THE ROCKWOOL FOUNDATION
Research

THE IMPACT OF PATERNAL INCARCERATION ON BOYS' DELINQUENCY: A NEW METHOD FOR ADJUSTING FOR MODEL-DRIVEN BIAS

LARS HØJSGAARD ANDERSEN,
ROCKWOOL FOUNDATION RESEARCH UNIT

KRISTIAN BERNT KARLSON,
UNIVERSITY OF COPENHAGEN

STUDY PAPER 120

JUNE 2017

The Impact of Paternal Incarceration on Boys' Delinquency:
A New Method for Adjusting for Model-Driven Bias

Study Paper No. 120

Published by:

Address:
The ROCKWOOL Foundation Research Unit
Soelvgade 10, 2.tv.
DK-1307 Copenhagen K

Telephone +45 33 34 48 00
E-mail forskningsenheden@rff.dk
web site: www.en.rff.dk

June 2017

The ROCKWOOL Foundation Research Unit

Study Paper No. 120

THE IMPACT OF PATERNAL INCARCERATION ON BOYS' DELINQUENCY: A NEW METHOD FOR ADJUSTING FOR MODEL-DRIVEN BIAS

Lars Højsgaard Andersen,
ROCKWOOL FOUNDATION RESEARCH UNIT

Kristian Bernt Karlson,
UNIVERSITY OF COPENHAGEN

Copenhagen 2017

The Impact of Paternal Incarceration on Boys' Delinquency: A New Method for Adjusting for Model-Driven Bias

ABSTRACT

Existing research is likely to have underestimated the mediating role of confounders, such as abilities, for the association between the delinquency of fathers and sons. Measures of sons' delinquency are often dichotomous, indicating whether a son experiences incarceration, for example, and scholars of criminology often apply nonlinear probability models to analyze such outcomes. But in so doing, we show, scholars involuntarily make their estimates vulnerable to model-driven bias. In this paper, we introduce to scholars of criminology a recent advance in the modeling of nonlinear probability models, the "KHB method", which corrects for model-driven bias. We use data from the NLSY97 to illustrate that existing strategies for estimating the impact of paternal incarceration on son's delinquency when mediating factors are taken into account produce biased estimates, and we use high quality registry data from Denmark to show that this is true even in administrative data and across length of incarceration. We also present and discuss the relationship between delinquency and abilities among sons, by paternal incarceration experience and by country, thereby adding to the burgeoning research on uneven consequences of paternal incarceration for children.

INTRODUCTION

One of the most prominent analytical strategies which scholars of criminology apply to prove the intergenerational transmission of delinquency is mediation analysis (see Walters and Mandracchia, 2017 for a review of mediation analysis in criminological research). In mediation analysis, we first present the unconditional association between parental incarceration and some outcome for the children (such as delinquency). Then, most often in a regression framework, we condition on other features of those parents and their children, such as background characteristics (e.g., Wildeman, 2014), parenting style (Thornberry, Freeman-Gallant, Lizotte, Krohn, and Smith, 2003), or the children's abilities (e.g., Maguin and Loeber, 1996). We do so because we know that the unconditional association between parental incarceration and children's outcomes could well be mediated by these characteristics. Conditioning the association between parental incarceration and children's outcomes on mediating factors offers a deeper understanding of how much of the association that runs through the mediating factors, thus alluding to the social mechanisms which drive the effects of parental incarceration for children. And, indeed, this analytical strategy has shown, on the one hand, that the significance of intergenerational transmissions of delinquency and contact with the criminal justice system persists even when we take mediation into account (e.g., Murray and Farrington, 2005). On the other hand, it has also shown that many mediating factors express substantially important mechanisms through which the consequences of parental incarceration for children unfold (e.g., Hjalmarsson and Linquist, 2012).

Yet, the association between parental incarceration and children's outcomes could in fact be even stronger than what has been established in existing studies which condition on other covariates. The reason for this does not pertain to mediation analysis per se, and mediation analysis should and will continue to be a powerful statistical tool for analyzing the intergenerational transmission of delinquency. The conditional association between parental incarceration and children's outcomes could be stronger because of the statistical properties of the class of models which scholars of criminology typically apply to their mediation analyses. Many of children's outcomes which are of interest to criminologists are binary, measuring, for example, whether the children experience contact with the criminal justice system or not (e.g., Farrington, Coid, and Murray, 2009), or whether they experience social exclusion (Foster and Hagan, 2007) and child homelessness (Wildeman, 2014). Such binary outcomes are most often analyzed using nonlinear probability models, most prominently binary logistic regression models (logit models). But methodological research on nonlinear probability models advises against comparing estimates across such models, as is done in mediation analysis, because this strategy does not allow us to separate the impact of mediation from the impact of a model-driven bias (Karlson, Holm, and Breen, 2012). Put directly, differences in results from bivariate and conditional nonlinear probability models—which are the differences that we pay special attention to in mediation analyses—could well be biased, thus posing a serious threat to advancements to research on the consequences of parental incarceration.

In this paper, we introduce to scholars of criminology one of the most recent methodological advances in the mediation analysis of binary outcomes: the "KHB Method" (Karlson, Holm, and Breen, 2012). This method effectively reveals the impact of model-driven bias when we condition the association between parental incarceration and children's outcomes on other variables, and it provides estimates of the conditional association between parental incarceration and children's outcomes that are unaffected by model-driven bias. An important result in this methodological research is that the degree of mediation will be understated if the model-driven bias is ignored.

As a general example of the strength of the KHB Method for the analysis of the intergenerational transmission of delinquency, we focus on the association between paternal incarceration and sons' contact with the criminal justice system. This association has been shown to be strong but decreases (yet is still substantial) when potential confounders are taken into account (e.g., Farrington, Coid, and Murray, 2009). We use the son's abilities as a general example of a potentially important mediator of this association (see Hagan and Foster, 2012), and we show that analyzing the association between paternal incarceration, sons' abilities, and sons' involvement with the criminal justice system in standard logit models leads to an amount of model-driven bias that underestimates the significance of abilities for how these negative consequences unfold.

Empirically, we make a number of important contributions to existing research. First, we use data from the NLSY97 to show what we just outlined in general terms, namely that model-driven bias in logit models is substantial and could misguide conclusions regarding the association between paternal incarceration, son's abilities, and sons' risk of experiencing incarceration (which is our main indicator of contact with the criminal justice system—we replicate our main results using conviction as outcome, however). Second, we replicate results from the NLSY97 on data from a very different context, registry data from Denmark, which is an administrative panel of all residents in Denmark. Using this data source, the just mentioned model-driven bias is found to be even larger. Third, Danish registry data are of such quality and detail that we can show results by length of paternal incarceration, which is an important contribution to existing research that typically treats paternal incarceration as a binary event (for a discussion, see Andersen, 2016). Again, we find that the impact of model-driven bias is substantial (although fairly stable) across length of paternal incarceration. Fourth, we contribute to the emerging literature on the unequal consequences of parental incarceration for children by presenting and discussing the relationship between incarceration and abilities among sons, by paternal incarceration experience and by data source.

In all, we advance the point that model-driven bias could well have led scholars of criminology to underestimate the importance of the channels through which the negative consequences of parental incarceration for children unfold. By introducing one of the most recent advances in the methodology of nonlinear probability models, the KHB Method, to scholars of criminology, it is our hope to point the growing research literature on parental incarceration (and other related research literatures) in the direction of new research frontiers.

BACKGROUND

Scholars of criminology have made tremendous advancements in research on the negative consequences of parental incarceration for children (for reviews of the literature, see Foster and Hagan, 2015a; Murray and Farrington, 2008). The association between the delinquency of parents and children—especially of fathers and sons—has long been well established (e.g., Farrington, Barnes, and Lambert, 1996). Recent studies not only show how mass incarceration contributes to social (and racial) inequality (e.g., Foster and Hagan, 2015b; Wakefield and Wildeman, 2014), they also show how family dynamics tend to concentrate delinquency and contact with the criminal justice system in family networks (e.g., Farrington, Joliffe, Loeber, Stouthamer-Loeber, and Kalb, 2001). Parenting styles, for example, play a crucial role in transmitting such behavior across generations (Thornberry, Freeman-Gallant, Lizotte, Krohn, and Smith, 2003). Importantly, similar results are reported across different countries, such as the Netherlands (e.g., Bijleveld and Wijkman, 2009), the United Kingdom (e.g., Farrington, Coid, and Murray, 2009), Sweden (e.g., Frisell, Lichtenstein, and Långström, 2011), and Denmark (e.g., Andersen, 2016), and expand to other domains of children's lives too, such as their risk of experiencing foster care (Andersen and Wildeman, 2014), their school performance (Turney and Haskins, 2014), and to health outcomes (Miller and Barnes, 2015; Turney, 2014). Other recent contributions have shown that not all parental incarceration (maternal incarceration, in particular) is harmful to children (e.g., Cho 2009a, 2009b; Giordano, 2010; Wildeman and Turney, 2014), and that the consequences of parental incarceration are unequally distributed even among children who suffer this experience (Turney, 2017). Despite this new research on heterogenous effects of parental incarceration, the overall conclusion from research is still clear: There are substantially important intergenerational transmissions of delinquency and contact with the criminal justice system.

The consequences of parental incarceration for children are in fact so far-reaching and important, that a recent review describes how parental incarceration has become an engine for systemic exclusion, implying that parental incarceration leads to disconnection from such key societal domains as civic, educational, economic, and family ones (Foster and Hagan, 2015a). More than 2.7 million children had a parent incarcerated in the United States in 2010 and far more children have ever experienced paternal incarceration in the United States (Wakefield and Wildeman, 2014). And considering the racial and social distribution of parental incarceration (Wildeman, 2009), the implications of the institutionalization of parental incarceration for social inequality can hardly be overstated (Wakefield and Wildeman, 2014). Not focusing on parental incarceration but on parental conviction, Hjalmarsson and Linquist (2012) find that whereas the intergenerational transmission of criminal conviction is weaker than of educational achievement (high school completion), it is indeed stronger than the intergenerational component of poverty, alluding to just how important contact with the criminal justice system is for families and for social inequality.

Research has also progressed greatly in understanding the channels through which parental incarceration matters for children. We now know, for example, that the consequences of separation of parents and sons during periods of parental incarceration outgrow the consequences of other types of separation up to age 32 (Murray and Farrington, 2005). And we know that parenting styles and financial hardship mediate the transmission of delinquency from one generation to the next (Thornberry, Freeman-Gallant, Lizotte, Krohn, and Smith, 2003). Another example is the finding that the intergenerational transmission of occupational status and poor educational achievement matters for the intergenerational transmission of conviction (Ramakers, Bijleveld, and Ruiter, 2011).

In this paper, we argue that because of the statistical properties of the models which criminologists tend to apply when we analyze the consequences of parental incarceration, and the channels through which these consequences unfold, we may in fact underestimate both the true extent of the consequences of parental incarceration for children and the significance of the potential channels through which these consequences unfold. We argue that the combination of the analytical approach that is most common to the analysis of the consequences of parental incarceration—mediation analysis—with the fact that many of the outcomes that interest criminologists are binary may have opened the backdoor for substantial bias in estimates. This methodological issue has not yet been addressed in criminology and the solution which we focus on in this paper has only recently been established in the sociological methodology literature (Karlson, Holm, and Breen, 2012). With this paper, we introduce the issue and its solution to the field of criminology, and we empirically show the margin of bias which permeates mediation analyses on binary outcomes using an example from the criminology literature (the association between paternal incarceration, sons' abilities, and sons' incarceration) and two excellent data sources, one from the United States (the NLSY97) and one from Denmark (Danish registry data).

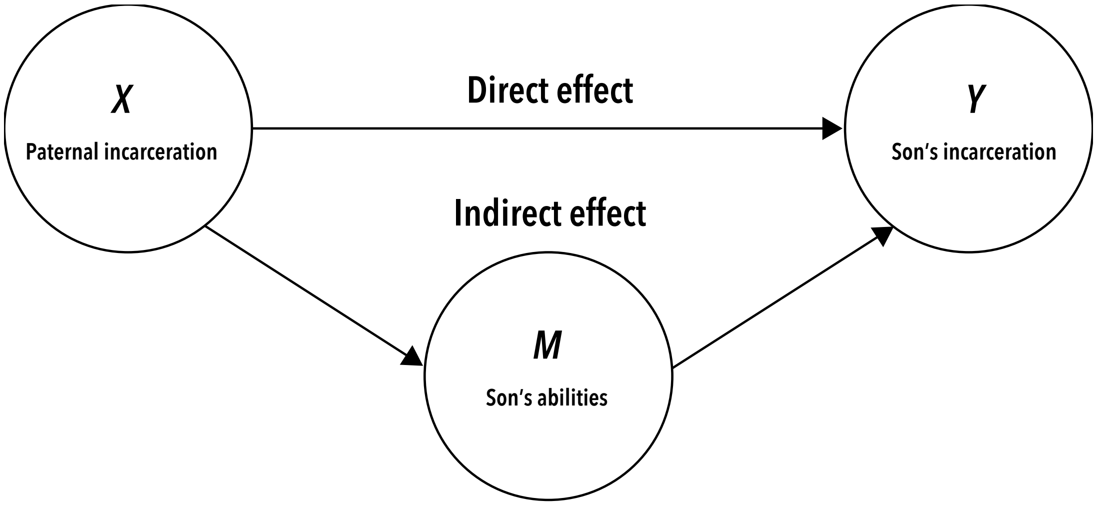
MEDIATION ANALYSIS

Mediation analysis estimates whether one or more variables (M) mediate the association between an independent (X) and dependent variable (Y) (for a review of mediation analysis in criminological research, see Walters and Mandracchia, 2017). Figure 1 depicts a simplified version of the causal pathways which call for mediation analysis. We may think of Figure 1 as decomposing the total association between paternal (X) and children's (Y) incarceration into a component explained or mediated by the child's abilities (M) and a component unexplained by these abilities. The explained component is often referred to as the indirect effect, i.e., the effect of X on Y running through M, whereas the unexplained component is termed the direct effect, i.e., the effect of X on Y that does not run through M. Finally, the two components (explained and unexplained) sum to the total effect of X on Y.

Studying the total association between paternal and children's incarceration has value because it gauges the consequences of paternal incarceration for children, i.e., it is a measure of the rate of intergenerational transmission of criminal behavior. However, often we are interested in not only empirically establishing these consequences, but also understanding why we observe this association. What are the underlying mechanisms driving the intergenerational association? To learn about these mechanisms, the standard practice in mediation analysis is to condition the total association on a third variable, the mediator.[1] In Figure 1, this third variable is the son's abilities (M). Thus, the son's academic abilities (or lack thereof) are taken to be an important mechanism through which the impact of paternal incarceration affects the likelihood of the child being incarcerated at a later point.

1 As we later return to, the conditioning is usually done by way of regression.

Fig. 1 Causal Pathways in Mediation Analysis



NOTES: Paternal incarceration, Son's abilities, and Son's incarceration refer to our empirical

By conditioning on the mediator, we can determine how much of the total or unconditional association that is channeled through the mediating factor, and hence learn whether the mediating factor is substantially important. Settling the importance of mediators is pivotal both to our empirical understanding of the association, to our efforts in developing theories, and to the likely impact of different policies targeted at the association in question.

Mediation analyses are common in criminological research, and over the years mediation analyses have pushed the research frontier in many important ways. Murray and Farrington (2005), for example, analyze the association between separation through parental incarceration and children's antisocial and delinquent behavior and add to their model other well-known risk factors for such behavior. They find that over and above those other risk factors, separation through parental incarceration has a larger impact on children's behavior than does other types of separation. Farrington, Coid, and Murray (2009) measure the intergenerational transmission of conviction among males and take into account the mediating role of a range of control variables. Their results prove a strong intergenerational transmission of conviction which decreases when controls are added to model, as one would expect.

Foster and Hagan (2007) measure the association between background characteristics and variables indicating social exclusion (homelessness, non-insurance, political disengagement), and then adds father's education and incarceration to the model. Results show that paternal incarceration and educational attainment matters greatly for social exclusion, a finding that was validated in Foster and Hagan (2009) and which also relies on mediation analyses. Hagan and Foster (2015b) further show that young adult children who have had a father or mother imprisoned are at increased risk of experiencing socioeconomic deprivation, including inadequate access to food. Importantly, results from their mediation analysis show that the effects of paternal and maternal incarceration decrease with educational attainment.

Examples which all focus on Sweden include Frisell, Lichtenstein, and Långström (2011) who measure strong familial risks of violent crime conviction, and their results show that these risks are modified by background. Hjalmarsson and Linquist (2012) analyze the father—child correlation in convictions and their results show that the intergenerational transmission in convictions is weaker than the intergenerational transmission of education but stronger than for poverty. In their mediation analysis, parental human capital and behavior account for 60-80 percent of uncontrolled association between fathers' convictions and sons' convictions. And comparing results from Sweden to ones from the United Kingdom, Murray, Jansson, and Farrington (2007) control the association between the incarceration of fathers and sons for the mediating role of parental conviction. They find that parental incarceration is a stronger risk factor for children's incarceration in the United Kingdom than in Sweden. But they also find that the link disappears in Sweden, and not in the United Kingdom, when the mediating role of conviction is taken into account (which the authors explain by structural differences between the UK and Sweden, such as differences in sentence lengths; sentences are much shorter in Sweden).

Other examples include Wildeman (2014) who analyzes the association between parental incarceration and child homelessness and takes the mediating role of parental drug and alcohol problems, family finances (e.g., parental inability to pay bills), loss of institutional and informal supports (e.g., losing public housing), and maternal capacities and capabilities (e.g., maternal depression) into account. Results show a) that recent paternal but not maternal incarceration increases homelessness, b) that results are driven by African Americans, c) that the effect of parental incarceration increases in economic hardship and decreases in institutional support. Also, Miller and Barnes (2015) analyze the impact of paternal incarceration on health-related outcomes, educational outcomes, and economic outcomes. Their results suggest that parental incarceration is significantly related to these outcomes in early adulthood even when controlled for the important mediating factor of self-control (and other factors too).

The last example which we will describe here uses data from the one of the contexts that we also use in this paper. Andersen (2016) uses Danish register data to measure the association between frequency and duration of paternal incarceration and children's educational outcomes and criminality in young adulthood. Results show that the association is important, but also that it decreases once factors that are known to mediate the association (such as the child's gender, the family type, and a host of characteristics of the mother and the father) are taken into account.

In our empirical example in this paper, we estimate whether the association between paternal incarceration (X) and sons' risk of experiencing incarceration (Y) is mediated by sons' abilities (M). This question is an important empirical one. The unconditional association between the incarceration of fathers and sons is of course interesting in itself, as it is a measure of the degree to which criminal behaviors are passed on from generation to generation. But if we wish to understand why we observe this unconditional association, it is important that we obtain a correct estimate of the mediating role of sons' abilities. Various studies (e.g., Murray and Farrington, 2005), including one especially rigorous causal study which appeared in a recent issue of this journal (Wildeman and Andersen, 2017), have shown that paternal incarceration matters for sons' risk of experiencing contact with the criminal justice system (X → Y). Other studies (e.g., Foster and Hagan, 2009) have shown that paternal incarceration matters for children's abilities and level of schooling (X → M). And yet other studies (e.g., Maguin and Loeber, 1996) have shown that abilities and school performance matter for the risk of engaging in criminal activities and come into contact with the criminal justice system (M → Y). Thus, our empirical example is ideal for the purposes of this paper because the effect of paternal incarceration on sons' risk of incarceration could well be mediated by the indirect effect of paternal incarceration on sons' abilities, which then has a separate effect on their incarceration risk.

A METHODOLOGICAL ISSUE IN MEDIATION ANALYSES OF BINARY OUTCOMES: RESCALING

The strengths and applicability of mediation analysis are thus many, and research has benefitted greatly from using this analytical tool. In this paper, however, we direct attention to one very important weakness of mediation analysis when analyzing binary outcomes: Rescaling.

Criminological research often studies binary outcomes, measuring whether some outcome was realized or not, such as whether a young adult who experienced paternal incarceration as child is convicted of crimes (e.g., Andersen, 2016), is incarcerated (e.g., Murray, Jansson, and Farrington, 2007), or experiences homelessness (e.g., Wildeman, 2014). Our preferred statistical model for analyzing binary outcomes is nonlinear probability models, since these models have the favorable attribute that they do not estimate probabilities outside the 0—1 interval. Foster and Hagan (2007), Murray, Janson, and Farrington (2007), Farrington, Coid, and Murray (2009), Ramakers, Bijleveld, and Ruiter (2011), Hjalmarsson and Linquist (2012), Wildeman (2014), Hagan and Foster (2015b), and Miller and Barnes (2015) all present results which implicitly or explicitly rely on mediation analyses and nonlinear probability models, for example.

But in doing so, we argue, scholars run the risk of not fully appreciating the degree of mediation. The reason for this is that, in nonlinear probability models, the regression coefficients cannot be estimated separately from the unexplained or error variance in the model. Thus, in contrast to linear models, in nonlinear probability models, the error variance cannot be estimated from the data, and one therefore most often assumes that the error term follows a given parametric distribution. (Karlson, Holm, and Breen, 2012). In logit models, researchers assume that the error term follows a logistic distribution. In probit regression models, researchers assume that the error term follows a standard normal distribution.

The methodological issue which we raise in this paper, and which has been raised in the sociological methodology literature in Karlson, Holm, and Breen (2012), stems from the fact the magnitude of an estimated coefficient in a nonlinear probability model depends on the unexplained or error variance in the model. The regression coefficients are said to be only identified up to a scale, and this scale is a direct function of the unexplained portion of the model. In this context, rescaling occurs whenever researchers successively add variables (i.e., mediators) to a nonlinear probability model to learn about mediation. The rescaling occurs because the added variables, per definition, explain some of the unexplained variance of the model, and the reduction in error variance consequently feeds into the estimates of the regression coefficients (by changing the scale of the coefficients).

Conventionally, mediation analysis has been performed on nonlinear probability models as on linear models. In this context, mediation analysis has been considered to yield unbiased estimates of the direct and indirect effects. But in fact, comparing coefficients from models with and without the mediators involves comparing the simultaneous impact of mediation and rescaling on the association between the dependent and independent variable, effectively leading to biased estimates of the degree of mediation (as captured by the indirect effect). In fact, the rescaling operates in such a way that the degree of mediation often will be understated in the conventional use of mediation analysis comparing estimates between models. In our Method section, we introduce a recent advance in the methodology of nonlinear models, the "KHB method" (Karlson, Holm, and Breen, 2012), which effectively produces mediation analysis estimates that are not biased by rescaling issues. Before doing so we, however, explicate the reasons why criminologists are in dire need of such a method.

POTENTIAL CONSEQUENCES OF RESCALING

Why is the distinction between mediation and rescaling important for empirical studies in the research field of criminology? Put directly, because inattention to this distinction could well lead to wrongful conclusions about the empirical associations which we rely on for deriving theories and making policies. If, for example, the association between paternal incarceration and sons' delinquency appears not to change once controlled for the mediating role of abilities, we may be inclined to conclude that the reason why we observe higher delinquency rates among boys who experience paternal incarceration is caused by paternal incarceration affecting all children in the same way. But, because of the rescaling properties which we have just outlined, this conclusion could be driven entirely by rescaling (causing downward bias), something pertaining to our choice of statistical model, not to any empirical mechanisms.

The same is true regarding upward bias in estimates stemming from rescaling issues. Assume, for example, that all studies of the consequences of paternal incarceration for sons' delinquency found statistically significant mediating effects of abilities. How much value would these findings have for theory development and policy issues if they were all driven exclusively by rescaling because of the researchers' implicit error term variance assumption and not because of any substantial mechanisms? At best, such a situation would imply a waste of time and money; at worst it would undermine the authority of empirical research.

METHOD

Karlson, Holm, and Breen (2012) offer a solution to the problem of rescaling in nonlinear probability models, the "KHB method". To explain this method, we first briefly introduce a way of motivating or deriving a nonlinear probability model such as a the logit or probit model, known as the latent variable approach.[2]

DERIVING THE LOGIT AND PROBIT MODEL

Assume that the binary outcome of interest, Y, in fact is a realization of an underlying continuous variable that measures the propensity to engage in the behavior that Y indicates. To fix ideas, assume that Y measures whether or not a child is incarcerated. The underlying continuous variable, which we denote Y*, would thus reflect the propensity of the child to engage in criminal behaviors that lead to incarceration. The point here is that Y* is not observed to us as researchers (i.e., it is unobserved). All we observe is the indicator for whether the child was incarcerated. To develop the logit model using this framework, we assume that there exists a certain link between the underlying propensity, Y*, and the observed indicator, Y. We assume that Y is a binary realization of Y* according to the following threshold rule

$$\begin{aligned} &Y = 1 \text{ if } Y^* > \tau \\ &Y = 0 \text{ if otherwise.} \end{aligned} \qquad (1)$$

Thus, we observe a child as incarcerated if he or she passes a threshold, τ, on the underlying propensity to engage in incarceration-prone criminal activities.

Under the assumption of the existence of an underlying continuous propensity, we may now take the next step in the derivation of the logit or probit model. We now assume that the underlying propensity, Y*, is linearly related to a predictor variable, X—e.g., whether or not the parent was incarcerated—by way of a regression model:

$$Y^* = \alpha + \beta X + u \qquad (2)$$

This model simply states that Y* is a linear function of X, and the model's parameters should be interpreted as those of any linear regression model. Ideally, we would like to estimate the model in (2), but as Y* is unobserved to us, we cannot estimate the model. However, because of the link between the observed indicator, Y, and the underlying propensity, Y*, we can obtain something that can be estimated. To see how, we make the further assumption that the error term in (2), u, is a logistically distributed variable for the logit or a normally distributed variable for the probit,as was already mentioned. This assumption is a fundamental assumption in the sense that (a) it cannot be tested and (b) is required for deriving the logit or probit model.

2 The models can also be derived by other approaches, namely the transformational or truly discrete approach. However, both approaches lead to the exact same issues, but the latent variable approach is particularly useful for understanding the biasing impact of rescaling.

Under the distributional assumption that we place on the error term, u, we can rewrite the error term in a way that suits the purpose of deriving the nonlinear probability model. We assume that the error term is a multiple of a standard distribution; that is, we can write the error term as

$$u = \omega s \qquad (3)$$

where ω is the standard logistic distribution in the logit case, and the standard normal distribution in probit case. Equation (3) simply states that the error term of the model in Equation (2) can be written as the product of a standard distribution and a scale factor, s. The scale factor is a proportionality factor, a number, which allows the variance of u to differ from ω . It basically scales the standard distribution to fit the error distribution.

The assumptions in Equations (1)-(3) are sufficient to derive the nonlinear probability model. We begin by observing that the probability of the child being incarcerated, $\Pr(Y = 1)$ is the chief outcome of interest. However, according to the threshold rule in Equation (1), we may write this probability as $\Pr(Y^{*} > \tau)$. Furthermore, according to the model in Equation (2), we may write the probability as $\Pr(\alpha + \beta X + u > \tau)$. Finally, from the definition in Equation (3), we can express the probability as $\Pr(\alpha + \beta X + \omega s > \tau)$, which then can be restated as

$$\Pr\left(\frac{\alpha - \tau + \beta X}{s} < \omega\right) \qquad (4)$$

This expression of the probability of the child being incarcerated can be converted into the logit or probit model, *depending on the assumption placed on the variable following a standard distribution*, ω . If ω is assumed to follow a standard logistic (normal) distribution, then we can obtain the logit (probit). The reason for this is that the standard distributions are known parametric distributions in the sense that the mean and variance are known. For the logit, the standard distribution has zero mean and 1.81 standard deviation, whereas for the probit, the standard distribution has zero mean and unit standard deviation. Under these assumptions and using some further algebra, it is possible to write the logit and probit model as

Logit: $$\operatorname{logit}\left(\Pr(Y=1)\right) = \frac{\alpha - \tau}{s} + \frac{\beta}{s} X \qquad (5a)$$

Probit: $$\Pr(Y=1) = \Phi\left(\frac{\alpha - \tau}{s} + \frac{\beta}{s} X\right), \qquad (5b)$$

where Φ is the standard normal cumulative distribution function. The major conclusion that arises from this exercise is that the coefficients estimated from logit or probit models reflect both the underlying effect, as measured by β, and the scale parameter, s, which is a direct function of the error variance of the model (in that s captures the deviation of the underlying error from that of the standard distribution). If we assume that b is the estimated logit or probit coefficient of X on Y, i.e., the coefficient that would be estimated by a statistical program, then

$$b = \frac{\beta}{s}. \qquad (6)$$

Put in plain terms, the coefficients from nonlinear probability model confound the estimated effects with residual model heterogeneity. The two cannot be separated. This is also why methodologists say that coefficients from these models are identified up to a scale. To fix ideas about this property of nonlinear probability models, we may think about β as reflecting the mean (specifically, β_{YX} times the corresponding variable returns the mean of that variable), and s as reflecting the variance. Mean and variance are not separately identifiable in binary variables, and using the latent variable approach is thus one way of illustrating the impact of this property on the nonlinear probability models.[3]

MEDIATION ANALYSIS IN LOGIT AND PROBIT MODELS

Because coefficients are identified up to scale in logit and probit models, the utility of these coefficients in applied research is greatly limited. This is particularly true in mediation analysis in which researchers compare coefficients of the same predictor variable across models successively including additional variables (i.e., mediators). The issue arises via the scale parameter, s. As Equation (3) shows, the scale parameter is a proportionality factor that ensures that the model's error variance is a multiple of the variance of assumed standard distribution. However, whenever researchers add more variables to a model, the model's error variance is reduced, and so, the scale parameter also reduces in order for the error to be a multiple of the standard distribution. Karlson, Holm, and Breen (2012) refer to this issue as rescaling. Rescaling leads to a bias in mediation analysis because changes in the coefficients of X between models with different covariates will reflect not only conventional confounding (caused by the underlying correlations between X and M), but also differences in scale parameters (insofar as the added covariates, M, reduce the error variance of the model).

To demonstrate the rescaling bias in more detail, we adopt the notation of Karlson, Holm, and Breen (2012), and assume that we have two linear regression models for the underlying propensity to be incarcerated, Y* [similar to Equation (2)]:

$$H_R : Y^* = \beta_{YX} X + e, \quad \text{where } e = s_R \omega \tag{7}$$

$$H_F : Y^* = \beta_{YX.M} X + \beta_{YM.X} M + v, \quad \text{where } v = s_F \omega \tag{8}$$

The reduced model, HR, refers to the uncontrolled model that gauges the total association between X and Y, $\beta_{YX.M}$. The full model, HF, refers to the model including the mediator, and it gauges the association between X and Y* conditional on M, $\beta_{YX.M}$. Under the assumptions stated in Equations (1)-(3), we may now derive the logit or probit models corresponding to the models in Equations (7)-(8) using the exact same derivations as in Equations (4)-(5). Before we do this, however, we want to note that the error terms in Equations (7) and (8) differ in one important way. The error term in the reduced model in Equation (7) includes per definition the mediator included in Equation (8) and, consequently, the error variance in the reduced model in Equation (7) will be larger than the error variance in the full model in Equation (8). This difference has important implications for the scale parameters in the two models. Because the scale parameter allows the error variance to differ from the standard logistic or normal distribution, the scale parameter in the reduced model, s_R, will be larger than the scale parameter in the full model, s_F.[4]

3 We reiterate that the scale identifiability property of these coefficients is not a consequence of us deriving the logit or probit from a latent variable approach. Derivation approaches that do not make this assumption lead to the same results (see Neuhaus et al. 1991). More generally, then, one can say that coefficients from nonlinear probability models are confounded by unmeasured heterogeneity or omitted covariates that are orthogonal to the predictor of interest.

4 This is of course only strictly true if the mediator, M, in the full model has an effect on Y* independent of X, but this would be the situation in many real-world applications.

To demonstrate the impact of the changing scale parameters on mediation analysis, we now state the corresponding nonlinear probability models of Equations (7) and (8) under the assumptions in Equations (1)-(3). For simplicity, we only do so for the logit model, but the same issue applies to the probit model. The corresponding logit models are

$$H_R^{Logit}: \text{logit}\left(\Pr(Y=1)\right) = \frac{\beta_{YX}}{s_R} X \tag{9}$$

$$H_F^{Logit}: \text{logit}\left(\Pr(Y=1)\right) = \frac{\beta_{YX.M}}{s_F} X + \frac{\beta_{YM.X}}{s_F} M \tag{10}$$

In conventional mediation analysis, researchers would compare the effect of X between Models (9) and (10) to gauge the magnitude of mediation, i.e., the magnitude of the mechanism, encoded in M, that links X to Y. However, following the latent variable derivation of the logit model stated in Equations (9) and (10), the difference between these two coefficients can be stated as

$$\delta^{CONV} = \frac{\beta_{YX}}{s_R} - \frac{\beta_{YX.M}}{s_F}, \tag{11}$$

where the superscript CONV refers to the conventional estimator of mediation. The difference in Equation (11) is conventionally taken as an estimate of the indirect effect, i.e., the effect of X on Y running through M. However, as Equation (11) shows, the coefficient difference confounds between-model differences in coefficients (mediation) with between-model differences in scale parameters (rescaling). Notice that, following Karlson, Holm, and Breen (2012), we refer to between-model differences in the underlying beta-coefficients as mediation in the sense that this is the change in the coefficients caused by the underlying correlation between X and M, whereas the rescaling operates independently of this correlation (but depends on the correlation between Y* and M conditional on X).

To further understand the biasing impact of the rescaling on the mediating impact of M, we reiterate that $s_R > s_F$, i.e., the scale parameter of the reduced model is always larger than that of the full model. This means that the coefficient of X in the full model in (10) will be divided with a smaller number than the coefficient of X in the reduced model in (9). Dividing with a smaller number means that the coefficient in the full model will be artificially larger, simply as a result of rescaling. This property has important consequences for applied research. In many applications, researchers expect the effect of X on Y to decline once they control for the mediator, M.[5] However, rescaling *offsets* this decline, meaning that the mediating impact of M on the XY-association generally will be *understated* in nonlinear probability models. The magnitude of this bias depends on how different the scale parameters of the reduced and full model are, a difference which is a direct function of the effect of M on Y* conditional on X.[6]

The offsetting effect of rescaling can have troubling effects on conventional practices of mediation analysis in nonlinear probability models. Karlson, Holm, and Breen (2012) give the example in which rescaling perfectly offsets the mediating impact of M. This would be the case if the decline in the underlying coefficients, β_{YX} - $\beta_{YX.M}$, would be perfectly offset by rescaling (i.e. s_R - s_F). Thus, in this example, adding M to the model would lead to no change in the coefficient of X. Researchers could then, wrongfully, conclude that M is not a mediator of the association between X and Y, when, in fact, the association is mediated by M—a challenge which we already discussed.

5 Conditional associations are generally unpredictable, but mainstream practice in the social sciences expects the effect to decline in magnitude when mediators are included.

6 To be more precise, it depends both on the effect of M on Y* conditional on X and the conditional variance in M conditional on X, i.e., it depends on the linear predictor of this part of the model.

SEPARATING MEDIATION FROM RESCALING

The impact of rescaling in nonlinear probability models has been largely ignored in applied mediation analysis, where the conventional approach of comparing estimates of the same predictor between models successively including covariates has proliferated (e.g., Foster and Hagan (2007), Murray, Janson, and Farrington (2007), Farrington, Coid, and Murray (2009), Ramakers, Bijleveld, and Ruiter (2011), Hjalmarsson and Linquist (2012), Wildeman (2014), Hagan and Foster (2015b), Miller and Barnes (2015). Karlson, Holm, and Breen (2012) suggest a general solution to this issue.[7] Their solution effectively separates mediation (in the underlying model) from rescaling, thus allowing researchers to gauge the mediating impact of M net of rescaling. To explain their method, we begin by defining an auxiliary regression model that specifies the effect of the predictor, X, on M:

$$M = \eta + \beta_{MX} X + \tilde{M} \qquad (12)$$

In this model, we are chiefly interested in the error term, $\tilde{M}$ (M-tilde). This error term can be thought of as a residualized mediator, residualized for the predictor X. Now, if we replace M with $\tilde{M}$ in the underlying linear model, we obtain

$$H_F^*: Y^* = \beta_{YX.\tilde{M}} X + \beta_{Y\tilde{M}.X} M + v^*, \quad \text{where } v^* = s_F^* \omega \qquad (13)$$

This model is simply just a model corresponding to the full model, but including the residualized mediator instead of the regular mediator. The model in (13) is key to the method developed by Karlson, Holm, and Breen (2012). They prove that the two following equalities hold:

$$\beta_{YX.\tilde{M}} = \beta_{YX} \qquad (14)$$

$$v^* = v \Leftrightarrow s_F^* = s_F \qquad (15)$$

Equation (14) states that the coefficient of X in the underlying linear model in the respecified full model in Equation (13) is the same as the coefficient in the reduced underlying linear model in Equation (7). This follows from the forced orthogonality between X and the residualized mediator, $\tilde{M}$, in Equation (13). Equation (15) further shows that the error term of the respecified full model in Equation (13) is equal to the error term in the full model in Equation (8). Indeed, the two models are equivalent models, just expressed in different ways. However, as Equation (15) also shows, because the error terms are the same, so are the scale parameters of the two models.

The equalities in Equations (14) and (15) allow researchers to separate mediation from rescaling. To see this, we first need to write the logit model corresponding to the underlying linear model in Equation (13):

$$H_F^{Logit*}: \text{logit}\left(\Pr(Y=1)\right) = \frac{\beta_{YX.\tilde{M}}}{s_F^*} X + \frac{\beta_{Y\tilde{M}.X}}{s_F^*} \tilde{M} = \frac{\beta_{YX}}{s_F} X + \frac{\beta_{Y\tilde{M}.X}}{s_F} \tilde{M} \qquad (16)$$

Notice that the second equality holds in this equation, because of the equalities in Equations (14) and (15). Thus, Equation (16) demonstrates the core of the method by Karlson, Holm, and Breen (2012). The coefficient of X in this model can be written as the coefficient of X in the reduced model, β_{YX}, but divid-

7 Other solutions also exist, but extensive simulation studies suggest that the method by Karlson, Holm, and Breen (2012) is always as good as or better than existing alternatives. See also Breen, Karlson, and Holm (2013).

ed with the scale parameter of the full model, s_F . Thus, it is possible to write the mediating impact of M net of rescaling a

$$\delta = \frac{\beta_{YX}}{s_F} - \frac{\beta_{YX.M}}{s_F} = \frac{\beta_{YX} - \beta_{YX.M}}{s_F} \qquad (17)$$

The result in Equation (17) shows that it is possible to identify mediation identified up to the full scale of the model, when stating the indirect effect in terms of a difference. However, the proportional reduction in coefficient of X that is caused by including M—known as the percent mediated—is a scale free measure altogether:

$$\delta^{PCT} = \frac{\frac{\beta_{YX}}{s_F} - \frac{\beta_{YX.M}}{s_F}}{\frac{\beta_{YX}}{s_F}} \cdot 100\% = \frac{\beta_{YX} - \beta_{YX.M}}{\beta_{YX}} \qquad (18)$$

The percent mediated is often taken as a measure of the magnitude of mediation and, following the method of Karlson, Holm, and Breen (2012), this measure can be recovered from nonlinear probability models without being distorted by rescaling.

Thus the method by Karlson, Holm, and Breen (2012)—which is based on residualizing predictors—allows researchers to study mediation net of rescaling in nonlinear probability models. In fact, their method appears to be generally consistent with the "rules of path analysis" known from linear models, and thus offers a unified framework for studying mediation in nonlinear probability models (see Breen, Karlson, and Holm 2013). Furthermore, there exists a user-written routine in STATA®, khb, which implements their method (Kohler, Karlson, and Holm 2011). This program also reports correct standard errors, using the formulas for the indirect effects provided in Karlson, Holm, and Breen (2012).

DATA AND ANALYTIC PLAN

DATA

We use two data sources in our empirical example, the NLSY97 and Danish registry data. We do so to display that the challenge of rescaling when applying nonlinear probability models does not depend on data source or context. Also, Danish registry data offer a high level of precision and detail—and a sufficiently large number of observations (because they are population data) to break down results by length of incarceration.

THE NLSY97

The National Longitudinal Survey of Youth 1997 (NLSY97) is a survey administered by the US Bureau of Labor Statistics designed to represent people living in the United States in 1997 who were born during 1980–1984. The NLSY97 covers a wide array of respondents' lives, primarily focusing on labor market behavior and educational experiences but also covering many other aspects of these youths' lives, such as their family background and measures of delinquency and criminal offenses. Data were collected using a computer-assisted personal interviewing system, which automatically checks the respondent's answers for obvious inconsistencies and the like, which is likely to improve the quality of the survey data. And importantly, when sensitive questions came up in the survey, such as questions related to criminal behavior, the interviewing system enabled respondents to enter their answers directly into the computer without sharing these with the interviewer, again most likely improving the quality of answers to such sensitive questions. For the full documentation on the NLSY97, see http://www.nlsinfo.org.

From the NLSY97 we use all the boys who were born between 1980 and 1984 and who were residents of the United States in 1997 (round 1). As outcome variable we add a binary marker of whether the boy had experienced any incarceration after turning 15 years of age and before the survey round in which the boy was 23 years old. This variable was constructed from questions on whether the boy had been incarcerated for each of the prior 12 months at each annual survey round. In a supplementary analysis we use a binary marker of whether the boy was convicted of a crime between age 15 and 23, measured from annual questions of whether the boy was convicted within the past 12 months.

As a measure of abilities, our mediating variable, we add each boy's math and verbal test score from the Armed Services Vocational Aptitude Battery (ASVAB). The full ASVAB covers abilities in 12 different domains, but the math and verbal test score which we use only compiles information on mathematical knowledge, arithmetic reasoning, word knowledge, and paragraph comprehension. In practice, NLS Program staff produced this variable to correspond to the Armed Forces Qualification Test, which is used by the US Department of Defense to screen potential enlisters. It measures each respondent's test score percentile within three month age groups in the NLSY97. The variable thus takes values between zero and 99, and a higher score implies higher abilities relative to the respondent's age group.

Paternal incarceration measures whether a respondent has ever experienced paternal incarceration before his 15th birthday. To obtain this variable, we rely on a self-reported measure of whether the respondent's father was ever incarcerated during the survey round corresponding to the boy's 15th birthday.

DANISH REGISTRY DATA

Danish registry data are administrative population data from Denmark. Each resident in Denmark has a unique personal number which identifies him or her in the many registries. Information is recorded by various agencies, such as the Prison and Probation Service and the Ministry of Education, and following formal approval, researchers may obtain access to the information through Statistics Denmark.[8] Researchers may then merge individual level information across the many registries and across years by using de-identified versions of the personal numbers. Effectively, Danish registry data are thus an individual level panel of administrative records, describing various aspects of people's lives.

From the Danish registry data, the population register, we select a dataset of all boys born in 1991. As outcome variable we add a binary marker of whether each boy experienced incarceration for more than 24 hours between his 15th birthday and age 23 (in 2015, the latest available year) from the criminal justice registers available from the Prison and Probation Service. In a supplementary analysis, we use a binary marker of criminal conviction within the same age bracket, including only convictions of violating the penal code (thus excluding minor offenses) and in another supplementary analysis we include even shorter stints of incarceration (i.e., arrest only) in our outcome measure.

8 For information on how to gain access to Danish registry data, see http://dst.dk/en/TilSalg/Forskningsservice

As a measure of abilities, our mediating variable, we add each boy's exam grades (GPA) in ninth grade, measured around age 16 (in 2007).[9] The exam grades are available from the Ministry of Education. GPA measures each boy's mean grade across six subjects measured on a 15-point scale (ranging from -3 to 12).[10] Contrary to the ability measure which we rely on from the NLSY97, our GPA measure from the Danish registry data is not relative (showing, for example, each person's percentile score within an age group as the measure in the NLSY97 does). Instead, our GPA measure from the Danish registry data is absolute in the sense that it simply records each boy's mean grade. Means below 2.0 are considered failed. In supplementary analyses we redefine the measure of abilities to include only GPA in math.

To obtain our measure of paternal incarceration we find their biological fathers' personal identification number in the population register. We then add each father's incarceration history from the criminal justice registries. The criminal justice registries include precise dates of admission and release of all incarceration spells in Denmark, allowing us to include only paternal incarceration spells which occurred before the son's 15th birthday and which were either in effect at the son's birth or began after it. We exclude incarcerations shorter than seven days, because seven days is the shortest prison sentence one could get in Denmark. We also include information on length of paternal incarceration, distinguishing between paternal incarceration spells of or exceeding seven days, two weeks, one month, three months, and one year.

CRIMINAL JUSTICE CONTACT IN THE UNITED STATES AND DENMARK

To criminologists, and probably to social scientists in general, both the United States and Denmark may be considered extreme observations when we look at the distribution of developed democracies. The United States has the highest incarceration rate; Denmark has one of the lowest. In fact, the difference in the incarceration rates of these two countries is more than ten-fold: 698 per 100,000 in the United States and 61 per 100,000 in Denmark, according to the latest available comparison (Walmsley, 2016). One reason for this difference comes from differences in sentence lengths. In Denmark, sentences are generally short (as mentioned, the shortest prison sentence in Denmark is seven days) and around two out of three sentences are three months or shorter and only one in four sentences exceed six months (Danish Prison and Probation Service, 2016). In the United States, the mean sentence length in 2008–2009 was around 4.7 years for federal prisoners and 2.1 years for state prisoners (Guerino, Harrison, and Sabol, 2011; Motivans, 2012).

Naturally, these vastly different incarceration rates matter for children's risk of experiencing the incarceration of a parent. Wildeman and Andersen (2015) estimate that if the risk of paternal incarceration among Danish children were to approach the risk that children in the United States experience paternal imprisonment (i.e., in the federal system typically following sentences of more than a year), all stints of incarceration in Denmark (even those that do not expand beyond 24 hours) would have to be included in the calculation. Children in the United States are thus much more likely to experience the incarceration of a parent and they are more likely to do so for longer periods than are Danish children.

Other important differences exist between the United States and Denmark. Denmark is an egalitarian country that has strong welfare institutions, for example. Social inequality in the United States outgrows

9 Because we measure incarceration from age 15 to age 23 our measure of abilities in the Danish data could for some be measured after incarceration has already happened. This would be problematic to the extent that incarceration at such a young age affects test scores (which seems likely). The incarceration rate at this young age is, however, very low and is not likely to matter much for our overall results.

10 The subjects are: written Danish, verbal Danish, Danish grammar, mathematics, science (physics and chemistry), and verbal English.

that of Denmark by more than 50 percent, as measured from the two countries' Gini coefficients (household disposable income adjusted for household size, United States: .394; Denmark: .254, OECD, 2016), expressing the universal availability of social security in Denmark. The availability of social security also means that newly released prisoners are entitled to social assistance, which could help them in achieving resocialization and could alleviate some of the pressure which newly released family members place on other family members in the United States (e.g., Comfort, 2016). Also relevant to questions of resocialization, efforts aimed at securing resocialization begin already during incarceration in Denmark, where a broad range of labor market training and education is offered. Human capital loss associated with incarceration might thus not be as important in Denmark because one has the chance to achieve formal training during incarceration. Again, stronger resocialization might alleviate some of the negative consequences of paternal incarceration for children.

In many ways, Denmark and the United States thus represent both extremes of developed democracies. But research from both of these contexts documents damaging effects of paternal incarceration on children and families (e.g., Andersen, 2016; Wildeman, 2014). In this paper, we add to this research by analyzing the amount of model-driven bias which permeates typical models of analyzing the role of mediating variables for the relationship between paternal incarceration and sons' outcomes—using data from both of these extreme contexts. We thereby document that the challenge of model-driven bias cuts across even the most extreme of developed democracies.

ANALYTIC PLAN

Our analytic plan follows four steps. First, we describe the two datasets which we use in our empirical example. We compare key attributes of each dataset and provide descriptive statistics on the variables that we use in the following analytic steps.

Second, in our main results we use data from the NLSY97 and the Danish registry data to present estimates of the association between paternal incarceration and sons' risk of incarceration while controlling for sons' abilities in the United States and Denmark. We compare estimates from logit models (estimates that may be biased by rescaling) to estimates derived from the KHB method (estimates that are not biased by rescaling).

Third, we exploit the sample size and availability of information in the Danish registry data to present results similar to those in our second analytic step but split up by length of paternal incarceration. Data from the NLSY97 do not allow us to present similar results from the United States simply because the sample is too small and the available information regarding length of paternal incarceration is insufficient to do so.

Fourth, we contribute to the emerging literature on the unequal consequences of parental incarceration for children by presenting and discussing the relationship between incarceration and abilities among sons, by paternal incarceration experience and by country/data source.

In supplementary analyses, we replicate the main results using son's risk of criminal conviction as outcome variable. Also, we exploit the richness of the Danish registry data to replicate the main findings including any incarceration (including incarcerations shorter than 24 hours) in our outcome variable, and using sons' GPA in math only. Results from the supplementary analyses are available on request from the authors.

RESULTS

DESCRIPTIVE STATISTICS

Table 1 summarizes and compares the two datasets which we use in our empirical example. The upper half of the table describes the attributes of each data source, and the bottom half of the table describes the distribution of the variables in each dataset, by paternal incarceration status.

Table 1. Descriptive Statistics of the NLSY97 and Danish Registry Data

	NLSY97		Danish Registry Data	
Country	United States		Denmark	
Host	US Bureau of Labor Statistics		Statistics Denmark	
Data type	Survey		Administrative	
Birth cohorts	1980-1984		1991	
Data frequency	Annual		Annual	
Retention	81.8 % (boys only)		86.8 %[a]	
Item non-response[b]	25.4 %		12.4 %	
Paternal incarceration	Self-reported: Father ever incarcerated		Precise dates of all admissions and releases	
Son's incarceration	Self-reported: Monthly incarceration status		Precise dates of all admissions and releases[c]	
Son's conviction	Self-reported: Convicted within past 12 months		Precise dates of all convictions	
Son's abilities	Armed Services Vocational Aptitude Battery (ASVAB)		Grade Point Average (GPA) [d]	
	Paternal incarceration	No paternal incarceration	Paternal incarceration	No paternal incarceration
Son incarcerated	.194	.054**	.105	.019***
	(.049)	(.006)	(.307)	(.138)
Son convicted	.355	.164***	.281	.096***
	(.057)	(.009)	(.450)	(.294)
Son's ability	35.656	55.011***	4.615	6.047***
	(3.033)	(0.742)	(2.245)	(2.396)
N	91	1,770	847	24,219

NOTES: Entries in the bottom half of the table are means conditional on sample and paternal incarceration status. Standard deviations in parentheses. Observations in NLSY97 are weighted by the NLSY97 Custom Weights for rounds 1-14.aOut of the 32,945 boys born in 1991 in Denmark, 262 (0.8 %) died before 2015; 560 (1.7 %) have fathers who died before their 15th birthday; 3,294 (10 %) had missing father ID in the population register either in 1991 or 2016. bItem non-response refers to the percentage not having valid information on one of the variables that we use in our analyses. cIn our main analyses we use incarcerations exceeding 24 hours. But we also replicate our results using any incarceration of the son, including incarcerations shorter than 24 hours. dWe also replicate our results using the GPA in math only.

ABBREVIATIONS: NLSY97 = National Longitudinal Study of Youth 1997. N = Number of observations.

SOURCES: Own calculations in data from the US Bureau of Labor Statistics (NLSY97) and Statistics Denmark (Danish Registry Data).
* p < .05; ** p < .01; p < .001.

The NLSY97 is a survey and Danish registry data are obtained from administrative records, as mentioned, which has implications for the nature of the data. The dataset from the Danish registries is much larger than the dataset from the NLSY97, for example, for the obvious reason that the administrative data covers the entire Danish male 1991 birth cohort. The NLSY97 consists of a sample drawn in such a way as to represent the US population born in 1980-1984. NLS staff provides a number of statistical weights to secure researchers the possibility of inferring results from the survey to the mentioned US population (and we apply these weights, see the notes to our tables).

Another difference relates to differential reliability of available information obtained from surveys and from administrative records. As mentioned, the NLS program staff went through great length to minimize the potential impact of, for example, the presence of an interviewer when the respondents answered sensitive questions, such as questions related to criminal activities. But even though procedures such as these increase the precision and validity of information obtained from self-reports, self-reports always come at the price of recollection bias and imprecision (Kirk, 2006). A respondent may not, for example, be aware of his father's incarceration, especially if this occurred when the respondent was a small child. One advantage of administrative data is that they represent official records, and hence do not suffer from recollection biases or imprecision; if a father was incarcerated, official records will prove it.[11] And, given the detailed information on admission and release dates available in the Danish registries, another advantage of administrative data is that they allow us to analyze only paternal incarcerations which were already in effect when the son was born or commenced after it. Another advantage of administrative data is that they hold highly reliable information on incarceration length. Again, asking respondents in a survey to recall for how long their father was incarcerated when they were, for example, small children, is not likely to produce as reliable data as those obtained from official records.

The two datasets also cover two different time periods. Respondents in the NLSY97 were all born in the early 1980s. People in the Danish dataset are all born in 1991. The reason for this difference is that detailed incarceration data are unavailable in Denmark before 1991, and using data from before 1991 would thus involve problems related to identifying sons who experienced paternal incarceration.

The retention rate exceeds 80 percent in both datasets. In Denmark, most of the boys from the 1991 birth cohort who exited the data before their 15th birthday did so because we have incomplete information on whom their father is. Turning to the individual variables which we use in our analyses (the rate of item non-response), around twice as high a share of the US data have insufficient information than in the Danish data.

In both the United States and in Denmark, sons who experienced paternal incarceration before their 15th birthday have higher incarceration rates, higher conviction rates, and lower abilities than sons who did not experience paternal incarceration. Rates of contact with the criminal justice system are more than four (incarceration) and two (conviction) times higher among sons who experienced paternal incarceration than among others. With regard to abilities, sons who experienced paternal incarceration have far lower scores than sons who did not. Again, these results mirror what we expected from the existing literature on the links between paternal incarceration, sons' abilities, and sons' contact with the criminal justice system. The level of criminal justice contact differs across the two countries, and the US rates of contact are much higher than those reported for Denmark. Yet, these differences come with little surprise in light of the differences between the criminal justice systems of the two countries.

11 Registration errors in incarceration records may occur. But such error is likely to be miniscule as incarceration information is third-party reported to Statistics Denmark from the Danish Prison and Probation Service.

BIAS-CORRECTING THE MEDIATION ANALYSIS

Table 2 presents our main results from the NLSY97 sample and the sample from the Danish registry data. We report three key results. First, estimates from logit models are fairly similar across the two data-sets, and both document higher incarceration rates among boys who experienced paternal incarceration before their 15th birthday than boys who did not.[12]

Table 2. Model-Driven Bias in Estimates from Binary Logistic Regression Models, Derived from Estimates of the Total, Direct, and Indirect Effects of the Association Between Paternal Incarceration, Sons' Abilities, and Son's Incarceration up to Age 23, by Sample and by Estimator.

Sample	NLSY97		Danish Registry Data	
Estimator	Logit	KHB	Logit	KHB
Total effect	1.436 (0.330)	1.525 (0.345)	1.787 (0.121)	1.976 (0.127)
Direct effect	1.110 (0.351)	1.110 (0.351)	1.323 (0.126)	1.323 (0.126)
Indirect effect	0.326 (0.077)	0.415 (0.099)	0.464 (0.036)	0.653 (0.126)
Confounding percentage	22.7 %	27.2 %	25.9 %	33.0 %
Model-driven bias in percentage	−21.5 %		−29.0 %	
N	1,861		25,066	

NOTES: Observations in NLSY97 are weighted by the NLSY97 Custom Weights for rounds 1-14.

ABBREVIATIONS: NLSY97 = National Longitudinal Study of Youth 1997. KHB = Karlson, Holm, and Breen (2012). N = Number of observations.

SOURCES: Own calculations in data from the US Bureau of Labor Statistics (NLSY97) and Statistics Denmark (Danish Registry Data)

Second, for both countries, the estimate of the total effect in the conventional logit model is substantially smaller in magnitude than the estimate of the total effect reported by the KHB method. This difference arises as a result of the rescaling bias that we focus on in this paper. Turning to the mediating impact of abilities, Table 2 also reports the percent mediated. Results obtained from the logit model, i.e., a con-

12 We notice that comparing the magnitude of logit coefficients between the two countries is potentially limited by the scale identifiability of these coefficients. If scale parameters differ between the two countries, something that cannot be tested empirically, then the comparison of underlying coefficients will be confounded by these scale differences. See Breen, Holm, and Karlson (2014) for a recent discussion of this issue. Notice, however, as we explained earlier, the percent mediated is not affected by this bias, as the scale parameter drops from this measure of mediation.

ventional mediation analysis disregarding the rescaling bias, show that ability mediates 23 percent of the intergenerational incarceration association in the NLSY97 and 26 percent in Danish registry data, while the corresponding mediation percentage from the KHB model is about 27 percent and 33 percent, respectively. These findings have the important interpretation that issues related to rescaling in logit models bias downward the mediating role of abilities in both datasets. We estimate the model-driven bias, defined in terms of the bias of the indirect effect, to amount to −22 percent in the NLSY97 and to −29 percent in the Danish data. Our findings thus show that model-driven bias most likely has led researchers to underestimate the importance of the mediating role of abilities for the association between paternal incarceration and children's abilities when they evaluate the consequences of paternal incarceration for boys in both the United States and in Denmark. Put differently, the magnitude of the bias suggests that scholars would underestimate the mediating role of ability by about 25 percent on average, a significant bias in relative terms.

Third, although results identify severe problems of model-driven bias in nonlinear probability models of paternal incarceration, bias still differs across the sample from the United States and from Denmark. Model-driven bias in the model on the Danish dataset is around 35 percent higher than in the model on the NLSY97, in relative terms. Given the many differences between the two contexts, which we already outlined in detail, this difference should come as little surprise, however. Also, as explained when comparing the two data sources, the Danish dataset is much newer (and much larger) and likely much more precisely measured than the NLSY97 data, which could also easily matter for how important model-driven bias is.

Findings from supplementary analyses show that the just mentioned results are similar when we analyze sons' conviction risk. And furthermore, results are again similar when we include any incarceration in our outcome measure, even incarcerations shorter than 24 hours, and when we measure abilities from GPA in math only instead of GPA in all subjects. Results from the supplementary analyses are available on request from the authors.

LENTGH OF PATERNAL INCARCERATION

Table 3 reports results which exploit the richness of Danish registry data to break down results by length of paternal incarceration. We show results for paternal incarceration spells equal to or longer than seven days, two weeks, a month, three months, and a full year. Results show that bias is remarkably stable across length of incarceration, falling in the area of −28 to −31 percent. There is no reason to expect, thus, that the main results could be driven by heterogeneity in bias across sentence lengths, which alludes to how general the issue of model-driven bias in nonlinear probability models is.

Table 3. Model-Driven Bias in Estimates from Logit Models, Derived from Estimates of the Total, Direct, and Indirect Effects of the Association Between Paternal Incarceration, Sons' Abilities, and Son's Incarceration up to Age 23, by Length of Paternal Incarceration, Sample and Estimator.

	Main results		≥ 2 weeks		≥ 1 month		≥ 3 months		≥ 1 year	
Estimator	Logit	KHB	Logit	KHB	Logit	KHB	Logit	KHB	Logit	KHB
Total effect	1.787	1.976	1.798	1.982	1.842	2.043	1.662	1.854	1.752	1.905
	(0.121)	(0.127)	(0.124)	(0.130)	(0.134)	(0.141)	(0.160)	(0.167)	(0.209)	(0.220)
Direct effect	1.323	1.323	1.329	1.329	1.385	1.385	1.212	1.212	1.332	1.332
	(0.126)	(0.126)	(0.130)	(0.130)	(0.140)	(0.140)	(0.167)	(0.167)	(0.220)	(0.220)
Indirect effect	0.464	0.653	0.469	0.654	0.456	0.657	0.450	0.642	0.421	0.573
	(0.036)	(0.049)	(0.038)	(0.050)	(0.041)	(0.054)	(0.048)	(0.059)	(0.070)	(0.077)
Confounding percentage	25.9 %	33.0 %	26.1 %	33.0 %	24.8 %	32.2 %	27.1 %	34.6 %	24.0 %	30.1 %
Model-driven bias in percentage	−29.0 %		−28.3 %		−30.6 %		−29.9 %		−26.6 %	
N	25,066		25,066		25,066		25,066		25,066	

NOTES: Standard errors in parentheses.
ABBREVIATIONS: KHB = Karlson, Holm, and Breen (2012). N = Number of observations.
SOURCES: Own calculations on data Statistics Denmark.

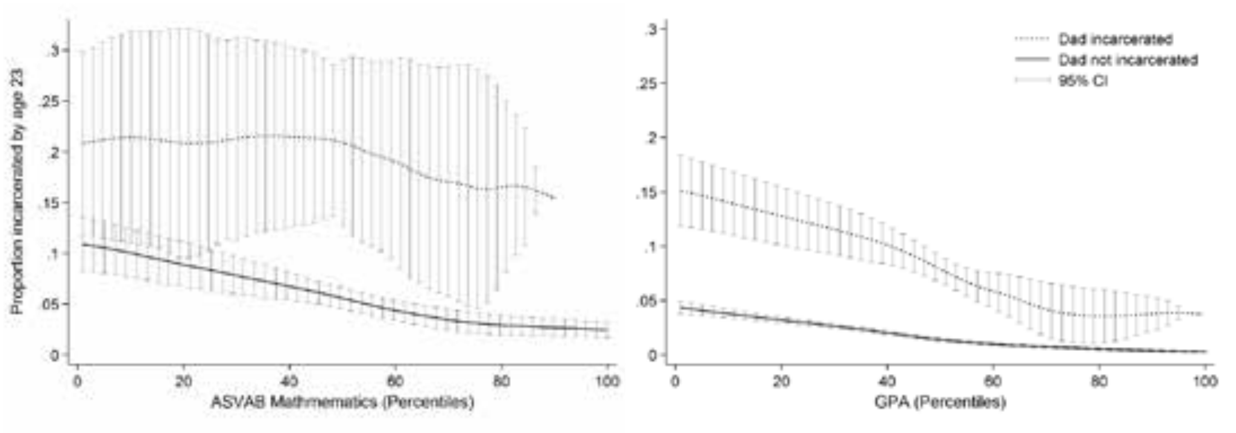
Results from supplementary analyses show that whereas we reach the same conclusion when we measure abilities from GPA in math only, some signs appear that model-driven bias increases with sentence length when we analyze convictions as outcome and when we include even brief periods of incarceration in our outcome. These increases are, however, small and it remains unclear whether they have any substantial interpretation.

UNEQUAL CONSEQUENCES OF PATERNAL INCARCERATION FOR SONS

Although logit models, with proper rescaling correction, are useful for summarizing mediation, the non-linear transformations of the probabilities involved in these models make them difficult to interpret in terms of the predicted probabilities. One solution to this issue would be to report marginal effects, i.e., the average effect of a unit increase in paternal incarceration on a son's probability of being incarcerated. However, in this section, we adopt a different strategy and let "the data speak".

In Figure 2, we report estimates, for each country, of the probability of the son being incarcerated by paternal incarceration status and ability, using regression smoothing techniques. Thus, Figure 2 effectively presents the two-way interaction between paternal incarceration and ability for each country. The solid lines show the proportion incarcerated among sons who did not experience paternal incarceration, and the dashed lines show the same for sons who experienced paternal incarceration as children.

Fig. 2 sProportion of Sons Incarcerated by Age 23, by Abilities and Paternal Incarceration Status, and by Sample.



(a) NLSY97 (b) Danish Registry Data

NOTES: Observations in NLSY97 are weighted by the NLSY97 Custom Weights for rounds 1-14. N = 1,861 in the NLSY97. N = 25,066 in the Danish Registry Data. Plots rely on local polynomial smoothing using the Epanechnikov kernel and bandwidth 20.

ABBREVIATIONS: NLSY97 = National Longitudinal Study of Youth 1997. ASVAB = Armed Services Vocational Aptitude Battery. GPA = Grade Point Average. CI = Confidence Interval.

SOURCES: Own calculations on data from the US Bureau of Labor Statistics (NLSY97) and Statistics Denmark (Danish Registry Data).

The figure shows four results. First, the rate of incarceration in the United States exceeds the one in Denmark across the entire distribution of abilities. Second, incarceration risks for sons of both incarcerated and non-incarcerated fathers generally decline with abilities in both countries, suggesting that ability lowers—independently of paternal incarceration—sons' incarceration risks.

Third, in both countries, the association between the incarceration of fathers and sons is positive across the entire distribution of sons' abilities (i.e., higher incarceration rates among sons who experienced paternal incarceration as children relative to those who did not). Again, this finding validates the claim that children who experience paternal incarceration fare worse on life course outcomes, irrespective of their abilities. This effect corresponds to the direct effect of paternal incarceration reported in Table 2 (i.e., the direct effect reported in Table 2 is an average of this difference stated in log odds ratios)

Fourth, Figure 2 provides insights into how ability differentially compensates the adverse effect of paternal incarceration on son's probability of being incarcerated across the two countries. For the US, see that the ability-slope is slightly steeper among sons of non-incarcerated fathers than among sons of incarcerated fathers. This means that the probability or risk difference—an absolute measure of effect—between the two groups of children slightly increases across the ability distribution. A similar increase is found if one converts these probabilities to odds ratios or risk ratios, i.e., relative measures of the effect. Thus, for the US, the effect of paternal incarceration, however measured, increases with child ability. In other words, the compensating effect of ability, in terms of reducing the gap between sons of incarcerated and non-incarcerated fathers, appears not to exist for the US, and high ability sons suffer more from the consequences of paternal incarceration.

For Denmark, however, the ability-slope is much steeper among sons of incarcerated fathers than among sons of non-incarcerated fathers across the majority of the ability distribution in Figure 2. Consequently, the effect of paternal incarceration decreases across the ability distribution, when reported in absolute terms as a risk difference. Nevertheless, if one converts these differences into odds ratios or risk ratios, these ratios increase across the ability distribution. This would suggest that, in relative terms, the effect in Denmark follows much the same trend across the ability distribution as in the US. In sum, the compensating effect of having above-median abilities appear to differ in Denmark depending on the effect measure adopted. In the Discussion section, we offer two positions on how to reconcile these seemingly contradictory results.

DISCUSSION

Research on the consequences of parental incarceration, and especially paternal incarceration, for children has grown in many important ways over the recent years. With this paper, we add to the growth by offering new methodological knowledge on how we should measure the role of mediating factors in the analysis of parental incarceration and children's outcomes when those outcomes are binary (measuring, for example, contact with the criminal justice system). Specifically, we introduce one of the most recent advances in the modeling of nonlinear probability models, the "KHB Method" (Karlson, Holm, and Breen, 2012).

Mediation analyses are common in criminological research (as well as in the social sciences more broadly) and have advanced knowledge on many important criminological topics over the years (see Walters and Mandracchia, 2017 for a recent review). Researchers often present results of the bivariate association between some outcome and a key variable of interest, such as parental incarceration, using regression techniques. They then add mediating factors to their statistical model. The change which they observe in the parameter estimate associated with the key variable of interest across these models is then said to express the role of the mediating factors for the association under study. But when outcomes are binary, we argue, the standard statistical models for performing mediation analyses risk being biased by the very statistical assumptions used to fit those models to the data. As we have shown in this paper, such bias is potentially large. Our results show that existing models for estimating the importance of mediators, such as abilities, for the association between paternal incarceration and children's binary delinquency outcomes underestimates this importance by a substantial margin.

To display the generality of the issue of model-driven bias which we introduce in this paper, we used two very different datasets to replicate our main results. One dataset is from the National Longitudinal Survey of Youth 1997 (NLSY97), which is a survey that represents people living in the United States in 1997 who were born during 1980–1984. The other dataset is from the administrative records in Denmark, Danish registry data, and here, we focus on birth cohort 1991. In both datasets we measure delinquency as whether the child (i.e., the son) experienced incarceration before age 23 (although we also replicate our main results using other measures of delinquency in supplementary analyses). The dataset from the Danish registries is much larger than the dataset from the NLSY97, and because those data are obtained from official records, they are also much more precisely measured than the self-reported information from the NLSY97. Still, we observe a substantial amount of model-driven bias using both of these datasets. To further show the generality of the issue under study, results across paternal length of incarceration show that the same relative amount of model-driven bias occurs across length of incarceration (because of data limitation this exercise could not be performed on data from the NLSY97).

The amount of model-driven bias which arises from applying nonlinear probability models to the case under study in this paper is in the same range in both of the datasets which we use—implying that model-driven bias could lead researchers to underestimate the mediating role of abilities by around 25 percent. Bias is, however, somewhat higher in Denmark than in the US. Some of this difference is likely to be driven by the differences in the data sources that we just briefly reiterated. Yet from comparing across Denmark and the United States the proportion of sons who experience incarceration by age 23 by abilities and paternal incarceration status, we also found that some of the difference could be driven by differences in the nature of the associations in those two contexts.

IMPLICATIONS FOR THE UNITED STATES AND DENMARK

The mediating role of abilities seems to be more important in Denmark than in the US, as was also confirmed by our regression analyses in the main results (28 percent of the total effect of paternal incarceration on sons' incarceration was attributable to the mediating role of abilities in the NLSY97; the corresponding number from Denmark is 33 percent). We found that in the NLSY97, the effect of paternal incarceration, however measured, increases with child ability. We thus find that ability fails to reduce the gap between sons of incarcerated and non-incarcerated fathers, and high ability sons suffer more from the consequences of paternal incarceration in the US.

In Denmark, we find another and interesting pattern. Here, the effect of paternal incarceration decreases across the ability distribution, when reported in absolute terms as a risk difference. But converting this absolute measure to a relative one—odds ratios—reveals that these ratios increase across the ability distribution. In relative terms the effect in Denmark thus follows much the same trend across the ability distribution as in the US, leading to the conclusion that the compensating effect of having above-median abilities differs in Denmark depending on the effect measure adopted

How are we to reconcile these seemingly different results? We offer two positions: one for policy-makers and one for mobility scholars. For policy-makers, the risk difference would be of primary interest. Policy-makers are chiefly interested in identifying individuals for whom interventions are most effective. To fix idea, we imagine an intervention that would decrease the number of incarcerated fathers (e.g., by non-custodial sentencing alternatives). Such interventions would impact children's probabilities of incarceration in two ways.

First, assuming that such intervention did not change the distribution of abilities (i.e., the direct effect as defined within fixed levels of ability), the intervention would be slightly more effective among above-median ability sons in the United States (although the effectiveness would be quite similar overall), whereas it would be most effective among below-median ability sons in Denmark. In a sense, the low risk difference among above-median sons in Denmark suggests that the Danish welfare state probably have other mechanisms in place that promote the educational achievement of bright children irrespective of which socio-economic circumstances they are born into. Second, we may imagine that such intervention would hypothetically increase the overall ability of children (i.e., the indirect effect). For the US, this would imply a perverse scenario in which a larger fraction of sons would experience the comparatively more adverse effects of having an incarcerated father.[13] For Denmark, however, this would imply an equalizing effect in that a larger fraction of sons would experience the comparatively less adverse effects of having an incarcerated father. Thus, the policy-implications of our findings are quite different for the two countries.

13 This is of course only under the assumption that the effect of ability on sons' probability of being incarcerated independently of paternal incarceration status would not change with such interventions.

In contrast to policy-makers, mobility scholars are interested in transmission processes across generations (i.e., processes of social mobility), and the convention within this area is therefore to use the odds ratio as the associational measure of interest, as this speaks to the level of relative, as opposed to absolute, mobility. In this situation, we would conclude that ability appears to have little compensating effect in terms of promoting relative mobility chances for both countries. Relative mobility levels simply appear to be higher (i.e., the odds ratios are lower) among below-median ability sons. Numerical evaluations (based on Figure 2 but not reported here) suggest that this trend is relatively stronger in the United States than in Denmark, yet the fact that the two patterns are close to identical in the two countries suggests that the mobility promoting effect of ability operates very similarly in these two otherwise quite dissimilar countries.

Future research should build upon these findings and analyze which institutional features and structural constraints of the two contexts might influence the consequences of paternal incarceration for children. Examples of potentially important features and structural constraints are the equalizing role of strong welfare institutions in Denmark and the demographic dispersion of incarceration in the US, where incarceration is especially predominant among disadvantaged social and racial groups.

In conclusion, theory development essentially comes from weighing the importance of different social mechanisms regarding the social phenomena which scholars observe. It is therefore crucial that empirical research produces reliable estimates of the importance of those mechanisms. Failure to do so could imply that the theories which we rely on for understanding why children of incarcerated parents fare worse on life course outcomes, for example, misrepresent the true importance of those mechanisms. Results presented in this paper show that if we continue to apply the most commonly used models (nonlinear probability models) for estimating the mediating role of abilities for the association between the incarceration of fathers and sons, model-driven bias could lead us to substantially underestimate the importance of that role. It is therefore pivotal that researchers in criminology pay attention to the statistical properties of the models which they apply—inattention to these properties could lead them to gravely underestimate the importance of substantially impactful mediation mechanisms.

REFERENCES

Andersen, Signe H. and Christopher Wildeman. (2014). The Effect of Paternal Incarceration on Children's Risk of Foster Care Placement. Social Forces 93(1):269–98.

Andersen, Lars H. (2016). How Children's Educational Outcomes and Criminality Vary by Duration and Frequency of Paternal Incarceration. The ANNALS of the American Academy of Political and Social Science 665(1):149–70.

Berger, Lawrence, Maria Cancian, Laura Cuesta, and Jennifer L. Noyes. (2016). Families at the Intersection of the Criminal Justice and Child Protective Services Systems. The ANNALS of the American Academy of Political and Social Science 665(1):171-94.

Bijleveld, Catrien C. J. H. and Miriam Wijkman. (2009). Intergenerational Continuity in Convictions: A Five-Generation Study. Criminal Behaviour and Mental Health 19:142–55.

Breen, Richard, Kristian B. Karlson, and Anders Holm. (2013). Total, Direct, and Indirect Effects in Logit and Probit Models. Sociological Methods & Research 42(2):164-91.

Breen, Richard, Anders Holm, and Kristian B. Karlson. (2014). Correlations and Nonlinear Probability Models. Sociological Methods & Research 43(4):571-605.

Cho, Rosa M. (2009). Impact of Maternal Imprisonment on Children's Probability of Grade Retention. Journal of Urban Economics 65(1):11–23.

Cho, Rosa M. (2009). The Impact of Maternal Imprisonment on Children's Educational Achievement: Results from Children in Chicago Public Schools. The Journal of Human Ressources 44(3):772–97.

Cho, Rosa M. (2010). Maternal Incarceration and Children's Adolescent Outcomes: Timing and Dosage. Social Service Review 84(2):257–82.

Comfort, Megan. (2016). "A Twenty-Hour-a-Day Job": The Impact of Frequent Low-Level Criminal Justice Involvement on Family Life". The ANNALS of the American Academy of Political and Social Sciences 665: 63-79.

Danish Prison and Probation Service. (2016). Statistik 2015. Copenhagen, Denmark: Danish Prison and Probation Service.

Farrington, David P., Geoffrey C. Barnes, and Sandra Lambert. (1996). The Concentration of Offending in Families. Legal and Criminological Psychology 1(1):47–63.

Farrington, David P., Jeremy W. Coid, and Joseph Murray. (2009). Family Factors in the Intergenerational Transmission of Offending. Criminal Behaviour and Mental Health 19:109–24.

Farrington, David P., Darrick Jolliffe, Rolf Loeber, Magda Stouthamer-Loeber, and Larry M. Kalb. (2001). The Concentration of Offenders in Families, and Family Criminality in the Prediction of Boys' Delinquency. Journal of Adolescence 24(5):579–96.

Foster, Holly and John Hagan. (2007). Incarceration and Intergenerational Social Exclusion. Social Problems 54(4):399–433.

Foster, Holly and John Hagan. (2009). The Mass Incarceration of Parents in America: Issues of Race/Ethnicity, Collateral Damage to Children, and Prisoner Reentry. The ANNALS of the American Academy of Political and Social Science 623(1):179–94.

Foster, Holly and John Hagan. (2015). Punishment Regimes and the Multilevel Effects of Parental Incarceration: Intergenerational, Intersectional, and Interinstitutional Models of Social Inequality and Systemic Exclusion. Annual Review of Sociology 41:135–58.

Foster, Holly and John Hagan. (2015). Maternal and Paternal Imprisonment and Children's Social Exclusion in Young Adulthood. The Journal of Criminal Law & Criminology 105(2):387–430.

Frisell, Thomas, Paul Lichtenstein, and Niklas Långström. (2011). Violent Crime Runs in Families: A Total Population Study of 12.5 Million Individuals. Psychological Medicine 41(1):97–105.

Giordano, Peggy. (2010). Legacies of Crime. A Follow-Up of the Children of Highly Delinquent Girls and Boys. Cambridge: Cambridge University Press.

Guerino, Paul, Paige M. Harrison, and William J. Sabol. (2011). Prisoners in 2010. Washington, DC: US Bureau of Justice Statistics.

Hagan, John and Holly Foster. (2012). Intergenerational Educational Effects of Mass Imprisonment in America. Sociology of Education 85(3):259–86.

Hjalmarsson, Randi and Matthew J. Lindquist. (2012). Like Godfather, like Son: Exploring the Intergenerational Nature of Crime. Journal of Human Resources 47(2):550–82.

Karlson, Kristian B., Anders Holm, and Richard Breen. (2012). Comparing Regression Coefficients between Same-Sample Nested Models using Logit and Probit: A New Method. Sociological Methodology 42(1):286-313.

Kirk, David S. (2006). Examining the Divergence across Self-Report and Official Data Sources on Inferences about the Adolescent Life-Course of Crime. Journal of Quantitative Criminology 22(2): 107-29.

Kohler, Ulrich, Kristian B. Karlson, and Anders Holm. (2011). Comparing Coefficients of Nested Nonlinear Probability Models. The Stata Journal 11(3):420-38.

Maguin, Eugene and Rolf, Loeber. (1996). Academic Performance and Delinquency, in Crime and Justice 20. Chicago: The University of Chicago Press: 145-264.

Miller, Holly V. and J. C. Barnes. (2015). The Association between Parental Incarceration and Health, Education, and Economic Outcomes in Young Adulthood. American Journal of Criminal Justice 40(4):765–84.

Motivans, Mark. (2011). Federal Justice Statistics 2009 - Statistical Tables Bureau of Justice Statistics Statistical Tables. Washington, DC:US Bureau of Justice Statistics.

Murray, Joseph and David P. Farrington. (2005). Parental Imprisonment: Effects on Boys' Antisocial Behaviour and Delinquency through the Life-Course. Journal of Child Psychology and Psychiatry, and Allied Disciplines 46(12):1269–78.

Murray, Joseph and David P. Farrington. (2008). The Effects of Parental Imprisonment on Children. Crime and Justice 37(1):133–206.

Murray, Joseph, Carl-Gunnar Janson, and David P. Farrington. (2007). Crime in Adult Offspring of Prisoners: A Cross-National Comparison of Two Longitudinal Samples. Criminal Justice and Behavior 34(1):133–49.

Neuhaus, J.M, Kalbfleish, J.D., and Hauck, W.W. (1991). A Comparison of Cluster-Specific and Population-Averaged Approaches for Analyzing Correlated Binary Data. International Statistical Review 59(1):25-35.

OECD. (2016). Income Inequality Remains High in the Face of Weak Recovery. Income Inequality Update. OECD: Centre for Opportunity and Equality:1-6.

Ramakers, Anke A. T., Catrien Bijleveld, and Stijn Ruiter. (2011). Escaping the Family Tradition: A Multi-Generation Study of Occupational Status and Criminal Behaviour. British Journal of Criminology 51(5):856–74.

Thornberry, Terence P., Adrienne Freeman-Gallant, Alan J. Lizotte, Marvin D. Krohn, and Carolyn a Smith. (2003). Linked Lives: The Intergenrational Transmission of Antisocial Behavior. Journal of Abnormal Child Psychology 31(2):171–84.

Turney, Kristin. (2014). Stress Proliferation Across Generations? Examining the Relationship between Parental Incarceration and Childhood Health. Journal of Health and Social Behavior 55:302–19.

Turney, Kristin. (2017). The Unequal Consequences of Mass Incarceration for Children. Demography 54(1):361-89.

Turney, Kristin and Anna Haskins. (2014). Falling Behind? Children's Early Grade Retention after Paternal Incarceration. Sociology of Education 87:241–58.

Wakefield, Sara and Christopher Wildeman. (2014). Children of the Prison Boom. Mass Incarceration and the Future of American Inequality. Oxford: Oxford University Press.

Walmsley, Roy. (2016). World Prison Population List, 11th ed. London, UK: Institute for Criminal Policy Research.

Walters, Glenn D. and Jon T. Mandracchia. (2017). Testing Criminological Theory through Causal Mediation Analysis: Current Status and Future Directions. Journal of Criminal Justice Online First: http://dx.doi.org/10.1016/j.jcrimjus.2017.02.002.

Wildeman, Christopher. (2009). Parental Imprisonment, the Prison Boom, and the Concentration of Childhood Disadvantage. Demography 46(2): 265–80.

Wildeman, Christopher. (2010). Paternal Incarceration and Children's Physically Aggressive Behaviors : Evidence from the Fragile Families and Child Wellbeing Study. Social Forces 89(1): 285–309.

Wildeman, Christopher. (2014). Parental Incarceration, Child Homelessness, and the Invisible Consequences of Mass Imprisonment. The ANNALS of the American Academy of Political and Social Science 651(1):74–96.

Wildeman, Christopher and Signe H. Andersen. (2017). Paternal Incarceration and Children's Risk of Being Charged by Early Adulthood: Evidence from a Danish Policy Shock. Criminology 55(1):32–58.

Wildeman, Christopher, Signe H. Andersen, Hedwig Lee, and Kristian B. Karlson. (2014). Parental Incarceration and Child Mortality in Denmark. American Journal of Public Health 104(3):428-33.

Wildeman, Christopher and Lars H. Andersen (2015). Cumulative Risks of Paternal and Maternal Incarceration in Denmark and the United States. Demographic Research 32(June): 1567–80.

Wildeman, Christopher and Kristin Turney. (2014). Positive, Negative, or Null? The Effects of Maternal Incarceration on Children's Behavioral Problems. Demography 51(3):1041–68.