2013

What is the effect of IQ on offending?

Daniel P. Mears and Joshua C. Cochran

PRINT VERSION CITATION: Mears, Daniel P., and Joshua C. Cochran. 2013. "What Is the Effect of IQ on Offending?" Criminal Justice and Behavior 40(11):1280-1300.

PRE-PRINT VERSION

What Is the Effect of IQ on Offending?*

Daniel P. Mears and Joshua C. Cochran

* Author's Note. Direct correspondence to Daniel P. Mears, Ph.D., Mark C. Stafford Professor of Criminology, Florida State University, College of Criminology and Criminal Justice, 634 West Call Street, Tallahassee, FL 32306-1127, e-mail (dmears@fsu.edu), phone (850-644-7376), fax (850-644-9614). Joshua C. Cochran, Ph.D., University of South Florida, Department of Criminology, 4202 East Fowler Avenue, Tampa, FL 33620-7200, email (jccochran@usf.edu), phone (813-974-9708). We thank Sonja Siennick and Brian Stults for serving as sounding boards in the development of this paper and the anonymous reviewers for constructive suggestions and ideas for strengthening it.

Biographical Sketches

Daniel P. Mears, Ph.D., is the Mark C. Stafford Professor of Criminology at Florida State University's College of Criminology and Criminal Justice, 634 West Call Street, Tallahassee, FL 32306-1127, phone (850-644-7376), fax (850-644-9614), e-mail (dmears@fsu.edu). He conducts basic and applied research on crime and justice. His work has appeared in *Criminology*, the *Journal of Research in Crime and Delinquency*, and other crime and policy journals and in a book, *American Criminal Justice Policy* (Cambridge University Press 2010).

Joshua C. Cochran, Ph.D., is an Assistant Professor at University of South Florida's Department of Criminology, 4202 East Fowler Avenue, Tampa, FL 33620-7200, email (jccochran@usf.edu), phone (813-974-9708). His research interests include criminological theory, imprisonment, and juvenile court sentencing. He has published in *Justice Quarterly, Journal of Quantitative Criminology, Crime and Delinquency*, and other crime and justice journals.

What Is the Effect of IQ on Offending?

Abstract

The aim of this study is to advance scholarship on the IQ-offending relationship by examining the functional form of this relationship and whether confounding introduced by socioeconomic status and other factors can be adequately addressed. Data from the National Longitudinal Survey of Youth are analyzed using generalized propensity score and propensity score matching analyses. The results suggest that the relationship is curvilinear, such that lower and higher levels of IQ are associated with lower levels of offending. They also indicate that the distribution of confounders, especially socioeconomic status, may limit the ability of statistical approaches to arrive at unbiased estimates of IQ effects.

KEYWORDS: IQ intelligence offending curvilinear relationship

The idea that low intelligence causes offending ascended into prominence with the publication of Goddard's (1914) book, *Feeblemindedness*, in which he argued that "at least 50 percent of all criminals are mentally defective" (p. 9). Although his work focused on individuals with very low levels of intelligence, subsequent research focused on the more general idea that intelligence is inversely related to offending. Then, concomitant with the rise in a more sociological view of criminal behavior, this belief fell into disrepute for several decades, only to be resuscitated by Hirschi and Hindelang's (1977) review, which showed that many studies consistently identified an inverse association between intelligence and offending. Subsequently, Wilson and Herrnstein (1985) and then Herrnstein and Murray (1994) forcefully and provocatively argued that poor cognitive abilities contributed to many adverse outcomes, including crime. This argument continues to receive considerable attention (Walsh, 2011).

A central reason for the contentious debate about offending and IQ, a term typically treated as synonymous with the concept of intelligence, stems in no small part from the implications that originally were associated with the idea of an IQ-offending relationship. Goddard (1914:582), for example, argued that colonizing the "mentally defective" constituted an appropriate and effective way, along with sterilization (p. 589), of managing the large number of feebleminded individuals, and, by extension, criminals. It also stems from the implications that Herrnstein and Murray (1994:251) articulated in their study: "In trying to understand how to deal with the crime problem, much of the attention now given to problems of poverty and unemployment should be shifted to another question altogether: coping with cognitive disadvantage."

Perhaps because of this controversy, the notion of an IQ-offending relationship has been subjected to increased scrutiny. Even so, basic questions about the IQ-offending relationship remain. Theoretical accounts of the relationship, for example, posit pathways that with rare exception have not been subject to rigorous assessment and that assume IQ and offending indeed are associated. The nature of the IQ and offending relationship remains, however, ambiguous. The bulk of studies to date rely on law enforcement, court, and prison records data to estimate the association of IQ and offending, thereby potentially conflating the possibility that IQ contributes to offending with the possibility that it leads to a greater likelihood of law enforcement detection or differential treatment. In addition, few studies employ methods that reduce the likelihood that identified effects of IQ on self-reported offending are spurious. Studies also have employed inconsistent approaches to coding IQ and offending. Not least, few studies have carefully investigated the possibility that, as some scholars have suggested, the relationship between IQ and offending is curvilinear.

Against this backdrop, the goal of this paper is to contribute to scholarship on the etiology of offending. In particular, the aim is to build on prior work on the IQ-offending relationship by assessing not only whether IQ and offending are associated but also whether the relationship holds after employing methods that can better address potential confounding and, not least, that allow for testing the argument that IQ and offending are curvilinearly related. We begin first by discussing prior theoretical and empirical research on IQ and offending, and then describe the data and methods used to assess the IQ-offending relationship. After presenting the findings, we discuss the implications of the study for theory and research.

The Theoretical Argument for an IQ-Offending Relationship

Scholars have identified several mechanisms through which IQ may affect offending. However, few studies have empirically examined these mechanisms. Accordingly, as Jensen (1998:298) has observed, there remains no consensus about which theories provide the most credible explanation for the relationship. Indeed, as Jolliffe and Farrington (2010:44) have emphasized, "the mechanism that connects low intelligence and offending is still hotly debated." Notably, any such debate proceeds from the assumption of an association between IQ and offending. However, as we discuss below, it may be that any observed relationship between IQ and official measures of offending stems from differential detection and treatment of low-IQ individuals by law enforcement, the courts, and correctional systems, or from potential confounding of IQ with other factors. Regardless, to the extent that a causal IQ-offending association relationship does exist, several theoretical mechanisms may account for it.

One centers on the idea that people who are lower in intelligence may lack sufficient moral awareness of how to behave and so are more likely to offend (Langdon, Clare, & Murphy, 2011). A related argument is that individuals with lower IQs may be less able "to foresee the consequences of their offending and to appreciate the feelings of victims" (Farrington & Welsh, 2008:41; see also Lynam, Moffitt, & Stouthamer-Loeber, 1993; Moffitt, 1993). It also may be that those with lower cognitive ability are less likely to develop self-control and thus be more likely to engage in criminal activity (Gottfredson & Hirschi, 1990; McGloin, Pratt, & Maahs, 2004; Felson & Staff, 2006).

Drawing on different theoretical traditions, scholars have identified still other pathways. For example, a lower IQ may lead to poor academic performance and school experiences. More generally, individuals with low IQs may be less able to successfully negotiate social relationships and situations, including relationships in school and with family and friends. These effects in turn may lead to weak social bonds, greater strain, greater association with delinquent peers, and negative labeling. In each instance, the anticipated end result is a greater likelihood of offending (Hirschi & Hindelang, 1977; Menard & Morse, 1984; Lynam et al., 1993; Ward & Tittle, 1994; Moffitt et al., 1995; McGloin et al., 2004; Farrington & Welsh, 2008).

One of the few studies to examine such possibilities empirically found that IQ influenced school performance, and, to a lesser extent, delinquent peer association and self-control, and, in turn, offending (McGloin et al., 2004). Ward and Tittle (1994) also found that IQ effects on offending arose through school performance and in turn social bonds. Similarly, Menard and Morse (1984) found that the association in their data between IQ and offending could be explained primarily by the effect of IQ on school performance and in turn perceived negative social labeling and delinquent peer association (p. 1369).

Such studies provide important insights into how IQ may affect offending. However, they proceed from the assumption that IQ does in fact affect offending and, typically, from a related assumption that the relationship is linear. Below, we review what is known about the IQ-offending relationship and discuss arguments that the effect of IQ on offending is curvilinear and, in turn, the implications of such arguments for theory and research.

Empirical Evidence on the IQ and Offending Relationship

Among the most widely cited assessments of the state of research on IQ and offending are Hirschi and Hindelang's (1977) and Wilson and Herrnstein's (1985) reviews. Both aimed to correct what was viewed as a sociological bias in criminology that downplayed the significance of IQ. Each one led to the conclusion that, contrary to what many criminology textbooks suggested, a large body of empirical studies clearly demonstrated that there is a negative relationship between IQ and offending. In a later work, Herrnstein and Murray (1994:242) stated that, "taking the literature as a whole, incarcerated offenders average an IQ of about 92, eight points below the mean." They noted, as did Hirschi and Hindelang (1977), that this relationship appears to hold with self-reported offending as well. Recent reviews echo their assessments and do so by citing these works along with more contemporary research on the topic (e.g., Moffitt, Caspi, Silva, & Stouthamer-Loeber, 1995; Ellis & Walsh, 2003; Farrington & Welsh, 2008; Walsh, 2011).

Juxtaposed against this general finding are caveats that make it difficult to arrive at an unequivocal statement about IQ and offending. First, the extent to which estimates of the relationship between the two are spurious remains largely unknown. Reviews note that the relationship may be spurious because of the possibility of differential police detection of low-IQ individuals or the failure to take into account a third variable, such as social class and race, that may influence both IQ and offending (Hirschi & Hindelang, 1977; Menard & Morse, 1984; Lynam et al., 1993; Herrnstein & Murray, 1994; Cullen, Gendreau, Jarjoura, & Wright,1997; McGloin et al., 2004; Jolliffe & Farrington, 2010). A large body of studies introduces controls for social class and race, but few studies address the possibility of differential detection or the related possibility that some individuals, such as mentally retarded offenders, may be diverted from criminal justice processing. The most notable exception is Moffitt and Silva's (1988) study of New Zealand 13-year-olds; they used official records data and self-report data to examine two groups of delinquents and found that delinquents who were detected by the police were nearly identical in IQ to delinquents who were not so detected (see also Moffitt et al., 1995; Ellis &

Walsh, 2003).

Second, some dispute exists about the strength of the IQ and offending relationship. Studies suggest that arrested and incarcerated individuals have lower IQs as compared to the general population (Hirschi & Hindelang, 1977; Herrnstein & Murray, 1994; Walsh, 2011). Few studies, however, assess the difference using general population samples and self-reported offending (Herrnstein & Murray, 1994:243; Walsh, 2011).[1] The studies that do so are less consistent in their estimates and some suggest that the effect is smaller (Hirschi & Hindelang, 1977:576; see, generally, West & Farrington, 1973; Herrnstein & Murray, 1994; Cullen et al., 1997; McGloin et al., 2004). Implicit in many accounts is the assumption that variation within the normal range of IQ (e.g., 90-109) is meaningful and can contribute to variation in offending.

Claims about the association between IQ and offending are complicated by variation in how the two are coded. For example, studies vary greatly in how they measure intelligence (e.g., Wechsler Intelligence Scale for Children, Stanford-Binet, Raven Progressive Matrices, Ammons Quick Test) and how they code it, with some using binary codings, quartiles, deciles, or continuous measures. In addition, many studies compare the average IQs of incarcerated youth and non-incarcerated youth or, conversely, they create low and high IQ groups or some other classification (see, e.g., Hirschi & Hindelang, 1977; Lynam et al., 1993; Herrnstein & Murray, 1994; McCord & Ensminger, 1997; Ellis & Walsh, 2003). They then compare the percentages of each group who are incarcerated with the percentages of those who are not. Measures of offending also vary greatly. Some studies use dichotomous measures, such as "incarcerated versus non-incarcerated," or use offending indices or counts that themselves sometimes are recoded into groups (e.g., the top decile of an index versus any lower decile).

There is, accordingly, no consistent estimate of, say, the effect of unit increases in IQ on the likelihood or amount of offending. Farrington (1993), for example, compared youth in a low IQ group, consisting of youth in the bottom 25 percent of an intelligence measure, with those in the top 75 percent. The approach is neither "correct" nor "incorrect" any more than is the use of a series of IQ groups (e.g., Hirschi & Hindelang, 1977) or a continuous measure of IQ (e.g., Cullen et al., 1997). The differences in approaches, however, make it difficult to make meaningful comparisons across studies. Even so, it appears defensible to claim based on extant work that, as a general matter, IQ and offending are negatively associated to some degree (Ellis & Walsh, 2003; Farrington & Welsh, 2008; Jolliffe & Farrington, 2010; Walsh, 2011).

Third, the bulk of studies to date imply that the IQ and offending relationship is linear, even though several scholars have suggested that it is curvilinear. There is, for example, the possibility that, per social learning theory, individuals learn to commit crime and that, in turn, they need a certain level of intelligence to undertake such learning or, more broadly, to be able to engage in offending. Indeed, Wilson and Herrnstein (1985) made this claim and in turn argued that "the relationship between intelligence and crime is thus curvilinear, with most offenders falling in the low normal or borderline subnormal range . . .; their relative frequencies decline on either side of this range" (p. 155). In making this argument, they pointed to one of Hirschi and Hindelang's analyses (1977:574) and, in particular, to the slightly higher percentage of white males in the second-to-the-bottom IQ quintile who engaged in more officially recorded acts of delinquency than youth in the quintiles below and above them. Notably, however, the curvilinearity was not reported as statistically significant nor was it evident in Hirschi and Hindelang's (1977:577) analyses of self-reported offending. The possibility of a curvilinear association has been expressed by other scholars as well (see, e.g., Jensen, 1998:570; Lindsay & Taylor, 2010:289), yet has not been subject to careful empirical scrutiny.

The importance of investigating the claim about a nonlinear relationship is two-fold. If the IQ and offending relationship is curvilinear in the manner described by Wilson and Herrnstein (1985), then IQ may be more strongly associated with offending than prior research suggests and there may be a need to consider more closely the precise way in which IQ influences offending. The mechanisms, for example, may differ for individuals at different parts of the IQ distribution. As Wilson and Herrnstein (1985) suggest, individuals with lower IQs may be less able to engage in offending; they may not understand sufficiently how to undertake it. According to moral reasoning theory, these individuals also may be at an "earlier" stage of development, one that "is associated with obeying rules" and thus that reduces the likelihood of offending (Langdon et al., 2011:105). By contrast, individuals in the middle of the IQ spectrum may possess sufficient intelligence to undertake crime or, because of their level of intelligence, hold a more egocentric perspective that puts a greater premium on satisfying personal needs that in turn contributes to offending. In addition, as compared to higher IQ individuals, they may lack as many legitimate opportunities for financial gain and social status (DiRago & Vaillant, 2007); thus, they may experience more strain and, in turn, be more likely to offend. Individuals with higher levels of intelligence not only may have more such opportunities, they also may have moral reasoning that has progressed away from an egocentric view (Langdon et al., 2011:105). Such explanations do not exhaust the possibilities, but they illustrate that evidence of a curvilinear association between IQ and offending holds the potential to inform efforts to develop more theoretically nuanced accounts that may improve our understanding of the IQ-offending relationship.

In short, there is a need for studies that can address several gaps in the IQ and offending literature. Research is needed, for example, that examines a wide range of IQ groups and that also examines IQ as a continuous measure. Research also is needed that employs a more diverse set of measures of offending and that examines the extent to which the IQ and offending relationship is linear or curvilinear. And, not least, research is needed that can better address potential confounding that may bias estimates of the IQ and offending relationship.

Data and Methods

Overview

Given these research gaps, the goal of this study is to examine the nature of the IQ and offending relationship by examining self-reported offending and by investigating whether the IQ and offending association is curvilinear. In so doing, we identify several nuances that complicate discussions of the IQ and offending relationship and, at the same time, assess the extent to which it is possible to obtain estimated IQ effects on offending that adjust for imbalance in variables typically included to address potential confounding. To this end, we conduct generalized propensity score (GPS) and propensity score matching (PSM) analyses using data from the

NLSY. There are several advantages in using these data. They are nationally representative; the different waves of data are used in studies that investigate the relationship between intelligence and a range of outcomes (see, e.g., Zagorsky, 2007); they include measures both of intelligence and of self-reported crime; and they have been used in research on offending (e.g., Herrnstein & Murray, 1994; Cullen et al., 1997; McGloin et al., 2004). In addition, analyses with widely used data helps to ensure that the results are replicable. Following Herrnstein and Murray (1994) and Cullen et al. (1997), we use the first two years, 1979 and 1980, of the NLSY survey. To facilitate comparisons with these studies, we focused only on white males (N=3,253). We then replicated their analyses to ensure comparability and then undertook the GPS and PSM analyses. Below we describe the measures used; Table 1 provides the descriptive statistics.

Dependent Variables

In this study, we examine self-reported offending in 1980 and so avoid the concerns that arise when using official records data to estimate the effect of IQ on offending. In addition, and in contrast to most prior studies of the IQ-offending relationship, we examine different types of offending, including general, violent, property, drug, chronic, and diversity of offending, to investigate whether IQ exerts a different effect depending on the type of crime.

Each is a standardized measure except for chronic offending, which is a binary measure. For general offending, we followed Herrnstein and Murray (1994) and Cullen et al. (1997) by creating an index using twenty self-reported delinquency items. Each ranged from 0 to 6, with higher values indicating more frequent delinquency in the year prior to the interview (0=never, 1=once, 2=twice, 3=3-5 times, 4=6-10 times, 5=11-50 times, 6-50+ times). The twenty offenses were as follows (means and standard deviations are in parentheses): runaway (0.16, 0.59), truancy (1.30, 1.69), drinking (2.08, 2.11), vandalism 0.36, 0.92), fighting (0.56, 1.09), shoplifting (0.54, 1.11), petty theft (0.37, 0.92), grand theft (0.11, 0.54), robbery (0.10, 0.52), assault (0.82, 1.33), aggravated assault (0.20, 0.73), marijuana use (1.86, 2.36), hard drug use (0.60, 1.41), selling marijuana (0.33, 1.10), selling hard drugs (0.07, 0.53), fraud (0.48, 1.09), auto theft (0.14, 0.59), breaking and entering (0.12, 0.56), fencing (0.23, 0.77), and gambling

(0.06, 0.45). For the violence index, the measures included robbery, assault, and aggravated assault. Property included shoplifting, petty theft, grand theft, fraud, auto theft, fencing, and gambling. Drug included hard drug use and selling marijuana or hard drugs. The indices were produced by standardizing each of the items and then taking the average.

Herrnstein and Murray (1994) and Cullen et al. (1997) used a "top decile" of offending measure. Here, we go beyond their approach by employing two additional measures of offending. (For ancillary analyses, we used the top decile measure as well.) The first, chronic offending, is a binary measure indicating whether individuals self-reported engaging in any of seventeen illegal behaviors six or more times. Three offenses—running away, truancy, and drinking—were excluded because they are status offenses. The second, diversity of offending, is based on the dichotomous codings for each separate non-status offense. Higher values indicate that individuals engaged in a wider variety of crimes. Sweeten's (2012) review and assessment of offending measures recommends this approach as a more valid way of operationalizing criminality than ones that emphasize frequency, type of offense, or offense severity alone.

Independent Variable

The NLSY included the Armed Forces Qualification Test (AFQT). Herrnstein and Murray (1994) used it as a measure of IQ because "it is what psychometricians call 'highly g-loaded,' meaning that it is a good measure of general cognitive ability" (p. 120) and it is highly correlated with other measures of general intelligence (pp. 584-585).[2] Zagorsky (2007:491) has noted that, notwithstanding valid critiques of the AFQT, "since AFQT scores are highly correlated with general intelligence, g, the research community has used AFQT as a proxy for intelligence."

We use it here for the same reason but caution that a robust debate exists in the scholarly community about the best ways to conceptualize and operationalize intelligence. We use the AFQT measure, too, because measures of general intelligence are among the most widely used in the literature on offending; use of the measure thus allows us to provide analyses consistent with those undertaken by Herrnstein and Murray (1994), Cullen et al. (1997) and other scholars. At the same time, it is general intelligence that has been argued to be the primary factor contributing to offending (Herrnstein, 1995:50). Some studies suggest that types of intelligence may be differentially associated with offending. For example, verbal intelligence may be more strongly associated with offending than is performance intelligence (e.g., Moffitt, 1993; Farrington & Welsh, 2008). However, this finding is not consistent across studies (see, e.g., Lynam et al., 1993; Isen, 2010; Willis, Dumont, & Kaufman 2011; cf. Walsh, 2011) and belies the fact that many other types of intelligence have been posited to exist that might also differentially affect offending (Sternberg & Kaufman, 2011). In addition, studies find that diverse types of intelligence typically are strongly correlated with general intelligence (Jensen, 1998; Herrnstein & Murray, 1994; Nyborg, 2003; Rindermann, 2007; Isen, 2010; Sternberg & Kaufman, 2011), providing further warrant for focusing on general intelligence and its potential relationship to offending.

For all analyses, we use the 2006 revised approach to creating a standardized measure of the 1980 AFQT raw scores, at the recommendation of the NLSY researchers (see Bureau of Labor Statistics, 2011). The measure ranges from 0 to 100; a given score within this range indicates the percent of individuals who scored below that score. For example, a score of 85 indicates that an individual scored comparably to or better than 85 percent of the individuals in the sample. For the bivariate and PSM analyses, we use the NLSY-recommended set of scores to create deciles. Because the analysis sample consists of white males, the IQ score ranges for the deciles differ from those that Herrnstein and Murray (1994:371) provided for the total sample.

As inspection of Table 1 shows, the mean percentile rank for the NLSY sample was 50 (standardized IQ=100) but for our sample of white males is 56 (IQ=105). The IQ ranges for the deciles are: D1 (77.8-83.0), D2 (83.0-88.3), D3 (88.3-93.5), D4 (93.5-98.7), D5 (98.7-103.9), D6 (103.9-109.1), D7 (109.1-114.3), D8 (114.3-119.6), D9 (119.6-124.8), and D10 (124.8-130.0). Herrnstein and Murray (1994:120) used five groups: very dull (IQ <75); dull (IQ=75-90); normal (IQ=90-110); bright (IQ=110-125); and very bright (IQ>125). The use of deciles here allows for comparison with the many studies that group individuals in this way; at the same time, it allows for more precise estimation of a curvilinear IQ-offending relationship in the PSM analyses. For the GPS analyses, described below, we use the continuous measure of IQ.

Matching (Control) Variables

For the matching analyses, discussed below, we include measures used in prior research as controls to address confounding that might bias estimates of the IQ and offending relationship. In so doing, we follow Cullen et al. (1997) by using a much wider range of measures than those used by Herrnstein and Murray (1994) as control variables. Below, we describe the measures used here; additional details can be found in Herrnstein and Murray (1994) and Cullen et al. (1997), respectively. Because the focus is on white males, we do not match on either race or sex.

Social class is perhaps the central factor that research suggests must be controlled to obtain a credible estimate of IQ's effect on offending (Menard & Morse, 1984; Hirschi & Hindelang, 1977; Lynam et al., 1993; McGloin et al., 2004). Using the approach detailed by Herrnstein and Murray (1994), we created a standardized index of socioeconomic status (SES) from four items (means and standard deviations in parentheses): parent occupational prestige, derived from a ten-category coding of the Duncan socioeconomic index, per Herrnstein and Murray (1994:574) (5.77, 4.76), average net family income ($24,446, $13,666), father's highest grade completed (11.91, 3.91), and mother's highest grade completed (11.75, 3.14).

In addition to SES, we match on other measures, such as residing in an urban or rural area. Respondents' age in 1979 is included, following Herrnstein and Murray (1994) as a standardized measure; the age range was 14 to 22. We also included measures of whether individuals lived with their mother or their father, respectively, at the age of 14, and a measure of religious participation, ranging from 1 to 6, with higher values indicating more frequent religious participation. Respondents were asked, "How often have you attended religious services?" Response options included: 1=not at all, 2=several times or less during the year, 3=about once a month, 4=two or three times a month, 5=about once a week, 6=more than once a week.

The measure of work ethic derives from the following question: "If, by some chance, you (and your husband/wife) were to get enough money to live comfortably without working, do you think you would work anyway?" An affirmative response was coded as "1=strong work ethic" and a negative response was coded as "0=not a strong work ethic." The study also included a question about academic aspirations and, in particular, the highest grade that respondents expected to complete; responses ranged from fifth grade to 18 or more years of education. And it included a question that tapped into respondents' stakes in conformity: "Do you expect to be working in a job 5 years from now?" Responses were coded as "1=yes" or "0=no."

Locus of control reflects the extent to which respondents reported feeling that they have control over their own circumstances. The NLSY question was: "Many times I feel that I have little influence over the things that happen to me . . . or it is impossible for me to believe that chance or luck plays an important role in my life." Responses were coded as "1=luck has no role" and a "0=luck has a big role." A tendency to rationalize criminal behavior was measured using responses to a question asking respondents whether they would be likely to shoplift if they were otherwise unable to support their family (1=probably would, 0=probably wouldn't).

Methodology

Prior research on IQ and offending has relied primarily on regression-based approaches to addressing confounding. However, as MacDonald et al. (2007:2574) have emphasized, "regression adjustments for potential confounding variables were originally developed by Fisher (1935) to adjust for small imbalances in randomized studies where, by design, the 'treatment' and 'control' distributions of x, the independent variables, are essentially equivalent." When larger imbalances or curvilinear relationships exist, such an approach may not provide credible estimates of the association between a given independent variable and outcome. Propensity score approaches provide an approach that can address this situation by creating balance, where possible, among observed confounders and an outcome (Rosenbaum & Rubin, 1983; Morgan & Harding, 2006). GPS extends the PSM approach by attempting to achieve balance among confounders across levels of an independent variable. At the same time, it imposes no specific functional form on the relationship between this variable and observed outcomes (Hirano & Imbens, 2005; Bia & Mattei, 2008; Loughran, Mulvey, Schubert, Fagan, Piquero, & Losoya 2009; Ertefaie & Stephens, 2010). This last feature is especially useful because it enables any functional form, if one exists, to surface. To test the idea that IQ will be related to offending in a U-shaped manner, with offending highest among individuals in the middle of the IQ spectrum, we employ GPS analyses, using Stata; the IQ measure for this analysis is continuous. We augment these analyses with PSM analyses to illuminate methodological limitations that attend to estimating an IQ and offending association.

Findings

Generalized Propensity Score Analysis

Before presenting results from the GPS analyses, we begin first by examining the bivariate relationship between IQ and offending. Figure 1 presents the mean levels of general offending for each of ten IQ deciles. As inspection of the figure highlights, the relationship is, indeed, curvilinear. Individuals in the middle deciles exhibit higher levels of offending. For these groups, the mean level of the general offending index ranges from approximately .08 (in the third decile, IQ range=88-93) to .05 (in the fourth, fifth, and sixth deciles, IQ=93-109). By contrast, the mean level of offending tapers off quickly to a low of almost -.04 for the lowest decile (IQ=77-83) and of almost -.08 for the highest decile (IQ=12-130). This inverted U-shaped pattern surfaced for the other five offending measures as well (results available upon request).

At first blush, then, there is support for the notion of a curvilinear association between IQ and offending. The bivariate analysis does not, however, address the potential confounding that may result from failing to control for factors such as SES. To address this issue, we employed GPS analyses. The first step consisted of using a set of matching variables to generate propensity scores predicting the likelihood of having a given level of IQ. The model used maximum likelihood (MLE) estimation that, in this instance, generated estimates identical to those produced using ordinary least squares (OLS) estimation (Bia & Mattei, 2008:362). Consistent with prior work (e.g., Cullen et al., 1997), SES, and several other measures (e.g., living with a mother only at age 14, religious participation, academic aspirations), were statistically significant predictors of IQ (see Table 2). The variance explained (adjusted R^2) was 42 percent.

With conventional propensity score matching that involves a binary independent variable, the goal is to create balance on covariates that typically are used as controls in models that predict the outcome of interest (Rosenbaum & Rubin, 1983). GPS proceeds using the same logic. That is, and as developed by Hirano and Imbens (2005), it attempts "to remove all biases in comparisons by treatment status by adjusting for differences in a set of covariates" (Bia & Mattei, 2008:354; see also Ertefaie & Stephens, 2010; Stuart, 2010). To this end, and as with conventional PSM analyses, the next step involved assessing covariate imbalance. Using the Hirano and Imbens (2005) approach, as implemented in Stata (Bia & Mattei, 2004), we created three equal-sized IQ intervals and compared the 11 covariate means for each interval group with the combined covariate means of the other two intervals. This step serves purely to assess covariate balance or imbalance; the subsequent analyses do not consist of three IQ groups.

Inspection of Table 3 indicates that before introducing the propensity score adjustments, 21 of the 33 possible comparisons had t-values exceeding 1.96 (p-value < .05, two-tailed test). The pre-adjustment mean differences in some instances was substantial. For example, prior to adjustment, the mean difference in SES, a standardized measure, for interval 1 as compared to the other intervals was .628 standard deviations. After introducing the adjustments using the GPS procedure (see Bia & Mattei, 2004:357-358), the covariate imbalance was substantially reduced, with only 8 of the 33 comparisons indicating a difference that remained statistically significant. Notably, too, the estimation substantially reduced the magnitude of mean differences in almost all covariates. The last column in Table 3 presents the percentage reduction, after the GPS adjustment, in the covariate mean differences. In a handful of cases, the imbalance worsened, but the mean differences in these cases, prior to adjustment, either were not statistically significant or were substantively trivial. For almost all other covariates, balance was improved. The SES imbalance in interval 1, for example, was reduced from .628 to .174, and the difference in age was entirely eliminated. To be sure, any evidence of imbalance creates cause for concern and is an issue to which we return below. Even so, the GPS adjustments provide greater confidence in the credibility of a comparison of offending across levels of IQ.

We turn now to the third step: modeling the conditional expectation of offending given both

IQ and the propensity score. Hirano and Imbens (2005:82) have emphasized that "there is no direct meaning to the estimated coefficients in [such models]" (see also Bia & Mattei, 2008:359). For this reason, we focus on presenting the predicted probabilities and confidence intervals. The GPS procedure produces averaged predicted values across each level of IQ, with standard errors and confidence intervals for the estimates produced using bootstrap methods (see Bia & Mattei, 2008:359-360). Recall that a key advantage of GPS is that it makes no assumptions about parametric form. The resulting predicted values for general offending, across each one-unit increase in the IQ percentile rank score, are presented in Figure 2. The use of the percentile ranks here facilitates comparison with the subsequent PSM analyses.

What emerges from the GPS analyses is a pattern that mirrors that in the bivariate analysis. On the left side of the curve, as IQ increases, so, too, does the level of offending. Then, among individuals with IQ percentile scores ranging from 30 to 40 percent (standardized IQ scores of 93-99), offending is relatively stable. Thereafter, as IQ increases, the level of offending decreases. More simply, individuals with IQs ranging from 93 to 99 exhibit the highest levels of offending. A similar inverted U-shaped pattern surfaced for the other types of offending and the top decile offending measure used by Herrnstein and Murray (1994) and Cullen et al. (1997).[3]

Propensity Score Matching Analyses

To further investigate the possibility of a curvilinear association between IQ and offending and the related concern about covariate imbalance when employing models that attempt to identify the effect of IQ, we introduce a series of comparisons that attempt to create covariate balance using a more traditional PSM approach. Table 4 presents the results of the analyses, which examine IQ deciles and differences in six types of offending. Three panels are presented.

In the first, Panel A, we focus on each decile and seek matches from any individual from any of the deciles below the given decile. This panel thus examines the type of effect estimated in many prior studies that divide individuals into two groups, such as low IQ and high IQ. Here, however, we systematically examine the extent to which we can obtain matches among the individuals in any given decile to individuals in all deciles below. Any identified effect can be interpreted as the difference in the likelihood of offending among the individuals in the higher-IQ decile with that of matched individuals below that decile.

In the second, Panel B, we present a parallel set of analyses, only here the matches are drawn exclusively from individuals in the two deciles immediately below a given decile. Panel B tests whether any difference in offending can be identified between individuals in a given decile and individuals whose IQs are relatively close to theirs. It thus avoids the problematic assumption that matches are possible from individuals much further down the IQ spectrum.

In the third and final panel, C, the focus shifts—here, we compare individuals in a given IQ decile with matched individuals from the bottom three deciles.[4] The logic here is to identify whether the types of effects identified in Panel A derive primarily from comparisons among individuals who lie far apart or closer together on the IQ distribution.

For each comparison, we created propensity scores, using the matching variables that are listed in Table 1 and that were employed in the GPS analyses, to predict the likelihood of belonging to a given decile versus the alternative.[5] Across all of the estimation models, there consistently was imbalance in the covariates when comparing the higher IQ group with the lower IQ group. After employing one-to-one nearest-neighbor matching, with the caliper set to .01 and not allowing for replacement—a conservative matching strategy used in prior studies (see, e.g., Massoglia, 2008; Mears & Bales, 2009; see, generally, Apel & Sweeten, 2010)—we obtained balance on all covariates.[6] However, as reported in the bottom row of each panel, this result came at the cost of losing some cases to off-support—that is, few matches could be found in some cases, thus delimiting the generalizability of the results in those instances.

Inspection of Table 4 highlights several noteworthy patterns. First, and ignoring, for the moment, tests of statistical significance, a curvilinear relationship appears to be evident.

For example, in Panel A, individuals in decile 2 are more likely to engage in general offending as compared to individuals in decile 1; individuals in decile 3 are more likely to engage in general offending as compared to individuals in deciles 1 and 2; individuals in decile 4 are more likely to engage in general offending as compared to individuals in deciles 1 through 3;

and so on, until decile 7 (standardized IQ scores=109-114). At that point, individuals become less likely to offend compared to individuals in all lower IQ deciles. By decile 10, the lower likelihood of offending is more pronounced; the reduction in the standardized offending scale is -.13. This pattern generally holds across all six types of offending and accords with what one would anticipate if the relationship between IQ and offending were curvilinear in the manner depicted in Figures 1 and 2. That is, as one moves from the lowest IQ levels to higher IQ levels, the likelihood of offending increases; then, approximately near the middle of the IQ distribution, the likelihood of offending decreases as one moves to the highest IQ levels. The pattern is not perfectly uniform but suggests support for a curvilinear relationship between IQ and offending.

Second, many of the comparisons are not statistically significant at conventional levels. For example, in Panel A, across 54 comparisons, only 10 are significant. The bulk of the differences come from the ninth and tenth deciles, which derived from comparing the offending of individuals in these deciles with all deciles below them, respectively. In Panel B, no statistically significant effects surface, and, in Panel C, only one statistically significant effect surfaces. Notably, however, the pattern of effects is similar across all three panels and each of the six types of offending. Even so, the lack of statistical significance in Panels B and C indicates that the IQ-offending association—when employing the immediately lower deciles and the bottom deciles as the basis for deriving counterfactual estimates—is not as strong as prior research suggests or, indeed, does not exist at all. For the first panel, if Bonferroni adjustments were used, then, with the 54 comparisons, none of the effects would be statistically significant. The Bonferroni adjustment is, however, a conservative approach. In addition, the magnitude of the differences, especially the consistent pattern for deciles 9 and 10, suggest that the highest IQ groups indeed have lower levels of offending as compared to individuals in lower IQ groups.

Third, estimates of IQ effects on offending that compare an extreme group, such as the highest IQ individuals, with all other individuals likely depend on comparisons from those individuals with IQs more proximate to them. Contrasting the effects across the different panels helps to illuminate this point. In Panel A, the strongest effects are identified among individuals in the highest IQ decile. However, in Panel B, there is minimal evidence of such an effect, suggesting that, when making comparisons between individuals in this IQ decile and those immediately below, there is in fact little difference in offending. In Panel C, matches were forced to come only from individuals in the bottom three deciles. Here, however, 74 percent of the individuals in the highest IQ decile (standardized IQ scores=125-130) could not be matched. By implication, then, the matching in Panel A, where almost no cases were lost off-support, derived from cases in the middle to upper end of the IQ spectrum. This finding highlights that comparisons of individuals at the opposite ends of the IQ spectrum likely can not be credibly made because of the inability to find appropriate matches in such instances.

To illustrate the reason for the inability at the highest deciles to identify matched counterparts at the lowest three deciles (those with IQs ranging from 77-93), we plot, in Figure 3, the bivariate relationship between IQ and SES. As inspection of the figure indicates, few individuals in the lowest IQ deciles were above the mean of the SES measure; at the same time, very few individuals in the highest IQ deciles were below the mean of this measure. It is this lack of overlap, or variance, that largely accounts for the inability to identify matches from individuals among the lowest IQ deciles to individuals with higher IQs. In short, although scholars have highlighted the importance of controlling for SES when examining IQ, to a large extent high IQ and low IQ individuals do not appear to be sufficiently similar with respect to SES to allow the effect of IQ to be isolated net of SES (see, generally, Dennis, Francis, Cirino, Schachar, Barnes, & Fletcher, 2009).

Conclusion

The goal of this paper was to examine the claim that an IQ-offending relationship exists and, in so doing, to shed light on several issues that complicate research on the relationship. To address limitations in prior work, this study used data from the NLSY, six measures of self-reported offending, continuous and categorical approaches to coding IQ, and generalized propensity score and propensity score matching analyses to take into account potential confounding and, more specifically, potential imbalance on covariates that have been used as control variables in prior work. Because the study focused on white males, caution should be exercised in generalizing the results to other populations. Briefly, several key findings emerged. First, both in bivariate analyses and in the GPS and PSM analyses, there was evidence of a curvilinear, inverted U-shaped relationship. Specifically, and regardless of the type of offending, individuals in the lowest and highest levels of the IQ continuum were less likely to commit crime as compared to individuals near the middle of the continuum.

Second, the PSM analyses produced different results depending on the groups used for identifying matches. For example, matching individuals in a given IQ decile with those from all lower IQ deciles was more likely to produce statistically significant differences in offending. Matching only to individuals in the deciles immediately below a given IQ decile produced few statistically significant results. Notably, this comparison may be the most credible because individuals in the immediately lower IQ percentiles may be more similar in other respects to those in the given decile, thus reducing the concern that the comparison fails to address important confounding. Matching high-IQ individuals to individuals at the lower end of the IQ distribution generally was not possible, suggesting that regression-based approaches to estimating IQ effects of individuals at the opposite ends of the IQ distribution likely cannot control for confounding because there is little overlap among the two groups in the confounders.

These results suggest several implications for scholarship. A central implication is the reinforcement of what prior research has demonstrated—as a general matter, IQ and offending are associated. However, this study consistently found evidence that the relationship is curvilinear. That suggests that prior research may have consistently misestimated the relationship and, in so doing, potentially understated the extent of association between IQ and offending. Consistent with the large bulk of prior research, we used a general measure of intelligence, but compelling arguments exist that intelligence may be best conceptualized as multidimensional (Sternberg & Kaufman, 2011). In the end, a larger body of studies, using other measures of intelligence and of offending, will be needed to determine if the IQ-offending relationship is curvilinear in the manner identified in this study. Further, research comparing data sets from different years may identify if the IQ-offending relationship varies over time.

The findings here suggest that a modest IQ-offending relationship holds regardless of how offending is measured or how IQ is coded. At the same time, both the GPS and PSM models raise questions about interpreting prior research. For example, a central criticism of studies of IQ and offending is that they may not address potential confounding. One solution is to control for potential confounders. However, that approach only works if the confounder is sufficiently distributed across levels of the variable of interest, in this case IQ (Oakes & Johnson, 2006; Dennis et al., 2009). In the present study, it is apparent that SES, an important potential confounder, is not evenly distributed across levels of IQ. Indeed, when comparing individuals in the lowest and highest deciles of the IQ spectrum, there was little overlap with respect to SES.

PSM approaches cannot solve this problem, but they can highlight that it exists. For example, when conducting PSM analyses, it is possible to assess common support—that is, the extent to which, after matching, two groups (e.g., a treatment group and comparison group) are similar with respect to observed characteristics. When the two groups differ, cases that cannot be matched are not included in the comparisons; the generalizability of the results thus is circumscribed. By contrast, and as Apel and Sweeten (2010:547) have emphasized, "standard regression . . . obscures this issue, and can, in some instances, extrapolate treatment effect estimates . . . when treated and untreated groups are actually not comparable at all." Based on the present study, this concern appears to be relevant to prior research on the IQ and offending relationship. It seems unlikely, given the results here and elsewhere, that a null relationship would be identified if confounding were better addressed. That said, given that many studies identify relatively small effects of IQ on offending and also that some studies identify no association, it seems conceivable that careful attention to this issue might result in research that identified more equivocal results than what has been suggested in reviews.

Finally, some scholars (e.g., McGloin et al., 2004), have emphasized a need for developing and testing theoretical arguments for an IQ effect on offending. We echo this call, especially given that the results of this study suggest warrant for the argument that IQ and offending are curvilinearly related. To date, the main explanation has been to suggest that some individuals have IQs sufficiently low as to render them incapable of offending. This argument does not strike us as especially compelling. Individuals in the NLSY were sufficiently intelligent to be able to participate in the study, and the identified effect emerged not just for individuals in the bottom decile or two of the IQ distribution. Among individuals at the upper end of the IQ spectrum, there may be a different explanation of offending that is needed. For these individuals, especially in contrast to individuals in the middle end of the IQ spectrum, there may be substantially greater opportunities for financial gain and social status (DiRago & Vaillant, 2007) and, at the same time, more costs associated with offending. If so, then, per strain and rational choice theories, we would anticipate less offending among this population. For individuals with middle-range IQs, there may be requisite levels of intelligence to engage in offending coupled with fewer costs, relative to higher-IQ individuals, to engaging in crime. There also is the possibility that moral reasoning is curvilinearly related to intelligence and that it in turn is what produces the observed curvilinear relationship between IQ and offending (Langdon et al., 2011). Such possibilities suggest the potential avenues along which a more theoretically nuanced understanding of the IQ-offending relationship might proceed.

Notes

[1] In many accounts, scholars report that self-report studies consistently find a negative association between IQ and offending; in making this claim, they cite the same reviews (e.g., West & Farrington, 1973; Hirschi & Gottfredson, 1977) or sources that draw on these reviews (e.g., Wilson & Herrnstein, 1985; Herrnstein & Murray, 1994). Hirschi and Hindelang (1977:576) aver that "most studies do find a relation between IQ and self-reported delinquency" but cite only two studies and their reanalysis of data that Hirschi collected for his previous study (Hirschi, 1969). In recent years, several studies have indeed investigated IQ and offending using self-report data and more rigorous methodologies (see, generally, Loeber et al. 2003; McGloin et al., 2004; Farrington & Welsh, 2007). However, the body of work is, as Herrnstein and Murray (1994) and McGloin et al. (2004) have emphasized, less robust than some reviews suggest.

[2] An extensive discussion of the AFQT and, more generally, of IQ and research on it can be found in Herrnstein and Murray (1994), especially pages 118-120 and the appendices.

[3] Because the pattern across all measures of offending were similar and because of space constraints, we have reported only the results for general offending. The results, however, for these other measures of offending are available upon request.

[4] Using the bottom three deciles was necessary to create a large enough pool of individuals from which to identify potential matches. For panel B, using the two deciles below a given decile was sufficient for identifying matches, reflecting the fact that the two groups typically were more similar than were the two groups in the panel C analyses.

[5] Two exceptions noted in the table are as follows: in panel B, decile 2 is matched to decile 1, and, in panel C, decile 3 is matched to deciles 1 and 2 and decile 2 is matched to decile 1.

[6] Because of the large number of PSM analyses and space constraints, it is not possible to present each predictive model and the balance statistics. These are all available upon request.

References

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

Bia, M., & Mattei, A. (2008). A Stata package for the estimation of the dose-response function through adjustment for the generalized propensity score. *The Stata Journal*, *8*, 354-373.

Bureau of Labor Statistics. (2011). *National Longitudinal Survey of Youth, 1979 cohort: Aptitude, achievement, and intelligence scores*. Washington, D.C.: Bureau of Labor Statistics. Available on-line: www.nlsinfo.org/nlsy79/docs/79html/79text/achtests.htm.

Cullen, F. T., Gendreau, P., Jarjoura, G. R., & Wright, J. P. (1997). Crime and the bell curve: Lessons from intelligent criminology. *Crime & Delinquency*, *43*, 387-411.

Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., & Fletcher, J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *Journal of the International Neuropsychological Society*, *15*, 331-343.

DiRago, A. C., & Vaillant, G. E. (2007). Resilience in inner city youth: Childhood predictors of occupational status across the lifespan. *Journal of Youth and Adolescence*, *36*, 61-70.

Ellis, L., & Walsh, A. (2003). Crime, delinquency, and intelligence: A review of the worldwide literature. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen* (pp. 343-365). Oxford, England: Elsevier.

Ertefaie, A., & Stephens, D. A. (2010). Comparing approaches to causal inference for longitudinal data: Inverse probability weighting versus propensity scores. *The International Journal of Biostatistics*, *6*, article 14, 14-22.

Farrington, D. P. (1993). Childhood origins of teenage antisocial behavior and adult social dysfunction. *Journal of the Royal Society of Medicine*, *86*, 13-17.

Farrington, D. P., & Welsh, B. C. (2007). *Saving children from a life of crime: Early risk factors and effective interventions*. New York: Oxford.

Felson, R. B., & Staff, J. (2006). Explaining the academic performance-delinquency relationship. *Criminology*, *44*, 299-320.

Fisher, R. A. (1935). *The design of experiments*. Edinburgh: Oliver and Boyd.

Goddard, H. H. (1914). *Feeblemindedness: Its causes and consequences*. New York: Macmillan.

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

Herrnstein, R. J. (1995). Criminogenic traits. In J. Q. Wilson & J. Petersilia (Eds.), *Crime* (pp. 39-64). San Francisco: ICS Press.

Herrnstein, R. J., & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Free Press.

Hirano, K., & Imbens, G. W. (2005). The propensity score with continuous treatments. In A. Gelman & X.-L. Meng (Eds.), *Applied Bayesian modeling and causal inference from incomplete-data perspectives* (pp. 73-84). West Sussex, England: Wiley.

Hirschi, T., & Hindelang, M. J. (1977). Intelligence and delinquency: A revisionist review. *American Sociological Review*, *42*, 571-587.

Hirschi, T. (1969). *Causes of delinquency*. Berkeley: University of California Press.

Isen, J. (2010). A meta-analytic assessment of Wechsler's P > V sign in antisocial populations. *Clinical Psychology Review*, *30*, 423-435.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.

Jolliffe, D., & Farrington, D. P. (2010). Individual differences and offending. In T. Newburn (Ed.), *The Sage handbook of criminological theory* (pp. 40-55). Thousand Oaks, CA: Sage.

Langdon, P. E., Clare, I. C. H., & Murphy, G. H. (2011). Moral reasoning theory and illegal behaviour by adults with intellectual disabilities. *Psychology, Crime & Law*, *17*, 101-115.

Lindsay, W. R., & Taylor, J. L. (2010). Mentally disordered offenders: Intellectual disability. In G. J. Towl & D. A. Crighton (Eds.), *Forensic psychology* (pp. 285-308). Malden, MA: Blackwell.

Loughran, T. A., Mulvey, E. P., Schubert, C. A., Fagan, J., Piquero, A. R., & Losoya, S. H. (2009). Estimating a dose-response relationship between length of stay and future recidivism in serious juvenile offenders. *Criminology*, *47*, 699-740.

Lynam, D., Moffitt, T., & Stouthamer-Loeber, M. (1993). Explaining the relation between IQ and delinquency: Class, race, test motivation, school failure, or self-control? *Journal of Abnormal Psychology*, *102*, 187-196.

MacDonald, J., Stokes, R. J., Ridgeway, G., & Riley, K. J. (2007). Race, neighbourhood context and perceptions of injustice by the police in Cincinnati. *Urban Studies*, *44*, 2567-2585.

Massoglia, M. (2008). Incarceration, health, and racial disparities in health. *Law & Society Review*, *42*, 275-306.

McCord, J., & Ensminger, M. E. (1997). Multiple risks and comorbidity in an African-American population. *Criminal Behaviour & Mental Health*, *7*, 339-352.

McGloin, J. M., Pratt, T. C., & Maahs, J. (2004). Rethinking the IQ-delinquency relationship: A longitudinal analysis of multiple theoretical models. *Justice Quarterly*, *21*, 603-635.

Mears, D. P., & Bales, W. D. (2009). Supermax incarceration and recidivism. *Criminology*, *47*, 1131-1166.

Menard, S., & Morse, B. J. (1984). A structuralist critique of the IQ-delinquency hypothesis: Theory and evidence. *American Journal of Sociology*, *89*, 1347-1378.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*, 674-701.

Moffitt, T. E., Caspi, A., Silva, P. A., & Stouthamer-Loeber, M. (1995). Individual differences in personality and intelligence are linked to crime: Cross-context evidence from nations, neighborhoods, genders, races, and age-cohorts. *Current Perspectives on Aging and the Life Cycle*, *4*, 1-34

Moffitt, T. E., & Silva, P. A. (1988). IQ and delinquency: A direct test of the differential detection hypothesis. *Journal of Abnormal Psychology*, *97*, 330-333.

Morgan, S. L., & Harding, D. J. (2006). Matching estimators of causal effects prospects and pitfalls in theory and practice. *Sociological Methods & Research*, *35*, 3-60.

Nyborg, H., ed. (2003). *The scientific study of general intelligence: Tribute to Arthur Jensen*. Oxford, England: Elsevier.

Oakes, J. M., & Johnson, P. J. (2006). Propensity score matching for social epidemiology. In J. M. Oakes & J. Kaufman (Eds.), *Methods in social epidemiology* (pp. 364-386). San Francisco: Jossey-Bass.

Rindermann, H. (2007). The g-factor of international cognitive ability comparisons: The homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations. *European Journal of Personality*, *21*, 667-706.

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

Sternberg, R. J., & Kaufman, S. B., eds. (2011). *The Cambridge handbook of intelligence*. New York: Cambridge University Press.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, *25*, 1-21.

Sweeten, G. (2012). Scaling Criminal Offending. *Journal of Quantitative Criminology*, *28*, 533-557.

Walsh, A. (2011). *Social class and crime: A biosocial approach*. New York: Routledge.

Ward, D. A., & Tittle, C. R. (1994). IQ and delinquency: A test of two competing explanations. *Journal of Quantitative Criminology*, *10*, 189-212.

West, D. J. & Farrington, D. P. (1973). *Who becomes delinquent?* London, England: Heinemann.

Willis, J. O., Dumont, R., & Kaufman, A. S. (2011). Factor-analytic models of intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.) *The Cambridge handbook of intelligence* (pp. 39-57). New York: Cambridge University Press.

Wilson, J. Q., & R. J. Herrnstein. (1985). *Crime and human nature*. New York: Simon and

Schuster.

Zagorsky, J. L. (2007). Do you have to be smart to be rich? The impact of IQ on wealth, income and financial distress. *Intelligence*, *35*, 489-501.

Table 1. Descriptive Statistics

	Mean	S.D.	Min.	Max.
Dependent variables				
General offending (z)	0.00	0.58	-0.59	4.58
Violent offending (z)	-0.01	0.77	-0.51	6.12
Property offending (z)	0.01	0.66	-0.43	5.45
Drug offending (z)	-0.01	0.82	-0.45	4.95
Chronic offending (1=yes, 0=no)	0.28	0.45	0.00	1.00
Diversity of offending (z)	0.01	0.54	-0.48	2.57
Independent Variable				
IQ (AFQT) (Full Sample)—Percentile Rank	50.22	28.74	0.05	100.00
IQ (AFQT) (Full Sample)—Standardized	100.00	15.00	77.84	130.01
IQ (AFQT) (Analysis Sample)—Percentile Rank	56.39	28.39	0.09	100.00
IQ (AFQT) (Analysis Sample)—Standardized	105.49	14.82	77.86	130.01
Matching (Control) Variables				
SES (z)	0.39	0.72	-3.34	2.24
Age (z)	-0.03	1.03	-1.67	1.80
Lived with mother at age 14 (1=yes, 0=no)	0.95	0.24	0.00	1.00
Lived with father at age 14 (1=yes, 0=no)	0.84	0.40	0.00	1.00
Religious participation (1=never, 6=1+ per week)	3.00	1.68	1.00	6.00
Strong work ethic (1=yes, 0=no)	0.84	0.35	0.00	1.00
Academic aspirations (grade level)	14.05	2.46	5.00	18.00
Stakes in conformity (1=yes, 0=no)	0.94	0.22	0.00	1.00
Locus of control (1=external, 0=internal)	0.58	0.49	0.00	1.00
Rationalization favoring crime (1=yes, 0=no)	0.03	0.18	0.00	1.00
Live in urban area (1=urban, 0=rural)	0.76	0.44	0.00	1.00

Figure 1. Bivariate Relationship Between IQ and Offending

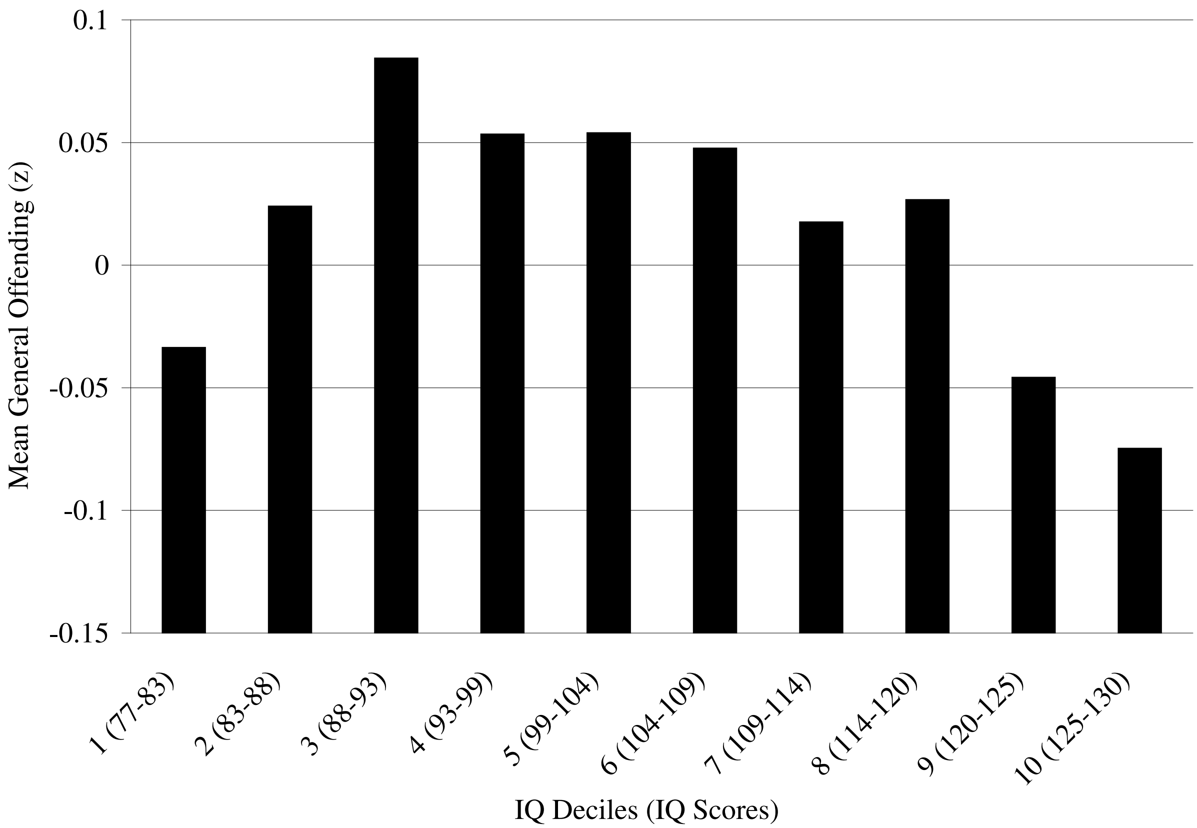

Table 2. Generalized Propensity Score IQ Prediction Model

	b	S.E.
SES (z)	8.586***	0.605
Age (z)	0.259	0.376
Lived with mother at age 14	-1.595***	0.866
Lived with father at age 14	0.541*	0.231
Religious participation	5.221*	1.554
Strong work ethic	1.893*	0.956
Academic aspirations	-2.277***	1.061
Stakes in conformity	-0.370	1.729
Locus of control	5.696	0.181
Rationalization favoring crime	1.350	0.753
Live in urban area	1.059	2.130
Constant	-33.272***	3.575
Adjusted R2 = 0.42		

* p < .05 ** p < .01 *** p < .001

Figure 2. Generalized Propensity Score Estimation of the IQ-Offending Relationship

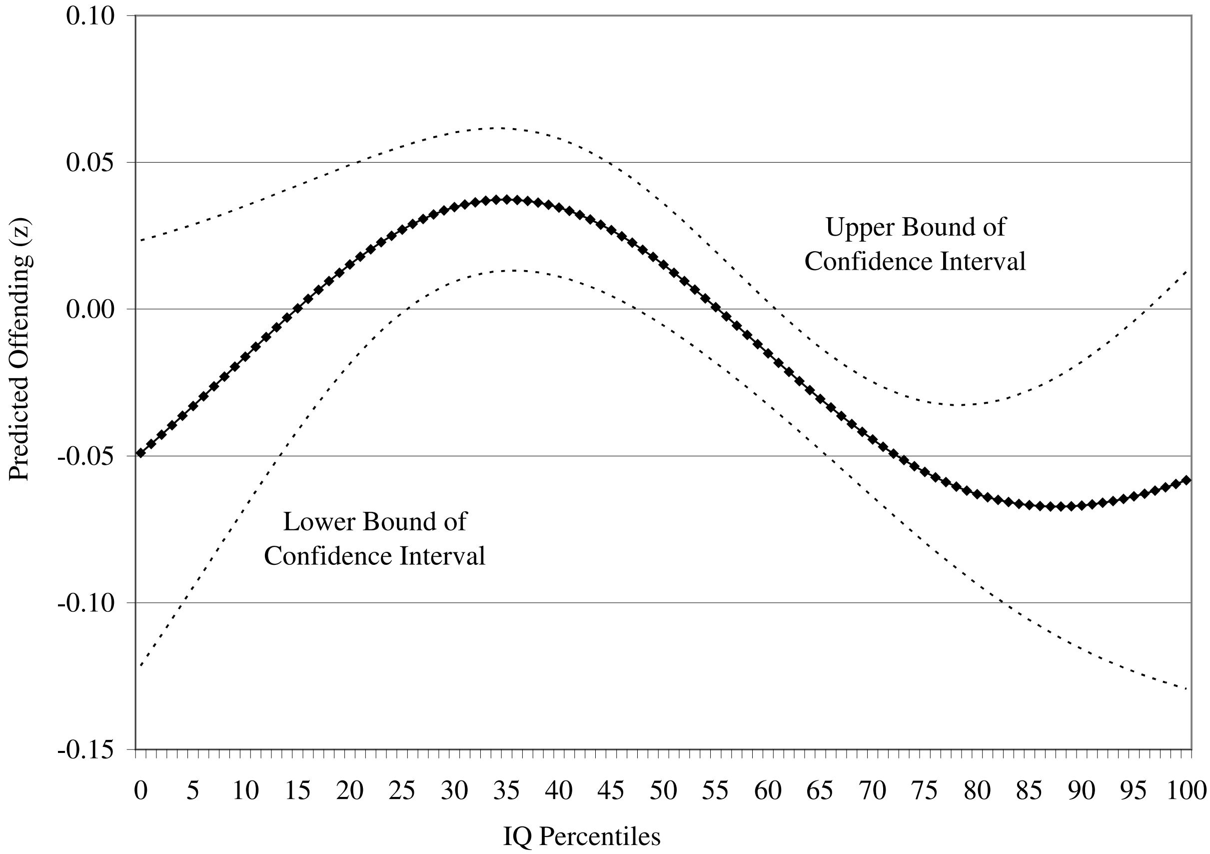

Table 3. Generalized Propensity Score Adjustment Statistics: Mean Differences in Covariates Between IQ Intervals

	Pre-adjustment Mean Diff.	Pre-adjustment t-value	Post-adjustment Mean Diff.	Post-adjustment t-value	Reduction in Diff. (%)
IQ Interval 1					
SES (z)	0.628	24.53	0.174	6.49	-72.2
Age (z)	0.176	4.42	-0.006	-0.12	-103.3
Lived with mother at age 14	0.047	5.06	0.015	1.39	-67.9
Lived with father at age 14	0.103	6.71	0.024	1.34	-76.7
Religious participation	0.477	7.34	0.136	1.72	-71.5
Strong work ethic	-0.003	-0.20	-0.006	-0.38	130.1
Academic aspirations	2.637	31.64	0.775	11.83	-70.6
Stakes in conformity	-0.033	-3.87	-0.015	-1.33	-54.3
Locus of control	0.058	3.04	0.007	0.29	-88.4
Rationalization favoring crime	-0.001	-0.18	0.000	-0.05	-65.1
Live in urban area	0.060	3.52	0.005	0.27	-90.9
IQ Interval 2					
SES (z)	-0.002	-0.06	-0.070	-2.54	3,982.8
Age (z)	-0.055	-1.42	-0.060	-1.55	9.6
Lived with mother at age 14	-0.007	-0.77	-0.009	-0.93	23.2
Lived with father at age 14	-0.008	-0.52	-0.013	-0.83	62.0
Religious participation	0.028	0.45	-0.045	-0.72	-259.7
Strong work ethic	0.031	2.31	0.031	2.37	3.2
Academic aspirations	0.242	2.63	-0.191	-2.05	-179.2
Stakes in conformity	-0.015	-1.82	0.001	0.11	-106.1
Locus of control	0.026	1.41	0.014	0.74	-47.3
Rationalization favoring crime	-0.002	-0.33	0.001	0.15	-145.1
Live in urban area	-0.003	-0.21	-0.009	-0.57	167.9
IQ Interval 3					
SES (z)	-0.542	-22.47	-0.124	-5.11	-77.1
Age (z)	-0.101	-2.73	0.029	0.66	-128.2
Lived with mother at age 14	-0.034	-3.95	-0.008	-0.75	-76.6
Lived with father at age 14	-0.082	-5.73	-0.027	-1.56	-67.3
Religious participation	-0.439	-7.27	-0.142	-2.02	-67.7
Strong work ethic	-0.026	-2.05	-0.024	-1.59	-7.5
Academic aspirations	-2.507	-32.56	-0.354	-7.33	-85.9
Stakes in conformity	0.043	5.37	0.001	0.14	-97.1
Locus of control	-0.075	-4.20	-0.025	-1.19	-66.6
Rationalization favoring crime	0.003	0.49	-0.001	-0.18	-145.3
Live in urban area	-0.048	-3.06	0.007	0.37	-114.5

Table 4. Propensity Score Analysis Estimation of the IQ-Offending Relationship

	IQ Deciles (and IQ Scores)								
	2nd (83-88)	3rd (88-93)	4th (93-99)	5th (99-104)	6th (104-109)	7th (109-114)	8th (114-120)	9th 120-125)	10th (125-130)
A. Matching to all lower deciles[a]									
Diff. in offending (z)	0.08	0.09	0.08	0.06	0.05	-0.04	0.02	-0.10*	-0.13*
Diff. in violent offending (z)	0.22	0.03	0.06	0.04	-0.04	-0.11	-0.01	-0.14*	-0.19*
Diff. in property offending (z)	-0.02	0.16	0.09	0.10	0.04	-0.05	0.02	-0.12*	-0.11*
Diff. in drug offending (z)	0.10	-0.01	0.14	-0.03	0.12	0.08	0.05	-0.07	-0.09
Diff. in chronic offending (1/0)	0.07	0.04	0.03	0.01	0.10*	0.02	-0.01	-0.06	-0.11*
Diff. in diversity of offending (z)	0.06	0.05	0.04	0.05	0.01	-0.02	0.03	-0.08*	-0.14*
Percent of cases lost to off-support	33%	9%	9%	4%	2%	2%	1%	3%	2%
B. Matching to next 2 lower deciles[a]									
Diff. in offending (z)	0.08	0.09	0.03	-0.03	-0.04	-0.04	0.01	-0.07	-0.04
Diff. in violent offending (z)	0.22	0.03	-0.10	-0.02	-0.07	-0.11	0.01	-0.02	-0.06
Diff. in property offending (z)	-0.02	0.16	0.05	-0.04	-0.04	-0.04	0.03	-0.09	-0.01
Diff. in drug offending (z)	0.10	-0.01	0.07	-0.04	0.03	0.03	-0.02	-0.09	-0.04
Diff. in chronic offending (1/0)	0.07	0.04	-0.02	-0.02	0.06	-0.01	-0.06	0.01	-0.06
Diff. in diversity of offending (z)	0.06	0.05	-0.01	-0.05	0.00	-0.03	0.04	-0.07	-0.05
Percent of cases lost to off-support	33%	9%	10%	4%	4%	3%	3%	7%	7%
C. Matching to bottom 3 deciles[b]									
Diff. in offending (z)	0.08	0.09	0.08	0.02	-0.05	-0.13*	0.08	-0.09	-0.10
Diff. in violent offending (z)	0.22	0.03	0.06	0.01	-0.11	-0.13	0.03	0.00	-0.10
Diff. in property offending (z)	-0.02	0.16	0.09	0.02	0.07	-0.12	0.08	-0.12	-0.08
Diff. in drug offending (z)	0.10	-0.01	0.14	0.04	0.03	-0.08	0.09	-0.13	-0.12
Diff. in chronic offending (1/0)	0.07	0.04	0.03	0.01	0.03	0.00	0.01	-0.07	-0.05
Diff. in diversity of offending (z)	0.06	0.05	0.04	-0.01	-0.04	-0.11	0.08	-0.11	-0.13
Percent of cases lost to off-support	33%	9%	9%	10%	26%	30%	46%	56%	74%

* = $p < .05$ (two-tailed)

a. Exception: Decile 2 is matched to decile 1.

b. Exceptions: Decile 3 is matched to deciles 1 and 2; decile 2 is matched to decile 1.

Figure 3. Bivariate Relationship Between IQ and SES