Adult criminal outcomes of juvenile justice involvement

William E. Copeland[1,2] (iD), Guangyu Tong[3], Elizabeth J. Gifford[4],
Michele M. Easter[2], Lilly Shanahan[5], Marvin S. Swartz[2] and Jeffrey W. Swanson[2]

[1]Department of Psychiatry, University of Vermont College of Medicine, Burlington, VT, USA; [2]Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Wilson Center for Science and Justice, Duke School of Law, Durham, NC, USA; [3]Yale Center for Analytical Sciences and Department of Biostatistics, Yale University, New Haven, CT, USA; [4]Center for Child and Family Policy and the Children's Health and Discovery Initiative, Duke University, Durham, NC, USA and [5]Jacobs Center for Productive Youth Development & Department of Psychology, University of Zurich, Zurich, Switzerland

Abstract

Background. The juvenile justice system in the USA adjudicates over seven hundred thousand youth in the USA annually with significant behavioral offenses. This study aimed to test the effect of juvenile justice involvement on adult criminal outcomes.

Methods. Analyses were based on a prospective, population-based study of 1420 children followed up to eight times during childhood (ages 9–16; 6674 observations) about juvenile justice involvement in the late 1990 and early 2000s. Participants were followed up years later to assess adult criminality, using self-report and official records. A propensity score (i.e. inverse probability) weighting approach was used that approximated an experimental design by balancing potentially confounding characteristics between children with *v.* without juvenile justice involvement.

Results. Between-groups differences on variables that elicit a juvenile justice referral (e.g. violence, property offenses, status offenses, and substance misuse) were attenuated after applying propensity-based inverse probability weights. Participants with a history of juvenile justice involvement were more likely to have later official and violent felony charges, and to self-report police contact and spending time in jail (ORs from 2.5 to 3.3). Residential juvenile justice involvement was associated with the highest risk of both, later official criminal records and self-reported criminality (ORs from 5.1 to 14.5). Sensitivity analyses suggest that our findings are likely robust to potential unobserved confounders.

Conclusions. Juvenile justice involvement was associated with increased risk of adult criminality, with residential services associated with highest risk. Juvenile justice involvement may catalyze rather than deter from adult offending.

CAMBRIDGE
UNIVERSITY PRESS

Introduction

Prior to 1899, youth who committed criminal offenses were charged, convicted, and jailed in the same system as adults. In 1899, the Juvenile Court Act took effect in Illinois, establishing a separate court for children 16 and under, with a rehabilitative rather than a punitive focus. Other states soon followed. Today, an estimated seven hundred thousand youth in the USA have cases handled in juvenile courts annually for person, property, drug, and public order offenses. Juvenile justice sentencing primarily consists of probation, community service, incarceration, and alternative schooling, reflecting the heterogeneous goals of juvenile justice which, in addition to public safety, include 'skill development, habilitation, rehabilitation, addressing treatment needs, and successful reintegration of youth into the community' (US Government, 2021). Despite its goals of rehabilitation and reintegration, the juvenile justice system has frequently issued sentences involving incarceration of minors. Although have fallen in the past 20 years, the USA continues to have the world's highest rates of juvenile incarcerations (McCarthy, Schiraldi, & Shark, 2016; Office of Juvenile Justice and Delinquency Prevention, 2020). Outcomes of youth involved in the juvenile justice system have often failed to support its rehabilitative aims (Bernburg, Krohn, & Rivera, 2006; Gatti, Tremblay, & Vitaro, 2009; Huizinga, Schumann, Ehret, & Elliott, 2003; Klein, 1986; Lambie & Randell, 2013; McCord, Widom, & Crowell, 2001). Indeed, the US juvenile justice system has been a target of substantial criticism, and has been characterized as 'ineffective, dangerous, unnecessary, obsolete, wasteful and inadequate' (Mendal, 2011).

Juvenile justice involvement may be detrimental to youth development for a number of reasons; for many youth in juvenile justice, multiple of these reasons apply. Firstly, at each step of the juvenile justice process, minority and poor youth receive disparate treatment in the context of similar circumstances, perpetuating racial and socioeconomic disparities

(Leiber & Fix, 2019; Poe-Yamagata, 2009). Secondly, involvement in this system may expose youth to verbal, physical, sexual, or emotional abuse (Wolff, Shi, & Siegel, 2009). For example, 12% of youth in juvenile facilities report past-year sexual victimization. Thirdly, the labeling of a child as 'a delinquent' may contribute to their increased subsequent involvement in deviance through increased identification with this status and also involvement in deviant peer groups (Dishion & Tipsord, 2011; Dodge, Dishion, & Lansford, 2006). *Peer contagion effects* are particularly problematic when programs group youth with a criminal record: Treatments that place peers exhibiting social deviance together are 30% less effective than individual programs. Indeed, 42% of group programs have adverse effects, with juveniles' behavior worsening with intervention (Lipsey, 2006). Fourthly, juvenile justice involvement is typically associated with existing psychiatric problems (Domalanta, Risser, Roberts, & Risser, 2003; Fazel, Doll, & Långström, 2008), and the essential features of incarceration – separation from family and loved ones, constant threat, isolation – may exacerbate these emotional and behavioral problems (Stokes, McCoy, Abram, Byck, & Teplin, 2015). Finally, juvenile incarceration often disrupts attainment of key educational and social milestones, leaving youth unequipped for successfully transitioning to adulthood. Taken together, many youth in juvenile justice are victimized and traumatized, socialized with deviant peers, often robbed of essential educational milestones, and, all in all, further marginalized from society.

Indeed, a few studies have shown that, rather than ameliorating offending or simply having no effect at all, juvenile justice involvement itself may exert a criminogenic effect (Eren & Mocan, 2017; Gatti et al., 2009; Huizinga et al., 2003; Kretschmar, Tossone, Butcher, & Marsh, 2018). Gatti et al. (2009) followed a community sample of 779 low socioeconomic status (SES) Canadian boys to examine the effects of juvenile justice involvement between ages 12 and 17 on adult criminal records up to age 25 after adjusting for possible childhood confounders. Juvenile justice involvement was associated with a sevenfold increase in the odds of adult criminality. Risk of adult criminality increased with restrictive interventions such as incarceration. Based on these results, the authors concluded that the effects of juvenile justice involvement were iatrogenic – that is, juvenile justice involvement itself is a risk factor for later criminal behavior.

The current study proposes to test the effectiveness of different levels of juvenile justice involvement in the USA from the 1990s to the early 2000s. To do this, this study uses a community-representative longitudinal study that assessed children *prior* to juvenile justice involvement on behaviors commonly associated with a juvenile justice referral like violence, property crimes, status offenses, and drug abuse. Children involved in the juvenile justice system also tend to differ from those without juvenile justice involvement on a number of additional childhood characteristics including attention problems, parental income, parental supervision, history of maltreatment, and emotional functioning (Gatti et al., 2009; Hawkins et al., 2000; Lipsey & Derzon, 1998). Therefore, we use a propensity score weighting approach that balances on behaviors that typically illicit juvenile justice referral as well as other potentially confounding characteristics between children with *v.* those without history of juvenile justice involvement. This approach approximates a randomized trial within an observational study by adjusting for non-randomization to treatment (here, juvenile justice involvement; VanderWeele, 2006).

We also propose two extensions to previous research. First, with different weighting schemes, we estimate the average effect of juvenile justice contact on later adult outcomes. This approach acknowledges the typical differences between those with and without juvenile justice contact that have previously been reported, providing estimates based on the counterfactuals of themselves (i.e. the group with juvenile justice contact). Second, test heterogeneity in juvenile justice outcomes based upon the level of involvement (i.e. residential *v.* non-residential). Specifically, we hypothesize that residential services – either in detention, jail, or a training school – will be associated with worse outcomes compared to less restrictive involvement (e.g. probation or seeing a court counselor). This has rarely been studied with respect to adult criminality.

Methods

Participants

The Great Smoky Mountains Study is a longitudinal, representative study of children in 11 mostly rural counties of North Carolina. Three cohorts of children, ages 9, 11, and 13 years, were recruited from a pool of some 12 000 children using a two-stage sampling design, resulting in *N* = 1420 participants (49% female). In all statistical analyses, sampling weights are applied to adjust for differential probability of selection and to allow results to generalize to the broader population of children from which the sample was drawn. See ascertainment online Supplementary eFig. 1 and (Copeland, Angold, Shanahan, & Costello, 2014) for additional detail.

Annual assessments were completed on the 1420 children from ages 9 to 16 (6674 total observations; 1993–2000), and again at ages 19, 21, 25, and 30 (4556 observations of 1336 participants; 1999–2015) for a total of 11 230 total assessments. Before all interviews, parent and child signed informed consent/assent forms. The study protocol and consent forms for each assessment were approved by the Duke University Medical Center Institutional Review Board. Participants received payment for their time ($100 for most recent wave).

Measures

Childhood Juvenile Justice involvement was identified using the *Child and Adolescent Services Assessment* (CASA; Ascher, Farmer, Burns, & Angold, 1996). The CASA was administered immediately after the CAPA/YAPA for all participants reporting at least one behavioral or emotional symptom. The interview began by reviewing all behavioral and emotional symptoms identified and then asked about specific settings in which the child may have received services. The focus of the current study was on juvenile justice services. The parent was asked whether the child 'had any contact with the court or juvenile justice services?' as well as specific questions about the types of juvenile justice contact (detention center/jail or probation/court counselor). Test-retest reliability of the CASA (self-report interclass correlation coefficient = 0.84; parent-report = 0.94) and concurrent validity with official mental health/substance center records was good (Farmer, Angold, Burns, & Costello, 1994). For this analysis, participants were characterized as having juvenile justice involvement at some point in childhood (up to age 16) or not. Residential services involved detention, jail, or a training school. This age was chosen, because during the time of this study,

individuals entered the adult criminal justice system at age 16 in North Carolina.

To estimate propensity score weights, those with and without a juvenile justice history were compared on behaviors assessed in the CAPA via parent and self-report that have previously been shown to result in juvenile justice referral: (1) violent behavior (e.g. physical fights, use of weapon); (2) property offenses (e.g. stealing, arson); (3) status offenses (e.g. truancy, running away); and (4) substance misuse (symptoms of abuse of dependence). Groups were also compared on the following variables often associated with service use and juvenile justice involvement (Farmer, Burns, Phillips, Angold, & Costello, 2003; Farmer, Stangl, Burns, Costello, & Angold, 1997; Jörg et al., 2012; Neufeld, Dunn, Jones, Croudace, & Goodyer, 2017): (1) sex; (2) race/ethnicity defined as American Indian, African American, or white; (3) low family SES; (4) unstable family structure (e.g. single parent family, divorce, presence of step-parent); (5) family dysfunction (e.g. inadequate parental supervision, domestic violence, maternal depression); (6) maltreatment including physical abuse, sexual abuse, and neglect; (7) number of total emotional (i.e. anxiety and depressive) disorder symptoms; (8) number of impairments secondary to psychiatric symptoms; (9) insurance status defined as private, public, or none; (10) number of DSM traumatic events (e.g. natural disasters, exposure to violence); (11) number of negative life events (e.g. moving, breakup with best friend).

Adult criminal outcomes were measured both, with self-reports using the Young Adult Psychiatric Assessment (YAPA; Angold et al., 1999) up to age 30 and with official charges harvested from the Administrative Offices of the Court's North Carolina database for all participants to age 25. Self-report items included *police contact*, having been placed on *probation*, and having been *incarcerated*. Official charges included *misdemeanor* and *felony* charges as well as felony charges that involve *violence*.

Statistical analysis

Juvenile justice history was represented by two grouping variables: (1) a dichotomous variable describing whether a participant had any juvenile justice history; (2) within the group with juvenile justice history, a dichotomous variable describing if the participant had experienced residential services. Each of these variables was used to predict adult criminal behaviors. All statistical analyses accounted for the two-stage sampling design using sampling weights.

Propensity score methods can improve the covariate balance through matching, weighting, or stratification, to make causal inferences or to reduce confounding in observational studies (Li & Li, 2019; Rosenbaum & Rubin, 1983; VanderWeele, 2006). Here, we used the inverse probability weights (IPW) derived propensity scores to balance covariates and to obtain the average treatment effect on the treated (ATT; Crump, Hotz, Imbens, & Mitnik, 2009; Li, Thomas, & Li, 2019). Additional trimming of observations was used to further improve the covariate balance (Crump et al., 2009). All covariates and interaction terms are displayed in Fig. 1 and online Supplementary eFig. 2. The propensity scores were then used to compute ATT weights. These weights were combined with the survey weight for all outcome models (Austin, Jembere, & Chiu, 2018). In addition, all variables used to derive the propensity scores were included as covariates in the outcome models. All analyses were performed using R 4.0.1 with adapted code from package 'PSweight' (Zhou, Tong, Li, & Thomas, 2020). With six adult crime outcomes, the Bonferroni-adjusted α level used for all analyses was 0.0083. Additional sensitivity analysis of the treatment effects using E values was performed as robustness checks (VanderWeele & Ding, 2017).

Missing data

Of the 1420 original participants, 1333 (93.9%) were followed up at least once in adulthood at ages 19, 21, 25, or 30. There were only two participants with missing juvenile justice history. Juvenile justice history was not associated with lower levels of participation in adulthood, suggesting no differential dropout. Missingness of individual adult crime outcomes within completed adult interviews was rare (~1%). Ten complete datasets were imputed to address missingness in both, outcomes and covariates. The R package 'mice' was employed for both, the imputation and synthesis of results in regression and propensity score analyses (van Buuren et al., 2020).

Results

Baseline differences related to juvenile justice involvement

In this sample, 226 of 1420 participants reported juvenile justice involvement. When weighted back to the population, this constitutes 10.6% of the population. Juvenile justice involvement was more common in males and in participants from a racial/ethnic minority background. Table 1 illustrates the difference in covariate balance in those with and without a juvenile justice contact before and after trimming and propensity score weighting. The left panel provides the number of participants as well as the prevalence (or mean) of each childhood covariate by juvenile justice history status and standardized mean difference (SMD). The largest between-groups differences were seen on behaviors that often result in juvenile justice contact: violence, delinquency, status offenses, and substance use. In addition, those with a juvenile justice history tended to have higher rates of various hardships, emotional symptoms, impairments, and life events/traumas, had public insurance or uninsured (*v.* private insurance) status, and were more likely to be from an American Indians or African American race/ethnic background. The right panel provides the same quantities (except the count of participants) after trimming ($d = 0.0175$) and weighting was applied. Between-groups differences between covariates were greatly attenuated with SMDs of all covariates below 0.10, suggesting a desirable covariate balance. Plots showing both balanced covariates and interaction terms can be found in Fig. 1. On average (across multiple imputed datasets), 225 of 226 participants with a juvenile justice history were retained after trimming.

Juvenile justice and adult criminality

Table 2 tests the effect of juvenile justice history on adult criminal outcomes. The first column shows the prevalence of each outcome within the sample; the subsequent columns provide the results from outcome regression models with weighting and trimming. All results are reported as odds ratios with 95% confidence intervals and *p* values. Odds ratio values above 1 indicate higher levels of adult crime in those with juvenile justice involvement. The first three rows show results predicting official criminal records, followed by rows for self-report measures of adult criminality. A parallel table with results from a logistic regression
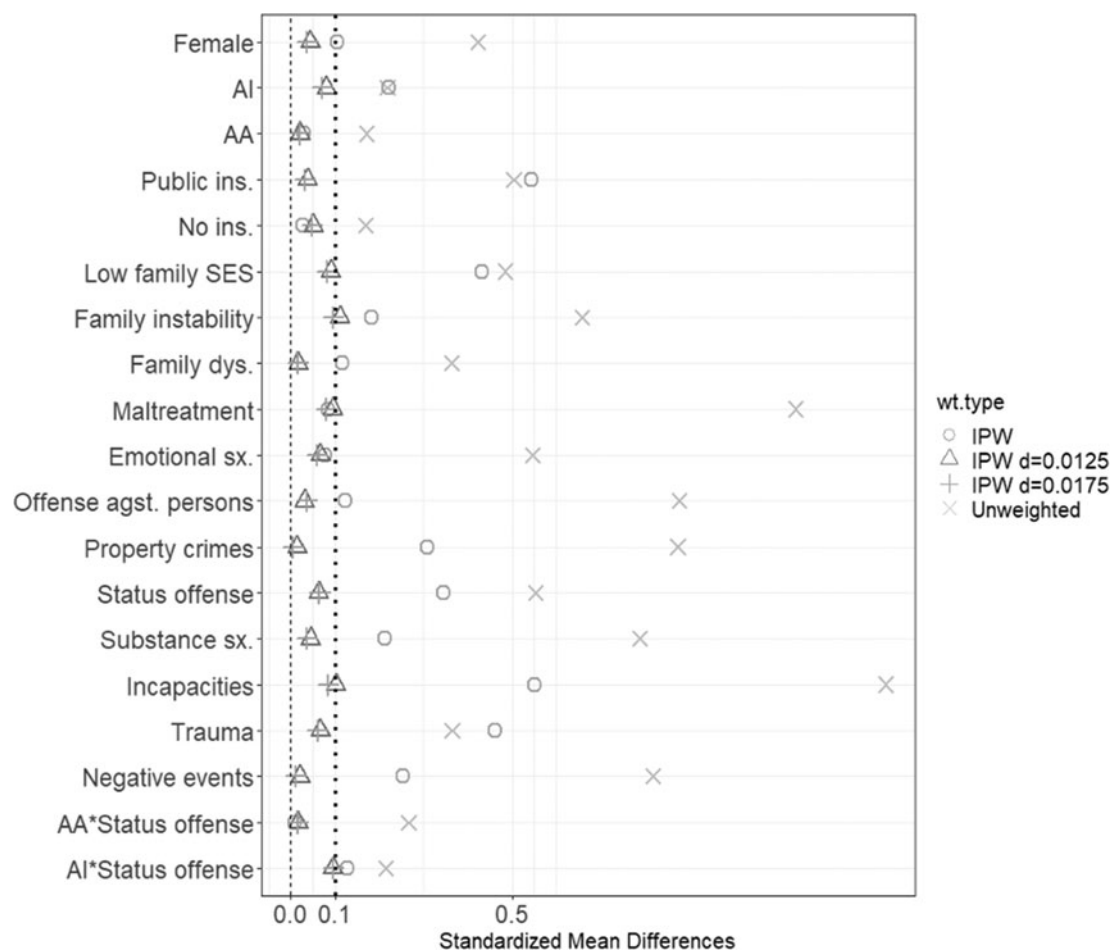
Fig. 1. Covariate balance plot for any juvenile justice *v.* no history in childhood using no weights and IPW with different trimming thresholds for ATT. Notes: AI, American Indian; AA, African-American; SES, socioeconomic status; Dys, dysfunction; Sx, symptoms; Ins, insurance. '*d*' indicates trimming thresholds. The trimming threshold of 0.0175 was adopted to obtain the best balance (SMD<0.1) and preserve the largest sample size. Results are average of 10 multiply imputed datasets.

model without IPW and trimming is available in online Supplementary eTable 1.

Juvenile justice involvement predicted *increased* risk of two of three official records outcomes and self-reported criminal outcomes each. All associations met the strict Bonferroni-adjusted *p* value. *E* values were calculated as a robustness check for significant effects. The effect of juvenile justice services on the likelihood of having an adult felony charge, for example, has an *E* value of 2.7, meaning that it would take a confounder with a risk ratio (RR) = 1.6 to alter our finding. This is equivalent to the strength of the strongest covariate (i.e. sex) which had an RR of 1.7. Similarly, *E* values for the violent felony charges, recent police contact and jail experience were calculated and described in the footnote of Table 2. These results indicate that our findings are likely robust to potential unobserved confounders.

Heterogeneity within juvenile justice involvement

Children who received juvenile justice services can be further divided into those with residential (detained/incarcerated) *v.* non-residential (probation or contact with a court counselor) involvement. The former has potentially higher risks of adult criminal outcomes given the behavior necessary to illicit detainment and

the greater disruption of detainment itself. Using the weighting and trimming described above, additional models were tested comparing each juvenile justice group to those without a juvenile justice history (online Supplementary eTable 2). The pattern of findings is generally similar to that of juvenile justice *v.* no juvenile justice overall, but the odds ratios for residential involvement were higher than those for non-residential involvement.

To test these differences directly, new sets of propensity score weights were derived. Online Supplementary eFig. 2 shows the covariate balance between the non-residential juvenile justice groups using IPW with different trimming thresholds. All covariates were balanced, with the result that all SMD were below 0.10 under the trimming threshold of *d* = 0.0175. Originally, there were 169 observations in the probation/court counselor group. After trimming, there were, on average (across multiple imputed datasets), 166 cases in the probation/court counselor group and 56 cases in the detention/jail/prison group. The left panel of Table 3 shows that a history of non-residential juvenile justice involvement predicts higher risk of a misdemeanor charge or violent felony charge in adulthood and higher likelihood of self-reported jail time as an adult compared to no juvenile justice involvement. None of these findings met the Bonferroni-adjusted *α* level. Stronger associations were observed in models comparing the residential juvenile justice involvement group *v.* the no

Table 1. Baseline differences on childhood covariates for juvenile justice involvement

	Unadjusted			After trimming and ATTW		
	JJ history	No JJ history	SMD	JJ history	No JJ history	SMD
Total *N*	226	1194		224–225	842–854	
	%	%		%	%	
Female	**30.7**	**51.1**	**0.4232**	30.6	28.9	0.0369
Race				0.0	0.0	
American Indian	**8.2**	**3.2**	**0.2179**	8.1	10.1	0.0699
African American	**11.3**	**6.4**	**0.1710**	11.3	11.9	0.0192
Insurance status				0.0	0.0	
Public	**36.6**	**15.1**	**0.5045**	36.6	38.0	0.0307
None	**14.3**	**8.8**	**0.1685**	14.2	15.9	0.0463
Low family SES	**49.5**	**26.5**	**0.4845**	49.4	53.5	0.0814
Family instability	**51.1**	**20.9**	**0.6591**	51.2	55.9	0.0954
Family dysfunction	**25.5**	**11.6**	**0.3625**	25.5	25.9	0.0143
Maltreatment	**64.1**	**15.4**	**1.1433**	64.2	68.0	0.0783
	M (S.D.)	M (S.D.)				
Violent behavior	**0.8 (0.6)**	**0.3 (0.5)**	**0.8784**	0.8	0.7	0.0359
Property crimes	**1.0 (0.7)**	**0.4 (0.6)**	**0.8756**	1.0	1.0	0.0044
Status offences	**0.4 (0.5)**	**0.1 (0.3)**	**0.5534**	0.4	0.3	0.0643
Substance symptom	**1.3 (0.5)**	**0.7 (0.8)**	**0.7889**	1.3	1.3	0.0352
Emotion symptoms	**5.9 (4.1)**	**3.9 (3.4)**	**0.5469**	5.9	6.2	0.0596
Impairments	**2.4 (1.0)**	**1.0 (1.1)**	**1.3464**	2.4	2.5	0.0840
Traumatic events	**1.5 (1.3)**	**1.1 (1.0)**	**0.3644**	1.5	1.6	0.0599
Negative events	**1.9 (1.1)**	**1.0 (1.0)**	**0.8195**	1.9	1.9	0.0105

JJ, juvenile justice contact; SES, socioeconomic status; Sx, symptoms; ATTW, average treatment effect on the treated weighting; SMD, standardized mean difference. All results in last three columns based on the average of 10 multiply imputed datasets. Rows in bold have standardized mean difference (SMD) >0.10. Variation in group sizes is due to trimming on imputed data. On average, the trimmed datasets have 848 in the no juvenile justice history group and 225 in the juvenile justice history group. After trimming (*d* = 0.0175) and weighting, all covariates and interaction terms (in Fig. 1) were balanced with SMD<0.10.

Table 2. Propensity score-adjusted model testing associations of juvenile justice history with adult criminality

	N (%)	OR	CI	*p* value
Official records				
Misdemeanor charge	404 (41.2)	1.6	[0.9–2.8]	0.1070
Felony charge	**126 (11.5)**	**2.8**	**[1.4–5.7]**	**0.0053**
Violent crime felony charge	**83 (9.5)**	**3.3**	**[1.5–7.2]**	**0.0036**
Self-reported				
Recent police contact	**200 (17.9)**	**2.5**	**[1.3–4.6]**	**0.0048**
Assault	40 (3.6)	1.9	[0.7–4.9]	0.2179
Jail	**183 (12.9)**	**3.2**	**[1.6–6.3]**	**0.0007**

Odds ratios combined from 10 multiply-imputed datasets with Rubin's formula. Models adjusted using average treatment effect on the treated weighting and all covariates and interactions used in the propensity score model (as shown in Fig. 1). For felony charge, the *E* value is 2.7 and RR is 1.7. For felony charge violent crime, the *E* value is 3.0 and RR is 1.8. For recent police contact, the *E* value is 2.5 and RR is 1.6. For jail, the *E* value is 3.0 and RR is 1.8.

involvement group (online Supplementary eFig. 3 for balance, and right panel of Table 3 for results). Residential involvement predicted higher risk of misdemeanor or felony charge and self-reported jail time. Similar findings were observed in the simple regression models (online Supplementary eTable 3). In sensitivity analyses, the *E* value for the effect of residential services on having

Table 3. Propensity score-adjusted models comparing residential and non-residential juvenile justice involvement to no juvenile justice history group.

	Non-residential *v.* no JJ contact			Residential *v.* no JJ contact		
	OR	CI	*p* value	OR	CI	*p* value
Outcomes						
Criminal records						
Misdemeanor charge	1.4	[0.8–1.2]	0.2712	**5.1**	**[1.2–10.7]**	**0.0008**
Felony charge	1.9	[0.9–4.1]	0.0834	**14.5**	**[3.2–66.2]**	**0.0006**
Violent crime felony charge	**2.4**	**[1.0–5.7]**	**0.0438**	3.0	[0.0–6142.5]	0.7806
Self-reported						
Recent police contact	**2.6**	**[1.3–5.1]**	**0.0046**	2.7	[0.9–8.3]	0.0748
Assault	2.0	[0.7–5.4]	0.1841	0.7	[0.0–40.7]	0.8732
Jail	**2.6**	**[1.3–5.5]**	**0.0090**	**6.5**	**[2.1–19.8]**	**0.0010**

Odds ratios combined from 10 multiply-imputed datasets with Rubin's formula. Models adjusted using average treatment effect for the treated weighting and all covariates and/or interactions used in the propensity score model (as shown in online Supplementary eFigs 2 and 3). Rows in bold have *p* value <0.05. In the right panel, the *E* value for misdemeanor charge is 3.9 and its corresponding RR value is 2.3. The *E* value for felony charge is 7.1 and its corresponding RR value is 3.8. The *E* value for jail time is 4.5 and its corresponding RR value is 2.5.

a felony charge was 7.1 and its corresponding RR value was 3.8. No covariates had an effect of this magnitude. Similarly, *E* values for the misdemeanor charges and a jail experience were calculated and described in the footnote of Table 3. These results indicate that our findings are likely robust to potential unobserved confounders.

Discussion

The goals of the juvenile justice system include public safety, rehabilitation, treatment, deterrence, and reintegration. The efficacy of juvenile justice services in meeting these aims has long been questionable (Eren & Mocan, 2017; Gatti et al., 2009; Huizinga et al., 2003; Kretschmar et al., 2018). The value of the current analysis was twofold: (1) combining causal modeling approaches with a 25+ year US population-based study to study the effects of juvenile justice involvement on adult crime; and (2) testing the effects of residential *v.* non-residential juvenile justice measures. In this sample, there was no evidence that juvenile justice services (received in the early 1990s to early 2000s in North Carolina) reduced the risk of adulthood criminality measured either using official criminal records or self-report. To the contrary, juvenile justice involvement was associated with increased risk of all types of adult criminality, and this increased risk was not small in size. These results were robust across the different statistical approaches used here (i.e. propensity score weighting and multivariate regression), and the size of the effects was moderate to large. Residential juvenile justice involvement was associated with the highest risk of adult criminal behavior. Rather than remediating the behavioral problems of youth, juvenile justice involvement may have catalyzed the transition to adult offending in this sample.

It is important to use caution in concluding iatrogenic effects of juvenile justice involvement. First, the Great Smoky Mountains study is representative of a mixed rural–urban area in the southeast, but not representative of the US population. The study includes a higher prevalence of white and American Indian participants and a lower prevalence of Black participants than the US population. Second, this study is also specific to the juvenile justice system in this area of North Carolina at a particular point in time (early 1990s to early 2000s). This was a period of transition

and reform in the state from a more punitive approach to juvenile crime to one with an increasing attention to rehabilitation, as evidenced by the N.C. Juvenile Justice Reform Act of 1998. Juvenile justice has continued to experience reforms to reduce incarceration, incorporate developmentally appropriate interventions, divert youth from the justice system, and improve equity (Abrams, 2013). Thus, while this study provides evidence of the limitations of the juvenile justice system at the time, it may or may not generalize to current juvenile justice efforts. Third, this study is a longitudinal-observational study that was not designed as a trial of the effectiveness of juvenile justice involvement. Fourth, we do not have specific information about the juvenile charges themselves; yet different charges could confer different risk for adult outcomes. This potential for unmeasured confounding also extends to other criminogenic variables not assessed in the study. We were able to account for multiple dimensions of conduct problems and substance use (as well as a bevy of other individual and family variables). Finally, trimming of propensity models improves covariate balance and reduces confounding, but it also can result in small differences from the original sample.

As has been reported before (Farmer et al., 1997; Farmer et al., 2003; Jörg et al., 2012; Neufeld et al., 2017), youth with juvenile justice involvement are strikingly different from other youth. This was also the case here, with differences in sex, racial/ethnicity, conduct symptoms, impairments, and family instability being detected between these groups. But the differences were more pervasive yet. Specifically, youth with juvenile justice involvement differed in insurance status, family SES, family dysfunction, history of maltreatment, emotional disorder symptoms, trauma exposure, and exposure to stressful life events. In this sample, 64.5% of juvenile justice-involved youth reported a history of physical abuse, sexual abuse, or neglect, compared to 15.4% in those without juvenile justice involvement. In theory, such pervasive differences could reasonably account for previously reported deleterious effects of juvenile justice involvement. Yet, in propensity score models – which balance such covariates – juvenile justice involvement predicted higher levels of both, felony and violent felony charges. In essence, by involving them into the juvenile justice system, the children with the highest levels of risk and adversity in this community sample were 'treated' with exposure to yet another potent risk factor: the juvenile justice system itself.

How might juvenile justice involvement increase later criminality? As noted in the introduction, several explanations have been offered including additional trauma, labeling, peer contagion, milestone disruption, and exacerbation of existing mental health problems. Our study suggests that many of these mechanisms are most likely, or exacerbated, among those with detention or incarceration. Indeed, participants with residential juvenile justice involvement in our study had the highest levels of later adult criminality. Those with less intensive juvenile justice involvement (e.g. probation) were, however, also at risk of adult criminality. The latter finding may implicate the stigmatizing effects of labeling by others or identification of the self as a 'juvenile offender' and secondary effects on family bonds, and peer associations as one of the mechanisms involved in predicting adult criminality (Bernburg et al., 2006).

The juvenile justice system has known problems of perpetuating racial disparities. Those referred to the juvenile justice system and those sentenced to residential interventions are more likely to be from a race/ethnic minority background (Poe-Yamagata, 2009). Any contact, and especially residential contact with the juvenile justice system then goes on to in and of itself, increases the risk of later criminality. Thus, the juvenile justice system has perpetrated and exacerbated systemic racism. The finding that with increasing degree of juvenile justice involvement, risk of adult criminality also increases is the epitome of a failed system: The system that is in part designed to protect society from crime actually contributes to more crime.

Any path forward to reform the juvenile justice system is likely multifactorial including drastically restricting the offenses that are punishable by incarceration (as some states already have), and a shift to empirically-supported, evidence-based non-residential alternatives (e.g. multisystemic therapy, functional family therapy, and multidimensional treatment foster care). It is a high priority to evaluate the long-term outcomes of these reforms. Additional reforms need to heed labeling and peer contagion effects that have compromised juvenile systems in the past. This is also an opportunity for an increased focus on interventions that prevent youth crime in the first place. Juvenile offending, like mental illness and other travails of youth, has strong foundations in early adversities like poverty and maltreatment. Policies targeting such pleiotropic risk factors like the Child Tax Credit in the USA afford the greatest potential to prevent juvenile delinquency and its long-term consequences in childhood and beyond.

References

Abrams, L. S. (2013). Juvenile justice at a crossroads: Science, evidence, and twenty-first century reform. *Social Service Review*, 87(4), 725–752.

Angold, A., Cox, A., Prendergast, M., Rutter, M., Simonoff, E., Costello, E. J., & Ascher, B. H. (1999). The Young Adult Psychiatric Assessment (YAPA). Retrieved from Durham, NC.

Ascher, B. H., Farmer, E. M. Z., Burns, B. J., & Angold, A. (1996). The Child and Adolescent Services Assessment (CASA): Description and psychometrics. *Journal of Emotional and Behavioral Disorders*, 4, 12–20.

Austin, P. C., Jembere, N., & Chiu, M. (2018). Propensity score matching and complex surveys. *Statistical Methods in Medical Research*, 27(4), 1240–1257.

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.

Copeland, W. E., Angold, A., Shanahan, L., & Costello, E. J. (2014). Longitudinal patterns of anxiety from childhood to adulthood: The great smoky mountains study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 53(1), 21–33. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0890856713006989?showall=true.

Crump, R. K., Hotz, V. J., Imbens, G. W., & Mitnik, O. A. (2009). Dealing with limited overlap in estimation of average treatment effects. *Biometrika*, 96(1), 187–199.

Dishion, T. J., & Tipsord, J. M. (2011). Peer contagion in child and adolescent social and emotional development. *Annual Review of Psychology*, 62, 189–214.

Dodge, K. A., Dishion, T. J., & Lansford, J. E. (2006). Deviant peer influences in intervention and public policy for youth. *Social Policy Report*, 20(1), 1–20.

Domalanta, D. D., Risser, W. L., Roberts, R. E., & Risser, J. M. H. (2003). Prevalence of depression and other psychiatric disorders among incarcerated youths. *Journal of the American Academy of Child & Adolescent Psychiatry*, 42(4), 477–484.

Eren, O., & Mocan, N. (2017). Juvenile punishment, high school graduation, and adult crime: Evidence from idiosyncratic judge harshness. *Review of Economics and Statistics*, 103(1), 34–47.

Farmer, E., Burns, B., Phillips, S., Angold, A., & Costello, E. (2003). Pathways into and through mental health services for children and adolescents. *Psychiatric Services*, 54(1), 60–66.

Farmer, E. M. Z., Angold, A., Burns, B. J., & Costello, E. J. (1994). Reliability of self-reported service use: Test-retest consistency of children's responses to the Child and Adolescent Services Assessment (CASA). *Journal of Child and Family Studies*, 3(3), 307–325.

Farmer, E. M. Z., Stangl, D. K., Burns, B. J., Costello, E. J., & Angold, A. (1997). Service use for children's mental health across one year: Patterns and predictors of use.

Fazel, S., Doll, H., & Långström, N. (2008). Mental disorders among adolescents in juvenile detention and correctional facilities: A systematic review and metaregression analysis of 25 surveys. *Journal of the American Academy of Child & Adolescent Psychiatry*, 47(9), 1010–1019.

Gatti, U., Tremblay, R. E., & Vitaro, F. (2009). Iatrogenic effect of juvenile justice. *Journal of Child Psychology and Psychiatry*, 50(8), 991–998.

Hawkins, J. D., Herrenkohl, T. I., Farrington, D. P., Brewer, D., Catalano, R. F., Harachi, T. W., & Cothern, L. (2000). Predictors of youth violence. *Juvenile Justice Bulletin*.

Huizinga, D., Schumann, K., Ehret, B., & Elliott, A. (2003). The effect of juvenile justice system processing on subsequent delinquent and criminal behavior: A cross-national study. *Final Report to The National Institute of Justice*, *Grant*(1999IJCX0037).

Jörg, F., Ormel, J., Reijneveld, S. A., Jansen, D. E., Verhulst, F. C., & Oldehinkel, A. J. (2012). Puzzling findings in studying the outcome of "real world" adolescent mental health services: The TRAILS study. *PLoS ONE*, 7(9), e44704.

Klein, M. W. (1986). Labeling theory and delinquency policy: An experimental test. *Criminal Justice and Behavior*, 13(1), 47–79.

Kretschmar, J. M., Tossone, K., Butcher, F., & Marsh, B. (2018). Examining the impact of a juvenile justice diversion program for youth with behavioral health concerns on early adulthood recidivism. *Children and Youth Services Review*, 91, 168–176.

Lambie, I., & Randell, I. (2013). The impact of incarceration on juvenile offenders. *Clinical Psychology Review*, 33(3), 448–459.

Leiber, M. J., & Fix, R. (2019). Reflections on the impact of race and ethnicity on juvenile court outcomes and efforts to enact change. *American Journal of Criminal Justice*, 44(4), 581–608.

Li, F., & Li, F. (2019). Propensity score weighting for causal inference with multiple treatments. *The Annals of Applied Statistics*, 13(4), 2389–2415.

Li, F., Thomas, L. E., & Li, F. (2019). Addressing extreme propensity scores via the overlap weights. *American Journal of Epidemiology*, 188(1), 250–257.

Lipsey, M. W. (2006). The effects of community-based group treatment for delinquency. In J. E. Lansford, K. A. Dodge, & T. J. Dishion (Eds.), *Deviant Peer Influences in Programs for Youth* (pp. 162–184). Guilford Publications.

Lipsey, M. W., & Derzon, J. H. (1998). Predictors of violent or serious delinquency in adolescence and early adulthood: A synthesis of longitudinal research. In R. Loeber & D. P. Farrington (Eds.), *Serious and violent juvenile offenders* (pp. 86–105). Thousand Oaks, CA: Sage Publications.

McCarthy, P., Schiraldi, V. N., & Shark, M. (2016). The future of youth justice: A community based alternative to the youth prison model.

McCord, J., Widom, C. S., & Crowell, N. A. (Eds.) (2001). *Juvenile crime, juvenile justice. Panel on juvenile crime: Prevention, treatment, and control*. Washington, DC: National Academy Press.

Mendal, R. A. (2011). No place for kids: The case for reducing juvenile incarceration.

Neufeld, S. A., Dunn, V. J., Jones, P. B., Croudace, T. J., & Goodyer, I. M. (2017). Reduction in adolescent depression after contact with mental health services: A longitudinal cohort study in the UK. *The Lancet Psychiatry*, 4(2), 120–127.

Office of Juvenile Justice and Delinquency Prevention (2020). *Juveniles in residential placement, 2017*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention.

Poe-Yamagata, E. (2009). *And justice for some: Differential treatment of minority youth in the justice system*. DIANE Publishing.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.

Stokes, M. L., McCoy, K. P., Abram, K. M., Byck, G. R., & Teplin, L. A. (2015). Suicidal ideation and behavior in youth in the juvenile justice system: A review of the literature. *Journal of Correctional Health Care*, 21(3), 222–242.

US Government. (2021). Juvenile Justice website. Retrieved from https://youth.gov/youth-topics/juvenile-justice.

van Buuren, S., Groothuis-Oudshoorn, K., Vink, G., Schouten, R., Robitzsch, A., Doove, L., … Arel-Bundock, V. (2020). mice: Multivariate Imputation by Chained Equations. Retrieved from https://CRAN.R-project.org/package=mice.

VanderWeele, T. (2006). The use of propensity score methods in psychiatric research. *International Journal of Methods in Psychiatric Research*, 15(2), 95–103.

VanderWeele, T. J., & Ding, P. (2017). Sensitivity analysis in observational research: Introducing the E-value. *Annals of Internal Medicine*, 167(4), 268–274.

Wolff, N., Shi, J., & Siegel, J. A. (2009). Patterns of victimization among male and female inmates: Evidence of an enduring legacy. *Violence and Victims*, 24(4), 469–484.

Zhou, T., Tong, G., Li, F., & Thomas, L. E. (2020). PSweight: An R package for propensity score weighting analysis. *arXiv preprint arXiv*:2010.08893.