.g., accelerated and/or prolonged rate of change) early adolescent arrest alters the developmental course of offending. Does early arrest alter the prevalence, frequency, and variety of offending (Research Question 1)? Does early arrest alter the rate of change in offending over time (Research Question 2)? Second, this study accounts for processes of cumulative disadvantage that an early arrest may spur, by controlling for the accumulation of subsequent arrests that youth experience. Does the effect of early arrest on the developmental course of offending remain after accounting for the accumulation of subsequent arrests over the life course (Research Question 3)? Third, this research tests for demographic variation in the effect of early arrest by comparing analyses across race/ethnicity and gender subsamples. Does the effect of early arrest on the developmental course of offending vary by race/ethnicity or gender (Research Question 4)? To minimize the influence of between-individual differences associated with early arrest, 42 covariates spanning a wide variety of domains of risk associated with offending such as features of the youth's family/household, school and neighborhood environment, as well as key correlates of delinquency, such as early risk behaviors (delinquent acts, gang involvement, substance use) and antisocial peers, are included in propensity score models.

Methods

Data

Data were derived from the National Longitudinal Survey of Youth (NLSY97), a representative household sample of people living in the United States in 1997 who were born from 1980 through 1984 and were 12 to 17 years of age at baseline ($N = 8984$; Bureau of Labor Statistics, U.S. Department of Labor 2019). Respondents complete a self-administered survey that collects information on sensitive topics that reflect antisocial behavior including crime, delinquency, and legal system contact. Youth were interviewed on an annual basis beginning in 1997 to 2011, and biennially thereafter. Data collection is currently ongoing. The study uses a total of 14 waves of data (1997 to 2011; excluding the 2004 wave, see analytic sample inclusion criteria discussion below) capturing the span of the life course from early adolescence to the late-20s. The selection of waves was influenced by a change to the questionnaire that reduced the number of offending questions asked after the 2011 wave. Descriptive inquiry into the data reveals that beyond this wave respondents self-report very little involvement in offending. Annual retention rates are above 80% for all waves included in this study. Two-thirds ($n = 849$) of the analytic sample had complete data at all waves and an additional quarter ($n = 339$) were missing data in three or fewer waves.

The analytic sample is a subsample of the NLSY97 data defined by two inclusion criteria. First, to isolate early life-course legal system contact, the sample was restricted to youth ages 12 (the youngest) to 14 at baseline ($n = 5419$; age as of 12/31/1996). Age 14 was selected because it distinguishes early- from mid-adolescence (Association of Maternal and Child Health Programs, n.d.) and is recognized as distinguishing early versus late onset offending (e.g., Patterson & Yoerger, 2002; Moffitt, 1993). Second, to focus on the development of self-reported offending over the life course, the sample was restricted to youth who were continually queried about their offending across all study waves. In 2005, the NLSY97 began restricting the collection of data on self-reported offending to a 10% randomly selected subsample as well as any respondent who previously reported an arrest at or after wave 4.[1] This leaves an analytic sample of 1293 respondents (13,059 person-year observations).

[1] The 2004 wave (round 8 interview) is dropped from the analysis. The random subsample responding to self-reported offending questions in this round was reselected in 2005; the sample selected in 2005 continues to answer self-reported offending questions in subsequent waves (confirmed via personal correspondence with NLS User Services). Thus, we use the 2005 subsample to define our analytic sample across all waves.

Because arrest is one defining feature of the longitudinal subsample, the analytic sample reflects a more delinquent subsample than the original NLSY97. Indeed, 60% of the analytic sample self-reported a delinquent act at baseline (1997) compared to 41% of the same-aged full sample. Although the selection criteria reduce the generalizability of the results back to the full sample, the difference in offending levels between the full and analytic samples does not impact the suitability of the modeling approach as the core concern is establishing equivalent comparison groups based on baseline delinquency and other covariates in the analytic sample.

These data are well suited to study linkages between criminal legal system contact and trajectories of offending for several reasons. First, the age distribution and length of follow-up allows for the observation of individuals during adolescent onset and peak years of offending into young adulthood, when most desist from offending (Doherty & Bersani, 2018). Second, self-reported offending and legal system contact are measured in every wave, allowing for the plotting of age-crime curves using annual data covering 15 years. Third, the data capture a wide array of information across key developmental domains including individual, family, peers, school, and neighborhood factors. As such, these characteristics and known correlates of offending can be used to match early-arrest and non-early-arrest youth on their likelihood of experiencing an early arrest. Finally, though the NLSY97 is a general population, household sample (i.e., not offender-based), this sample reflects a wide variety of offending behavior, including acts that are serious in nature. For instance, Brame et al. (2014) found that, on average, a fifth of youth in the NLSY97 reported at least one arrest by age 18, many of whom have also been incarcerated (Apel, 2016).

Measures

To ensure the correct temporal ordering of study variables, early arrest, the focal explanatory variable, was measured at waves 1, 2, and 3 (years 1997, 1998, and 1999) and the outcome variable, self-reported offending, was measured at waves 4 through 15 (years 2000 to 2011). All matching covariates used in the propensity score matching (described later) were measured at wave 1. Each wave reflects a 3-year age span such that by the final wave, respondents were 26–28 years of age (Fig. S1).

Self-reported offending

The dependent variable is a time-varying indicator of self-reported offending. In each wave, respondents were asked if they were involved in a series of delinquent/criminal acts since the date of the last interview including: purposely damaged or destroyed property; stolen something worth less than 50 dollars; stolen something worth 50 dollars or more including stealing a car; committed other property crimes; attacked someone with the idea of seriously hurting them or had a situation end up in a serious fight or assault of some kind; sold or helped to sell drugs. With this information, three outcome measures were constructed. First, a binary indicator measured the prevalence of offending; coded 1 at waves in which respondents reported engaging in any of the behaviors and 0 at waves in which they reported no delinquent behavior. Second, frequency of offending was measured by summing responses to the question asking respondents to report the number of times they were involved in each of the six offense types, since the date of the last interview. Item responses ranged from 0 to 999 for drug sales and 0 to 99 for all other items. Because drug sales had a distinctly higher frequency than the other five items the drug sales item was censored to 99 in each wave to ensure that results were not driven by drug sales. Third, a variety score represented the sum of a set of binary variables, each indicating whether the respondent engaged in a given offense type. Higher values represent a greater diversity of offending behavior in each wave (0 = no delinquent involvement; 6 = engaged in all six behaviors). Variety scores are highly correlated with measures of seriousness and frequency (Monahan & Piquero, 2009) yet are less sensitive to the influence of high frequency non-serious behaviors (Sweeten, 2012) allowing for the examination of heterogeneity in offending as it relates to the variety of offending.

Early arrest

The key independent variable is a binary indicator of early arrest at age 14 or younger. In each wave, respondents were asked to report their history of arrest (arrested by police or taken into custody for an illegal or delinquent offense, excluding minor traffic violations) since the last interview. Those who were age 14 or younger in waves 1, 2, or 3, who reported an arrest in any of those waves were coded 1, and others of the same ages with no arrests in those waves were coded 0. Although 60% (n = 770) of the sample self-reports involvement in at least one delinquent act at baseline, only a quarter of the sample (n = 309) reported experiencing an early arrest. It is important to note that a few of these early-arrest respondents (n = 29) experienced an arrest a year or more prior to baseline, and some of these (n = 11) were arrested at or before age 10. Sensitivity analyses (described below) were conducted to examine the influence of these pre-baseline arrests on the findings.

Cumulative arrest

For waves 4 through 15, a time-varying cumulative arrest measure was calculated where the total number of self-reported arrests was added to the cumulative sum from the previous wave. Thus, the measure of cumulative arrest at the first subsequent wave (wave 4, year 2000) is equal to the number of arrests since the date of the previous interview, and at each subsequent wave the cumulative arrest number either increases or remains the same. At the last wave in the analyses (wave 15, year 2011), the cumulative arrest measure is equal to the total number of arrests the respondent experienced after wave 3. The measure of cumulative arrest is used as a time-varying control variable to test the influence of early adolescent arrest net of the accumulation of arrests on self-reported offending in subsequent waves. At wave 4, the mean cumulative arrest is 0.36, and by wave 15, it reaches 2.35 and ranges from 0 to 61 total arrests. Mean cumulative arrest reaches 1.91 among the non-early arrest group and 3.89 among the early arrest group.

Matching covariates

The breadth of information collected in the NLSY97 was leveraged by utilizing 42 matching covariates. Specific measurement definitions and response categories are shown in Table 1, organized into five domains: demographics, family, school, adolescent (risk), and neighborhood. All covariates were measured at wave 1.

Demographics Respondent gender, race, socioeconomic indicators of poverty ratio and net worth (NLSY97 created variables, Bureau of Labor Statistics, U.S. Department of Labor 2019), parental education level, mother's age at first birth, and family size.

Family The family domain includes structure and process measures. A binary indicator for family structure distinguished youth living with both biological parents from other living situations. Indicators of parental religiosity, positive outlook, and monitoring capture family processes (NLSY97 created scales, Bureau of Labor Statistics, U.S. Department of Labor 2019).

School The school domain includes indicators of school type, attendance, academic success, victimization and fear of safety at school, and attachment. Binary indicators distinguish the type of school (public versus other) and if a respondent ever attended a Headstart program. Attendance is measured using indicators of times tardy, absent, and school suspension. Grade repetition and Peabody Individual Achievement Test capture elements of academic success. Victimization experiences include experiencing theft, threat of aggression and/or physical fight at school, and fear for one's safety at school. Youth also reported their perception of the quality of teachers.

Adolescent (risk) The adolescent behavior domain includes indicators for the focal youth as well as their peers. Indicators capture a youth's history of running away from home, gang membership, delinquent behavior for property and person offenses, drug sales, and substance use. Youth reported the average percent of their peers involved in antisocial behavior (e.g., substance use, cutting classes) and the percent of their peers involved in prosocial activities (e.g., school clubs, sports, volunteer work). The behavior and emotional problem scale captures frequency of problem behaviors (NLSY created scale; Bureau of Labor Statistics, U.S. Department of Labor 2019).

Neighborhood To address contextual features, indicators of environmental and home risks as well as vicarious and direct experiences with crime are included. Physical environmental risk (e.g., absence of electricity and/or heat in the past month), enriching environment (e.g., dictionary and/or computer in the home), and home risk (e.g., access to basic needs, family routines) measure proximal and distal contexts (NLSY97 created scales). Youth report how often they hear gun shots in their neighborhood, witness gun violence, their history of household break-in, and any experiences with repeated bullying in childhood. Finally, a binary variable distinguishes urban from rural neighborhoods.

The majority of the 42 covariates were missing fewer than 5% of observations. Covariates with the most missing cases include whether the respondent repeated a grade (13%), net worth (23%), and income-to-poverty ratio (23%). Multiple imputation with chained equations was used and the analyses were performed using 20 imputed datasets (Allison, 2000) to check for consistency in outcomes. Time-varying self-reported offending and arrest measures were included in the imputations but the analyses were limited to observations with valid data on these outcome measures (von Hippel, 2007).

Analytic Plan

To examine associations between early arrest and subsequent development of offending, the analytic plan relied on a combination of propensity score matching (PSM) and multilevel models (MLM). First, PSM was used to minimize differences between early-arrest youth and non-early-arrest youth. PSM approximates a simulated random design. Early-arrest youth comprise the treatment group and non-early-arrest youth represent the control group. The procedure involved calculating a propensity score representing the predicted probability of

Table 1 Matching covariates measured at baseline (1997) when respondents were ages 12–14

Measure	Description
Demographics	
Sex[a]	A binary variable: male, female
White[a]	A binary variable: non-Hispanic white, other
Black[a]	A binary variable: non-Hispanic Black or mixed race, other. Note, fewer than 1% of cases were mixed race
Hispanic[a]	A binary variable: Hispanic, other
Poverty Ratio	NLSY created variable[b]: Ratio of household income to the poverty level
Net worth	NLSY created variable: household net worth according to the parent. Divided into tertiles (three categories, with lowest as reference)
Teenage at Birth	Mother was 19 years of age or younger at the birth of the respondents (0 = no; 1 = yes)
Parental Education	Mean of residential mother and father education levels (based on highest grade completed, 0 = none to 20 = 8 or more years of college; dichotomized (0 to 12 year = 0, more than 12 years = 1)
Family Size	NLSY created variable: The number of children 18 or younger living in the household (range 1–9)
Family	
Family Structure	NLSY created variable: child lives with two biological parents (0 = no; 1 = yes)
Religiosity	NLSY created variable: parent religiosity scale, higher scores indicate greater religiosity (divided by 100; range 0–6)
Positive Outlook	Four item mean scale: Parent expects the best, rarely count on good things happening (reverse coded), always optimistic about the future, hardly ever expects things to go my way (reverse coded); range 1–4
Monitoring	NLSY created variable: Composite score of mother reported monitoring where higher scores indicate greater parental monitoring (range 0–16)
School	
Public School	A binary variable: public, other
Head Start	Ever attended an official, government sponsored Headstart program (0 = no, 1 = yes)
Repeat	Ever repeated any grade (0 = no, 1 = yes)
Tardy	Number of times late for school without an excuse in the fall of the current school year (range 0–99)
Absent	Number of times absent from school in the fall of the current school year (range 0–90)
Suspension	Ever suspended from school (0 = no, 1 = yes)
PIAT Score	NLSY created variable: Peabody Individual Achievement Test normed math score (range 0–100)
School Victimization	Ever had something stolen, threatened to be hurt, and/or got into a physical fight at school (0 = no, 1 = yes)
Fear at School	Feels safe at school (0 = Strongly agree/agree; 1 = Strongly disagree/disagree)
Teacher Attachment	Respondent perceives teachers are good (1 = strongly agree, 4 = strongly disagree)
Adolescent (Risk)	
Antisocial Peers	Average percent of peers in your grade/last grade enrolled in school who engaged in: smoking, getting drunk, belonging to a gang, using illegal drugs, cutting classes or school, having sex (range 1–5)
Prosocial Peers	Average percent of peers in your grade/last grade enrolled in school who engaged in: going to church regularly, involved in sports, clubs, school activities, plan to go to college, doing volunteer work (range 1–5)
Behavioral and Emotional Problem Scale	NLSY created variable: higher scores indicate more frequent and/or numerous behavior problems (range 0–8)
Runaway	Ever runaway, that is, left home and stayed away at least overnight without your parent's prior knowledge or permission (0 = no, 1 = yes)
Gang	Ever belong to a gang (0 = no, 1 = yes). Gang is defined as a group that hangs out together, wears gang colors or clothes, has set clear boundaries of its territory or turf, protects its members and turf against other rival gangs through fighting or threats

Table 1 (continued)

Measure	Description
Property	4 items: ever destroy property, steal anything worth $50 or less, steal anything worth $50 or more, other property crimes. Additive scale of the count of different property crimes (0 = none, 1, 2, 3 or more)
Violence	Ever attacked someone with the intent to seriously hurt them or have a situation end up in a serious fight or assault of some kind (0 = no, 1 = yes)
Sell drugs	Ever sold or help sell marijuana, hashish, or other hard drugs such as heroin, cocaine, or LSD (0 = no, 1 = yes)
Smoking	Ever smoked a cigarette (0 = no, 1 = yes)
Alcohol Use	Ever had a drink of an alcoholic beverage (0 = no, 1 = yes). Drink defined as a can or bottle of beer, a glass of wine, a mixed drink, or a shot of liquor. Not including childhood sips they might have had from an older person's drink.
Marijuana Use	Ever used marijuana, for example: grass or pot (0 = no, 1 = yes)
Neighborhood	
Physical Environment Risk	NLSY created variable: absence of electricity and heat in past month, dilapidated housing in neighborhood, interviewer reports feeling concerned for safety in neighborhood/home. Higher scores indicate a higher physical environment risk (scale divided by 100; range 0–7)
Enriching Environment	NLSY created variable: presence of a computer, presence of a dictionary, involvement in extracurricular activities. Higher scores indicate a more enriching environment (range 0–3)
Urban	NLSY created variable: Residence in an urban or a rural area as of the survey date (0 = rural, 1 = urban)
Gun Shots	In a typical week, how many days do you hear gunshots in your neighborhood. (0 = 0) (1 to 7 = 1)
Witness	Before you turned 12, did you ever see someone get shot or shot at with a gun (0 = no, 1 = yes)
Neighborhood victimization	Before you turned 12, did you ever have your house or apartment broken into (0 = no, 1 = yes)
Bullied	Before your turned 12, were you ever the victim of repeated bullying (0 = no, 1 = yes)
Home Risk	NLSY created variable: combining respondent, parent, and interviewer assessments of the youth's physical home environment and neighborhood including access to basic needs, adequate housing, study materials, family routines. Higher scores indicate a higher home risk environment (scale divided by 100; range 0–15)

[a]These variables were used as matching criteria in the full sample analyses. In subsample analyses for gender, race/ethnicity (white/nonwhite) was a matching covariate; in subsample analyses for race/ethnicity, gender was a matching covariate. Bold covariates were included in propensity score matching for gender and race subsample analyses

[b]NLSY97 contains a number of created variables derived from different survey items, many computed by area experts (e.g., Child Trends, Inc.) who assess data quality, reliability and validity of scales and indexes (NLSY97 Codebook Supplement)

experiencing an early arrest given the observed covariates, and then matching control with treatment cases on this propensity score (Apel & Sweeten, 2010a; Guo & Fraser, 2014; Rosenbaum & Rubin, 1983). A key advantage of this approach is the ability to "balance" or "equate" treatment and control observations on a large number of covariates, prior to regression modeling with the outcome variable. It also relaxes assumptions about the functional form of associations between the covariates and the probability of treatment. Furthermore, it allows for the exclusion of cases where there is inadequate overlap between treatment and control groups, thus making comparisons more appropriate (Stuart, 2010). Like standard regression methods, PSM does not account for all unobserved differences between treatment and control cases but concerns with selection bias are minimized with the inclusion of many observed covariates.

For each imputed dataset, the propensity score was constructed with a logistic regression model in which the treatment variable (early arrest) was the outcome variable and the matching covariates were explanatory variables. In the first imputed dataset, the median propensity score for the treatment group was 0.40 (range 0.03 to 0.99), and the median for the control group was 0.13 (range 0.01 to 0.94). Distributions were similar across each imputed dataset. The particular PSM strategy was kernel matching (kernel = Epanechnikov; bandwidth = 0.06), which uses all control cases that fall within a given bandwidth of the propensity score for each corresponding treatment case. Control cases are weighted based on their distance from the treatment case, with greater weight given to closer matches. We restrict analyses to regions of common support, resulting in the loss of a few treatment cases (ranging from 6 to 11 cases across imputed datasets; 1.9% to

3.6% of the arrested subsample). This strategy retains a larger number of cases than other matching strategies, but consistency was also checked using nearest-neighbor matching[2] in sensitivity analyses and results were consistent across matching strategies.

After PSM, analyses used a series of MLMs to examine differences between matched groups in their trajectories of offending over time. MLMS are particularly powerful for accommodating unbalanced data (when the number and spacing of observations varies across respondents) and repeated-measures data that violate the independence assumption of ordinary least-squares regression (see Osgood, 2010; Raudenbush & Bryk, 2002). For examining differences between treatment and control groups, models include the kernel weights in the regressions, which weighted control cases by their distance from treatment observations. The examination of individual change involved two levels of analysis. The analytic details here correspond to negative binomial regression (used for the frequency and variety of offending outcomes). The basic elements (i.e., growth parameters, treatment variable, cumulative arrest, random effects) for logistic regression (used for the prevalence of offending outcome) were consistent across analyses. Level 1 captured within-individual change over time and estimated each person's development with a unique growth trajectory. The basic elements of the level 1, within-individual model are:

$$\begin{aligned} \log\left(l_{ij}\right) &= \eta_{it} \\ \eta_{it} &= \pi_{0i} + \pi_{1i}\mathrm{Age}_{it} + \pi_{2i}\mathrm{Age}_{it}^2 + \mathrm{e}_{it}, \end{aligned} \tag{1}$$

where η_{it} is the expected crime count for individual i at age t, is equal to the intercept (π_{0i}) plus the growth parameters (π_{1i} and π_{2i}) and the residual (e_{it}). The intercept is the estimated number of (count for frequency, types for variety) offenses when age is set to zero. For this study, the growth trajectory followed a quadratic function of age (age_{it}, $\mathrm{age}^2{}_{it}$). We tested for fit with a cubic age term; the addition of this parameter did not improve the model fit (based on the likelihood ratio test) nor was the term significant indicating that a quadratic functional form was sufficient. The age and age^2 parameters estimate the average rate of change and rate of acceleration or deceleration, respectively (Raudenbush & Bryk, 2002:163). To allow for more stable estimation due to collinearity between age terms, age and age^2 were grand mean centered. Grand mean centering yields an interpretation of π_{1i} as the average rate of growth across the entire observation period. In the models, age and age^2 were then group mean centered. A key advantage of group mean centering prior to modeling is that it separates the within-individual change from any differences due to cohort or differential attrition (see Jacobs et al., 2002:515; Bryk & Raudenbush, 1992). The parameters from level 1 become the outcome variables in level 2 of the model:

$$\pi_{0i} = \beta_{00} + \beta_{01}(\text{early arrest}) + \mathrm{r}_{0i} \tag{2}$$

$$\pi_{1i} = \beta_{10} + \beta_{11}(\text{early arrest}) + \mathrm{r}_{1i} \tag{3}$$

$$\pi_{2i} = \beta_{20} + \beta_{21}(\text{early arrest}) \tag{4}$$

To test whether the development of offending over time differed as a function of early arrest, the basic level 2 model was expanded to include a series of cross-level interactions as shown in Eqs. (2)–(4) that add the propensity score matched early arrest variable to the intercept and slope parameters. The coefficients for early arrest represent the difference in the expected crime count (β_{01}), slope (β_{11}), and curvature (β_{21}). Models were specified to allow for variation between individuals in the offending (π_{0i}) and age (π_{1i}) parameters as indicated by the error terms r_{0i} and r_{1i}. Further, to isolate the contribution of early arrest to the development of offending, beyond continued legal system engagement, Eq. (1) was augmented to include the time-varying cumulative arrest measure:

$$\eta_{it} = \pi_{0i} + \pi_{1i}\mathrm{Age}_{it} + \pi_{2i}\mathrm{Age}_{it}^2 + \pi_{3i}\mathrm{Cumulative\ Arrest}_{it} + \mathrm{e}_{it} \tag{5}$$

Results

Covariate Balance

Because early-arrest and non-early-arrest youth differ not only in their risky behavior but also on other characteristics, the aforementioned differences may be biased by selection whereby the likelihood of experiencing an early arrest is driven by pre-existing differences in the propensity to be arrested. These differences are minimized in the propensity score matching, as shown in the covariate balance statistics in Table 2. The means of each covariate are compared across treatment and control groups before and after matching on the propensity score. Columns labeled "Unadjusted" compare the treatment and control groups prior to matching. The unadjusted values in Table 2 show the descriptive composition of the sample separated by early-arrest and non-early arrest groups. Looking at demographic composition, both groups of youth are majority male (72%) and nearly half are non-white. Prior to matching, between group comparisons across all domains show that the early-arrest group was more likely to be male, Black, disadvantaged, live in risky environments, experience school challenges, and engage in more delinquent activity than

[2] This alternative approach involved assigning a given number of control observations (neighbor = 10) to the treatment case with the closest propensity score.

Table 2 Descriptive summary statistics and covariate balance before and after propensity score matching

Covariate	Imputed data set = 1						2	3	4	5	19	20
	Unadjusted		Adjusted		SB	*p*	*p*	*p*	*p*	*p*	*p*	*p*
	Treated	Control	Treated	Control								
Demographics												
Male	0.728	0.600***	0.728	0.735	−1.3	0.861	0.960	0.836	0.893	0.634	0.899	0.942
Black	0.338	0.260**	0.338	0.349	−2.4	0.775	0.824	0.887	0.938	0.925	0.991	0.959
Hispanic	0.159	0.215*	0.159	0.146	3.3	0.657	0.814	0.880	0.878	0.906	0.816	0.921
Poverty Ratio	224.818	275.343**	224.820	212.560	4.8	0.534	0.620	0.748	0.663	0.879	0.710	0.735
Net Worth 2	0.331	0.334	0.331	0.314	3.6	0.654	0.690	0.595	0.862	0.703	0.911	0.826
Net Worth 3	0.298	0.399**	0.298	0.301	−0.7	0.933	0.654	0.914	0.757	0.888	0.585	0.537
Teenage at Birth	0.192	0.125**	0.192	0.178	4.0	0.648	0.532	0.503	0.623	0.560	0.835	0.513
Parental Education	0.404	0.468*	0.404	0.373	6.3	0.435	0.298	0.600	0.443	0.417	0.399	0.416
Family Size	2.583	2.529	2.583	2.575	0.6	0.938	0.987	0.666	0.768	0.808	0.849	0.858
Family												
Family Structure	0.325	0.491***	0.325	0.329	−0.9	0.912	0.727	0.978	0.820	0.836	0.775	0.881
Religiosity	3.592	3.640	3.592	3.627	−2.2	0.785	0.850	0.734	0.979	0.693	0.656	0.889
Positive Outlook	2.812	2.921**	2.812	2.821	−1.7	0.832	0.739	0.869	0.947	0.548	0.673	0.674
Monitoring	9.182	9.985***	9.182	9.269	−2.6	0.755	0.534	0.973	0.395	0.359	0.623	0.666
School												
Public School	0.934	0.928	0.934	0.944	−4.1	0.591	0.624	0.750	0.548	0.680	0.631	0.553
Head Start	0.301	0.206***	0.301	0.314	−2.9	0.741	0.693	0.974	0.883	0.812	0.826	0.740
Repeated Grade	0.225	0.153**	0.225	0.275	−12.7	0.158	0.485	0.436	0.462	0.590	0.542	0.346
Tardy	3.096	1.426***	3.096	2.543	7.3	0.322	0.459	0.449	0.264	0.525	0.580	0.667
Absent	5.904	3.902***	5.904	5.735	1.9	0.764	0.716	0.902	0.741	0.646	0.933	0.960
Suspension	0.589	0.300***	0.589	0.595	−1.2	0.887	0.856	0.830	0.780	0.841	0.941	0.824
PIAT Score	31.858	38.009***	31.858	32.525	−2.6	0.743	0.977	0.770	0.715	0.958	0.978	0.961
School Victimization	0.705	0.482***	0.705	0.709	−0.7	0.925	0.895	0.940	0.849	0.962	0.896	0.971
Fear at School	0.258	0.149***	0.258	0.262	−1.0	0.913	0.815	0.924	0.798	0.985	0.858	0.857
Teacher Sttachment	2.119	1.894***	2.119	2.100	3.2	0.699	0.875	0.839	0.808	0.954	0.697	0.877
Adolescent (Risk)												
Behavioral and Emotional Problems	2.861	2.276***	2.861	2.812	3.0	0.704	0.960	0.888	0.894	0.824	0.932	0.924
Antisocial Peers	2.301	1.865***	2.301	2.339	−4.1	0.624	0.646	0.592	0.690	0.884	0.629	0.637
Prosocial Peers	2.963	3.077*	2.963	2.989	−3.4	0.681	0.956	0.976	0.842	0.765	0.683	0.762
Runaway	0.242	0.080***	0.242	0.194	13.2	0.155	0.559	0.547	0.529	0.307	0.397	0.446
Gang	0.159	0.046***	0.159	0.159	0.0	0.999	0.982	0.686	0.927	0.541	0.982	0.961
Property	1.464	0.744***	1.464	1.414	4.9	0.584	0.544	0.517	0.609	0.597	0.469	0.484
Attack	0.454	0.173***	0.454	0.444	2.1	0.816	0.568	0.639	0.846	0.778	0.733	0.928
Sell Drugs	0.182	0.040***	0.182	0.191	−3.0	0.771	0.946	0.472	0.775	0.974	0.977	0.761
Smoking	0.616	0.351***	0.616	0.608	1.6	0.848	0.917	0.924	0.847	0.664	0.825	0.865
Alcohol Use	0.556	0.364***	0.556	0.554	0.4	0.960	0.833	0.687	0.701	0.913	0.667	0.794
Marijuana Use	0.368	0.122***	0.368	0.388	−5.0	0.600	0.599	0.551	0.595	0.662	0.775	0.681
Neighborhood												
Urban	0.801	0.752	0.801	0.773	6.8	0.394	0.991	0.607	0.901	0.668	0.826	0.498
Physical Environment Risk	1.880	1.457***	1.880	1.858	1.5	0.869	0.933	0.951	0.643	0.917	0.847	0.885
Enriching Environment	1.526	1.704***	1.527	1.489	4.9	0.552	0.307	0.469	0.492	0.353	0.334	0.276
Gun Shots	0.291	0.239	0.291	0.278	3.1	0.709	0.821	0.901	0.965	0.996	0.855	0.860
Witness	0.238	0.129***	0.238	0.259	−5.3	0.567	0.337	0.524	0.615	0.921	0.420	0.570

Table 2 (continued)

Covariate	Imputed data set = 1						2	3	4	5	19	20
	Unadjusted		Adjusted		SB	*p*	*p*	*p*	*p*	*p*	*p*	*p*
	Treated	Control	Treated	Control								
Neighborhood Victimization	0.252	0.163***	0.252	0.264	−3.2	0.719	0.865	0.598	0.928	0.919	0.773	0.742
Bullied	0.334	0.234***	0.334	0.335	−0.1	0.990	0.848	0.980	0.958	0.914	0.764	0.983
Home Risk	3.958	2.891***	3.958	3.900	2.4	0.779	0.826	0.808	0.767	0.766	0.985	0.896
Mean SB						3.3	2.9	2.9	2.8	2.8	2.8	2.8

National Longitudinal Survey of Young, 1997 Cohort. Propensity score matching based on kernel matching. For the first imputed dataset, $n = 302$ treatment cases and $n = 984$ control cases. Asterisks represent statistical significance for *t*-tests of unadjusted data

SB standardized bias, $p = p$ value for independent samples *t*-test between treatment and control groups

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$ (two-tailed)

their non-early-arrest counterparts. Although early arrest youth have higher rates of involvement in risky behavior (e.g., running away, association with antisocial peers, gang affiliation, property and person delinquency, and drug sales/use) they do not represent a distinct group characterized by early violent behavior (e.g., life course persister, Moffitt, 1993). Both samples have a majority of youth who have not engaged in more serious or violent forms of risky behavior that would be expected in non-normative offending histories, such as attacking someone or having prior gang affiliation.

Columns labeled "Adjusted" show these comparisons after matching. Whereas youth in the treatment and control groups differed significantly on nearly every covariate prior to matching, these differences are minimized after matching, resulting in groups that are statistically similar on each of the observed covariates within each of the 20 multiply imputed datasets. Additionally, levels of standardized bias (SB) were within the conventional standard of SB < 20 (Rosenbaum & Rubin, 1985). As a sensitivity check, in the regression models described later, the principle of double robustness was applied to further reduce residual covariate imbalance between groups and increase the precision of the estimates by including as control variables any matching covariates for which the standardized bias was greater than 10 (Stuart, 2010). These sensitivity analyses reveal very similar results to those presented.

The Effect of Early Arrest on the Developmental Course of Offending

Table 3 presents the MLM regression results that include the matched early arrest measure from the PSM analyses and better isolate the impact of early arrest.[3] Reported coefficients reflect the values at the mean-centered age of 21.5 years.

Prevalence, frequency, and variety of self-reported offending

These results suggest that early arrest is significantly associated with later offending. The percent change in offending was calculated using the following equation (Long & Freese, 2006): percent change = (exponentiation (early arrest coefficient) − 1) * 100. Youth who were arrested by age 14 have a 56% higher odds of subsequent offending ($p < 0.01$; Model 1a), a 96% increase in the expected frequency of offending ($p < 0.01$; Model 1b), and a 30% increase in the expected number of offenses ($p < 0.01$; Model 1c) compared to non-early arrested youth.

Rate of change in self-reported offending

To examine if early arrest alters the curvature or rate of change in the self-reported offending trajectory, the early arrest measure was interacted with the age and age-squared slopes (Eqs. (3) and (4)). The cross-level interactions of early arrest with age and age-squared are not significant (models 2a, 2b, and 2c). This means that while early arrest alters aspects of the developmental course of offending by amplifying the risk, frequency, and diversity of subsequent offending, there is no evidence in these data that early arrest significantly alters the rate of change in offending with age (e.g., the shape of the trajectory).

Subsequent arrests

While prior models match treatment and control youth on the probability of experiencing an early arrest, subsequent legal system contact can impact involvement in crime and may alter the developmental course of offending. In these data, 85% of early-arrest youth accumulated at least one

[3] Results of MLM prior to PSM (not shown) are consistent with those of Table 3 but the latter are tempered by the PSM. For example, the effect of early arrest on the prevalence of offending declines by more than half (from $b = 1.110$; se = 0.136 to $b = 0.454$; se = 0.143) in the matched-groups models.

Table 3 Multilevel models after propensity score matching: long-term impact of early arrest on the developmental course of offending

Model	1a				2a				3a			
	Coeff	SE	CI		Coeff	SE	CI		Coeff	SE	CI	
Panel A. Full sample—prevalence of offending (logistic regression)												
Intercept	−1.798***	(0.121)	−2.035	−1.561	−1.787***	(0.132)	−2.046	−1.528	−2.092***	(0.162)	−2.409	−1.775
Within individual												
Age	−0.289***	(0.020)	−0.329	−0.250	−0.285***	(0.028)	−0.340	−0.229	−0.333***	(0.026)	−0.384	−0.282
Age^2	0.012**	(0.004)	0.004	0.019	0.012*	(0.005)	0.002	0.022	0.013**	(0.004)	0.005	0.020
Cumulative arrest									0.117**	(0.043)	0.033	0.201
Between individual												
Early arrest	0.444**	(0.144)	0.161	0.726	0.424*	(0.167)	0.097	0.750	0.372**	(0.141)	0.095	0.649
Cross-level interactions												
Age × Early arrest					−0.009	(0.034)	−0.076	0.058				
Age^2 × Early arrest					−0.001	(0.007)	−0.015	0.014				
Variance Components												
Intercept	2.871	(0.337)	2.280	3.615	2.867	(0.337)	2.278	3.610	2.830	(0.339)	2.238	3.578
Age	0.060	(0.012)	0.040	0.090	0.060	(0.012)	0.040	0.090	0.065	(0.012)	0.045	0.095
Model	1b				2b				3b			
	Coeff	SE	CI		Coeff	SE	CI		Coeff	SE	CI	
Panel B. Full sample—frequency of offending (negative binomial regression)												
Intercept	−1.729***	(0.212)	−2.144	−1.313	−1.819***	(0.251)	−2.312	−1.327	−2.443***	(0.251)	−2.935	−1.951
Within Individual												
Age	−0.553***	(0.044)	−0.639	−0.468	−0.580***	(0.064)	−0.704	−0.455	−0.657***	(0.048)	−0.751	−0.563
Age^2	0.006	(0.007)	−0.020	0.009	−0.005	(0.011)	−0.027	0.017	−0.002	(0.007)	−0.017	0.012
Cumulative arrest									0.282**	(0.066)	0.152	0.412
Between individual												
Early arrest	0.675**	(0.220)	0.243	1.107	0.840**	(0.276)	0.299	1.380	0.506*	(0.215)	0.086	0.927
Cross-level interactions												
Age × Early arrest					0.049	(0.065)	−0.079	0.177				
Age^2 × Early arrest					−0.002	(0.014)	−0.028	0.025				
Variance components												
Intercept	9.475	(0.768)	8.083	11.109	9.517	(0.774)	8.114	11.163	8.744	(0.772)	7.354	10.397
Age	0.181	(0.029)	0.132	0.247	0.181	(0.030)	0.132	0.248	0.192	(0.033)	0.136	0.270
Model	1c				2c				3c			
	Coeff	SE	CI		Coeff	SE	CI		Coeff	SE	CI	
Panel C. Full sample—variety of offending (negative binomial regression)												
Intercept	−1.672***	(0.093)	−1.854	−1.490	−1.696***	(0.105)	−1.901	−1.491	−1.868***	(0.124)	−2.110	−1.625
Within Individual												
Age	−0.238***	(0.016)	−0.271	−0.206	−0.246***	(0.023)	−0.292	−0.200	−0.270***	(0.022)	−0.314	−0.226

Table 3 (continued)

Model	1c				2c				3c			
	Coeff	SE	CI		Coeff	SE	CI		Coeff	SE	CI	
Age^2	0.004	(0.003)	−0.001	0.009	0.004	(0.004)	−0.004	0.011	0.005	(0.003)	<−0.001	0.010
Cumulative arrest									0.076*	(0.037)	0.005	0.148
Between individual												
Early arrest	0.259*	(0.102)	0.059	0.460	0.304*	(0.121)	0.066	0.541	0.219*	(0.101)	0.020	0.417
Cross-level interactions												
Age × Early arrest					0.014	(0.024)	−0.033	0.061				
Age^2 × Early arrest					0.001	(0.005)	−0.009	0.010				
Variance components												
Intercept	1.501	(0.142)	1.246	1.806	1.505	(0.142)	1.250	1.812	1.418	(0.139)	1.171	1.718
Age	0.025	(0.005)	0.018	0.036	0.025	(0.005)	0.018	0.036	0.028	(0.005)	0.020	0.039

National Longitudinal Survey of Young, 1997 Cohort. Results combined across 20 imputed datasets. Kernel matching; kernel = Epanechnikov; bandwidth = 0.06. Observations outside region of common support are excluded producing variation in sample size across imputed datasets (N = 13,043–13,099 obs. from 1280–1286 youth)

SE standard error, *CI* 95% confidence intervals

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$ (two-tailed)

additional arrest from mid-adolescence to their late 20s. In contrast, those with an early arrest and no subsequent arrests are a rare phenomenon (3% of the analytic sample). Lack of any legal system contact is fairly limited in this sample with a third (32%) reporting no history of arrest from early adolescence through to their late 20s. The prevalence of arrest is driven in part by the sample restriction beginning in the 2005 wave (see data description above). After including the time-varying cumulative arrest measure, the effect of having an early arrest is tempered but remains a significant factor for the development of offending (Table 3), increasing the average odds of offending by 45% ($p < 0.01$; Model 3a), the expected frequency of offending by 66% ($p < 0.05$; Model 3b), and the expected number of different offenses by 25% ($p < 0.05$; Model 3c).

To ease interpretation of the findings, the estimates from Table 3 are used to calculate the difference in self-reported offending comparing the early arrest with subsequent arrest (s), early arrest but no subsequent arrest(s), and non-early arrest but subsequent arrest(s) groups with a never arrested group. These estimates are visually represented in Fig. 1. Differences are reported at ages 15 (mid-adolescence and soon after early-arrest), 18 (adult status in most states), and 21 (mean-centered age in this study). While there is visual evidence that the offending trajectories for the early-arrest and non-early-arrest groups are converging in young adulthood (differences between bars are reduced), the significant effect in the models represents the effect of early arrest at the mean-centered age of 21.5 demonstrates an effect of early arrest that lingers into adulthood and well beyond the traditional demarcation between juvenile and adult legal systems (age 18).

To observe the development of offending over time, the estimates from Table 3 are also used to calculate average trajectories of self-reported offending for early-arrest youth and non-early-arrest youth in Fig. 2. The trajectories are plotted in excel varying response values for early arrest (0 no early arrest; 1 early arrest) and the group specific mean for cumulative arrests by age for the early arrest and non-early arrest groups. Age by group specific values were computed based on the following calculation: Age "x" predicted value = intercept + (age coeff. * age "x" value) + (age2 coeff. * age2 "x" value) + (early arrest coeff. * early arrest value) + (cumulative arrest coeff. * cumulative arrest value).

The error terms are set to zero for the purposes of plotting the trajectories. Two patterns are particularly noteworthy. First, the inclusion of cumulative arrest (Models 3a, 3b, and 3c, Table 3) produces a self-reported offending trajectory for a never arrested group whose expected odds, frequency, and variety of offending is not zero but instead is characterized by a heightened level of offending in adolescence that declines soon thereafter. This

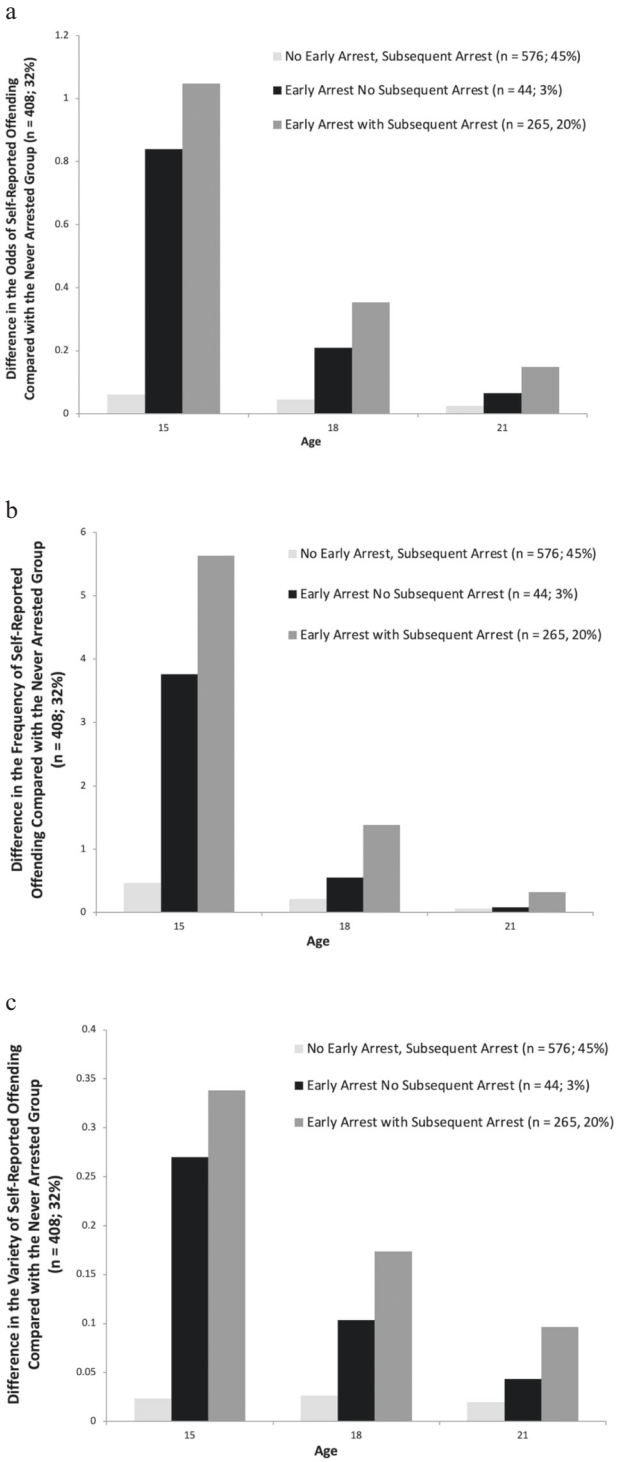


Fig. 1 Difference in self-reported offenses compared with the never arrested ($n = 408$; 32%) by age among matched samples. **a** Prevalence of self-reported offending. **b** Frequency of self-reported offenses. **c** Variety of self-reported offenses

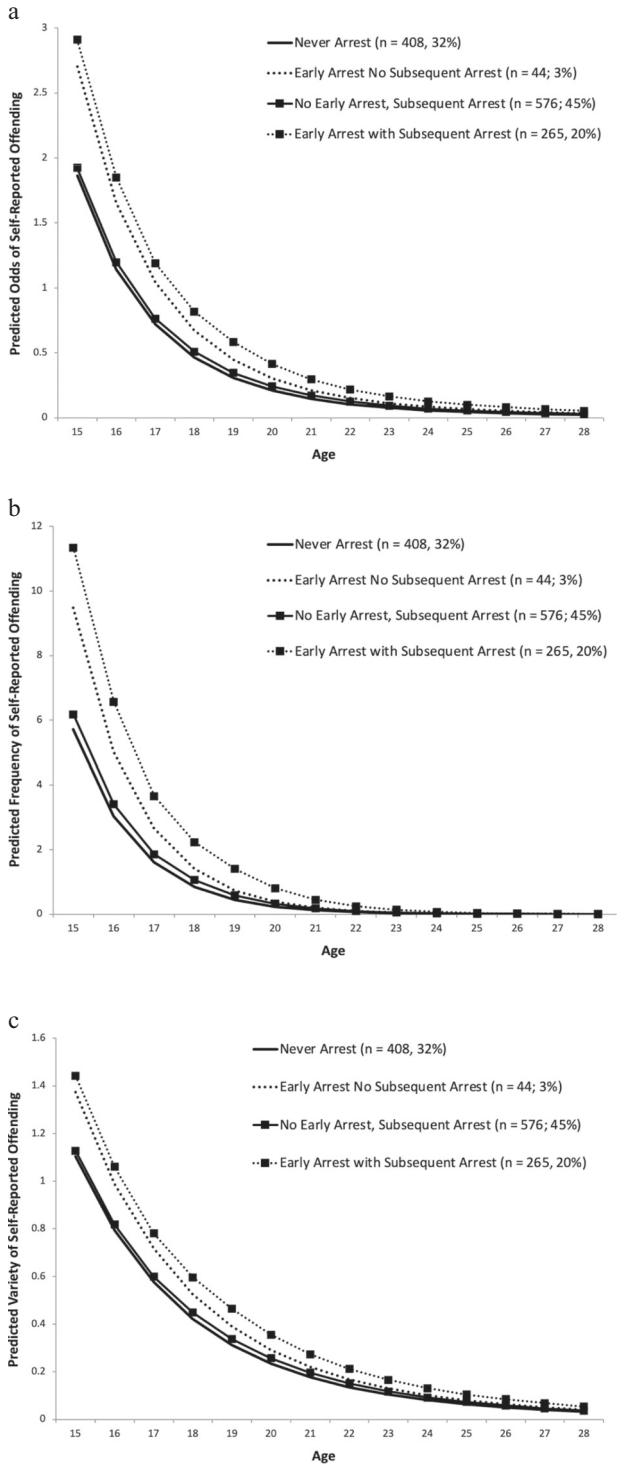


Fig. 2 Self-reported offending by early arrest status among matched samples. **a** Odds of self-reported offending. **b** Frequency of self-reported offenses. **c** Variety of self-reported offenses

normality of deviance is also supported by the finding that less than one-fifth (18%) of youth report no arrests or any engagement in self-reported offending across the entire study period.

Second, the developmental impact of early arrest is observed when comparing trajectories of those with and without an early arrest. Those who experience an arrest at 14 years of age or younger have a higher risk of involvement in crime, engage in crime more often, and are involved

in a greater variety of offenses; metrics associated with a greater severity of offending behavior. In this sense, involvement in crime is accelerated for early-arrest youth compared to their similarly-situated, non-early-arrest peers. This is particularly true among youth who experience an early arrest and subsequent arrests.

Demographic heterogeneity

The NLSY97's large sample size allowed us to test for demographic differences in the impact of early arrest on the development of offending. We performed a series of analyses parallel to those described above for subsamples stratified by race and gender. Propensity score matching was performed within each race- and gender-stratified subsample using a similar set of covariates excluding race in the White and non-White subsamples and gender in the Male and Female subsamples. To achieve balance and maintain consistency in matching covariates across the 20 imputed datasets for all gender and race subsamples, the set of matching covariates was modified slightly. For example, "tardy" and "absent" covariates were combined into a single additive scale. Substance use items (marijuana, alcohol, smoking) were combined into one binary indicator (0 = none, 1 = one or more). The reduced set of matching covariates remains diverse with respect to life-course domains and maintain a variety of items strongly associated with arrest.

Results of the race- and gender-subsample analyses that mirror Models 1a, 2a, and 3a in Table 3 for the full sample show that having an early arrest is detrimental for youth of all racial and gender groups. Having an early arrest significantly increases the risk of offending compared to their same race and gender grouped counterparts without an early arrest (results available upon request). Table 4 shows the results of these models when time-varying cumulative arrest is included as a control (Figs. S2 and S3 present the trajectories for race and gender subsamples, respectively). The odds of offending increase 50% for males, 52% for females, 55% for non-White youth, and 69% for White youth. The effect of early arrest remains significant and is associated with a higher prevalence, frequency and variety of offending for the White subsample (Models 1a, 1b, and 1c). With the inclusion of subsequent arrests, early arrest remains a significant factor for the prevalence of offending for the non-White subsample (Model 2a). Considering gender, early arrest is significant for all outcomes in the male subsample (Models 3a, 3b, and 3c) but statistical significance is attenuated for the female subsample (Models 4a, 4b, and 4c; Table 4).

Multiple approaches were used to check for significant differences across male and female as well as White and non-White subsamples including comparing confidence intervals to look for evidence of non-overlap (indicating significant differences) and conducting an equality of regression coefficient tests (Paternoster et al., 1998). Each was consistent in finding no significant differences in the effect of early arrest across subsamples. In addition, supplementary analyses checked for differences in the effects of early arrest across categories of race and gender in the full sample. For this, interaction terms (early arrest*race and early arrest*gender) were added to a version of the MLMs that relied on the unmatched rather than matched data, with the other matching covariates included as control variables. These analyses relied on unmatched data because gender and race variables had been used in the construction of the propensity scores. Results of the interaction tests indicated that differences in the effect of early arrest on the development of offending were not statistically significant across race or gender categories.

Sensitivity Analyses

Two sets of sensitivity analyses not yet described were also performed. The first had to do with the timing of arrest. The primary focus of this research is on understanding if and how early arrest alters the developmental trajectory of offending. However, youth varied in the timing of their first arrest relative to the measurement of their baseline characteristics. To capture the influence of heterogeneity in the timing of early arrest and minimize bias due to temporal ordering, three binary indicators of first arrest were created that become incrementally more inclusive of youth. These include: 1) an arrest at or before age 10 ($n = 11$); 2) an arrest more than one year prior to baseline, which includes an additional 20 youth ($n = 31$); and 3) an arrest that may have occurred any time prior to baseline, which adds 24 more youth ($n = 55$). The final specification accounts for the fact that the difference between age at the interview and age at first reported arrest was 1 year or more (e.g., age at interview was 14 and age of reported first arrest was 13); however, the difference in calendar months could have been less than 12 for these 24 youth). Each indicator was included separately as a control variable in the multilevel models. With the addition of these controls, results for the association of early arrest with subsequent offending were similar in direction and statistical significance to those reported in the results tables.

The second sensitivity test had to do with the distribution of survey items that were used to measure self-reported offending frequency. In the later waves when respondents were older, the responses to these items were increasingly skewed toward zero. To examine whether this skewness influenced the results for the models of the frequency of offending, all items were censored to a maximum of 20 times for a given offense. This resulted in a final set of item

Table 4 Multilevel models after propensity score matching, showing long-term impact of early arrest on development of offending, by race and gender

Model	White (1a)				Non-white (2a)				Male (3a)				Female (4a)			
	Coeff	SE	CI		Coeff	SE	CI		Coeff	SE	CI		Coeff	SE	CI	
Panel A. Prevalence of offending (logistic regression)																
Intercept	−1.936***	(0.218)	−2.363	−1.508	−2.383***	(0.209)	−2.792	−1.974	−1.841***	(0.173)	−2.179	−1.502	−2.779***	(0.253)	−3.275	−2.283
Within Individual																
Age	−0.300***	(0.033)	−0.364	−0.236	−0.341***	(0.034)	−0.409	−0.274	−0.318***	(0.027)	−0.371	−0.265	−0.306***	(0.039)	−0.383	−0.229
Age^2	0.015**	(0.005)	0.005	0.024	0.007	(0.005)	−0.003	0.017	0.010*	(0.004)	0.002	0.020	0.025***	(0.007)	0.011	0.039
Cumulative arrest	0.070	(0.047)	−0.022	−0.161	0.151**	(0.015)	0.051	0.252	0.091*	(0.038)	0.016	0.165	0.233***	(0.057)	0.122	0.344
Between individual																
Early arrest	0.524*	(0.208)	0.117	0.931	0.438*	(0.210)	0.027	0.850	0.404**	(0.152)	0.106	0.703	0.420	(0.286)	−0.141	0.981
Variance components																
Intercept	2.746	(0.514)	1.903	3.048	2.958	(0.493)	2.133	4.101	2.429	(0.349)	1.832	3.220	3.085	(0.083)	2.093	4.549
Age	0.060	(0.015)	0.036	0.098	0.055	(0.017)	0.030	0.101	0.058	(0.014)	0.036	0.093	0.032	(0.015)	0.012	0.082
Model	White (1b)				Non-white (2b)				Male (3b)				Female (4b)			
	Coeff	SE	CI		Coeff	SE	CI		Coeff	SE	CI		Coeff	SE	CI	
Panel B. Full sample—frequency of offending (negative binomial regression)																
Intercept	−2.098***	(0.413)	−2.908	−1.288	−3.083***	(0.307)	−3.684	−2.481	−2.000***	(0.367)	−2.720	−1.280	−3.405***	(0.429)	−4.246	−2.564
Within individual																
Age	−0.616***	(0.069)	−0.752	−0.480	−0.662***	(0.055)	−0.769	−0.555	−0.653***	(0.055)	−0.761	−0.544	−0.578***	(0.078)	−0.732	−0.424
Age^2	<−0.000	(0.012)	−0.023	0.023	−0.011	(0.008)	−0.026	0.005	−0.010	(0.013)	−0.035	0.015	0.033**	(0.012)	0.010	0.056
Cumulative arrest	0.248	(0.134)	−0.014	0.509	0.316***	(0.064)	0.190	0.441	0.266***	(0.076)	0.116	0.415	0.466***	(0.086)	0.297	0.635
Between individual																
Early arrest	0.723*	(0.309)	0.117	1.329	0.591	(0.304)	−0.005	1.187	0.581*	(0.256)	0.079	1.084	0.314	(0.398)	−0.466	1.095
Variance components																
Intercept	9.100	(1.443)	6.669	12.418	8.325	(1.004)	6.573	10.545	8.381	(0.981)	6.663	10.542	8.525	(1.527)	6.000	12.113
Age	0.159	(0.055)	0.080	0.314	0.223	(0.040)	0.157	0.318	0.176	(0.058)	0.092	0.337	0.111	0.036	0.059	0.209
Model	White (1c)				Non-white (2c)				Male (3c)				Female (4c)			
	Coeff	SE	CI		Coeff	SE	CI		Coeff	SE	CI		Coeff	SE	CI	
Panel C. Full sample—variety of offending (negative binomial regression)																
Intercept	−1.814***	(0.164)	−2.136	−1.492	−2.145***	(0.146)	−2.432	−1.859	−1.659***	(0.137)	−1.928	−1.391	−2.603***	(0.197)	−2.989	−2.217
Within individual																
Age	−0.249***	(0.027)	−0.301	−0.196	−0.283***	(0.024)	−0.330	−0.236	−0.256***	(0.022)	−0.299	−0.213	−0.272***	(0.035)	−0.340	−0.203

Table 4 (continued)

Model	White (1c)				Non-white (2c)				Male (3c)				Female (4c)			
	Coeff	SE	CI		Coeff	SE	CI		Coeff	SE	CI		Coeff	SE	CI	
Age^2	0.007*	(0.004)	0.000	0.014	<0.001	(0.004)	−0.007	0.007	0.004	(0.003)	−0.003	0.010	0.017***	(0.005)	0.007	0.027
Cumulative arrest	0.044	(0.036)	−0.027	0.114	0.114***	(0.028)	0.058	0.169	0.061	(0.033)	−0.003	0.126	0.200***	(0.034)	0.134	0.267
Between Individual																
Early arrest	0.380**	(0.146)	0.093	0.666	0.272	(0.144)	−0.010	0.554	0.246*	(0.107)	0.036	0.457	0.301	(0.190)	−0.072	0.675
Variance components																
Intercept	1.422	(0.215)	1.058	1.912	1.579	(0.199)	1.233	2.023	1.248	(0.147)	0.990	1.572	1.645	(0.309)	1.139	2.377
Age	0.026	(0.006)	0.017	0.040	0.026	(0.006)	0.017	0.040	0.024	(0.005)	0.016	0.036	0.020	(0.007)	0.010	0.040

National Longitudinal Survey of Young, 1997 Cohort. Results combined across 20 imputed datasets. Kernel matching; kernel = Epanechnikov; bandwidth = 0.06. Observations outside region of common support excluded producing sample size variation across imputed datasets (White: $N = 6837$–6853 obs. from 666 to 668 youth; Non-White: $N = 6153$–6209 obs. from 608 to 614 youth; Male: $N = 8082$–8113 obs. from 809 to 812 youth; Female: $N = 4892$–4964 obs. from 465 to 472 youth)

SE standard error, *CI* 95% conf. intervals, *TV* time-varying

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$ (two-tailed)

responses ranging from 0 to 20 in each wave for each self-reported offending item (compared with response ranges of 0–99 in the main analyses). Results using the truncated frequency measure are substantively similar to those presented in the main tables.

Discussion

In 2008, a systematic summary of the literature on the effect of arrest and justice system sanctions on subsequent behavior noted that despite robust and growing attention to the topic, rarely has research examined the effect of arrest from a developmental perspective spanning different and multiple phases of the life course (Huizinga & Henry, 2008). With rare exceptions, over a decade later, this critique remains accurate. This study contributes to this gap in the research by examining if and how arrest alters the developmental course of offending. Beginning from the perspective that much delinquent behavior is normal during adolescence and soon wanes as youth make the transition to adulthood, the current study aimed to understand if and how early arrest altered the typical developmental course of offending. Two groups of youth - those exposed to arrest early in the life course (age 14 or younger) and those with no early arrest – were matched on their likelihood of an early arrest, based on a wide range of factors including early engagement in risky behavior and family, school, peer, and neighborhood disadvantages, to account for selection into arrest. Then, trajectories of offending for the two groups were compared to assess for differences in the magnitude and rate of change in prevalence, frequency, and variety trajectories over time.

The underlying premise was that if early arrest bears developmental influence, then early arrest should significantly affect the expected pattern of offending even after matching early-arrest and non-early-arrest groups on the likelihood of experiencing an early arrest. In support of this premise, the results suggest that early arrest alters the development of offending in important ways. First, consistent with prior research, most youth in these data self-reported offending, but tended to do so outside the purview of the legal system. Their developmental course of offending followed a general pattern characteristic of normative involvement in delinquency during adolescence followed by desistance in young adulthood. Second, the findings suggest that experiencing an arrest in early adolescence alters the development of offending by accelerating the likelihood, severity, and frequency of offending throughout adolescence and lasting at least into early adulthood. Moreover, these associations with later offending remained even when accounting for the accumulation of subsequent arrests. This evidence suggests that rather than

fostering desistance (e.g., deterrence), early arrest impedes and may even disrupt the natural processes of offending and desistance that are expected by heightening and extending the risk of offending into young adulthood.

Second, the evidence from these analyses suggests that an early arrest is detrimental for all youth regardless of race or gender, but they provide little indication that these effects vary significantly across White and Non-White youth or between males and females. In models of two of the three outcomes in the non-White subsample and in models of all three outcomes for females, the effect of early arrest is reduced to non-significance when controlling for subsequent arrests. For these groups, the accumulation of subsequent arrests is a strong and significant factor associated with self-reported offending over the life course. Thus, the effects of early arrest may be partly explained by differences in rates of subsequent arrests for Non-White and female youth; however, females accumulated fewer arrests over time relative to males and mean rates of subsequent arrest were virtually identical for White and Non-White groups. While, in general, arrest begets arrest, recent research demonstrates that this connection is stronger for Black and Hispanic youth (Raphael & Rozo, 2019; McGlynn-Wright et al., 2020). The patterns observed in this study in combination with the findings from recent studies highlight the importance of accounting for demographic variation when examining outcomes of legal system contact such as arrest and the need for data capable of providing a more nuanced look at the potential mechanisms underlying this variation.

Third, results of the cross-level interactions between age and early arrest were not significant and indicate that while early arrest amplified the rate and seriousness of offending, the distribution of self-reported offending across age was similar in shape comparing early arrest and non-early arrest youth. In other words, we do not observe evidence that early arrest initiates a significantly elevated pattern of offending that persists throughout young adulthood. Instead, by the mid-20s, trajectories for all groups trend toward zero regardless of arrest history.

Implications

Taken together, the results hold import for policies and practices. Early arrest has been conceptualized both theoretically and practically as a key factor distinguishing those at particularly high risk for serious, chronic offending. Yet, the results of this study suggest that official responses in the form of arrest early in the life course may be criminogenic in their developmental consequences at least when considering short-term detriments for self-reported offending. In line with prior research, offending was common for all youth in this sample. However, among youth characterized by similar risk and propensity to offend, early arrest distinguished a higher prevalence, frequency, and variety of offending behavior that lingered into young adulthood. The finding that early arrest distinguishes offending trajectories well beyond age 18 is meaningful because it suggests that the consequences of early arrest may have a resonant influence on the life course. Most states use age 18 to delineate between juvenile and adult legal system jurisdiction; thus, at this age the consequences of offending and legal system contact become more punitive and permanent (e.g., criminal record). Moreover, while disparities in offending appear to converge by the mid-20s for early-arrest and non-early-arrest groups, the effects of legal system contact may have diffused to other key life-course domains. The transition to adulthood is a particularly pivotal, demographically dense moment in the life course when developmental opportunities linked to desistance are widespread. The potential to become ensnared in the adult legal system, incurring its indelible and public imprint, limits access to these opportunities and diffuses disadvantage through a decreased likelihood of marriage (e.g., Apel & Sweeten, 2010b), educational and employment prospects (e.g., Kirk & Sampson, 2013), and increased detriments to general health and well-being, and negative attitudes toward police (e.g., Schmidt et al., 2015).

Whereas findings suggest that early arrest bears developmental influence in terms of amplifying and diversifying involvement in offending, results also fail to find evidence that early arrest is associated with persistent criminal pursuits into young adulthood. Even contending with cumulative consequences following arrest, self-reported involvement in crime among the early-arrest group showed a swift decline during the transition to adulthood. In this sense, early arrest is not a risk factor distinguishing those with unique criminal potential but instead may function to manufacture this potential. The consequences of this notion are profound given that early arrest is often used as an aggravating factor in legal proceedings and is a common metric in risk assessments that distinguish youth who receive formal versus informal processing (see Cottle et al., 2001). As an alternative, evidence from this research suggests that the question of how to best hold youth responsible for their actions may be answered by turning to legal system alternatives with "minimal intervention and maximum diversion" (McAra & McVie, 2007:337), to assuage potential developmental ramifications of legal system contact and allow youth to remain tethered to conventional society having available to them all its opportunities for successful life-course development.

Limitations and Future Directions

It is important to note limitations pertaining to the data and research scope to place the findings of this research in context. First, this study took as a starting point that early adolescent offending in this sample captured normative involvement. Though the data used are culled from a general population and descriptive inquiry suggests that most delinquency at baseline did not involve extreme forms of behavior characteristic of more pathological or persistent offending, we cannot rule out that individuals whose offending is nonnormative are included in the data. Moreover, NLSY97 sample selection decisions across the longitudinal data collection process influence the determination of our analytic sample and the generalizability of the results of the study. Whereas our analytic strategy increases internal validity for comparing differences across statistically similar groups distinguished by early arrest histories, future research on generalizable samples should replicate these analyses.

Second, although the potential of early arrest to alter the developmental course of offending was positioned as part of the process of labeling and cumulative disadvantage it was beyond the scope of this research to formally test this mechanism. The inclusion of the cumulative arrest measure, which reduces but does not negate the strength of the effect of early arrest, suggests that cumulative disadvantage, in the form of subsequent system contact, is occurring but is not a sufficient explanation. While early arrest may initiate a process of cumulative disadvantage that amplifies involvement in crime, alternative explanations may also account for this relationship (see e.g., Doherty et al., 2016; Liberman et al., 2014). This research provides a roadmap for future studies into the influence of various theoretical factors such as the relationship between early adolescent arrest and identity, social bonds, participation in conventional institutions, and peer associations to understand how, why, and for how long the developmental course of offending is altered beyond what is expected.

Third, this research is based on observational data and cannot rule out the possibility that results are driven by differences between early-arrest and non-early-arrest youth that were not observed, such as their degree of penetration into the system. Though arrest is a seemingly straightforward variable to measure one's legal system contact history, in reality the event of arrest taps into a diverse range of experiences from deep formal system processing to informal processes that divert youth from the system. Prior research indicates that this variation is associated with youth outcomes (see e.g., Wiley & Esbensen, 2016). Moreover, the life-course framework acknowledges that "lives are embedded and shaped by context" (Laub & Sampson, 2003:33) and that the impact of turning points, including police contact, may depend on these contexts (e.g., Chenane et al., 2020; Doherty & Bersani, 2016). Future research should continue to disentangle how the effect of different types of early life-course experiences with the legal system and their situational context influence the developmental course of offending.

Early adolescent arrest is only one event early in the life course with the potential to redirect the development of offending. Future research should continue to expand the notion of turning points beyond their positive potential in adulthood and include all stages of the life course to advance understanding of continuity and change in the development of offending (see e.g., Widom et al., 2018; Sampson & Winter, 2018). In a similar vein, the effect of early arrest was observed on the outcome of self-reported offending; yet the developmental effect of early arrest is likely widespread and multiplicative in its non-criminal consequences. Future research should delve into the developmental impact of early arrest on the accumulation of disadvantages across and between life-course domains.

Finally, while we sought to examine how system contact early in the life course may differentially influence the developmental course of offending for males compared with females, and White respondents compared with their non-White peers, increasingly small sample sizes limited our ability to speak to differences at the intersection of gender and race or incorporate ethnicity. Though we find evidence that differences in the magnitude of the effect of early arrest may exist for persons of different race and gender groups the differences were not statistically significant. Importantly, the lack of statistical significance does not mean that the effect of early arrest is similarly experienced across gender and racial groups. Moreover, it may be that differences are masked in the global race and gender demographic groupings used here. Evidence that the imprint of legal system contact is more indelible for youth of color with compounding consequences across the life course (e.g., Del Toro et al., 2019) demands future inquiry into the differential harms borne by early arrest and the ways in which development may be altered depending on one's social position(s).

Conclusion

Engaging in delinquency is common during adolescence and quickly subsides with the transition into adulthood. Yet, for some this behavior results in arrest which can set in motion a series of cumulative disadvantages and collateral consequences. Findings from this research suggest that early arrest holds developmental implications influencing the magnitude and severity of offending over the early life course. The ramifications of this contact are observed across

adolescence and into early adulthood, for all youth regardless of race or gender, and remain even when accounting for the accumulation of subsequent arrests. Yet, the results also show that early arrest does not distinguish a persistent pattern of offending across the life course. Instead, involvement in crime wanes in young adulthood for all youth in these data. These findings support policies advocating for alternatives to formal legal system responses that have the potential to mitigate the developmental consequences of early arrest into emerging adulthood.

Acknowledgements We wish to thank John Laub and D. Wayne Osgood for their insights and comments on earlier versions of the manuscript.

Authors' Contributions B.E.B. conceived of the study, participated in its design and coordination, and drafted the manuscript; W.C.J. performed statistical analyses, participated in the interpretation of the data and writing of the manuscript; E.E.D. participated in the conception of the study, modeling approach, and writing of the manuscript. All authors read and approved the final manuscript.

Data Sharing and Declaration The datasets generated and/or analyzed during the current study are available in the National Longitudinal Surveys repository, https://www.nlsinfo.org/investigator.

Compliance with Ethical Standards

Conflict of Interest The authors declare no competing interests.

Ethical Approval This study utilized publicly accessible, deidentified data that does not constitute research with human subjects. Ethical standards were followed during original data collection efforts.

Informed Consent No identifiable information was obtained by the authors of this manuscript. Researchers collecting data for the National Longitudinal Surveys datasets obtain informed consent from participants.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Allison, P. D. (2000). Multiple imputation for missing data: A cautionary tale. *Sociological Methods & Research*, *28*(3), 301–309

Apel, R. (2016). The effects of jail and prison confinement on cohabitation and marriage. *The Annals of the American Academy of Political and Social Science*, *665*(1), 103–126

Apel, R. J., & Sweeten, G. (2010a). Propensity score matching in criminology and criminal justice. In A. R. Piquero & D. Weisburd (Eds), *Handbook of quantitative criminology* (pp. 543–562). Springer

Apel, R., & Sweeten, G. (2010b). The impact of incarceration on employment during the transition to adulthood. *Social Problems*, *57*(3), 448–479

Association of Maternal and Child Health Programs. (n.d.). Project area: Adolescent development. Retrieved December, 2020

Augustyn, M. B., Ward, J. T., Krohn, M. D., & Dong, B. (2019). Criminal justice contact across generations: Assessing the intergenerational labeling hypothesis. *Journal of Developmental and Life-course Criminology*, *5*(2), 137–175

Beaver, K. M., DeLisi, M., Mears, D. P., & Stewart, E. (2009). Low self-control and contact with the criminal justice system in a nationally representative sample of males. *Justice Quarterly*, *26* (4), 695–715

Bernburg, J. G. (2019). Labeling theory. In M. Krohn, N. Hendrix, G. Penly Hall & A. Lizotte (Eds), *Handbook on crime and deviance* (pp. 179–196). Springer

Bernburg, J. G., & Krohn, M. D. (2003). Labeling, life chances, and adult crime: The direct and indirect effects of official intervention in adolescence on crime in early adulthood. *Criminology*, *41*(4), 1287–1318

Bernburg, J. G., Krohn, M. D., & Rivera, C. J. (2006). Official labeling, criminal embeddedness, and subsequent delinquency: A longitudinal test of labeling theory. *Journal of Research in Crime and Delinquency*, *43*(1), 67–88

Bersani, B. E., & Doherty, E. E. (2018). Desistance from offending in the twenty-first century. *Annual Review of Criminology*, *1*, 311–334

Brame, R., Bushway, S. D., Paternoster, R., & Turner, M. G. (2014). Demographic patterns of cumulative arrest prevalence by ages 18 and 23. *Crime & Delinquency*, *60*(3), 471–486

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Sage Publications, Inc

Bureau of Labor Statistics, U.S. Department of Labor. (2019). *National Longitudinal Survey of Youth 1997 cohort, 1997-2011 (rounds 1-15). Produced and distributed by the Center for Human Resource Research (CHRR)*. The Ohio State University

Bushway, S. D., & Tahamont, S. (2016). Modeling long-term criminal careers: What happened to the variability? *Journal of Research in Crime and Delinquency*, *53*(3), 372–391

Chenane, J. L., Wright, E. M., & Wang, Y. (2020). The effects of police contact and neighborhood context on delinquency and violence. *Victims & Offenders*, *16*(4), 1–24

Chesney-Lind, M., & Eliason, M. (2006). From invisible to incorrigible: The demonization of marginalized women and girls. *Crime, Media, Culture*, *2*(1), 29–47

Chiricos, T., Barrick, K., Bales, W., & Bontrager, S. (2007). The labeling of convicted felons and its consequences for recidivism. *Criminology*, *45*(3), 547–581

Cottle, C. C., Lee, R. J., & Heilbrun, K. (2001). The prediction of criminal recidivism in juveniles: A meta-analysis. *Criminal Justice and Behavior*, *28*(3), 367–394

Crutchfield, R. D., Skinner, M. L., Haggerty, K. P., McGlynn, A., & Catalano, R. F. (2012). Racial disparity in police contacts. *Race and Justice*, *2*(3), 179–202

Del Toro, J., Lloyd, T., Buchanan, K. S., Robins, S. J., Bencharit, L. Z., Smiedt, M. G., Reddy, K. S., Pouget, E. R., Kerrison, E. M., & Goff, P. A. (2019). The criminogenic and psychological effects of police stops on adolescent black and Latino boys. *Proceedings of the National Academy of Sciences*, *116*(17), 8261–8268

Doherty, E. E., & Bersani, B. E. (2016). Understanding the mechanisms of desistance at the intersection of race, gender, and neighborhood context. *Journal of Research in Crime and Delinquency*, *53*(5), 681–710

Doherty, E. E., & Bersani, B. E. (2018). Mapping the age of official desistance for adult offenders: Implications for research and policy. *Journal of Developmental and Life-Course Criminology*, *4*(4), 516–551

Doherty, E. E., Cwick, J. M., Green, K. M., & Ensminger, M. E. (2016). Examining the consequences of the "prevalent life events" of arrest and incarceration among an urban African-American cohort. *Justice Quarterly*, *33*(6), 970–999

Duell, N., Steinberg, L., Icenogle, G., Chein, J., Chaudhary, N., Di Giunta, L., Dodge, K. A., Fanti, K. A., Lansford, J. E., Oburu, P., Pastorelli, C., Skinner, A. T., Sorbring, E., Tapanya, S., Tirado, L. M. U., Alampay, L. P., Al-Hassan, S. M., Takash, H. M. S., Bacchini, D., & Chang, L. (2018). Age patterns in risk taking across the world. *Journal of Youth and Adolescence*, *47*(5), 1052–1072

Erosheva, E. A., Matsueda, R. L., & Telesca, D. (2014). Breaking bad: Two decades of life-course data analysis in criminology, developmental psychology, and beyond. *Annual Review of Statistics and Its Application*, *1*, 301–332

Ge, X., Donnellan, M. B., & Wenk, E. (2003). Differences in personality and patterns of recidivism between early starters and other serious male offenders. *Journal of the American Academy of Psychiatry and the Law Online*, *31*(1), 68–77

Goffman, A. (2009). On the run: Wanted men in a Philadelphia ghetto. *American Sociological Review*, *74*(3), 339–357

Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Stanford University Press

Gottfredson, M., & Hirschi, T. (2020). *Modern control theory and the limits of criminal justice*. Oxford University Press

Guo, S., & Fraser, M. W. (2014). *Propensity score analysis: Statistical methods and applications (Vol. 11)*. SAGE Publications, Inc

Hirschfield, P. J. (2008). The declining significance of delinquent labels in disadvantaged urban communities. *Sociological Forum*, *23*(3 Sep), 575–601

Huizinga, D., & Henry, K. L. (2008). The effect of arrest and justice system sanctions on subsequent behavior: Findings from longitudinal and other studies. In A. M. Liberman (Eds.), *The long view of crime: A synthesis of longitudinal research* (pp. 220–254). Springer

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development*, *73*(2), 509–527

Jacobsen, W. C., Ragan, D. T., Yang, M., Nadel, E. L., & Feinberg, M. E. (2021). Arrested friendships? Justice involvement and interpersonal exclusion in rural schools. *Journal of Research in Crime and Delinquency*

Jones, N. (2014). "The regular routine": Proactive policing and adolescent development among young, poor black men. *New Directions for Child and Adolescent Development*, *2014*(143), 33–54

Kirk, D. S. (2009). Unraveling the contextual effects on student suspension and juvenile arrest: The independent and interdependent influences of school, neighborhood, and family social controls. *Criminology*, *47*(2), 479–520

Kirk, D. S., & Sampson, R. J. (2013). Juvenile arrest and collateral educational damage in the transition to adulthood. *Sociology of Education*, *86*(1), 36–62

Laub, J. H, Rowan, Z. R., & Sampson, R. J. (2019). The age-graded theory of informal social control. In D. P. Farrington, L. Kazemian, & A. R. Piquero (Eds.) *The Oxford handbook of developmental and life-course criminology* (pp. 295–322). Oxford University Press

Laub, J. H., & Sampson, R. J. (2003). *Shared beginnings, divergent lives*. Harvard University Press

Laub, J. H., & Sampson, R. J. (2020). Life-course and developmental criminology: Looking back, moving forward—ASC division of developmental and life-course criminology inaugural David P. Farrington Lecture, 2017. *Journal of Developmental and Life-Course Criminology*, *6*, 1–14

Le Blanc, M., & Loeber, R. (1998). Developmental criminology updated. *Crime and Justice*, *23*, 115–198

Liberman, A. M., Kirk, D. S., & Kim, K. (2014). Labeling effects of first juvenile arrests: Secondary deviance and secondary sanctioning. *Criminology*, *52*(3), 345–370

Liu, J., Francis, B., & Soothill, K. (2011). A longitudinal study of escalation in crime seriousness. *Journal of Quantitative Criminology*, *27*(2), 175–196

Long, J. S., & Freese, J. (2006). *Regression models for categorical dependent variables using stata*. 2nd ed. Stata Press

Lopes, G., Krohn, M. D., Lizotte, A. J., Schmidt, N. M., Vasquez, B. E., & Bernburg, J. G. (2012). Labeling and cumulative disadvantage: The impact of formal police intervention on life chances and crime during emerging adulthood. *Crime & Delinquency*, *58*(3), 456–488

McAra, L., & McVie, S. (2007). Youth justice? The impact of system contact on patterns of desistance from offending. *European Journal of Criminology*, *4*(3), 315–345

McGlynn-Wright, A., Crutchfield, R. D., Skinner, M. L., & Haggerty, K. P. (2020). The usual, racialized, suspects: The consequence of police contacts with black and white youth on adult arrest. *Social Problems*

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: a developmental taxonomy. *Psychological Review*, *100*(4), 674

Moffitt, T. E. (1994). Natural histories of delinquency. In E. G. Weitekamp, & H. Kerner (Eds.), *Cross-national longitudinal research on human development and criminal behavior* (pp. 3–61). Springer

Monahan, K. C., & Piquero, A. R. (2009). Investigating the longitudinal relation between offending frequency and offending variety. *Criminal Justice and Behavior*, *36*(7), 653–673

Morris, R. G., & Piquero, A. R. (2013). For whom do sanctions deter and label? *Justice Quarterly*, *30*(5), 837–868

Mowen, T. J., Brent, J. J., & Bares, K. J. (2018). How arrest impacts delinquency over time between and within individuals. *Youth Violence and Juvenile Justice*, *16*(4), 358–377

Natsuaki, M. N., Ge, X., & Wenk, E. (2008). Continuity and changes in the developmental trajectories of criminal career: Examining the roles of timing of first arrest and high school graduation. *Journal of Youth and Adolescence*, *37*(4), 431–444

Nguyen, H., & Loughran, T. A. (2018). On the measurement and identification of turning points in criminology. *Annual Review of Criminology*, *1*, 335–358

Osgood, D. W. (2010). Statistical models of life events and criminal behavior. In A. R. Piqeruo, & D. Weisburg (Eds.), *Handbook of quantitative criminology* (pp. 375–396). Springer

Paternoster, R., Brame, R., Mazerolle, P., & Piquero, A. (1998). Using the correct statistical test for the equality of regression coefficients. *Criminology*, *36*(4), 859–866

Patterson, G. R., & Yoerger, K. (2002). A developmental model for early-and late-onset delinquency. In J. B. Reid, G. R. Patterson, & J. Snyder (Eds.), *Antisocial behavior in children and adolescents: A developmental analysis and model for intervention* (pp. 147–172). American Psychological Association

Pettit, B., & Western, B. (2004). Mass imprisonment and the life course: Race and class inequality in US incarceration. *American Sociological Review*, *69*(2), 151–169

Piquero, A. R. (2008). Taking stock of developmental trajectories of criminal activity over the life course. In A. M. Liberman (Ed.), *The long view of crime: A synthesis of longitudinal research* (pp. 23–78). Springer

Puzzanchera, C., & Hockenberry, S. (2021). *Trends and characteristics of delinquency cases handled in Juvenile Court, 2019*. Office of Juvenile Justice and Delinquency Prevention. https://www.ojjdp.gov/ojstatbb/snapshots/DataSnapshot_JCS2019.pdf

Raphael, S., & Rozo, S. V. (2019). Racial disparities in the acquisition of juvenile arrest records. *Journal of Labor Economics*, *37*(S1), S125–S159

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (Vol. 1)*. SAGE Publications, Inc

Rengifo, A. F., & Pater, M. (2017). Close call: Race and gender in encounters with the police by black and Latino/a youth in New York City. *Sociological Inquiry*, *87*(2), 337–361

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*, 41–55

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, *39*, 33–38

Rutter, M., Giller, H., & Hagell, A. (1998). *Antisocial behavior by young people*. Cambridge University Press

Sampson, R. J., & Laub, J. H. (1997). A life-course theory of cumulative disadvantage and the stability of delinquency. *Developmental Theories of Crime and Delinquency*, *7*, 133–161

Sampson, R. J., & Winter, A. S. (2018). Poisoned development: Assessing childhood lead exposure as a cause of crime in a birth cohort followed through adolescence. *Criminology*, *56*(2), 269–301

Schmidt, N. M., Lopes, G., Krohn, M. D., & Lizotte, A. J. (2015). Getting caught and getting hitched: an assessment of the relationship between police intervention, life chances, and romantic unions. *Justice Quarterly*, *32*(6), 976–1005

Steffensmeier, D., Zhong, H., & Lu, Y. (2017). Age and its relation to crime in Taiwan and the United States: Invariant, or does cultural context matter? *Criminology*, *55*(2), 377–404

Steinberg, L. (2004). Risk taking in adolescence: What changes, and why? In R. E. Dahl, & L. P. Spear (Eds.) *Annals of the New York Academy of Sciences: Vol. 1021. Adolescent brain development: Vulnerabilities and opportunities* (pp. 51–58). New York Academy of Sciences

Stevens, T., Morash, M., & Chesney-Lind, M. (2011). Are girls getting tougher, or are we tougher on girls? Probability of arrest and juvenile court oversight in 1980 and 2000. *Justice Quarterly*, *28*(5), 719–744

Stolzenberg, L., & D'Alessio, S. J. (2004). Sex differences in the likelihood of arrest. *Journal of Criminal Justice*, *32*(5), 443–454

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science: A Review Journal of the Institute of Mathematical Statistics*, *25*(1), 1

Sweeten, G. (2012). Scaling criminal offending. *Journal of Quantitative Criminology*, *28*(3), 533–557

Tapia, M. (2011). Gang membership and race as risk factors for juvenile arrest. *Journal of Research in Crime and Delinquency*, *48*(3), 364–395

von Hippel, P. T. (2007). Regression with missing Y's: An improved strategy for analyzing multiply imputed data. *Sociological Methodology*, *37*(1), 83–117

Ward, J. T., Krohn, M. D., & Gibson, C. L. (2014). The effects of police contact on trajectories of violence: A group-based, propensity score matching analysis. *Journal of Interpersonal Violence*, *29*(3), 440–475

Widom, C. S., Fisher, J. H., Nagin, D. S., & Piquero, A. R. (2018). A prospective examination of criminal career trajectories in abused and neglected males and females followed up into middle adulthood. *Journal of Quantitative Criminology*, *34*, 831–52

Wiesner, M., & Capaldi, D. M. (2003). Relations of childhood and adolescent factors to offending trajectories of young men. *Journal of Research in Crime and Delinquency*, *40*(3), 231–262

Wiley, S. A. (2015). Arrested development: does the grade level at which juveniles experience arrest matter? *Journal of Developmental and Life-Course Criminology*, *1*(4), 411–433

Wiley, S. A., & Esbensen, F. A. (2016). The effect of police contact: Does official intervention result in deviance amplification? *Crime & Delinquency*, *62*(3), 283–307

Yun, I., Cheong, J., & Walsh, A. (2014). The relationship between academic achievement and likelihood of police arrest among delinquents. *International Journal of Offender Therapy and Comparative Criminology*, *58*(5), 607–631

Bianca E. Bersani is an Associate Professor at the University of Maryland. Her research examines persistence and desistance from criminal offending, life course trajectories, emerging adult justice, and immigration and crime.

Wade C. Jacobsen is an Assistant Professor at the University of Maryland. His research investigates the roles of schools and the criminal justice system in shaping child wellbeing and inequality.

Elaine Eggleston Doherty is a Visiting Research Professor at the University of Maryland. Her research focuses on examining stability and change in risk behaviors across the life course.