

A Quasi-Experimental Study on the Effects of Community versus Custodial Sanctions in Youth Justice

Youth Violence and Juvenile Justice
2022, Vol. 0(0) 1–24
© The Author(s) 2022



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/15412040221133094

journals.sagepub.com/home/yvj



Gwendolyn J. Koops-Geuze¹ , Hilde T. Wermink², and Frank M. Weerman^{1,3}

Abstract

Although community sanctions have become a popular alternative to custodial sanctions in youth justice, primary questions about the recidivism effects of community sanctions remain unanswered. The current study aims to fill this gap through a quasi-experimental analysis of 2-year recidivism differences between 4,425 youth subject to community sanctions versus custodial sanctions in the Netherlands in 2015 and 2016. Recidivism was analyzed in terms of overall, serious, and very serious recidivism for the full sample, a low risk subsample, and a medium-high risk subsample. Findings indicate that youth subject to community sanctions are less likely to recidivate overall, and less like likely to recidivate seriously than youth subject to custodial sanctions. Community sanctions were found to be particularly beneficial for preventing very serious recidivism among low risk youth. Additionally, it was found that medium-high risk youth subject to community sanctions are less likely to recidivate overall, and less seriously than medium-high risk youth subject to custodial sanctions. Implications of these findings for future research and practice are discussed.

Keywords

community sanctions, quasi-experiment, recidivism, sanction effects, youth justice

Since the 1970s, most western youth justice systems have increasingly applied punishment alternatives to imprisonment (Cox & Godfrey, 2020; Winterdyk, 2015). Youth court statistics from the United States (U.S.) and Canada indicate that in 2018, approximately 72% (Hockenberry & Puzanchera, 2020) and 88% (Statistics Canada, 2020) of youth offenders received an alternative

¹Department of Law, Society & Crime, Erasmus University Rotterdam, Rotterdam, Netherlands

²Department of Criminal Law and Criminology, Leiden University, Leiden, Netherlands

³Netherlands Institute for the Study of Crime and Law Enforcement, Amsterdam, Netherlands

Corresponding Author:

Gwendolyn J. Koops-Geuze, Department of Criminology, Erasmus University Rotterdam, Burgemeester Oudlaan 50, Rotterdam, South-Holland 3062 PA, Netherlands.

Email: geuze@law.eur.nl

sanction that did not involve custody. Similar statistics exist in Europe, as in Germany this rate is 83%, and in the Netherlands, 72% (Jehle, 2019; Van der Laan and Beerthuis, 2021). One reason for this trend is Article 40 of the United Nations Convention on the Rights of the Child (UNCRC), which mandates that youth justice responses must minimally interfere with the lives of youth offenders given their immature developmental state, and thus diminished moral blameworthiness (Feld, 2013; UN General Assembly, 1989). Another reason for the increased application of punishment alternatives to imprisonment is a significant body of research which indicates that custodial sanctions (involving confinement or detention) may fail to adequately address the underlying causes of crime, and rather have unintended harmful consequences for the effective reintegration and rehabilitation of youth offenders (Cullen and Gendreau, 2001; Fagan & Kupchik, 2011; McGuire, 1995; UN General Assembly, 1989; Walgrave, 1998).

One frequently imposed alternative to custodial sanctions with varying elements and punishment aims across youth justice systems is the ‘community sanction’, which typically contains elements of learning, therapy, mandatory labor, supervision, or out of court diversion (Aebi et al., 2021; Uit Beijerse, 2019; Winterdyk, 2015). Moreover, the punishment aims of community sanctions range from retribution, restoration, rehabilitation, reintegration, and deterring youth from subsequent offending (Bateman, 2017; Cullen & Gendreau, 2001). Noteworthy, the punishment aims of community sanctions in youth justice contain less emphasis on retribution, but rather emphasize rehabilitation, reintegration, and preventing recidivism (Aebi et al., 2021; Farrington et al., 2012; Piquero et al., 2013). Community sanctions as defined in this study (community sanctions imposed in the Netherlands), are limited to unconditional hours of community service, short behavioral intervention, or a combination of both.

Nevertheless, a persistent, remaining question is whether community sanctions are as successful or more successful in reducing recidivism as custodial sanctions, as community sanctions were initially developed as a viable, alternative punishment option to custodial sanctions in youth justice (Cullen & Gendreau, 2001; Fagan & Kupchik, 2011). Unfortunately, the evidence on whether community sanctions in youth justice are at least as effective in reducing recidivism as custodial sanctions is incomplete, especially compared to the available evidence in the adult justice domain (Koops-Geuze & Weerman, 2021; Latimer, 2001; Loeffler & Nagin, 2021; Nagin et al., 2009; Villettaz et al., 2015).

Moreover, existing studies on this topic have two significant shortcomings. Firstly, the methodological quality of existing studies comparing differences in recidivism between community versus custodial sanctions is limited. Studies of this nature rarely contain sufficient methodological rigor to draw robust conclusions about the effects of community sanctions compared to custodial sanctions (Koops-Geuze & Weerman, 2021; Nagin et al., 2009; Wong et al., 2016). Secondly, there is a relative paucity of studies that provide meaningful insights into for whom, under which conditions, and when community sanctions are most effective (Lipsey, 2009; Mears et al., 2015; Wong et al., 2016). This is a significant gap given that sanction effects may be heterogenous rather than homogenous depending on various dimensions, such as, prior criminal history, post-release conditions, the counterfactual, and the demographic and social characteristics of populations (Mears et al., 2015).

In this article, the afore-mentioned shortcomings will be addressed by means of analyzing a large-scale, administrative dataset with criminal history, recidivism, and various recidivism risk factors on all youth offenders convicted in the Netherlands in 2015 and 2016. Sophisticated matching techniques are applied to create two rigorous, comparable groups to analyze differences in long-term recidivism effects among youth offenders who served either a community or a custodial sanction. Moreover, as the dataset includes enough unique individuals to examine heterogeneity of effects, we analyze differences between youth of low versus medium-high risk of recidivism.

Theorizing about Youth Justice Sanction Effects

Criminological theories provide contradicting expectations about the recidivism effects of community sanctions versus custodial sanctions in youth justice. Arguably, community sanctions involve less extensive labelling of the offender (Becker, 1963), and are less detrimental to mechanisms of social control (Hirschi, 1969). In essence, community sanctioned youth are enabled to maintain or strengthen social bonds, such as their attachment to intimate personal relations (e.g., residing at home), as well as legitimate commitments and involvements (e.g., playing sports, volunteering). Conversely, custodially sanctioned youth may lose access to various types of attachments, and therefore suffer a weakening of social bonds. Moreover, community sanctioned youth possibly receive less exposure to deviant others and positive associations towards offending, than youth subject to custodial sanctions (Sutherland, 1939). The latter typically involves placement among a population of other offenders and thereby increased exposure to opportunities for associating with new offenders or learning new offending behavior. Moreover, community sanctions appear more fit to allow for reintegrative shaming of the offender and restoration of imposed harms, which yields greater effectiveness according to the restorative justice (RJ) or reintegrative shaming perspective (Braithwaite, 1989). Given these perspectives, one may argue that community sanctions, rather than custodial sanctions, would be more effective in terms of reducing recidivism.

Nonetheless, some theoretical perspectives predict an opposite direction of effects, namely that community sanctions would be less effective than custodial sanctions in terms of reducing recidivism. According to deterrence theory (Beccaria, 1986), punishment must be severe enough to adequately deter individuals from offending, and according to the rational choice perspective (Clarke & Cornish, 1986), perceived costs must outweigh the perceived benefits of offending, in order to adequately deter individuals from engagement in offending behavior. Yet, community sanctions are objectively listed as less severe punishment than custodial sanctions (Cochran et al., 2014; Uit Beijerse, 2019), and therefore may have fewer deterrent effects than custodial sanctions, also in terms of rational, cost-benefit calculations (Beccaria, 1986; Clarke & Cornish, 1986; Nagin, 2013). Moreover, youth subject to custodial sanctions and thus confined in custody, may receive greater exposure to rehabilitation, training, and other types of programming. Typically, these programs are aimed at developing alternative mechanisms for coping with strain or resolving the underlying causes of strain leading to offending behavior (Agnew, 1992). Considering these perspectives, one would expect custodial sanctions, rather than community sanctions to be more effective in terms of reducing recidivism.

Prior Research on Youth Justice Sanction Effects

Although few youth justice studies specifically assessed the effects of community versus custodial sanctions, overall attempts to assess the effects of alternative sanctions, and diversion from traditional criminal justice system processing have been ongoing for decades (Klein, 1979; Krisberg et al., 1995; Latimer, 2001; Schwalbe et al., 2011; Wilson & Hoge, 2013; Wong et al., 2016; Wilson et al., 2018). In general, existing reviews on this topic have produced mixed results in terms of (not) finding significant differences in recidivism between different types of diversion programs and sanctions. Overall, police diversion programs (Wilson et al., 2018), and sanctions focused on behavioral intervention, family intervention, skills training, and restorative justice were found to yield larger reductions in recidivism than other types of diversion programs and sanctions (Krisberg et al., 1995; Latimer, 2001; Schwalbe et al., 2011; Wilson & Hoge, 2013; Wong et al., 2016). Moreover, in a quasi-experimental comparison of sanctions with varying types of severity, Cochran et al. (2014) overall found that less severe sanctions yielded lower recidivism

rates than sanctions of greater severity. Altogether, these reviews suggest that in youth justice, the administration of (less severe) diversion programs and sanctions may be more effective than traditional criminal justice system processing (of greater severity).

Additionally, some reviews report that the potential recidivism effects of diversion differ depending on the pre-defined recidivism risk level of youth offenders. In their review of police diversion programs for youth, [Wilson et al. \(2018\)](#) found that police diversion programs were especially beneficial for low-risk youth with no prior record. Yet, an extensive review conducted by [Lipsey \(2009\)](#), analyzed the effectiveness of treatment programs for youth at varying stages of the criminal justice system and found that high risk youth benefited more from treatment interventions than low risk youth, and that low-risk youth benefited more from non-treatment type of interventions than high risk youth, in terms of recidivism. Likewise, [Wilson and Hoge's \(2013\)](#) review of caution programs versus treatment interventions found that treatment interventions were especially beneficial for high-risk youth, while caution programs were found to be more beneficial for low-risk youth ([Wilson & Hoge, 2013](#)). In sum, these studies suggest that the effects of diversion may differ depending on the recidivism risk-level of offenders upon which they are imposed.

Yet, none of the above-mentioned reviews assessed the type of community sanctions as defined in this study, or directly compared the recidivism effects of specifically, community sanctions to custodial sanctions. A recent meta-analysis conducted by [Koops-Geuze & Weerman \(2021\)](#) indicated that only seven existing studies have directly compared the recidivism effects of community sanctions versus custodial sanctions in youth justice. Although most studies did not yield significant differences in overall recidivism ([Brandau, 1992](#); [Schneider, 1986](#); [Van der Laan & Essers, 1990](#); [Van der Laan & Essers, 1993](#)), the combined effect sizes of existing studies indicated significantly lower recidivism for youth subject to community sanctions compared to youth subject to custodial sanctions ([Koops-Geuze & Weerman, 2021](#)). Moreover, multiple studies found that youth subject to community sanctions engaged in significantly less serious recidivism than their custodial counterparts ([Bontrager-Ryon et al., 2017](#); [Brandau, 1992](#); [Essers et al., 1995](#); [Schneider, 1986](#); [Van der Laan & Essers, 1990](#); [Van der Laan & Essers, 1993](#)).

Current Study

The aim of this study was to examine recidivism differences between community sanctions versus custodial sanctions using an overall, low risk, and medium-high risk sample of youth offenders sentenced by a youth court judge in the Netherlands. As such, we utilized a large-scale, administrative dataset and applied matching techniques to conduct a quasi-experimental analysis of differences in overall, serious, and very recidivism between youth subject to a community or custodial sanction in 2015 or 2016.

While high-quality effect studies rather employ a randomized controlled trial (RCT) design ([Farrington et al., 2002](#)), conducting RCTs to study sanction effects in youth justice is difficult. A research design whereby youth offenders are randomly assigned to different sanctions may violate existing sentencing guidelines, and essentially also constitutes a violation of one's right to a fair trial as per Article 40 of the UNCRC ([UN General Assembly, 1989](#)). Alternatively, quasi-experimental methods such as matching techniques can be employed, which are more feasible to conduct in youth justice settings yet minimize selection bias concerns. Moreover, well-designed matching studies are comparable to RCTs in terms of conducting valid measurements of sanction effects ([Loeffler & Nagin, 2021](#); [Stuart, 2010](#)).

Overall, well-designed matching studies are defined as having large sample sizes ($N > 1000$), and diverse covariates available for analysis to reduce selection bias because of unobserved confounders ([Loeffler & Nagin, 2021](#); [Shadish, 2013](#)). However, existing studies that employed

(quasi-)experimental designs to assess differences in recidivism between community versus custodially sanctioned youth typically had (much) smaller sample sizes ($N = 140\text{--}245$) and may have lacked access to important information related to sanction assignment, such as case characteristics. Additionally, prior studies may have lacked access to essential information related to the outcome of recidivism, such as recidivism risk in relation to specific domains of life, like family situation, mental health, and school performance (Brandau, 1992; Bontrager-Ryon et al., 2017; Van Wormer, 2018).

Nonetheless, this study aimed to expand upon prior research in multiple ways, as it employed a large sample size ($N = 4425$) and rich dataset, including a wide variety of covariates. Among others, the dataset contained important background characteristics that play a crucial role in sanction assignment, and the outcome of recidivism on ten domains namely, substance use, attitudes, mental health, family, relationships, school, skills, leisure, and work. Additionally, the present sample only involved youth sentenced by a youth court judge. In the Netherlands, custodial sanctions can only be imposed at the court stage, and not at the police or prosecutorial stage of the criminal justice system (Uit Beijerse, 2019). Thus, in theory, youth in the present sample, as all sentenced by a youth court judge, had comparable chances of receiving either a community, or a custodial sanction. Moreover, as outlined in greater detail below, the ratio of successfully matched youth ranged from 97 to 99%, indicating most youth were kept, rather than excluded from the sample, the latter which would have caused problematic selection bias (Stuart, 2010). As such, this study clearly met the requirements for high-quality matching and consequently, reliable effect estimations (Loeffler & Nagin, 2021; Shadish, 2013).

Altogether, this study was the first analysis of community sanctions versus custodial sanctions in youth justice utilizing sophisticated matching techniques and an extensive dataset with valuable characteristics to create reliable comparison groups and examine meaningful variation. As outlined above, criminological theories provide contradicting expectations about the potential effects of community versus custodial sanctions. Yet, the overall message from existing studies was that less severe sanctions, and community alternatives are at least as effective, and possibly more effective in terms of recidivism when compared to custodial sanctions (Cochran et al., 2014; Koops-Geuze & Weerman, 2021). Additionally, youth subject to community sanctions typically engaged in less serious recidivism than youth subject to custodial sanctions. Prior research has also found that less intense alternatives to traditional criminal justice sanctions are particularly effective for low-risk youth. Therefore, our first hypothesis states that youth subject to community sanctions yield less recidivism than youth subject to custodial sanctions. Our second hypothesis states that youth subject to community sanctions engage in less serious recidivism than youth subject to custodial sanctions. Finally, our third hypothesis asserts that low-risk youth subject to community sanctions yield less recidivism than low-risk youth subject to custodial sanctions, and that higher risk youth subject to community sanctions yield more recidivism than higher risk youth subject to custodial sanctions.

Methods

Data and Sample

To investigate the outcomes of interest in this study, we analyzed data received from the Dutch Scientific Research and Documentation Centre (DSRDC). This dataset contained encrypted and anonymized data from the Judicial Documentation System and included nationwide data on all youth aged 12 to 18 who were sentenced to a community sanction or custodial sanction by a youth court judge in 2015 or 2016, in the Netherlands. Among others, the dataset contained information about criminal history, such as prior justice system contacts, prior community sanctions, and prior

custodial sanctions, as well as the prevalence, frequency, and type of recidivism. At the DSRDC, the Judicial Documentation System data was linked with data obtained from the Dutch Child Protection Council, which contained individual scores on each of the ten assessments domains of the National Youth Justice System Risk Assessment Tool (NYJSRAT).

The NYJSRAT is a national risk assessment tool used to assess recidivism risk and protective factors in youth justice proceedings in the Netherlands (Mensink et al., 2021). In general, risk assessments aim to predict recidivism risk among youth offenders on a variety of static factors (such as criminal history) and dynamic factors (such as antisocial attitudes) (Vincent et al., 2012). According to Baird et al. (2013), the predictive validity of risk assessments is most accurate when factors that are evidentiary related to the outcome of recidivism are assessed. Among others, significant risk factors in relation to recidivism among youth offenders include, prior criminal history, having delinquent family members or peers, the presence of antisocial attitudes and behavior, substance use, age at first offense, and poor school performance (Cottle et al., 2001; Ortega-Campos et al., 2016; Scott & Brown, 2018).

Overall, the NYJSRAT generates static and dynamic risk factor scores expressed in 'low', 'medium' or 'high' risk terms for each of the ten risk assessment domains which include the risk factors reported by Baird et al. (2013) (see Figure 1). In addition, it calculates an overall recidivism risk score, which is a weighted average of the total, individual risk scores. The NYJSRAT can be administered by various criminal justice system actors, and prior to sentencing this assessment is conducted as part of a mandatory, presentence report. Consequently, the results of the presentence report, including the risk assessment outcomes, are used to assist youth judges in making sentencing decisions (Vincent et al., 2012; Van Wingerden et al., 2014). Thus, the NYJSRAT plays an important role in the sanction assignment mechanism.

In total, the DSRDC successfully linked 86% of the NYJSRAT data with the JDS data, which yielded a final, linked dataset with 6716 unique cases in total. Various selections were made to this dataset that inevitably reduced the sample size for subsequent analysis. Firstly, any youth that were deceased during the observation period were removed from the sample ($n = 27$). Secondly, in order to study sanction effects among youth who truly experienced the sanction we excluded individuals with fully conditional sentences ($n = 979$). Thirdly, youth sentenced to both a community and custodial sanction in 2015 or 2016 were excluded, as for these youth we could no longer isolate the effect of having received either a community, or custodial sanction ($n = 848$). To furthermore ensure that community sanctioned youth did not experience custody, we also excluded any community sanctioned youth subject to pre-trial custody longer than 24 hours in length ($n = 168$). Fourthly, we excluded youth convicted of truancy offenses, because custody cannot be imposed for these types of offenses, meaning such youth were not truly 'at risk' of custody in the Netherlands ($n = 16$). Fifthly, given that the maximum possible length of community sanctions in our sample was 240 hours, we only compared youth subject to custodial sanctions of similar maximum length. In the Netherlands, 2 community sanction hours are equal to 1 day in custody (Meijer et al., 2021). Using this conversion rate, the maximum sentence length of custodial sanctions in our sample was limited to 120 days, or 4 months. Therefore, in essence we only compared youth subject to community sanctions with youth subject to (short) custodial sanctions of maximally 4 months. As a result, youth sentenced to a custodial sanction longer than 4 months were excluded from our sample ($n = 194$). Finally, cases with data quality problems due to missing information were removed ($n = 59$). Consequently, the final sample that was used as the basis for our matching procedure consisted of 4425 reference cases, with 4045 community sanctioned youth and 380 custodially sanctioned youth.

Table 1 displays the unmatched sample characteristics, dividing between youth subject to community versus custodial sanctions, with counts (n) and percentages (%) for categorical variables, and means (M) and standard deviations (SD) for continuous variables.

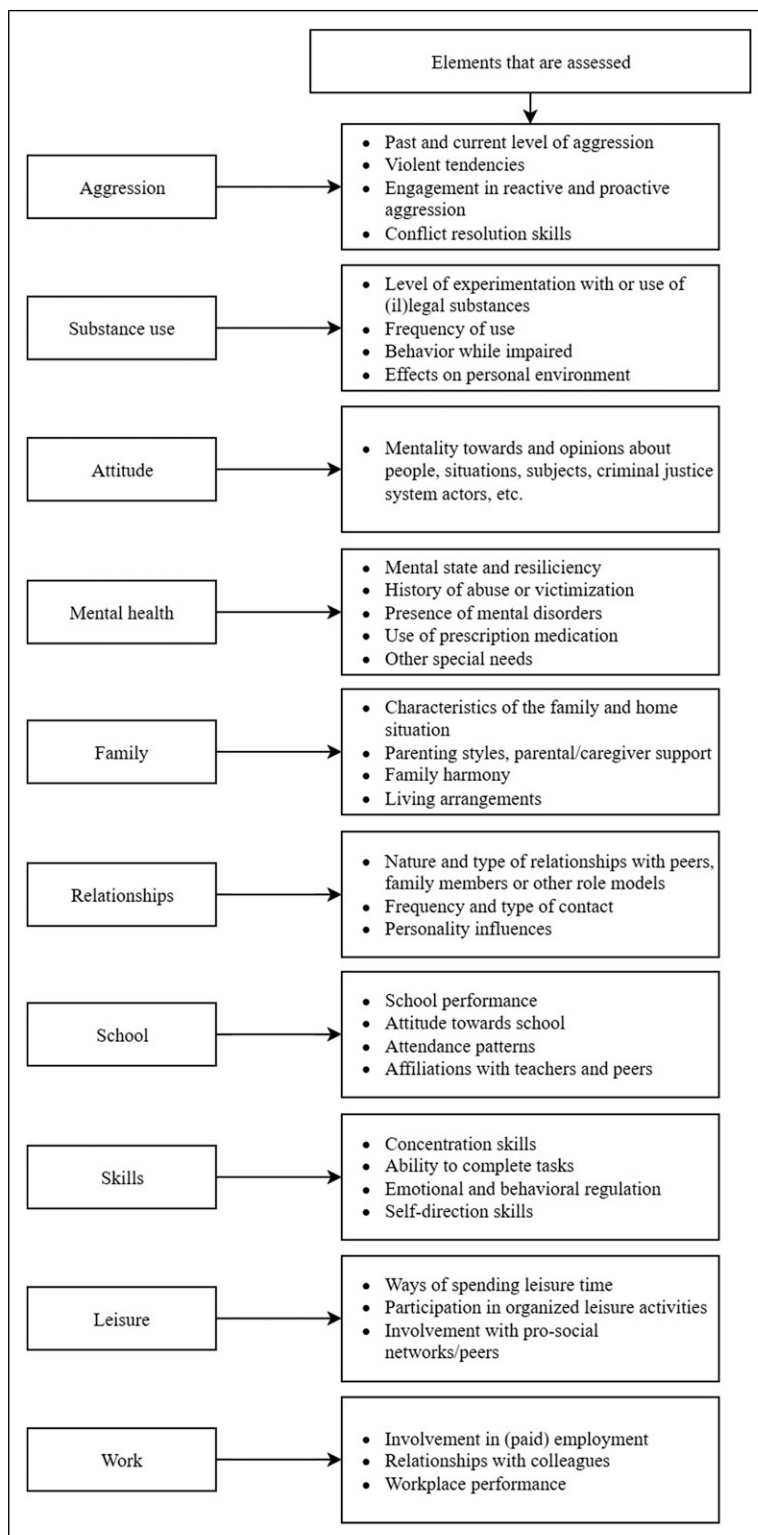


Figure 1. Detailed Explanation of the Risk Assessment Domains.

Table 1. Sample Characteristics before Matching.

Baseline characteristic	Community(<i>n</i> = 4045)		Custody (<i>n</i> = 380)	
	<i>n</i>	%	<i>n</i>	%
Sex				
Male	3680	87.4	352	92.4
Female	532	12.6	29	7.6
CO.B.				
Netherlands	3775	89.6	286	75.1
Other	437	10.4	95	24.9
Court district				
North	140	3.3	23	6.0
East	1075	25.5	64	16.8
South	494	11.7	25	6.6
West	2503	59.4	269	70.6
Offense type				
Violent	857	20.3	42	11.0
Sexual	58	1.4	13	3.4
Property w violence	398	9.4	132	34.6
Property w/o violence	1706	40.5	144	37.8
Vandalism/ Public order	854	20.3	27	7.1
Drugs	118	2.8	18	4.7
Traffic	88	2.1	0	0
Other	133	3.2	5	1.3
Maximum imposable penalty				
Mild	1127	26.8	29	7.6
Moderate-severe	2469	58.6	188	49.3
Severe	616	14.6	164	43.0
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Age	16.3	1.3	16.6	1.3
Criminal history				
Prior contacts	1.2	1.7	2.3	2.7
Prior community sanctions	0.3	0.7	0.7	1.0
Prior custodial sanctions	0.1	0.3	0.2	0.6
Sanction characteristics				
Number of charges in case	1.7	1.1	2.2	1.6
Length of community sanction (hours)	43.1	29.9	N/A	N/A
Length of custodial sanction (days)	N/A	N/A	45.8	34.3
Risk assessment scores				
Aggression	3.6	3.0	4.8	2.9
Substance use	1.4	2.0	1.7	2.3
Attitude	4.2	2.5	5.8	2.3
Mental health	1.9	2.2	2.5	2.3
Family	2.5	1.9	3.5	1.8
Relationships	3.1	2.4	4.1	2.5
School	3.2	2.3	3.9	2.2
Skills	3.1	2.2	4.2	2.0
Leisure	3.4	2.5	4.6	2.2
Work	0.7	1.7	1.1	2.1
Total risk score	2.7	1.4	3.6	1.4

Overall, the final sample predominantly consisted of males, specifically 87% among the community sanction group, and 92% among the custodial sanction group. In total, 90% of the community sanctioned youth versus 75% of the custodially sanctioned youth were born in the Netherlands with mean ages of 16.3 (M_{com}) and 16.6 (M_{cust}) years. Among both groups, cases were predominantly handled in court districts in the Western region of the Netherlands (59% com, 71% cust), and the most common conviction offense type was property crime (41% com, 38% cust). In total, 73% of the community sanction group, versus 92% of the custodial sanction group fell within the 'moderate-severe' and 'severe' category in terms of the maximum imposable penalty given the conviction offense. The average sentence length for youth subject to community sanctions was 43.1 hours compared to 45.8 days for youth subject to custodial sanctions. In terms of criminal history, youth subject to community sanctions had fewer prior community sanctions ($M = 0.33$) and prior custodial sanctions ($M = 0.05$), than youth subject to custodial sanctions ($M_{com} = 0.69$, $M_{cust} = 0.26$). Table 1 also displays the average risk assessment scores for all ten domains within the risk assessment tool. Youth subject to community sanctions scored lower on all ten domains, including the total risk assessment score with an average risk score of 2.7 (low) for the community sanction group, versus 3.6 (medium) for the custodial sanction group.

In sum, female youth, youth born in the Netherlands, youth with less severe conviction offenses, and youth with less significant criminal history were overrepresented among the community sanction group. This highlights the potential presence of selection effects, and thus demonstrates the necessity of controlling for such bias by means of sound empirical methods, such as the application of matching techniques.

Measures

Outcome Measures. The primary outcome variable in our study is the prevalence of recidivism, which was defined as any new, registered justice system contact resulting in a conviction, excluding technical acquittals and not-guilty verdicts. To assess robustness and heterogeneity within the outcomes, we also investigated recidivism seriousness, differentiating between serious recidivism and very serious recidivism. Serious recidivism was defined as any new, registered justice system contact with an imposable penalty of at least 4 years (e.g., participation in a criminal organization, forgery, common assault). Very serious recidivism was defined as any new, registered justice system contact with an imposable penalty of at least 8 years (e.g., attempt murder, rape, violent property crime).

For community sanctioned youth, the observation period commenced at the date of conviction, while the observation period for custodially sanctioned youth started upon release from custody. Nevertheless, the observation period for both groups, across all three outcome measures was exactly 2 years (730 days). The effects of incapacitation were not included in the main analyses of this study, since the main area of interest was the effect of the sanction itself, and youth in custody have few opportunities to recidivate.

As mentioned, we were also interested in whether recidivism outcomes differed depending on pre-defined recidivism risk levels. For this purpose, we used a total, average risk score variable, which was a weighted average of individual risk scores on each of the ten NYJSRAT risk assessment domains. In accordance with the official interpretation of the risk assessment tool, scores ranging from 0-3 indicated 'low' risk youth, scores $>3<6$ indicated 'medium' risk youth, and scores $>6<9$ indicated 'high' risk youth. As less than 2% of the sample fell within the 'high' category (with a maximum score of seven) we could not analyze the high-risk group separately, and therefore we combined these youth with the medium-risk subsample. Moreover, this naturally almost evenly divided the youth in our sample between these two risk categories. In turn, two measures were constructed indicating whether a youth was at 'low' or 'medium-high' risk of recidivism.

Independent Variables. As this study sought to assess the effect of specifically, community sanctions compared to custodial sanctions in youth justice, the focal independent variable (treatment group) was whether youth in the sample received an unconditional, community sanction. In the Netherlands, community sanctions can consist of mandatory community service, mandatory behavioral intervention, or a combination of both (Meijer et al., 2021; Uit Beijerse, 2019). Community service involves non-paid labor, such as working in a nursing home, a thrift store, or landscaping public grounds. Behavioral interventions can exist of a short emotional regulation or short social skills training ranging from 25 to a maximum of 40 hours of intervention (Meijer et al., 2021). Thus, the definition of community sanctions, as the treatment group in this study, encompassed youth subject to unconditional hours of community service, unconditional hours of short, behavioral intervention, or a combination of both.

Furthermore, quasi-experimental effect studies typically have a comparison (control) group in addition to the treatment group, which is referred to as the 'counterfactual' (Holmes, 2013). In essence, the counterfactual constitutes a comparison group that could have received the intervention but did not. The counterfactual must be chosen carefully, especially in quasi-experimental research which does not allow for the natural observation of individuals who did versus did not receive an intervention (Holmes, 2013). In the Dutch youth justice system, there are three major categories of punishment options namely, fines (2%), community sanctions (68%) and custodial sanctions (27%) (Meijer et al., 2021). The intervention of interest in this study, namely the community sanction, was initially developed as an alternative to custodial sanctions, and is still frequently utilized and referred to as such (Aebi et al., 2021; Winterdyk, 2015). Given this, youth subject to custodial sanctions were chosen as the counterfactual used for analysis in this study.

Thus, the comparison group chosen as counterfactual in this study included any youth in the sample that received an unconditional, custodial sanction. In the Netherlands, custodial sanctions are comprised of placement in a juvenile prison, where in case of pre-trial detention the number of days spent in pre-trial custody is subtracted from the number of days subject to placement in a juvenile prison, using a 1:1 conversion ratio (Meijer et al., 2021). In the present sample, 76% of custodially sanctioned youth had a pre-trial custody length of greater than 24 hours, and less than 3% experienced a pre-trial custody length greater than 14 days. In sum, the definition of custodial sanctions, as the counterfactual in this study, entailed sanctions predominantly comprised of incarceration (sometimes also referred to as imprisonment or institutionalization) combined with pre-trial custody.

Covariates. A rich set of covariates were included in the matching model that control for factors related to sanction assignment of receiving either a community or custodial sanction, and risk of recidivism. These covariates must adequately reduce selection bias, meaning that the chance for biased or invalid results because the sample does not adequately reflect the broader population, is minimized (Loeffler & Nagin, 2021; Shadish, 2013). In youth justice, various risk factors determine whether judges impose custodial rather than community sanctions. A recent study conducted in the Netherlands found that offense type, pre-trial custody, number of charges per case, convictions offense severity, and criminal history are sentencing determinants related to applying custodial sanctions in youth justice (Wermink et al., 2015). Likewise, studies on recidivism risk factors in youth justice indicate that criminal history, family issues, delinquent family members or peers, antisocial behavior, substance use, and school performance as factors affecting recidivism (Cottle et al., 2001; Ortega-Campos et al., 2016; Scott & Brown, 2018).

In this study, the following control variables were available for inclusion: age, sex, country of birth, number of charges per case, sentence length, conviction offense type, maximum imposable penalty, criminal history (prior justice system contacts, prior community sanctions, prior custodial sanctions), aggression, substance use, attitude, mental health, family, relationships, school, skills,

leisure, and work. Age was a continuous indicator of a youth's age upon the court case registration date, and sex was a binary indication of whether a youth was officially registered as male or female. Country of birth indicated whether a youth was born in the Netherlands, which was coded binary, namely, born in the Netherlands versus not born in the Netherlands. Number of charges per case indicated whether the total number of offenses charged per case was one, or more than one as some youths are charged with multiple offenses per case. Sentence length referred to the sanction's duration in hours, which entailed a comparable sentence length for both groups that was created using a 2:1 conversion key used in the Dutch youth justice system, where 2 community service hours translate into 1 day (24 hours) in custody (Uit Beijerse, 2019). Moreover, conviction offense type was classified into violent, sexual, property with violence, property, drugs, traffic, and other offense categories. Maximum imposable penalty referred to the maximum sentence that could be imposed according to the Criminal Law (for adults) given the conviction offense. This construct was divided into two variables, namely moderate-severe (penalty = >4<8 years), and severe (penalty = >8 years). Moreover, the criminal history variables were binary indicators of none, versus any prior criminal justice system contacts, prior community sanctions, and prior custodial sanctions.

Finally, we included ten covariates that were based on the NSJRAT recidivism risk assessment domains, as described in Figure 1. Categories indicated whether youth scored low, medium, or high on each of the ten assessment domains. In the analysis phase, sex (male), country of birth (native), conviction offense type (property), sentence length (long), number of charges per case (more than one) and the criminal history (at least one) variables were included in dummy format. All other matching variables were included in continuous format.

Analytic Strategy

Type of Quasi-Experimental Design. As mentioned, well-designed matching studies can be comparable in strength and validity as RCTs, when the requirements of large sample sizes (>1000), and a diverse availability of covariates are met, which is the case in this study (Loeffler & Nagin, 2021; Shadish, 2013). However, there are different types of matching techniques that can be resorted to, such as (coarsened) exact matching, Mahalanobis distance matching, or propensity score matching (PSM) (Stuart, 2010). Perhaps the best matching technique involves exact matching, where cases in both groups are matched based on exactly similar covariate values. This implies that two individuals who are exactly similar on all potentially relevant characteristics (covariates) are compared to each other (Rosenbaum, 2010). However, a well-known problem with exact matching is high drop-out rates, as many cases in one group may not be exactly matched to at least one case in the other group, especially when there are multiple covariates in the model (Rosenbaum, 2010). Yet, as recommended by Shadish (2013), well-designed matching models include as many relevant covariates into the matching model as possible.

Therefore, in this study we utilized PSM techniques that match cases between the community and custodial sanction groups, based on an overall distance between the cases in either group. The overall distance is known as the propensity score distance, which is the chance that a youth (would have) received a community sanction, given the scores on each of the covariates included in the model (Rosenbaum, 2010; Shadish, 2013). Thus, the propensity score for each case/youth is a total value that is calculated based on the combined scores of each covariate included in the model (Rosenbaum, 2010). Individuals from the community sanction group are then paired to individuals in the custodial sanction group in a way that minimizes the differences between the propensity scores of each pair. Overall, PSM has been found to yield reliable decreases in covariate imbalance upon matching, despite its inherent limitations (Ripollone et al., 2019; Stuart, 2010; Shadish, 2013). Thus, PSM techniques are capable to pair cases in such a way that the two groups to be compared are very similar to each other regarding the relevant characteristics (covariates), and as such reliable outcomes analyses can be conducted.

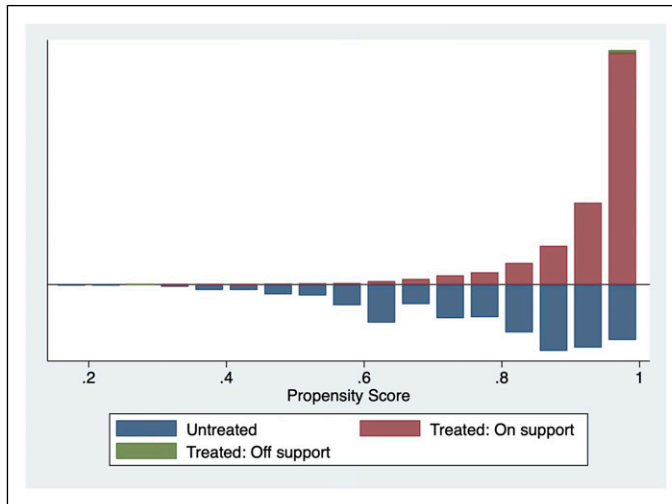


Figure 2. Common Support Graph.

Generating Matches. In this study, we utilized the matching algorithm *psmatch2* in Stata 17 to generate matches (Leuven & Sianesi, 2003; Stata Corp, 2021). Following recommendations from Green & Stuart (2014), the matching and outcome analysis process as described below is conducted separately for (1) the full sample, (2) a low-risk subsample, and (3) a medium-high risk subsample.

In generating matches, we only allowed cases to be matched that fell within the region of ‘common support’, which is the region where the individual propensity scores of cases between both groups overlap (Rosenbaum, 2010). In this way, it was ensured that cases in either group had a reasonable chance of receiving the intervention, or not receiving the intervention (Caliendo & Kopeinig, 2008; Rosenbaum, 2010). As demonstrated in Figure 2, an assessment of the common support graph indicated there was good overlap between the community and custodial sanction cases eligible for inclusion in matching.

Moreover, we used nearest neighbor (NN) matching because this algorithm obtained the highest level of balance compared to other matching algorithms. NN matching operates through selecting a case from the control group that is closest to a case in the treatment group in terms of the combined propensity score (Caliendo & Kopeinig, 2008). Given the small size of the custodial sanction group ($n = 380$) relative to the size of the community sanction group ($n = 4014$), we conducted NN matching with replacement, which allowed for the repeated use of control group cases, up to a maximum of four times, in generating matches. Additionally, we used a caliper distance of 0.02, meaning that all cases within a difference distance of 0.02 in terms of the difference in propensity scores between two cases in either group were included, and matched to one another. Consequently, cases that were outside this caliper distance were deemed too different and thus not included.

Although NN matching was most robust in terms of balance, we conducted radius and kernel matching and interpreted their output for sensitivity purposes. Radius matching is like NN matching, as in that it also uses a pre-determined caliper distance. However, in radius matching all cases that fall within the pre-determined caliper distance are used, rather than only those cases that are both NN and fall within the caliper distance (Caliendo & Kopeinig, 2008). Kernel matching operates using weights and weighs the propensity score of each case to construct matches. Matches that are close approximates in terms of the propensity scores receive more weight than

matches that are far-off (Caliendo & Kopeinig, 2008). The band width (the maximum allowed difference in weights between two cases in either group) used was 0.10.

Finally, the quality of matching was evaluated statistically through assessment of the obtained balance on each of the covariates between both groups. In the Stata output, we evaluated the obtained balance using the ‘standardized % bias’ (% bias) (Leuven & Sianesi, 2003), which is the difference in sample means between the two subsamples, as a percentage of the square root of the mean of the sample variance, between both groups. This statistic was displayed in percentages and generated for each covariate included in the matching model. For interpretation purposes, we used Cohen’s (1977) suggestion that covariates with a remaining % bias of >10% may be considered as having a problematic remaining imbalance. In addition, we interpreted overall balance statistics provided by the Stata output, namely Rubin’s B, Rubin’s R, and VAR. Rubin’s B provides the absolute standardized difference of means from a linear index of the propensity score between both groups. Ideally, Rubin’s B is <25 after matching. Rubin’s R is the ratio of variance between the matched groups, given the propensity score index. Ideally, Rubin’s R is >0.5<2.0 after matching. Lastly, %VAR is the variance ratio for continuous variables in the model, whereby exactly 100 indicates good balance (Leuven & Sianesi, 2003).

Estimating Treatment Effects. To estimate treatment effects, the Stata algorithm *psmatch2* generates probit regression models to estimate the average treatment effect (ATE) on the full sample for the specified outcome measures. Thus, we calculated the effects of community sanctions on youth who were and were not subject to community sanctions in relation to overall recidivism, seriousness of recidivism, and pre-defined recidivism risk levels. As we expected a specific direction of effects the output of a one-sided t-test, as well as Cohen’s *d* are used to interpret the outcomes. The significance of the one-sided t-test is interpreted as weak for $p = .10$, moderate for $p = .05$ and strong for $p = .01$. Cohen’s *d* is interpreted as a small effect for $d = 0.02$, a medium effect for $d = 0.05$, and a large effect for $d = 0.08$.

Results

Quality of Matching

Table 2 outlines the (im)balance on the covariates included in the matching model, before and after matching on (1) the full sample, (2) the low-risk subsample and (3) the medium-high risk subsample. Among the full sample, 99% of the community sanction group ($n = 4014$) was matched to 100% of the custodial sanction group ($n = 380$). As demonstrated in Table 2, the remaining % bias after matching was insignificant on almost all covariates in the model. Prior to matching, 20 covariates were imbalanced among the full sample, whereas after matching all covariates were balanced, except for the recidivism risk factor ‘relationships’ (12.2). Upon further examination it appeared the ‘relationships’ covariate was neither predictive of treatment assignment, nor recidivism. Moreover, the overall mean bias before matching (32.2) was significantly reduced to 4.4, Rubin’s B (125.7) was reduced to 28.5, Rubin’s R (1.1) remained under 2, and % VAR (100) indicated there was good balance.

Among the medium-high risk subsample, 99% of the community sanction ($n = 1806$) was matched to 100% of the custodial sanction group ($n = 276$). Again, the recidivism risk factor ‘relationships’ (16.4) was the only covariate with a residual % bias of >10% after matching. As in the full sample, this covariate was neither predictive of treatment assignment nor recidivism. Mean bias before matching (19.6) was significantly reduced to 3.2, Rubin’s B (115.8) was reduced to 28.2, Rubin’s R (1.2) remained under 2, and % VAR (100) indicated there was good balance.

Table 2. (Im)Balance on the Control Variables Before and After Matching.

Control Variable	Overall Sample		Med-High Risk Subsample		Low Risk Subsample	
	% bias before matching	% bias after matching	% bias before matching	% bias after matching	% bias before matching	% bias after matching
Sex (Male)	-18.3	-0.4	-19.6	-2.5	-21.6	-8.0
C.O.B. (Netherlands)	37.9	6.0	36.0	8.6		
Offense type (property crime)	7.3	4.1	11.7	4.1		
Maximum imposable penalty						
Moderate-severe	48.8	-8.0	50.5	-0.3		
Severe	65.3	0.2	64.6	0.9		
Age	-19.4	9.3	-24.9	-1.9		
Criminal history						
Prior contacts	-23.2	-2.2	-8.3	-7.7	-17.4	5.4
Prior community sanctions	-31.1	-1.5	-20.2	-4.6	-32.2	-2.3
Prior custodial sanctions	-40.3	-4.6	-38.9	2.3	-31.2	-1.8
Sanction characteristics						
Number of charges per case	-32.6	3.6	-27.0	3.2	-25.0	-4.3
Length of sentence	-47.4	0.2	-52.8	-1.0		
Risk assessment scores						
Aggression	-31.7	3.4	-8.6	10.1	5.1	-4.8
Substance use	-11.3	5.7	-5.2	7.0	16.3	-9.4
Attitude	-38.8	6.6	-18.9	-0.7	-7.8	-12.3
Mental health	-15.3	1.7	1.6	3.6	5.4	11.2
Family	-46.5	6.1	-29.0	1.7	-11.2	-0.8
Relationships	-30.0	12.2	-22.2	16.4	34.2	2.9
School	-19.4	-7.3	8.1	-6.3	-2.1	-2.7
Skills	-47.5	-4.9	-13.6	-5.5	-17.1	1.7
Leisure	-46.1	-4.1	-9.9	1.4	-34.0	-0.8
Work	-18.7	-0.4	-5.9	-2.9	-6.5	2.3
Summary Statistics						
MeanBias	32.2	4.4	19.6	3.2	17.7	4.9
Rubin's B	125.7	28.5	115.8	28.2	75.1	26.5
Rubin's R	0.9	1.1	0.9	1.2	1.0	1.3
% VAR	0	100	0	100	0	100

Among the low-risk subsample, 97% of the community sanction group ($n = 2167$) was matched to 100% of the custodial sanction group ($n = 104$). However, we were unable to obtain good balance utilizing the full covariate model, as was done for the other two samples. Therefore, for the low-risk subsample we only matched on the actual predictors of recidivism among this sample

(age, sex, prior contact, prior custodial sanctions, and prior community sanctions), in addition to the ten recidivism risk factors. Upon matching using this model, there are two covariates that remained out of balance, namely the recidivism risk factor 'attitude' (12.3) and 'mental health' (11.2). Yet, mean bias before matching (17.7) was significantly reduced to 4.9, Rubin's B (75.1) was reduced to 26.5, Rubin's R (1.3) remained under 2, and % VAR (100) indicated there was good balance. Moreover, 'attitude' was not predictive of treatment assignment or recidivism. However, 'mental health' was a predictor of recidivism, but upon further investigation it appeared that 87% of the low-risk subsample scored on the lower end of this risk factor, and as such we regarded the small imbalance as not problematic.

In sum, the obtained balance upon matching all three samples indicated we may be confident that observed differences in recidivism in the outcome results are not due to pre-existing differences in observable characteristics between the community and custodial sanction groups. Therefore, we proceed with interpretation of the results.

Results Overall Sample

As demonstrated in Table 3, the direction of the effect was negative ($t(21) = -1.46, p = .08$), meaning youth subject to community sanctions were less likely to recidivate than youth subject to custodial sanctions. This effect was significant at $p < .10$, whereas the results for serious recidivism ($t(21) = -1.95, p = .03$) were moderately significant ($p < .05$). Thus, in addition to a lower likelihood of overall recidivism, youth subject to community sanctions were less likely to recidivate seriously than youth subject to custodial sanctions. Yet, the results on very serious recidivism were not significant ($t(21) = -0.63, p = .27$), although the direction of results was in favor of the community sanction group. According to Cohen's D ($d = -0.64$), the reducing effect of community sanctions, as opposed to custodial sanctions was medium in terms of overall recidivism, and large ($d = -0.85$) in terms of serious recidivism.

Results Low-Risk Subsample

Among the low-risk subsample, the overall recidivism outcome was in favor of the community sanction group ($t(16) = -1.39, p = .09$), which was significant at $p < .10$. Thus, we found that low risk youth subject to community sanctions were less likely to recidivate than low risk youth subject to custodial sanctions. The results for serious recidivism were comparable ($t(16) = -1.67, p = .06$), and interestingly, the results for very serious recidivism among the low-risk youth were also significant ($t(16) = -1.77, p = .05$), and according to Cohen's D, large of magnitude ($d = -0.89$). Thus, for low-risk youth, community sanctions have a large reducing effect on the prevalence of very serious recidivism. Furthermore, the reducing effect of community sanctions as opposed to custodial sanctions on overall recidivism was medium ($d = -0.70$), and large in terms of serious recidivism ($d = -0.84$).

Results Medium-High Risk Subsample

Table 3 demonstrates that the overall recidivism outcome for the medium-high risk subsample was significant ($t(21) = -1.78, p = .04$), and in favor of the community sanction group. Thus, medium-high risk community sanctioned youth are less likely to recidivate than medium-high risk custodially sanctioned youth. Moreover, the results for serious recidivism ($t(21) = -2.46, p = .01$) were strongly significant ($p < .01$), yet the results for very serious recidivism were not significant ($t(21) = -1.29, p = .11$), however the direction is in favor of the community sanction group. According to Cohen's D the reducing effect of community sanctions, as opposed to custodial

Table 3. Outcome Results.

Outcome Type	Overall sample			Low risk youth			Med-high risk youth								
	t(21)	p	d	% com	% cust	t (16)	p	d	% com	% cust	t (21)	p	d	% com	% cust
Overall Recidivism	-1.46	.08*	-0.64**	42	50	-1.39	.09*	-0.70**	35	45	-1.78	.04**	-0.78**	51	60
Serious Recidivism	-1.95	.03**	-0.85***	37	47	-1.67	.06*	-0.84***	30	41	-2.46	.01***	-1.07***	46	58
Very Serious Recidivism	-0.63	.27	-0.27*	7	9	-1.77	.05**	-0.89***	5	13	-1.29	.11	-0.50*	9	14

* $p < .10$. ** $p < .05$. *** $p < .01$.* d = small. ** d = medium. *** d = large.

sanctions, was large in magnitude among medium-high risk youth for overall recidivism ($d = -0.78$) and serious recidivism ($d = -1.07$). This implies that medium-high risk youth subject to community sanctions were less likely to recidivate, and less likely to recidivate seriously than youth subject to custodial sanctions.

Sensitivity Analysis

For sensitivity purposes, we conducted radius and kernel matching to check the robustness of our findings using different, nonparametric matching algorithms. Radius matching returned very similar results as NN matching, which was the matching estimator that was used to interpret the main outcome analyses. However, radius matching yielded slightly less balance on the covariates in the model than NN matching. Conversely, kernel matching yielded outcome results in the same direction, but with much larger effect sizes. Given this, the remaining imbalance using kernel matching was much larger than when radius or NN matching was used, thus the results of kernel matching are also the least robust. Among the three different matching algorithms the outcomes are significant, and in the same direction, namely in favor of the community sanction group, thereby providing greater confidence in the current study's findings.

Conclusion & Discussion

This study involved a quasi-experimental analysis of 2-year recidivism differences between 4425 youth subject to a community, or a custodial sanction in the Netherlands in 2015 and 2016. In addition to a large sample size, a wide variety of important background characteristics were available for analysis, and as such, were able to analyze meaningful differences beyond overall recidivism. Namely, differences in serious and very serious recidivism, as well as differences between low-risk youth versus medium-high risk youth.

Overall, the results of the full sample match our first hypothesis that we expected to find less overall recidivism among youth subject to community sanctions, than youth subject to custodial sanctions. In relative terms, the prevalence of overall recidivism is 42% for the community sanctioned youth, compared to 50% for the custodial sanctioned youth. In essence, this finding is contrary to the deterrence (Beccaria, 1986) and rational choice (Clarke & Cornish, 1986) perspective, which would expect custodial sanctions to yield less recidivism as they are more severe and have higher perceived costs. Yet, this finding is in congruence with the labelling (Becker, 1963), and differential association (Sutherland, 1939) perspective, which suggest that youth subject to community sanctions would yield better effects as it may involve less extensive labelling, and less exposure to (opportunities to learn from) deviant others. Moreover, it also follows social control (Hirschi, 1969) theory's notions on the necessity to maintain positive social bonds to society, and Braithwaite's (1989) emphasis on the necessity of restoration and reintegrative shaming. However, of note is that the underlying mechanisms of these theories were not empirically tested in this study.

As mentioned, prior studies specifically comparing the effects of community sanctions versus custodial sanctions tend to not find significant differences in overall recidivism (Bontrager-Ryon et al., 2017; Brandau, 1992; Schneider, 1986; Van der Laan & Essers, 1990; Van der Laan & Essers, 1993; Van Wormer, 2018). However, the sample size in this study was much larger than any of the previously discussed studies, which certainly yielded more power for effect analyses. Therewith, the overall result in this study provides additional evidence which strengthens the existing evidence-base of the positive effect of community sanctions, as opposed to custodial sanctions, in youth justice.

Furthermore, given prior research, the second hypothesis was that youth subject to community sanctions would engage in less serious recidivism than youth subject to custodial sanctions. This hypothesis is confirmed in terms of serious recidivism ($t(21) = -1.95, p = .03$), however not in terms of very serious recidivism ($t(21) = -0.63, p = .27$). Specifically, the prevalence of serious recidivism is 37% for the community sanctioned youth, compared to 47% for the custodially sanctioned youth. The prevalence of very serious recidivism is 7% for the community sanctioned youth compared to 9% for the custodially sanctioned youth.

Further noteworthy are the results on serious recidivism, which are in congruence with previous study findings, who typically did find that community sanctioned youth recidivated less seriously than custodially sanctioned youth (Bontrager-Ryon et al., 2017; Brandau, 1992; Essers et al., 1995; Schneider, 1986; Van der Laan & Essers, 1990; Van der Laan & Essers, 1993). Moreover, finding that youth subject to community sanctions are less likely to recidivate (seriously) than youth subject to custodial sanctions yields further support for the continued, and perhaps increased application of community sanctions, as opposed to custodial sanctions in youth justice.

The third hypothesis implied that low-risk youth would benefit more from community sanctions, and that higher risk youth would benefit more from custodial sanctions. Indeed, low-risk youth subject to community sanctions are less likely to recidivate than low-risk youth subject to custodial sanctions, in terms of overall ($t(16) = -1.39, p = .09$), serious ($t(16) = -1.67, p = .06$), and very serious recidivism ($t(16) = -1.77, p = .05$). In relative terms, this is 35% (*com*) versus 45% (*cust*) for overall recidivism, 30% (*com*) versus 41% (*cust*) for serious recidivism, and 5% (*com*) versus 13% for very serious recidivism. As expected, low-risk youth benefit more from community sanctions than custodial sanctions, and custodial sanctions are rather harmful in terms of the increased risk of overall, serious, and very serious recidivism. Instead of preventing, or reducing recidivism, custodial sanctions administered to low-risk youth increase the odds of committing future offenses, in particular the odds of committing very serious offenses upon sanction administration. As such, community sanctions, rather than custodial sanctions should be administered to low-risk youth. Although this may already be common in practice, low risk youth do at times receive custodial sanctions, as there was a sample of low-risk youth with custodial sanctions available for analysis in this study. Yet, the findings of this study suggest that the administration of custodial sanctions to low-risk youth should perhaps be avoided.

However, contrary to our expectations, it was found that among medium-high risk youth, community sanctions also yield less overall recidivism ($t(21) = -1.78, p = .04$), and less serious recidivism ($t(21) = -2.46, p = .01$), than custodial sanctions. The overall recidivism rate is 51% for medium-high risk community sanctioned youth versus 60% for medium-high risk custodial sanctioned youth. In terms of serious recidivism, this rate is 46% (*com*) versus 58% (*cust*), and 9% (*com*) versus 14% (*cust*) for very serious recidivism. These findings also seem contrary to deterrence and rational choice theory, whereby one may expect that especially, higher risk youth are more likely to be deterred following severe sanctions, with higher costs (Beccaria, 1986; Clarke & Cornish, 1986). Perhaps, higher risk youth subject to community sanctions are exposed to greater opportunities for rehabilitation aimed at resolving the underlying sources of strain as causes of offending behavior (Agnew, 1992), than higher risk youth subject to custodial sanctions. This could explain finding lower recidivism among higher risk, community sanctioned youth rather than custodially sanctioned youth, as previous studies have found that especially high-risk youth benefit more from treatment interventions (Lipsey, 2009; Wilson & Hoge, 2013).

Nonetheless, this finding may have significant implications for the application of risk assessments in youth justice. Delinquent youth with higher risk scores may receive more severe sanctions given their increased recidivism risk, and therefore, presumed increased risk to society. However, the current findings indicate this line of reasoning increases the risk of both overall, and serious recidivism when short custodial sanctions are resorted to. In essence, the most severe type

of sanction, namely custodial sanctions, do not lead to its goal of reduced recidivism for higher risk youth, at least not in the short-term format of maximally 4 months, which was the case in this study. As such, it appears that among higher risk youth, community sanctions pose a better alternative than custodial sanctions. Alternatively, it may also be the case that higher risk youth need more intensive rehabilitation treatment and programming than possible within the timeframe of short custodial sanctions (RSJ, 2021).

Limitations & Directions

Several limitations of the current study's design should be discussed. Firstly, in our analysis of effects we only assessed different types of recidivism following sanction administration. However, alternative measurements of effectiveness upon sanction administration may be just as relevant as, or even more so than recidivism (Rosenfeld & Grigg, 2022). For example, community and custodial sanctions may differ in their effect on future employment, housing, social support, and (mental) health status. Unfortunately, the data in this study only provided information on the status of such measurements prior to, and not post sanction administration; the latter of which is necessary to assess whether any improvements in relation to these domains occurred.

Secondly, the custodial sanctions in this study were limited to a maximum length of 4 months, as this length had to be comparable to the maximum possible length of community sanctions. This implies that the current findings do not directly apply to longer custodial sanctions utilized in youth justice. Overall, short custodial sanctions significantly reduce the opportunities for rehabilitation (programs) in prison, meaning the underlying causes of offending behavior may remain unresolved, which in turn can lead to reduced effectiveness in terms of recidivism. A recent analysis of short, custodial sanctions in the Dutch youth justice system also emphasized that short custodial sanctions provide limited opportunities for rehabilitation, and thus meaningful alternatives should be resorted to (RSJ, 2021). In essence, while short custodial sanctions may not encompass enough time to allow for sufficient rehabilitation, long(er) custodial sanctions might.

Indeed, sanction environments with sufficient opportunities for rehabilitation may be more effective than sanction environments without that, regardless of the type of sanction, which emphasizes another limitation of the present study. Although there were extensive attempts to control for selection bias by means of including high quality covariates and the application of matching techniques, the possibility of unobserved characteristics related to sanction assignment or recidivism that were not accounted for remains, which could have impacted the results. Such as, a youth's attitude in court, or environmental differences within the (institutional) setting at which the community or custodial sanction is served. Additionally, we examined community sanctions existing of either community service, behavioral intervention, or a combination of both, without differentiating between these three types. Yet, community sanctions entirely comprised of community service hours may yield different effects than behavioral interventions specifically focused on inducing cognitive behavioral change. Future research should thus aim to provide a more detailed examination of heterogeneity in sanction environments and sanction types, as such differences can play an important role in sanction effects (Mears et al., 2015).

Future research with the possibility of an experimental setting, such as random assignment would be beneficial to overcome methodological limitations like unobserved selection bias. Alternatively, one can apply different types of matching techniques or analyses, such as instrumental variable analysis to analyze differences between the effects of community versus custodial sanctions in youth justice. Moreover, the effects of longer custodial sanctions, in comparison to community sanctions or short custodial sanctions could be further investigated, as well as the specific effects of community sanctions in relation to other domains besides recidivism. Provided that such measures are not always quantifiable, further research into the effects of

community sanctions in youth justice could entail (longitudinal) qualitative approaches and emphasize the perspective of youth subject to these sanctions themselves (the emic perspective), or their immediate environment.

In sum, when assessing the recidivism effects of community sanctions, compared to custodial sanctions in youth justice, community sanctions present a good alternative to custodial sanctions that lead to reduced overall, serious, and in some cases, very serious recidivism. Given the increased use of community sanctions, versus the decreased use of custodial sanctions one may state that perhaps, custodial sanctions have become the alternative to community sanctions, as opposed to vice versa. Considering this, the current study's findings, and the special legal position of youth, practical implications include that when short custodial sanctions may be imposed, the application of community sanctions, or more intense (custodial) sanctions with greater opportunities for rehabilitation should be considered instead. Additionally, community sanctions, rather than custodial sanctions were found to be beneficial for both low and medium-high risk youth in terms of reducing recidivism, meaning that overall, the application of short custodial sanctions should be carefully (re)considered.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Gwendolyn J. Koops-Geuze  <https://orcid.org/0000-0001-5776-304X>

References

- Aebi, M. F., Caneppele, S., Harrendorf, S., Hashimoto, Y. Z., Jehle, J., Khan, T.S., Kühn, O., Lewis, C., Molnar, L., Þórisdóttir, R., & Smit, P (2021). *European sourcebook of crime and criminal justice statistics 2021*. 6th ed. https://wp.unil.ch/europeansourcebook/files/2021/08/ESB_Data_Final_210731.pdf
- Agnew, R. (1992). Foundation for a general strain theory of crime and delinquency. *Criminology*, 30(1), 47–88. <https://doi.org/10.1111/j.1745-9125.1992.tb01093.x>
- Baird, C., Healy, T., Johnson, K., Bogie, A., Dankert, E. W., & Scharenbroch, C. (2013). *A comparison of risk assessment instruments in juvenile justice*. National Council on Crime and Delinquency. https://www.njcn.org/uploads/digital-library/NCCD_risk-assessment-comparison_August.pdf
- Bateman, T. (2017). *The state of youth justice 2017 an overview of trends and developments*. National Association for Youth Justice. <https://uobrep.openrepository.com/bitstream/handle/10547/622241/State-of-Youth-Justice-report-forweb-Sep17.pdf?sequence=3>
- Beccaria, C. (1986). *On crimes and punishment*. Hackett Publishing Company
- Becker, H. S. (1963). *Outsiders: Studies in the Sociology of deviance*. Free press Glencoe
- Braithwaite, J. (1989). *Crime. Shame and reintegration*. Cambridge University Press
- Brandau, T. J. (1992). An alternative to incarceration for youth offenders: The Delaware Bay marine institute (Doctoral Dissertation). *ProQuest Dissertations & Theses Global Database*. UMI No. 9232615
- Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1), 31–72. <https://doi.org/10.1111/j.1467-6419.2007.00527.x>

- Clarke, R. V., & Cornish, D. B. (1986). *The reasoning criminal: Rational choice perspectives on offending*. Springer-Verlag
- Cochran, J. C., Mears, D. P., & Bales, W. D. (2014). Assessing the effectiveness of correctional sanctions. *Journal of Quantitative Criminology*, 30(2), 317–347. <https://doi.org/10.1007/s10940-013-9205-2>
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge
- 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. <https://doi.org/10.1177/0093854801028003005>
- Cox, P., & Godfrey, B. (2020). The ‘great decarceration’: Historical trends and future possibilities. *The Howard Journal of Crime and Justice*, 59(3), 261–285. <https://doi.org/10.1111/hojo.12377>
- Cullen, F. T., & Gendreau, P. (2001). From nothing works to what works: Changing professional ideology in the 21st century. *The Prison Journal*, 81(3), 313–338. <https://doi.org/10.1177/0032885501081003002>
- Essers, A., Laan van der, P. H., & Veer van der, P. N. (1995). *Cashba: Een intensief dagprogramma voor jeugdige en jongvolwassen delinquenten*. WODC. <https://www.wodc.nl/onderzoeksdatabase/w00145-cashba-een-intensief-dagprogramma-voor-jeugdige-en-jongvolwassen-delinquenten.aspx>
- Fagan, J., & Kupchik, A. (2011). Juvenile incarceration and the pains of imprisonment. *Duke Forum for Law and Social Change*, 3(1), 29–61. <https://doi.org/10.2139/ssrn.1772187>
- Farrington, D. P., Gottfredson, D. C., Sherman, L. W., & Welsh, B. C. (2002). The Maryland scientific methods scale. In D. P. Farrington, D. C. Gottfredson, L. W. Sherman, & B. C. Welsh (Eds.), *Evidence-based crime prevention* pp. 13–21. Routledge
- Farrington, D. P., Loeber, R., Howell, J. C., Loeber, R., & Howell, J. C. (2012). Young adult offenders: The need for more effective legislative options and justice processing. *Criminology & Public Policy*, 11(1), 729–750. <https://doi.org/10.1111/j.1745-9133.2012.00842.x>
- Feld, B. C. (2013). Adolescent criminal responsibility, proportionality, and sentencing policy: Roper, graham, miller/jackson, and the youth discount. *Law and Inequality: Journal of Theory and Practice*, 31(2), 263–330
- Green, K. M., & Stuart, E. A. (2014). Examining moderation analyses in propensity score methods: Application to depression and substance use. *Journal of Consulting and Clinical Psychology*, 82(5), 773–783. <https://doi.org/10.1037/a0036515>
- Hirschi, T. (1969). *Young Adult Offenders*. University of California Press.
- Hockenberry, S., & Puzanchera, C. (2020). *Youth court statistics, 2018 (report No. 254798)*. <https://ojdp.ojp.gov/library/publications/youth-court-statistics-2018>
- Holmes, W. M. (2013). *Using propensity scores in quasi-experimental designs*. Sage Publishing. <https://us.sagepub-com.eur.idm.oclc.org/en-us/nam/using-propensity-scores-in-quasi-experimental-designs/book237237#resources>
- Jehle, J. (2019). Federal ministry of justice and Consumer protection. In *Criminal justice in Germany* (7th ed.). ■■■
- Klein, M. W. (1979). Deinstitutionalization and diversion of juvenile offenders: A litany of impediments. *Crime and Justice*, 1, 145–201. <https://doi.org/10.1086/449061>
- Koops-Geuze, G. J., & Weerman, F. M. (2021). Community sanctions in youth justice compared to other youth crime responses: A meta-analysis. *European Journal of Criminology*. Advance online publication. <https://doi.org/10.1177/1477708211035305>
- Krisberg, B., Currie, E., & Onek, D. (1995). What works with juvenile offenders? A review of “graduated sanction” programs. *Criminal Justice*, 10(2), 20–61
- Latimer, J. (2001). A meta-analytic examination of youth delinquency, family treatment, and recidivism. *Canadian Journal of Criminology*, 43(2), 237–253. <https://doi.org/10.3138/cjcrim.43.2.237>
- Leuven, E., & Sianesi, B. (2003). *PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing*. RePEc. <http://ideas.repec.org/c/boc/bocode/s432001.html>

- Lipsey, M. W. (2009). The primary factors that characterize effective interventions with youth offenders: A meta-analytic overview. *Victims and Offenders*, 4(1), 124–147. <https://doi.org/10.1080/15564880802612573>
- Loeffler, C. E., & Nagin, D. S. (2021). The impact of incarceration on recidivism. *Annual Review of Criminology*, Advance online publication. 030920–112506.
- McGuire, J. E. (1995). *What works: Reducing reoffending. Guidelines from research and practice*. Wiley
- Mears, D. P., Cochran, J. C., & Cullen, F. T. (2015). Incarceration heterogeneity and its implications for assessing the effectiveness of imprisonment on recidivism. *Criminal Justice Policy Review*, 26(7), 691–712. <https://doi.org/10.1177/0887403414528950>
- Meijer, R. F., Moolenaar, D. E. G., Choenni, R., & Van den Braak, S. W. (2021). *Criminaliteit en re-chthandhaving 2020: Ontwikkelingen en samenhangen (Cahier 2021-22)*. WODC. <https://repository.wodc.nl/handle/20.500.12832/3112>
- Mensink, K., Hill, J., & Weijters, G. (2021). *Op zoek naar profielen van jeugdige verdachten: Een verkenning op basis van risico- en beschermende factoren uit het Landelijk Instrumentarium Jeugdstrafrechtketen*. WODC. <https://repository.wodc.nl/handle/20.500.12832/3061>
- Nagin, D. S. (2013). Deterrence in the twenty-first century. *Crime and Justice*, 42(1), 199–263. <https://doi.org/10.1086/670398>
- Nagin, D. S., Cullen, F. T., & Jonson, C. L. (2009). Imprisonment and reoffending. *Crime and Justice*, 38(1), 115–200. <https://www.jstor.org/stable/10.1086/599202>
- Ortega-Campos, E., Garcia-Garcia, J., Gil-Fenoy, M. J., & Zaldivar-Basurto, F. (2016). Identifying risk and protective factors in recidivist juvenile offenders: A decision tree approach. *Plos One*, 11(9), 1–16. <https://doi.org/10.1371/journal.pone.0160423>
- Piquero, A. R., Hawkins, J. D., Kazemian, L., & Petechuk, D. (2013). *Bulletin 2: Criminal career patterns (study Group on the Transitions between youth Delinquency and adult crime) (No. 242932)*. <https://www.ncjrs.gov/pdffiles1/nij/grants/242932.pdf>
- Raad voor Strafrechttoepassing en Jeugdbescherming (2021). *Korte detenties nader bekeken*. <https://www.rsj.nl/documenten/rapporten/2021/10/29/korte-detenties-nader-bekeken>
- Ripollone, J. E., Huybrechts, K. F., Rothman, K. J., Ferguson, R. E., & Franklin, J. M. (2019). Evaluating the utility of coarsened exact matching for pharmacoepidemiology using real and simulated claims data. *American Journal of Epidemiology*, 189(6), 613–622. <https://doi.org/10.1093/aje/kwz268>
- Rosenbaum, P. R. (2010). *Design of observational studies*. Springer Link.
- Rosenfeld, R., & Grigg, A. (2022). *The limits of recidivism: Measuring success after Prison*. The National Academies Press. <https://doi.org/10.17226/26459>
- Ryon, S., Winokur Early, K., & Kosloski, A. E. (2017). Family-based and family-focused alternatives to incarceration: A quasi-experimental evaluation of interventions for delinquent youth. *Journal of Criminal Justice*, 51, 59–66. <https://doi.org/10.1016/j.jcrimjus.2017.06.002>
- Schneider, A. (1986). Restitution and recidivism rates of youth offenders: Results from four experimental studies. *Criminology*, 24(3), 533–552. <https://doi.org/10.1111/j.1745-9125.1986.tb00389.x>
- Schwalbe, C. S., Gearing, R. E., MacKenzie, M. J., Brewer, K. B., & Ibrahim, R. (2012). A meta-analysis of experimental studies of diversion programs for juvenile offenders. *Clinical Psychology Review*, 32(1), 26–33. <https://doi.org/10.1016/j.cpr.2011.10.002>
- Scott, T., & Brown, S. L. (2018). Risks, strengths, gender, and recidivism among justice-involved youth: A meta-analysis. *Journal of Consulting and Clinical Psychology*, 86(11), 931–945. <https://psycnet.apa.org/doi/10.1037/ccp0000343>
- Shadish, W. R. (2013). Propensity score analysis: Promise, reality, and irrational exuberance. *Journal of Experimental Criminology*, 9(1), 129–144. <https://doi.org/10.1007/s11292-012-9166-8>
- StataCorp (2021). *Stata statistical Software: Release 17*. College Station, TX: StataCorp LLC
- Statistics Canada (2020). *Table 35-10-0041-01 Youth courts, guilty by cases by type of sentence [Database record]*. CANSIM. <https://doi.org/10.25318/3510004101-eng>

- Stuart, E. A. (2010). Matching methods for casual inference: A review and look forward. *Stat Sci*, 25(1), 1–21. <https://doi.org/10.1214/09-STS313>
- Sutherland, E. H. (1939). *Principles of Criminology* (3rd ed.). J. B. Lippincott
- Uit Beijerse, J. (2019). *Jeugdstrafrecht: Beginselen, wetgeving en praktijk*. Maklu-uitgevers
- UN General Assembly (1989). *Convention on the rights of the Child*. UN General Assembly
- Van der Laan, P. H., & Essers, A. A. M. (1990). De Kwartaalkursus en recidive. *En onderzoek naar de effecten van het experiment Kwartaalkursus*. WODC. <https://www.wodc.nl/onderzoeksdatabase/w00099-de-kwartaalkursus-en-recidive>
- Van der Laan, A. M., Beerthuizen, M. G. C. J., & Beerthuizen, N. C. (2021). *Monitor jeugdcriminaliteit 2020 (Cahier 2021-9)*. WODC. <https://repository.wodc.nl/Bootbitstream/handle/20.500.12832/3058/Cahier2021-09-volledige-tekst.pdf?>
- Van der Laan, P. H., & Essers, A. A. M. (1993). Helpt DTC? Over recidive en andere effecten SWP. In J. Leeuwen van & A. Maris (Eds), *Vast of Zeker Een kansrijke aanpak buiten de gevangenis: Het dagtrainingscentrum Eindhoven* pp.73–79
- van Wingerden, S., van Wilsem, J., & Moerings, M. (2014). Pre-sentence reports and punishment: A quasi-experiment assessing the effects of risk-based pre-sentence reports on sentencing. *European Journal of Criminology*, 11(6), 723-744. <https://doi.org/10.1177/1477370814525937>
- Van Wormer, J. (2018). Developing an alternative juvenile programming effort to reduce detention over-reliance. *OJJDP Journal of Juvenile Justice*, 5(2), 12-30
- Villetaz, P., Gillieron, G., & Killias, M. (2015). The effects on re-offending of custodial vs. non-custodial sanctions: An updated systematic review of the state of know ledge. *Campbell Systematic Reviews*, 11(1), 1-92. <https://doi.org/10.4073/csr.2015.1>
- Vincent, G. M., Guy, L. S., & Grisso, T. (2012). Risk assessment in juvenile justice: A guidebook for implementation. MacArthurFoundation
- Walgrave, L. (1998). *Restorative justice for youths: Potentialities, risks, and problems for research*. Leuven University Press
- Wermink, H., Johnson, B. D., Nieuwbeerta, P., & de Keijser, J. W. (2015). Expanding the scope of sentencing research: Determinants of juvenile and adult punishment in The Netherlands. *European Journal of Criminology*, 12(6), 739–768. <https://doi.org/10.1177/1477370815597253>
- Wilson, D. B., Brennan, I., & Olaghere, A. (2018). Police-initiated diversion for youth to prevent future delinquent behavior: A systematic review. *Campbell Systematic Reviews*, 14(1), 1–88. <https://doi.org/10.4073/csr.2018.5>
- Wilson, H. A., & Hoge, R. (2013). The effect of youth diversion programs on recidivism: A meta-analytic review. *Criminal Justice and Behaviour*, 40(5), 497–518. <https://doi.org/10.1177/0093854812451089>
- Winterdyk, J. A. (2015). *Youth justice: International perspectives and trends*. CRC Press
- Wong, J. S., Bouchard, J., Gravel, J., Bouchard, M., & Morselli, C. (2016). Can at-risk youth be diverted from crime? A meta-analysis of restorative diversion programs. *Criminal Justice and Behavior*, 43(10), 1310–1329. <https://doi.org/10.1177/0093854816640835>

Author biographies

Gwendolyn J. Koops-Geuze is a PhD candidate at the Department of Law, Crime & Society at the Erasmus University of Rotterdam. Her PhD research examines the effects of community sanctions utilized in youth justice, utilizing a mixed methods design. Her main research interests are sanction effects, youth justice, recidivism, desistance, and mixed methods research.

Hilde T. Wermink, PhD, is an associate professor at the Institute of Criminal Law and Criminology at Leiden University, and PhD dean of the Law faculty. Her research interests include the

determinants of consecutive sentencing decisions, and the effects of imprisonment (and other punishments) on criminal behavior.

Frank M. Weerman, PhD, is an endowed professor in Youth Criminology at the the Department of Law, Crime & Society at the Erasmus University of Rotterdam, and senior researcher at the NSCR (Netherlands Institute for the Study of Crime and Law Enforcement). His research interests are focused on youth crime, in particular the role of peers and groups in the etiology of delinquent behavior.