Spring 2007

Causal Effect of Corporal Punishment on Children's Internalizing and Externalizing Behavioral Outcomes: Results from a Propensity Score Matching Analysis

Sara Zane Morris

Follow this and additional works at: https://digitalcommons.georgiasouthern.edu/etd

THE CAUSAL EFFECT OF CORPORAL PUNISHMENT ON CHILDREN'S

INTERNALIZING AND EXTERNALIZING BEHAVIORAL OUTCOMES: RESULTS

FROM A PROPENSITY SCORE MATCHING ANALYSIS

by

SARA Z. MORRIS

(Under the Direction of Chris L. Gibson)

ABSTRACT

Substantial research supports the association between corporal punishment and children's behavioral problems, suggesting that punishment leads to both internalizing and externalizing behaviors. However, this research has not been able to discern the direction of causality in this relationship for many reasons, including a lack of experimental data. The purpose of this study is to assess the causal effect of corporal punishment on internalizing and externalizing behaviors by using a Propensity Score Matching (PSM) analysis. This study uses data from waves 1 and 2 of the longitudinal cohort study of the Project on Human Development in Chicago Neighborhoods (PHDCN), including 6, 9, 12, and 15 year old cohorts. Results reveal two major findings as they relate to the effects of corporal punishment on children's behavioral outcomes. First, results show that a large amount of selection bias exists before matching occurred. Second, once matching was implemented no causal effect remained for corporal punishment on the proposed outcome measures. Limitations and ways of overcoming them in future research are also discussed.

INDEX WORDS:. Corporal punishment, Propensity score matching, Internalizing behavior, Externalizing behavior

THE CAUSAL EFFECT OF CORPORAL PUNISHMENT ON CHILDREN'S

INTERNALIZING AND EXTERNALIZING BEHAVIORAL OUTCOMES: RESULTS

FROM A PROPENSITY SCORE MATCHING ANALYSIS

by

SARA Z. MORRIS

B.A., Emory University, 2004

A Thesis Submitted to the Graduate Faculty of Georgia Southern University in Partial

Fulfillment of the Requirements for the Degree

MASTER OF ARTS

STATESBORO, GEORGIA

2007

THE CAUSAL EFFECT OF CORPORAL PUNISHMENT ON CHILDREN'S

INTERNALIZING AND EXTERNALIZING BEHAVIORAL OUTCOMES: RESULTS

FROM A PROPENSITY SCORE MATCHING ANALYSIS

by

SARA Z. MORRIS

Major Professor:	Chris L. Gibson
Committee:	William L. Smith
	Pidi Zhang

Electronic Version Approved:

May 2007

DEDICATION

I would like to dedicate this thesis to Ms. Suzanne Clonts. She was the first teacher (but fortunately not the last) to show me what the true meaning of the word "teacher" is. She inspired me to teach others, an inspiration I still have fifteen years after being in her class. She provided the guidance, support, and caring that I desperately needed at a very scary time in my life and I will always be indebted to her for that.

ACKNOWLEDGEMENTS

There are many people that I owe thanks and acknowledgement to for help throughout the process of writing this thesis. First and foremost, I must thank my parents and my three siblings, all of whom have had to hear me ramble, complain, and sometimes gush about this research. I could never have finished it without their love and support.

Second, I would like to acknowledge and thank Dr. Chris Gibson for everything he has done for me in the past two years. For being a mentor, a scholarly resource, a teacher, and most of all a friend—I thank you. Your support during the conceptualization, research, and writing of this thesis was invaluable. You have taught me many things over the past two years, but I am most thankful for the opportunity you have given me to experience "academia" as a profession, without which I might have made a different career choice. You have believed in me even when I wasn't sure I believed in myself, and for that I will always be grateful.

I would also like to acknowledge my other committee members, Dr. William L. Smith and Dr. Pidi Zhang, for their help and input during this process. I appreciate the time and energy you both put into helping me with this project as well as your instruction in classes over the past two years. I would also like to acknowledge the help of Dr. Tamara Leech from Butler University, who provided important information and resources pertaining to the methodology used in this thesis.

Thank you to the rest of the Sociology and Anthropology department for supporting (and employing!) me for the past two years. My time at Georgia Southern has been a great experience, and I hope to stay in touch with many of you in the future as colleagues. In particular, the guidance, education, and friendship of Dr. Danny Dixon

and Dr. Peggy Hargis were essential in completing this degree. Last, but certainly not least, thank you to my fellow graduate students, especially Laura Swainston. I definitely would not have been able to finish this thesis without advice and support from you. Most of all, I appreciate you being there to listen.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The use of corporal punishment as a discipline strategy is a hotly debated topic in the United States. Corporal punishment is defined as "the use of physical force with the intention of causing a child to experience pain, but not injury, for the purpose of correction or control of the child's behavior" (Straus, 1994: 4). Views range from a belief that corporal punishment is normal and acceptable to the belief that any type of physical punishment is abuse. Regardless of these views, 94% of parents use some type of corporal punishment to discipline their toddlers. Forty percent of parents still use corporal punishment to discipline their children as old as age 13 (Straus & Stewart, 1999). In addition to this evidence, a substantial number of caregivers report using corporal punishment on children as young as six months (Slade & Wissow, 2004).

A two-sided debate exists among social scientists regarding the effectiveness and appropriateness of corporal punishment by parents. One group agrees with Murray Straus (1994) when he says, "Used by authority figures who tend to be loved or respected as a way to achieve a morally correct end, it [corporal punishment] carries a powerful message aside from the immediate effect intended. The message is that if someone is doing something outrageous and other methods of getting the person to listen to reason have failed, it is ok to use physical violence (pg. 9)." These researchers believe that corporal punishment in any form is harmful and wrong, and many have published studies with evidence supporting this (Straus, Sugarman, & Giles-Sims, 1997). Others argue that under some circumstances (i.e. when mild forms are used to back up other disciplinary practices); corporal punishment can enhance positive child outcomes (Larzelere, 2000).

Since many parents feel that corporal punishment is an effective or appropriate way to discipline their children, it is important to understand what effects it may have on children. Studies show that several negative outcomes are associated with corporal punishment—for example, antisocial behavior (Grogan-Kaylor, 2005a), depressive symptoms (Straus & Kantor, 1994; Turner & Muller, 2004), suicidal ideation (Straus & Kantor, 1994), behavior problems (Brennar & Fox, 1998), and psychological distress (Turner & Finkelhor, 1996). However, most people in this country are not aware of these effects. Because of this and the fact that it is so widespread and accepted, corporal punishment is seen as a normal and almost expected way to discipline children. Caregivers may not realize the other ways (besides as an impetus to change behavior) it may be affecting children. Many parents assume that corporal punishment is an effective form of disciplining children (Straus, 2001).

McCord (2005) discusses the unintended consequences that often occur when children are punished. She postulates that if children grow up thinking that their parents are good, and then are punished by their parents, there are only bad assumptions that can stem from this. Two scenarios are possible. The first is that children believe that punishment is warranted and it is therefore okay to cause pain to another. The other is that children believe even though it is wrong to cause pain, it is alright to do something wrong in some instances. Regardless of which of these occurs, a child who is physically punished begins to believe that causing physical pain is acceptable in some circumstances.

Another unintended consequence is that children learn that behaving badly is wrong when they are caught, but not always. Since most children are not caught for

every wrong thing they do, and are only punished when they are caught, they learn that as long as they are not seen by parents or authority figures those bad behaviors happen without any repercussions. In contrast, if they were taught not to behave in a certain way and the reasons why, they might be more likely to understand that behavior is never acceptable (McCord, 2005).

In addition to these, a third unintended consequence of punishment is that it makes the bad behavior more attractive. Children who are told not to do something "because you will get in trouble" are more likely to display that behavior than if they are told not to do something for a legitimate reason. When a child learns the reasons behind why they should not engage in a certain behavior they can understand the rationale and are more likely to not exhibit the behavior. Children who are physically punished learn that the consequence for bad behavior is the physical punishment and therefore may still commit these behaviors because they do not understand why it is wrong (McCord, 2005).

Many parents are not aware of these unintended consequences listed above, or others that may be present. Given this and the high prevalence rates discussed above, it is crucial that social scientists continue to study the effects corporal punishment has on children. The research that exists now is varied and lengthy, but a definitive answer has still not been found as to whether or not this type of discipline is harmful or beneficial. This study will attempt to add another piece of evidence towards finding this answer.

Although the studies noted above, as well as others (Simons, Wu, Lin, Gordon, & Conger, 2000; Grogan-Kaylor, 2005b) have described the detrimental outcomes that are associated with corporal punishment, many of these studies share a similar methodological limitation. That is, most are unable to discern direction of causality. Are

children who exhibit more behavioral problems and aggressive temperaments among other characteristics more likely to experience corporal punishment? If so, can this explain why studies observe that corporal punishment is correlated with negative behavioral outcomes of children?   For instance, it may be that kids who are subject to corporal punishment are more at risk for engaging in externalizing behavior problems, rendering the relationship between corporal punishment and behavior problems spurious. Most studies are unable to clearly explain this relationship because of several reasons. These include only having observational data, a lack of experimental conditions, using recall data, and not being able to establish time-ordering in analyses (Wissow, 2001; Molnar, Buka, Brennan, Holton, & Earls, 2003; Grogan-Kaylor, 2005a; Grogan-Kaylor, 2005b; Simons et al, 2000; Straus & Kantor, 1994; Turner & Muller, 2004; Rodriguez, 2003). The current study attempts to overcome some of these past limitations.

Given the ethical restriction of researchers not being able to randomly assign children to experience corporal punishment, one way to investigate the spurious nature of the relationship between corporal punishment and child behavioral outcomes is through propensity score matching. This methodological technique allows children with similar probabilities of being punished, with some having been punished and others not, to be grouped together and then calculate whether or not there is a statistically significant effect of corporal punishment on behavioral outcomes. This method can substitute for random assignment and parsimoniously condense many factors that affect the likelihood of a child experiencing corporal punishment into one number expressed as a conditional probability (Hahs-Vaughn & Onwuegbuzie, 2006). In doing so, this will minimize selection bias so that children who are punished and those who are not will be similar

with respect to many covariates that might increase the likelihood of actually being

punished.

Purpose of the Study

The purpose of this study is to assess the treatment effect of corporal punishment

on several childhood outcomes by using a propensity score matching technique.

Observed covariates that have an impact on the likelihood of a child experiencing

corporal punishment will be analyzed to create a conditional probability, or propensity

score, of receiving punishment, and children will be matched with those who have similar

conditional probabilities. These propensity scores will then be used to evaluate the

differences on internalizing and externalizing behaviors. Including children who were

subjected to corporal punishment and those that were not punished, where both groups

have similar propensities of experiencing punishment, will produce a context in which

the causal effect of corporal punishment can be discerned, thus going beyond prior

research on this topic.

Data for the current study will be drawn from the Project on Human Development

in Chicago Neighborhoods (PHDCN), a large-scale, interdisciplinary study. It was

developed to investigate multiple aspects of child development, and, specifically, how

children develop in different neighborhood contexts. It contains a large sample size of

children that vary on important demographic characteristics, as well as a wealth of

behavioral and psychological data ("PHDCN", n.d.).

This data set, combined with the statistical method to be used, will provide a

stronger answer to the question this study proposes: what is the causal effect of corporal

punishment on childhood behavioral outcomes? It will go beyond the previous research

outlined in the next chapter and add valuable knowledge to several social science

disciplines.

      The second chapter will discuss relevant literature on corporal punishment,

followed by a chapter discussing in detail the methodology to be used, propensity score

matching. Chapter 4 will describe the methods and analytic strategy to be used in this

study. The last two chapters will explain the results and then offer discussion and

conclusions.

CHAPTER 2

LITERATURE REVIEW

Historically, many studies have investigated the impact of corporal punishment on the behavioral and psychological functioning of children. Probably the most prominent person within this area of research is Murray Straus. In several studies, he and others have observed a link between corporal punishment and lowered self-esteem, feelings of powerlessness, and more aggressive tendencies towards others (Straus & Kantor, 1994; Straus & Gimpel, 1994). This chapter will outline the history of corporal punishment, and then describe empirical findings from many studies regarding the effects of corporal punishment on children. The second part discusses literature focusing on the predictors of corporal punishment, and the third section of this chapter will describe laws in the United States and some other countries regarding corporal punishment. The next section includes limitations of previous research and how the current study will attempt to overcome them. The last section contains the research question and hypotheses for the current study.

History of Corporal Punishment

Punishment of children is assumed to have always existed in some form, which makes it a topic that is not always focused on, but usually covered in most historical accounts. However, most of the accounts (diaries and memoirs are most common) that exist contain information that is largely anecdotal and may not be indicative of general practice at the time it was written. Even so, the generalizations in regards to different time periods in history that follow in this section are agreed upon by the majority of Western historians (Donnelly, 2005).

In biblical times, the physical punishment of children appears to have been used and accepted. The biblical phrase "spare the rod and spoil the child" is often used to justify corporal punishment even today. This biblical or religious view holds that children are born with inherent sin and must be disciplined in order to be saved. Some historians cite examples such as infanticide during this time; infants that were not seen as suitable in strength, health, etc. were sometimes killed after birth (Biller & Soloman, 1986).

Children, and even infants, were assumed to have a will that needed to be "broken" in the sixteenth and seventeenth centuries, leading to strictness and harsh punishment. According to some historians, children during this time were bullied intentionally and sometimes sexually abused. A change in attitudes began to occur in the early eighteenth century; children were more likely to be seen as small humans that needed to be nurtured and trained by their parents. At this time the development of items such as books, games, and clothes just for children appeared. The goal became to instill a sense of shame to control behavior rather than to use harsh physical discipline (Donnelly, 2005).

In the 1830s this change continued and became more distinct; the focus shifted from physical punishment of children to psychological discipline. Since the late nineteenth century to the present there have been many influences (psychology, social work, etc.) that have bearing on how childhood is viewed and the place of physical punishment. Childhood has continued to be seen as a special time of life and emotional ties between parents and children have strengthened (Donnelly, 2005).

The sources of this information are largely diaries and memoirs, as well as advice manuals generated for parents. The information must therefore, be taken with some caution. Other historians argue that these types of evidence provide a biased view of history, and emphasize the context in which corporal punishment has happened historically. These scholars point out that there have been different times in history when corporal punishment was more or less of an issue and this reflects the social context of the times—it cannot be interpreted the same across time periods (Donnelly, 2005).

There appears to be several different periods in history when corporal punishment was labeled a "social problem." These periods of reform began as outcries against child abuse, but later developed to also include corporal punishment. The interest that is present right now started around the 1960s, but there were also similar levels of interest in the 1830s and 1840s, as well as from the 1870s to 1920s (Donnelly, 2005). The earliest research on corporal punishment appeared in the 1950s, but became more frequent from the mid 1960s to present (Giles-Sims, Straus, & Sugarman, 1995).

Several studies in the past fifty years have all found rates of corporal punishment to be high, between 95—99% of parents report using it, depending on the study (Straus & Stewart, 1999). Prevalence rates still remain high more recently for use of corporal punishment in American families. Straus and Stewart (1999) discuss literature on prevalence, chronicity, severity, and duration of corporal punishment. In spite of the high rates of use, corporal punishment is largely ignored in textbooks on child development (see Straus & Stewart, 1999). Straus & Stewart (1999) also obtained data from a Gallup Organization telephone poll taken in 1995. Results showed that a little more than a third of parents reported hitting their infants (under 1 year old), 94% reported hitting toddlers

(ages 3-5), and the rates declined after this age, although as late as age 13 over 40% still reported hitting their children to discipline them. The chronicity of corporal punishment was shown to be the greatest for 2 year old children, with a steady decline after this age. Their results also showed that the milder forms of corporal punishment were most prevalent among children ages 2-4, while the more severe forms became more common among middle aged children (5-12) (Straus & Stewart, 1999).

Even though corporal punishment is not seen as a problem by most American parents, it has been studied by many social scientists and is a controversial and hotly debated topic among social scientists. Several researchers have attempted to summarize past studies and compile existing research into meta-analyses. Results are contradictory, and therefore have not provided conclusive evidence for either side of the argument. For example, Larzelere and Kuhn (2005) reviewed the effect sizes of twenty-six past studies on corporal punishment. Findings showed, in contrast to many others (Grogan-Kaylor, 2005a; Straus & Kantor, 1994; Turner & Muller, 2004; Straus & Mouradian, 1998), that customary corporal punishment does not have a significant effect on behavior problems; they argue that in many studies more severe abuse (punishment that could be classified as child abuse) is combined together with customary punishment (i.e. punishment intended to change behavior and cause pain without injury), indicating a stronger effect than what is actually present. According to these authors, corporal punishment had more adverse consequences when it was used more severely or as the primary method of punishment.

Paolucci and Violato (2004) found similar results in their meta-analysis. Larzelere and Kuhn (2005) used some of the same studies, but their restrictions for inclusion were somewhat different than those used by Paolucci and Violato (2004),

leading to differences in the studies included. Larzelere and Kuhn (2005) required

studies in their analysis to investigate an alternative discipline tactic in addition to

corporal punishment, and children in the study had to be less than 13 years old. These

two criteria led to the difference in the studies analyzed. In the analysis of Paolucci and

Violato (2004) the studies had to meet three criteria to be included; they had to define

corporal punishment as a non-abusive form of punishment administered by a parent or

guardian, they had to focus on an outcome of interest (behavioral, affective, and/or

cognitive), and only empirical studies were included. The analysis of effect sizes in

thirty-five studies done by Paolucci and Violato (2004) showed that those children who

experience corporal punishment are at a very small increased risk for emotional or

behavioral problems, and no increased risk for cognitive problems. However, because of

insufficient data in many of the studies they used, they were unable to assess these effect

sizes based on important control variables such as age, gender, and quality of parent-child

relationship. Several studies did not include enough information on these variables,

leading them to be excluded from the analysis. The authors point out that there is most

likely useful information that could have been obtained if they had been able to do this.

There have been many studies (Day, Peterson, & McCracken, 1998; Grogan-Kaylor &

Otis, 2007; Straus et al, 1997) that show how these variables affect how often and how

severely a child is spanked, which may lead to differing effects.

 Gershoff (2002) also conducted a meta-analysis that included eighty-eight articles

investigating the association between corporal punishment and childhood "behavioral

experiences," a term used instead of "outcomes" to indicate that the majority of the

studies used are lacking in at least one criterion needed to prove causality (i.e. lack of

time-ordering). Her findings contradict those of the two meta-analyses discussed above.

Results show that corporal punishment is related to a large number of negative behaviors.

These include increased antisocial behavior, increased risk of being a victim of child

abuse, increased risk of being a perpetrator of abuse as an adult, increased aggression,

and decreased quality of the parent-child relationship. These results were present even

though she specifically did not use studies that aggregate together more serious abuse

with more common corporal punishment, the caution given by the authors above. The

only desirable outcome that her analyses showed to be associated with corporal

punishment was immediate compliance, but even this was inconsistent across studies.

In sum, corporal punishment has been present as a form of discipline throughout

the past several centuries. Nonetheless, it is still a topic of research for social scientists

and debate among parents. The meta-analyses described show the large amount of

variation that exists regarding adverse or positive consequences of corporal punishment.

Researchers evaluating large groups of studies often do not agree on results, let alone

individual studies investigating relationships between corporal punishment and specific

behavioral characteristics. The next sections will present some of these individual

studies.

Corporal Punishment and Externalizing Behaviors

A substantial amount of research has investigated the relationship between

corporal punishment and externalizing behaviors such as antisocial or oppositional

behavior. It is suggested that using violence to discipline children may teach them that

aggressive and violent behaviors are the correct way to reach their desired goals and lead

to an increased occurrence of these behaviors (Straus, 1996). The studies below vary in

their type of sample and ways of measurement, but all assess the relationship between corporal punishment and some sort of externalizing behavioral outcome.

Using a sample of over 3,000 children from the National Longitudinal Survey of Youth (NLSY), Straus and colleagues (1997) found that corporal punishment teaches children that violence is normal. They used a measure of antisocial behavior from the Behavioral Problems Index, a short scale adapted primarily from the Child Behavioral Checklist; this scale asks parents how often their child "cheats or tells lies," "bullies or is cruel/mean to others," along with several other questions. Corporal punishment was measured by asking parents "about how many times, if any, in the past week did you spank your child?" This study used multiple waves of data to control for previous behavior problems displayed by the child. The children in their sample that experienced corporal punishment showed an increase in antisocial behavior even when previous behavior and several other demographic characteristics were controlled.

In addition, Grogan-Kaylor (2004) assessed the relationship between corporal punishment and antisocial behavior using the same measure of antisocial behavior but in a different statistical model. This scale was modeled as an outcome of corporal punishment by using a sample of 1811 children between the ages of 4 and 14 from the NLSY. Corporal punishment was measured using a question from the Home Observation for Measurement of the Environment (HOME) instrument that asked whether or not the child had been spanked in the past week. A fixed-effects model was used to estimate the effect of corporal punishment across three waves of data, which allowed for the control of unobserved characteristics as long as they are time-invariant, resulting in a more rigorous design and analysis. Results showed a strong association

between corporal punishment and antisocial behavior even when controlling for behavior problems and demographics at the first time of measurement.

In a similar study, Grogan-Kaylor (2005b) used NLSY data to investigate the relationship between corporal punishment and antisocial behavior across neighborhood contexts. He again used a fixed-effects model and found that corporal punishment was associated with increases in externalizing behaviors such as antisocial behavior, hyperactivity, and peer problems. However, there was not a significant interaction effect between corporal punishment and neighborhood quality (i.e. perceptions of crime level, whether or not it was a good place to raise a child, whether neighborhood residents followed rules), indicating that this relationship did not vary by neighborhoods in which children reside.

A cross-cultural study of the effects of corporal punishment on antisocial behavior was conducted by Simons et al (2000). They investigated the use of corporal punishment in a sample of Iowa families as compared to a sample of Taiwanese families. The authors investigated the interaction between several parental characteristics, including warmth, support, control, and inductive reasoning. The Iowa families exhibited a lower prevalence of corporal punishment, and milder forms were more often used. In contrast, the Taiwanese sample had a high rate of physical disciplining practices, and the discipline was more likely to be severe. In this study, the delinquency checklist adapted from the National Youth Survey was used to measure delinquent acts in both samples. A three item scale was used to measure corporal punishment. The adolescents were asked how often their parents spanked or slapped them when they did something wrong, how often they were hit with an object such as a belt or paddle when they were punished, and how

often they were pushed, hit, grabbed or shoved by a parent when spending time with them in the previous month. Responses were on a five-point continuum ranging from never to always. The oppositional/antisocial behavior measure came from 19 items taken from the Revised Behavior Problems Checklist.

Taiwanese families in their sample used corporal punishment more frequently (i.e. when asked how often a child gets punished when they do something wrong, they were more likely to choose "almost always" or "always"). Results showed that lack of parental warmth was the strongest predictor of antisocial behavior in children, but corporal punishment had a larger effect on behavior problems when it was used without parental warmth. For example, among those families who used corporal punishment, the adolescents whose parents also showed lower levels of warmth were more likely to exhibit behavior problems than those who experienced corporal punishment but had parents that showed high levels of warmth. This was especially true for mother's punishment of their daughters. The authors suggest that either the relationship between corporal punishment and antisocial behavior is spurious because they both are associated with parental warmth/control, or that the effects of corporal punishment are offset by parents who offer high levels of warmth/control (Simons et al, 2000).

Straus and Mouradian (1998) also found a relationship between corporal punishment and antisocial behavior. Using a sample of 933 Minnesota mothers obtained by random digit dialing, the authors investigated corporal punishment and also impulsive corporal punishment and its effect on children. Corporal punishment was measured by how often they had spanked, slapped, or hit their child in the past six months. Impulsive corporal punishment was measured by asking how often they had spanked because they

were so angry they "lost it." Also included in the analysis was a measure of the mother's nurturance towards the child. They found that the more corporal punishment used, the more likely that it will be done impulsively, and the more corporal punishment used, the more antisocial behavior displayed. This relationship was even stronger when impulsive corporal punishment was used. The relationship between corporal punishment and antisocial behavior was the strongest when the mother used impulsive corporal punishment more than half the time.

Slade and Wissow (2004) found that even when children who are less than 2 years old are spanked there can still be negative outcomes. Using data from the NLSY, they found that in white non-Hispanic families the use of corporal punishment around age 2 was associated with significant behavior problems once they entered school. Behavior problems were measured in two ways, both developed from the Behavior Problems Index—one was a dichotomous measure, in which a value of "1" indicates the child is in the top 10% of children in the United States by ratings of behavior problems. The other measure was also dichotomous, with a "1" indicating that the mother had ever been called to her child's school because of a behavioral issue. The author's results held up after controlling for several demographic factors, child temperament, and aspects of parent-child interaction.

A similar finding was shown by Brennar and Fox (2001) in their study of children aged 1-5. Using a sample of 1,056 mothers that were recruited from day-care centers and preschools, their results showed that parental discipline techniques accounted for 13% of the unique variance in their behavioral problems outcome variable, and almost 20% of the total variance. Children who were frequently punished by their parents exhibited

more behavioral problems. The authors used the Parent Behavior Checklist to measure discipline strategies; this assesses parenting practices in three dimensions—discipline, nurturing, and expectations. The discipline subscale measured how often parents used verbal and physical punishment methods. Behavioral outcomes were measured by the Behavior Screening Questionnaire, which measures the frequency of nine problem behaviors, including having tantrums, not listening, disobeying, and acting aggressively.

Using a subset of the NLSY, McLoyd and Smith (2002) showed a relationship between corporal punishment and later behavior problems. They investigated the interaction between parental emotional support and spanking and the subsequent effects on behavior. Behavior was assessed using the Behavior Problems Index (BPI), which uses items drawn from the Child Behavior Checklist and other behavior problem checklists. Spanking was measured in the first wave by observing whether or not the mother hit the child, and in later years by questions asking how many times the mother had spanked the child in the past week. Emotional support was operationalized by using five items from the HOME scale, such as whether or not the mother caressed or hugged the child during the interview. Their results showed that being spanked increased the level of problem behaviors over time, with results being similar across ethnic groups (i.e. Hispanic, European American and African American). There were also increases in problem behaviors for those children that experienced an increase in spanking over time. Finally, emotional support decreased the negative impact of spanking over the 6 year period they studied.

As these studies illustrate, the relationship between corporal punishment and externalizing behavioral outcomes is well-documented and empirically supported. This

relationship is observed across many samples and even across cultures. There also

appears to be other family or parental factors that affect this relationship, such as the

amount of warmth or support given by the parents. While the behaviors addressed by the

previous studies are true of some children, some children experience different types of

problems that are characterized by depression or withdrawal. The next section will look

more closely at studies investigating the relationship between corporal punishment and

internalizing behaviors such as these.

Corporal Punishment and Internalizing Behaviors

There is also substantial research investigating the relationship between corporal

punishment and internalizing behaviors such as depression and psychological distress.

Corporal punishment can have negative effects on a child's self-concept and self-esteem

as well (Turner & Finkelhor, 1996). The studies that follow examine the relationship

between corporal punishment and internalizing behavior problems using a variety of

samples and methods. Empirical findings show that corporal punishment is related to

several different types of internalizing behavior problems in children.

Straus and Kantor (1994) found that corporal punishment in adolescence was

related to an increase in depressive symptoms, suicidal thoughts, probability of alcohol

abuse, and probability of child abuse. They used a subset of a sample of 6,002 families

that were part of the National Family Violence Survey. They note that a significant

difficulty with this study, however, is the use of recall data because they asked adults to

report on their childhoods rather than using a longitudinal design.

Parents who hold more physically abusive attitudes also have children with more

anxious and depressive symptoms. In a study done with forty-two New Zealand children

ages 8-12 and their parents, the parents were given vignettes and asked what type of discipline they would use in particular situations. Those parents that were more likely to name physical methods of punishment were said to have a higher "abuse potential," and their children were more likely to show these symptoms, as measured by the Children's Depressive Inventory, the Children's Manifest Anxiety Scale, and the Children's Attributional Style (Rodriguez, 2003).

Turner and Finkelhor (1996) also show that corporal punishment has a significant impact on psychological distress even after controlling for sociodemographic factors, abuse, and parental support. They used the National Youth Victimization Prevention Study (NYVPS) to analyze a sample of 1,042 boys and 958 girls between the ages of 10 and 16. Those that were frequently punished by parents showed a stronger likelihood of psychological distress symptoms such as "feeling sad," "feeling alone," and "felt they were doing things wrong." Additionally, as frequency of punishment increased the effect of parental support on decreasing these psychological distress symptoms lessened.

Afifi, Brownbridge, Cox, and Sareen (2006) found that experiencing physical punishment led to an increased likelihood of having psychological disorders in adulthood. Using a sample taken from the National Comorbidity Survey, a study on the mental health of the United States general population, they divided subjects into three groups— those that had never been physically punished, those that had experienced at least some physical punishment, and those that had been abused as a child. The variables used to divide subjects into these groups came from the physical punishment and physical abuse subscales from the Conflict Tactics Scale. Results showed that those who were physically punished were more likely to develop a psychiatric disorder in relation to

those who were not punished, although the group that was abused showed the greatest likelihood. These effects were present even when controlling for demographic information and parental bonding.

It has also been shown that corporal punishment increases the likelihood of depression later in life (Turner & Muller, 2004). Using a sample of 649 college students, Turner and Muller (2004) found a relationship between depressive symptoms and corporal punishment, even when more severe abuse was controlled for. Depressive symptoms were measured using the Center for Epidemiologic Studies Depression Scale, which asks how often over the preceding two weeks a subject had experienced each of 20 symptoms such as feeling lonely, having restless sleep, poor appetite, and had trouble keeping their mind on a task. Additionally, results showed that having an angry parent inflict the punishment on the child resulted in a stronger relationship between corporal punishment and these negative outcomes listed above. The effects of corporal punishment in this study were seen to be similar to the effects of physical abuse on depression.

As evidenced by the numerous studies reviewed above, there is a significant relationship between corporal punishment and depression, psychological distress, the likelihood of developing a psychological disorder, and level of anxiety. Not only has the effect of corporal punishment on these behaviors been observed in children, but some of these effects are also found in adults. If these behavior problems discussed in the previous two sections, both internalizing and externalizing, are caused by experiencing corporal punishment, then it is also crucial to know what characteristics will cause a child

to be more likely to experience that punishment. The next section discusses these characteristics.

Predictors of Parental Use of Corporal Punishment

Not only is it important to understand what the consequences of corporal punishment can be, it is also necessary to understand the factors that put children at an increased risk for experiencing corporal punishment. As mentioned previously, it is often argued that behavior problems and other characteristics found in children who have been spanked were already present and led to an increased incidence of punishment. Because of this belief it is important to understand what these behaviors and characteristics are. The current study attempts to take into account these differences among children that exist prior to experiencing corporal punishment to control for this problem. As outlined below, there are many characteristics of children and families that are associated with higher or lower rates of corporal punishment. Empirical results from several studies illustrating these characteristics and how they are related to the use of corporal punishment are presented in the following section.

Using a nationally representative sample gathered by the Gallup Organization, Dietz (2000) explores individual and family characteristics that predict the use of corporal punishment. She conducted a logistic regression analysis using the measure of ordinary corporal punishment and severe corporal punishment from the Conflict Tactics Scale for Parent-Child. These two concepts were measured dichotomously to indicate whether or not they had been used at least once in the past year. Parents were asked a series of questions over the phone about their discipline tactics during that reference period. Findings indicated that several variables significantly increased the use of

corporal punishment. Those parents referring to a child six years old or younger were more likely to use corporal punishment. Male children and African-American children were also more likely to be spanked. Mothers were more likely to use corporal punishment, but it should be noted that this model did not take into account the amount of time spent with the child; this finding may be explained by the lesser amount of time that most fathers spend with their children. In addition to these, those adults that indicated they had been abused by a parent were less likely to spank their own children. In the analyses on severe corporal punishment, it was found that parents who had less than a high school diploma, were Southern residents, and had an income of less than $15,000 were more likely to display this behavior (Dietz, 2000).

Giles-Sims et al (1995) also investigated characteristics associated with the use of corporal punishment using data from the 1990 NLSY. The authors used multivariate analyses of covariance (MANCOVA) and analysis of covariance (ANCOVA) to assess what demographic variables (including age and sex of the child, age of mother, marital status, education, race, religion, etc.) were related to the prevalence (whether or not it was used) and chronicity (how often it was used) of spanking. Spanking was measured using questions from the HOME scales. This includes both a measure of whether or not the interviewer observed the mother spanking the child, as well as questions asked to the mother about discipline practices. Their findings showed that several characteristics were associated with spanking. Children who were between the ages of 3 and 5 were most likely to be spanked, and more frequently spanked. The authors found several other variables were still related to receiving corporal punishment after controlling for SES. There was a larger percentage of boys that had been spanked, although there were not

significant differences in chronicity. Older mothers were less likely to spank their

children, and unmarried mothers showed a higher frequency of spanking. African

American parents were also more likely to spank their children, although there were not

significant differences in chronicity. Protestants, residents of a rural area, and those that

live in the South were also more likely to spank their children (Giles-Sims et al, 1995).

Another possible influencing factor is the use of child care. If children spend less

time around their parents (because they are in some sort of daycare), they may experience

less corporal punishment. This was explored by Magnuson and Waldfogel (2005). They

conducted a series of analyses to investigate whether or not Head Start programs or other

types of child care programs had an effect on parental disciplinary practices. They

showed that when controlling for child and family characteristics, those parents who had

children enrolled in a Head Start program were more likely to report never spanking their

child. However, within the group of parents that had ever spanked their children, being

enrolled in Head Start did not have an effect on the number who had spanked their child

in the past week. Head Start programs were also associated with a lower number of

domestic violence reports. Additionally, results showed that low income, single parent,

and less educated parents were more likely to spank their children. The data from this

study were drawn from the Early Childhood Longitudinal Study, Kindergarten Class of

1998-1999 (ECLS-K).

An analysis on the predictors of corporal punishment was recently done by

Grogan-Kaylor and Otis (2007). They used NLSY data from 2000 to estimate a tobit

model assessing the effects of many different variables on corporal punishment. Their

independent variables included age, race, and gender of the child, cognitive stimulation

(how many books were in the home, the frequency of educational outings, etc.), and measures of internalizing and externalizing behaviors from the Behavior Problems Index (BPI). Results showed gender was not associated with receiving corporal punishment, however, race of the child was—African-American children were more likely to be spanked by their parents as compared to Whites. In addition, younger children were more likely to receive corporal punishment. The authors also showed effects for religious affiliation (Protestants were more likely than Catholics to spank) and cognitive stimulation (higher levels were associated with lower levels of spanking). Children who exhibited more externalizing problems were more likely to be spanked, and those that exhibited more internalizing problems were less likely to be spanked. Additional analyses assessed the effect of these independent variables on the frequency of punishment. The results were essentially identical; the same independent variables that predicted whether or not a child was spanked at all also predicted how often they were spanked (Grogan-Kaylor & Otis, 2007).

Day et al (1998) also assessed predictors of corporal punishment using a research design that divided parents into separate groups based on race and marital status using a sample taken from the National Survey of Families and Households (NSFH). Their design incorporated six different groups of independent variables, including characteristics of the parent, characteristics of the child, parental ideologies, economic status, the relationship between the parent and child, and community attributes. This study differed from many others because the question on spanking asked if and how much the parent had spanked their child in the last week. Most other studies investigating spanking ask about a longer time period (usually one year previous). This

shorter time period may lead to more accurate results because it is easier for parents to remember one week rather than one year, but it also may not capture a typical week in any one family. If a child is spanked often, but had not been spanked during the week of the interview, this could lead to biased results. However, a child who is not typically spanked could have also been spanked during the week of the interview—leading the authors to argue that across the large sample these two groups average each other out (Day et al, 1998).

The results showed that influential predictors varied to some degree based on the race and marital status of the parent. Overall, however, findings showed that boys are spanked more often than girls, older children are more likely to be spanked, and mothers are more likely to spank than fathers. They also found that parents who perceived their children as difficult and incompetent were more likely to spank them, and many of the attributes of parents (education, age, mental health, and conservative religious beliefs) had an effect on spanking. Finally, context of the parent-child relationship, including how much they argued, household size, and the amount of social support available influenced the likelihood of being spanked (Day et al, 1998).

Laws Regulating Corporal Punishment

The research and information presented above clearly shows that corporal punishment is a prevalent discipline practice in the United States that can have adverse consequences for children. However, most states do not have any formal regulations or laws that prohibit corporal punishment. This section will briefly describe the laws that regulate corporal punishment in the United States and in several other countries.

Minnesota is the only state in the United States that has a law specifically outlawing corporal punishment by parents. However, this law is widely unknown. It is also only evident when a number of statutes are read together. No parents have been formally prosecuted for breaking this law (Bitensky, 1998). All other forty-nine states allow corporal punishment by parents, although some have more stringent restrictions on its use than others. Use of corporal punishment in public and private schools is prohibited in twenty-nine states. There are twenty-three states that have state laws allowing corporal punishment. The remaining states have differing laws based on school districts or more specific restrictions on the punishment (Couture, 2006).

There are seventeen countries that prohibit spanking by parents, teachers, or caretakers, with Sweden being the first to enact such a law in 1979 (Couture, 2006). These laws are often not intended to criminalize and incarcerate parents, but nevertheless they are intended to stop spanking. There have been a few reported cases of parental prosecution in some of these countries, but not as many as critics of anti-spanking laws predicted. The possibility does exist for legal punishment, or for the use of corporal punishment to be used to help win child custody battles by the non-spanking parent. The countries that outlaw spanking tend to encourage the use of other types of discipline, including time-outs and the loss of privileges (Bitensky, 1998).

Summary and Limitations of Past Studies

As shown by the research above, both the relationship between corporal punishment and subsequent behavior as well as predictors of receiving punishment are popular topics for study by social scientists. A large body of research shows that various negative outcomes are associated with corporal punishment, however, there is still

disagreement over the effects corporal punishment have on a child and variation among states and countries as to how it is regulated. For instance, some research shows (Larzelere & Kuhn, 2005) that the effects might not be as detrimental as others have presented (Gershoff, 2002). Individual and family predictors of the use of corporal punishment were outlined in the previous section, with most researchers agreeing on how these predictors affect the use of spanking. The current study will add to this research by investigating the relationship between corporal punishment and both internalizing and externalizing behavior problems, and by overcoming as much as possible the limitations presented below.

Although studies discussed in this chapter reveal that corporal punishment is related to a host of negative outcomes, these studies have several limitations. While many use statistical controls in their models, they cannot rule out that the relationship shown between corporal punishment and behavioral outcomes is nonspurious (Wissow, 2001; Molnar et al, 2003; Grogan-Kaylor, 2005a; Grogan-Kaylor, 2005b; Simons et al, 2000; Straus & Kantor, 1994; Turner & Muller, 2004; Rodriguez, 2003; Straus & Mouradian, 1998; Straus, 1996; Turner & Finkelhor, 1996). The best way to achieve an assessment of this would be to conduct a randomized experiment with time-ordered measurements. It is possible that children who are exposed to more corporal punishment experience this because they are more likely to misbehave and in some sense "cause" the punishment. Many of the studies described previously have attempted to make up for this limitation by including time-ordering in their analyses (Grogan-Kaylor, 2004; Grogan-Kaylor, 2005b; McLoyd & Smith, 2002). In these studies, the previous behavior problems were included as control variables and/or the measurement of behavior was

taken after the measurement for corporal punishment. However, this does not fully deal with the possible spurious relationship because these children who had behavior problems in the earlier measurement could still be "causing" the punishment to happen more frequently. This is arguably the most serious limitation of the majority of studies that assess the effects of corporal punishment (Paolucci & Violato, 2004; Benjet & Kazdin, 2003). The current study addresses this limitation by using a rigorous statistical method that matches children based on background characteristics and variables predicting the likelihood of receiving punishment (i.e. demographics, behavior problems, parental characteristics) before assessing the relationship between corporal punishment and later behavior problems. The result is a comparison of behavioral problems between children who received corporal punishment and those that did not, but are very similar on all other variables.

In addition, some studies use recall data, which can be biased because such data collection methods rely on the memory of subjects (Straus & Kantor, 1994; Turner & Muller, 2004). Subjects are asked to remember what they experienced as children, at a time that is many years after their childhood. For example, Straus & Kantor (1994) asked adult respondents from the National Family Violence Survey to remember how often their parents physically punished them when they were teenagers. While prevalence estimates for the use of corporal punishment were similar to current rates using other measurement methods, the ideal way of assessing rates of punishment would ask about a much more recent time period. The current study uses data gathered from parents that asks only about the year previous to the interview, which is superior to recall data from adults about their childhood.

Finally, cross-sectional data are used in several studies, which doesn't allow for temporal ordering (Turner & Finkelhor, 1996; Brennar & Fox, 1998). When the measurements analyzed are taken at the same time, it could just as easily be argued that the outcome variable (aggressive behavior, antisocial behavior, whatever it may be) occurred first (Benjet & Kazdin, 2003). Brennar and Fox (1998) included only one time of measurement in their study, leading to less support for their results. The measurement for behavior problems and discipline techniques were taken at the same time. In the current study two waves of data taken approximately three years apart will be used to help overcome this limitation. The measurements for corporal punishment as well as the control measurements for behavior problems are taken from the first wave of data and the dependent variables of internalizing and externalizing behaviors are taken from the second wave of data approximately three years later.

Research Questions and Hypotheses

The main research question for this study is: what is the causal effect of corporal punishment on internalizing and externalizing behavior outcomes in children? Previous studies have indicated that there is a relationship between corporal punishment and many behavioral outcomes, including antisocial behavior, depression, anxiety, and psychological distress. These behaviors correspond with many of the behaviors measured by the internalizing and externalizing scales of the Child Behavior Checklist (CBCL) and their sub-scales. There are several hypotheses that will be tested in this study based on these findings:

Hypothesis 1: Children who experience corporal punishment will have higher scores on the externalizing behavior scale of the CBCL than those who do not experience corporal punishment.

Hypothesis 2: Children who experience corporal punishment will have higher scores on the aggressive behavior scale of the CBCL than those who do not experience corporal punishment.

Hypothesis 3: Children who experience corporal punishment will have higher scores on the delinquent behavior scale of the CBCL than those who do not experience corporal punishment.

Hypothesis 4: Children who experience corporal punishment will have higher scores on the internalizing behavior scale of the CBCL than those who do not experience corporal punishment.

Hypothesis 5: Children who experience corporal punishment will have higher scores on the withdrawn behavior scale of the CBCL than those who do not experience corporal punishment.

Hypothesis 6: Children who experience corporal punishment will have higher scores on the depressive/anxious behavior scale of the CBCL than those who do not experience corporal punishment.

Hypothesis 7: Children who experience corporal punishment will have higher scores on the somatic complaints scale of the CBCL than those who do not experience corporal punishment.

CHAPTER 3

PROPENSITY SCORE MATCHING

This chapter will outline in detail Propensity Score Matching (PSM), including its

strengths and weaknesses in relation to other methodological approaches. Several

different matching techniques will be described and when it is appropriate to use them.

Also included in this chapter is a brief overview of studies that have used PSM across

disciplines to estimate causal effects when random assignment could not be achieved.

Description of PSM

In observational studies it is difficult to infer causality for several reasons.

Subjects are not randomly assigned to a condition or treatment in observational studies

and one of the most serious of these problems is the possibility of selection bias. As a

result, it is very likely that systematic differences between subjects exist that lead some to

be exposed to an event, condition, or treatment and others to not be exposed. For

instance, children can not be assigned at random to join a gang; therefore, it is likely that

kids who join a gang will be quite different than kids who do not across several

characteristics (i.e. family condition, parental characteristics, behavior, personality,

demographics, etc.). As such, selection bias can often result in inaccurate estimates of

causal effects of a treatment (or exposure) on an outcome (e.g., gang memberships causal

effect on violent victimization) because other characteristics are responsible for

differences found between groups that are exposed to a condition or treatment versus

those not exposed, rather than the treatment under investigation. Using a PSM strategy

allows observed characteristics that affect the probability of being exposed to a treatment

condition to be controlled for and condensed into one number for each subject (Hahs-Vaughn & Onwuegbuzie, 2006).

PSM is often used to balance or match two groups when an experimental design is not possible. It can be a way to infer causality when experimental conditions cannot be met. In a classical experimental design, subjects are randomly assigned to either a treatment or control group, which serves to eliminate systematic bias between the two groups (Campbell & Stanley, 1966). However, this type of design is rarely possible in social research because of the nature of society and ethical constraints imposed by ethics boards. It is often inappropriate for a social scientist to randomly assign subjects to a treatment or control group (such as belonging to a gang or experiencing corporal punishment), and even if it was, the act of this assignment can still affect results. Since children can not be randomly assigned to receive corporal punishment, one way to control for the consequential selection bias is to group subjects by their conditional probability of receiving punishment, which can eliminate the influence of extraneous variables that might increase the likelihood that some kids will be exposed to corporal punishment and others not (Rosenbaum & Rubin, 1984).

By using PSM, the influences of observed covariates on the likelihood of receiving corporal punishment are combined together into a conditional probability. In doing this, PSM is parsimonious while still taking into account many different covariates (Rubin, 1997). Multivariate models such as regression are based on several assumptions, one of which is termed the "ignorability of treatment assignment." This implies that after controlling for confounding variables the only differences that remain between treatment and control subjects are due to randomization. However, this is usually not verifiable in

observational studies and depends on the variables that are present in the data set. These methods are used to make causal inferences even though these underlying assumptions cannot be verified. PSM serves as a proxy for randomization by matching subjects on covariates that may affect their likelihood of receiving treatment. This method improves upon more common multivariate methods by matching the subjects before comparing them on treatment outcomes, to ensure that other variables are not responsible for the differences observed (Newgard, Hedges, Arthur, & Mullins, 2004).

Although regression models control for differences that exist prior to treatment, they still have limitations. As mentioned above, propensity score matching techniques also allow for a much larger number of independent variables to be considered while still being parsimonious because the effects are combined into one score. This large number of covariates can be included without danger of multicollinearity problems found in regression models (Hahs-Vaughn & Onwuegbuzie, 2006). In a multiple regression, multicollinearity can be a problem if the intent is to estimate how each independent variable affects the dependent variable separately and in relation to one another. However, in PSM the independent variables are used only to create the propensity score, eliminating this as a potential problem. These covariates are used only to estimate the propensity score and not to make any inferential statements (Newgard et al, 2004).

In addition, another problem can arise in regression models if there are specific groups of people not present in the analysis. For example, if there was not a representative group of poor or uneducated people within the sample, a regression model would still estimate outcomes for this group by extrapolating the model. The regression model uses information available to make predictions for this group even though they are

not present in the sample. Since the actual subjects are not present, a true estimation of the treatment effect on this group is not possible. This would be a misinterpretation of the treatment effect because of the lack of subjects fitting this description (Hill, Waldfogel, Brooks-Gunn, & Han, 2005). In PSM, these characteristics would be taken into account during the matching process. Treated subjects are matched with untreated subjects very similar to them, and if a match is not present in the sample the subject may be discarded.

Typical regression models also rely on the assumption of linearity between the covariates and outcomes. If this relationship is not present, or if the distribution of the treatment and control groups are very different, then the results estimated by a regression model can be unreliable. With PSM, each subject is matched based on the values for covariates and if there is not an adequate match the subject is dropped, helping to overcome this limitation that regression presents (Newgard et al, 2004).

Another method, fixed effects, has been proposed to make up for some of the shortcomings of linear regression. However, this method has problems that can be partially overcome with PSM as well. The biggest problem with this type of model is that although it can control for unobserved characteristics, it can only control for those unobserved characteristics that are "fixed" and do not change over time. Fixed-effect models work best when the data used contain a large amount of variation and measurements taken at many different time points (Hill et al, 2005). Another disadvantage of fixed-effect models is that they cannot estimate separate coefficients for time-invariant characteristics such as gender or ethnicity (Grogan-Kaylor, 2004). PSM shares some of the limitations of these models, in that it can only account for observed

differences, and once the propensity score is estimated the effects of each covariate cannot be disentangled from one another. It improves upon them, however, because the covariates can be either "fixed" or change over time.

As discussed, a propensity score is a conditional probability which combines the effects of observed covariates on the likelihood of receiving treatment or not. In the current study, the treatment condition is defined as having received corporal punishment. There will be kids that have similar probabilities of receiving corporal punishment, but some will have received corporal punishment and some will have not. If the kids that received corporal punishment are matched with a child that did not receive corporal punishment but has a similar propensity score, then those that do receive the treatment are basically chosen at random from each group. This process creates a matched sample that approximates two randomly selected groups, one that has received treatment (corporal punishment) and one that has not. If the sample consists of treatment and control units that have been randomly assigned to each group, the effect of treatment on outcome variables can more easily be assessed (Rubin, 1997).

PSM represents a counterfactual approach. This approach conceptualizes causality in terms of a treatment effect on some outcome (in the case of this study, corporal punishment's effect on behavioral outcomes). Each individual has the potential for two different outcomes, either an outcome under the treatment condition or one under the control condition, but only one of these can be observed at the same time (Sampson, Laub, & Wimer, 2006). The counterfactual question becomes—if a treated child did not receive that treatment, what would have happened to them (and vice-versa)? The answer

to this question can only be estimated, and propensity score matching is one way to make up for selection bias in creating this estimate (Guo, Barth, & Gibbons, 2004).

Matching Techniques in PSM

Once propensity scores are estimated subjects must be matched so that balance can be achieved for those who have and have not been exposed to corporal punishment. There are several different types of matching techniques that are used in PSM. The technique that should be chosen depends on the data set, especially how much overlap in propensity scores exist between treatment and control groups. The different matching techniques will generally provide the same results if the overlap is substantial. However, if there are large differences between the two groups, the decision about which technique to choose becomes more critical (Dehejia & Wahba, 2002).

Matching techniques include nearest-neighbor, caliper, and kernel matching. Nearest-neighbor matching is done by randomly ordering the list of treated subjects, selecting the first treated subject, and then pairing them with the untreated subject with the closest propensity score. These two subjects are then removed from the list and the process continues. Caliper and kernel matching are similar; they begin with choosing the first treated subject on the list. Then, a range around the propensity score is defined (i.e. +/ - .05) and an untreated subject is randomly selected from all untreated subjects that have a propensity score within that range (Guo et al, 2004). Kernel matching differs from caliper because it uses weighting to determine the best match, based on the average outcome difference between pairs (Sianesi, 2001).

Nearest-neighbor, caliper, and kernel matching can be used with slight variations depending on the specific dataset. Matching with replacement allows untreated subjects

to be matched with more than one treated subject, which increases the accuracy of the match. It allows units with the closest scores to be matched to each other because untreated units are used multiple times, which also reduces bias. When replacement is not used, the last treated subjects to be matched are sometimes left without an untreated subject with a similar propensity score, or else a treated unit may be matched with an untreated one that has a substantially different propensity score. However, if replacement is not used in conjunction with "trimming" the list of treated units based on whether they fall outside the minimum and maximum propensity score of the untreated units, the matches will be more accurate. In addition, these techniques are sometimes used with replacement to match more than one untreated case with each treated case. This is done by defining a specific number of "neighbors" to match, or by including all untreated cases that fall within the propensity score range defined by the calipers (Dehejia & Wahba, 2002). Using replacement may provide each treated unit with more opportunities to find a match, but it is not always used. By matching a control child with more than one treated child, that control child is essentially being compared to many different treated children. The effects seen from these comparisons could then be the result of comparing one small group of children that has not experienced corporal punishment to a much larger group of children that has. If this is the case, the effects become less generalizable to other groups of children.

Each matching strategy has limitations. PSM requires a large sample, because the sample size is substantially reduced after the matching occurs since not all kids will have a suitable match within the sample. Another reason sample size is decreased is because cases at either extreme end of propensity scores (e.g., those having the lowest conditional

probability of being punished vs. those having the highest conditional probability of being punished) are often excluded because there is not a suitable match (Guo et al, 2004). The beginning sample size to be used in the current study is approximately 2,800 children, which should leave a sufficient sample size after matching occurs.

Applications of PSM Across Different Disciplines

PSM has been used in several different disciplines to match individuals that have been exposed to some condition to those that have not. It is becoming somewhat more apparent in the social sciences, especially as an additional method in some studies to provide more rigorous testing for causal effects on particular outcomes. This section will review some recent studies that have applied the PSM technique to various questions across various disciplines concerning causality.

PSM and Criminology. Bingenheimer, Brennan, and Earls (2005) used a PSM method to assess the impact of exposure to firearm violence on later serious violent behavior. With data from 78 neighborhoods in Chicago, they use propensity score matching to subclassify their sample, which included 1,517 children who were either 12 or 15 at the beginning of the study, into 10 different strata. Their original analysis includes 153 covariates, although this is reduced to 48 that remain significant after a maximum-likelihood logistic regression is used. Those youth that were exposed to firearm violence were more likely to be perpetrators of serious violent acts.

Sampson et al (2006) used PSM to investigate the "treatment effect" of being married in relation to criminal activity. Their data were from a cohort of 500 men who were originally the subjects of a longitudinal study on juvenile delinquency that began in the 1940s. The authors searched for the men from this study and acquired offending

information about them for the time period between the end of the original study up to

age 70. They also conducted interviews with 52 of the men. Because of the nature of

this sample, their analyses include different sample sizes depending on the information

that was available about each person. They used PSM to adjust for covariates that

influence whether or not a man will get married, and then looked at the effect of marriage

on criminal activity. Using a hierarchical linear model, their results indicated that being

married was related to a significant decrease in criminal activity over the life course of

these men.

The relationship between drug use in high school and job-related outcomes later

in life was also studied using PSM. Ringel, Ellickson, and Collins (2007) used PSM to

match subjects on the likelihood of being a hard drug user in high school. The authors

used covariates such as race, educational attainment of parent, whether or not the subject

lived with both parents, two sub-scales assessing deviance and rebelliousness, academic

orientation, use of alcohol, marijuana, and cigarettes, and access to drugs. The later

measurement of job-related outcomes was taken at age 29, so a measurement of current

hard drug use and educational attainment at this time was also included. Both PSM and a

probit model with controls were used in this study, but the results for each were very

similar. Their results showed overall that drug use in high school led to poorer outcomes

in jobs later in life, but the most interesting part of their results was that they varied by

gender. Hard drug users in high school that are male were more likely to have benefits of

their jobs affected by this (i.e. health insurance, vacations) in contrast to the type of job

they held. Females were more likely to hold low skill and low status jobs, but the only

job benefit affected was a lesser likelihood of having health insurance.

PSM and Public Health. El-Bassel, Gilbert, Wu, Go, and Hill (2005) found a bidirectional relationship between frequent drug use and intimate partner violence in a sample of women recruited from a methadone clinic. They used PSM to match women with similar values on confounding variables before testing their hypotheses. The data used contained three different waves, allowing them to test opposite hypotheses between wave 1 and 2 and between wave 2 and 3. They found that the relationship between drug use and intimate partner violence varied based on which type of drug was being considered and frequency of use. By using PSM, the authors were able to eliminate bias that existed between their treatment and control group and ensure that observed covariates were taken into account that preceded the "treatment," (drug use) which preceded outcome variables (intimate partner violence).

PSM and Medical Research. PSM has been used in medical research more extensively than social research. Foster (2003) used PSM to investigate dose response in relation to symptomatology in a group of children being treated at a mental health hospital. Dose response refers to the frequency of visits for treatment to the mental hospital. PSM was used in this study to adequately control for factors leading to differing use of services at the mental health hospital. Children were matched on many covariates that could influence the number of visits they had to the hospital. He found that a child's frequency of visiting the hospital was unrelated to scores on the Child Behavioral Checklist, but there was a significant relationship between this frequency and the Child and Adolescent Functioning Assessment Scale. An interesting finding from his study was that children who were in the middle group in terms of frequency (between 3 and 6 visits) worsened during the follow-up period. After further analysis he discovered this

was mostly due to the group of children who had 3 or 4 visits. The author does not provide much explanation for this counterintuitive finding, but it is consistent with past research.

PSM has also been used to investigate the differential costs in medical care between patients receiving palliative care and those receiving usual care. Penrod, Deb, Luhrs, Dellenbaugh, Zhu, Hochman et al (2006) used a number of covariates including patient age, comorbidity, whether the stay was surgical or medical, and principal diagnosis to predict the likelihood of receiving palliative care. They used both generalized linear models and PSM, and found similar results with both. Patients who received palliative care were less likely to be admitted into ICU than usual care patients, and the direct and ancillary costs were significantly lower for palliative care patients.

PSM and Developmental Psychology. PSM was used by Hill et al (2005) to explore the effect of maternal employment on child behavioral development. The authors used a sample from the NLSY to assess differences in behavioral outcomes of children based on whether or not their mothers worked full-time, part-time, or not at all during the first three years after they were born. They used a number of covariates including the child's age, marital status of the mother at the time of birth, poverty status of the household at the time of birth, whether or not the child was the first born, and mother's age at birth. Some negative effects were found, but they were mostly small and many were nonsignificant. One comparison between children showed that those aged 5-8 whose mothers worked full-time during the first year of their life had higher levels of externalizing behaviors (measured by the BPI) and 5-6 year olds also had lower scores on the Peabody Individual Achievement Test—Math (PIAT-M) and the Peabody Individual

Achievement Test—Reading (PIAT-R) than those children whose mothers did not work at all in the first three years of life. When comparing children of mothers who worked full-time in the first year and those who worked part-time in the fist year, children aged 3-4 whose mothers worked full-time had lower scores on the Peabody Picture Vocabulary Test—Revised (PPVT-R), and 5-6 year olds had lower scores on the PIAT-M and the PIAT-R.

   PSM and Sociology. Gangl (2006) studied the effects of unemployment using PSM. He used data from both the Survey of Income and Program Participation (SIPP) and the European Community Household Panel (ECHP) to compare unemployment trajectories of those in the United States and 12 other countries. The author matched subjects based on background characteristics that might influence them to experience unemployment, as well as country and survey year. He found that the prospects of a person after unemployment was affected by institutional factors even taking into account macroeconomic differences that existed between countries. Not surprisingly, he also showed that the policies in place for labor market institutions played a large role. The author's main point was that those subjects that became unemployed in a country where the labor market was not highly institutionalized fared significantly worse than those who lived in an markets with stronger institutional power.

   Using a sample from the NLSY, Lundquist (2004) analyzed marriage rates by race for a sample of civilians and a sample of military personnel. She argues that the military is a controlled environment that does not have many of the problems of racial stratification and socioeconomic status found in the rest of society. By matching subjects on covariates that influence their likelihood of enlisting in the military, she compares the

rates of marriage for blacks and whites. Her results showed that black enlisted subjects

were three or four times more likely to marry than their civilian matches. White military

members were also more likely to marry than white civilians, but the magnitude of the

effect was much smaller. The author's main arguments for why these results are present

are because of the income and job stability provided by the military as compared to

opportunities in the larger society.

PSM is a statistical method that allows for an estimation of causal effects when

only observational data are present. It has been used across several disciplines, including

medicine, sociology, psychology, criminology, and public health. Several different

matching techniques and their variations were discussed above. In the current study PSM

is used to control for observed characteristics of children that may lead them to

experience corporal punishment. By matching children with similar propensity scores

who have experienced corporal punishment with others who have not, a more accurate

estimation of the causal effects of corporal punishment on behavioral outcomes can be

discerned.

CHAPTER 4

METHODS & ANALYTIC STRATEGY

This chapter describes the methodology and analytic strategy to be used in this study. The data used are first described in detail, with the sampling strategy and methods used to collect data included. The analysis sample is also described in this chapter. The dependent and independent variables are described, as well as the list of covariates. Throughout this chapter items for many scales are discussed, as well as the reliability coefficients of these scales. The full set of items on every scale used can be found in Appendix A. The last part of this chapter is the analytic strategy, which explains, step-by-step, how the analysis will proceed in testing hypotheses of the current study.

Methods

      Data. Data used for the current study are from the Project on Human Development in Chicago Neighborhoods (PHDCN). The PHDCN is an interdisciplinary study that was initiated to investigate the pathways by which children develop both positive and negative behaviors. Starting in the early 1990's, the PHDCN was funded by the MacArthur Foundation, the National Institute of Justice, the National Institute of Mental Health, the U.S. Department of Education, and the Administration for Children, Youth and Families. The principal investigator on the project is Felton J. Earls, affiliated with Harvard Medical School. This project assesses children's development based on a variety of influences including family, neighborhood, psychological, and behavioral effects ("PHDCN", n.d.).

      Many reasons exist as to why Chicago, Illinois was chosen as the city to conduct the most ambitious sociological study in history. First, while other cities were considered,

e.g., Los Angeles, Chicago was chosen as the study site because of its rich neighborhood diversity with respect to race, socioeconomic conditions, and ethnicity. Second, there is also a long history of social science research in Chicago (Shaw & McKay, 1942; Ford, 1950; Mowrer, 1950; Lewis & Maxfield, 1980; Lopata, 1978), as well as a large amount of support in the city for the PHDCN project ("PHDCN", n.d.).

The PHDCN sampling design began with identifying the 847 census tracts of Chicago which were combined to create 343 neighborhood clusters; each cluster consists of geographically contiguous census tracts that are very similar in key census indicators such as demographics. This project combines community survey data, census data, and official crime data, providing a range of variables to assess differences between neighborhoods. Because of this it is particularly well suited for investigating neighborhood effects on child development, although neighborhood measures will not be used in the current study (Molnar et al, 2003).

For the longitudinal, developmental portion of the study, a stratified probability sample was selected from the 343 neighborhood clusters to achieve a representative sample of 80 neighborhood clusters that reflected the racial, ethnic, and socioeconomic status of the 343 original neighborhood clusters. The longitudinal study of the PHDCN is fulfilled by the existence of three different waves of data collection on multiple children and adolescent cohorts residing in the 80 selected neighborhood clusters, each collected approximately two and a half years apart from one another (Sampson, 1997).

Data collection for the longitudinal study began in 1994-1995, the second wave of collection occurred between 1997 and 1999, and the third wave of data was collected between 2000 and 2001. After the random selection of 80 neighborhood clusters, a

random sample of block groups from within each cluster was drawn, and then each

dwelling and a list of occupants was enumerated from these block groups. From this list

of approximately 40,000 dwellings, children and adults were screened for participation.

Data were collected from seven different cohorts of children; approximately ages birth

(0), 3, 6, 9, 12, 15, and 18 and their primary caregivers. These data were collected by in-

home interviews of subjects and their caregivers. The accelerated longitudinal design

allows an approximation of tracking a single birth cohort for 25 years (Sampson, 1997).

The current study uses data collected during waves 1 and 2 for the 6, 9, 12, and 15

year old cohorts. For wave 1, response rates were as follows for each cohort used: 6 year

old, 75% (980 out of 1,307 screened eligibles), 9 year old, 75.9% (828 out of 1,091

screened eligibles), 12 year old, 74.3% (820 out of 1,103 screened eligibles), and 15 year

old, 71.6% (696 out of 972 screened eligibles). For wave 2, response rates were slightly

higher: 6 year old, 88%, 9 year old, 85.6%, 12 year old, 86.2%, and 15 year old, 82.7%.

There was a total of 6,212 eligible participants for wave 2, and 5,338 participated,

resulting in an overall response rate of 85.9%.

The original sample size for children in cohorts 6, 9, 12, and 15 was 2,850. After

eliminating missing data using listwise deletion for all variables in the analyses the

sample size became N=1,346 subjects, resulting in a retention of 47%. The large amount

of missing data is partially due to the large number of variables used because only

subjects with valid answers to all variables could be included. This sample was used for

all analyses. As mentioned in previous chapters, propensity score matching requires a

large sample to be implemented. The current study's sample size was large enough to

provide matches for a large number of the treated subjects.

Descriptive statistics for the sample are in Table 1. All variables used, including outcomes, treatment indicator, and covariates are included in this table. The sample is 50% male and 50% female, and has an average age of 10 years old. The majority of subjects (86%) are cared for primarily by their mother, and 66% of the primary caregivers are married. Eighteen percent of the sample is white, 48% is Hispanic, 31% is black, and 3% is classified as "other" ethnicity. The average family size is five members, and the average number of siblings is 2. Twenty-six percent of the sample received public assistance in the past year. Approximately 68% of the sample has experienced corporal punishment at least once in the past year. Of these children who experienced corporal punishment, 55% had been "slapped or spanked," 41% had been "pushed or grabbed," and 12% "had something thrown at them." Approximately 9% of the sample experienced more than one of these acts.

Because of the large amount of missing data, comparisons were made between the analysis sample and the original sample, which has between 2500 and 2850 subjects for each variable. The distributions of demographic characteristics were very similar in both of these samples. For instance, 49% of this sample is male, and the average age is 10 years old. Additionally, 55% of the primary caregivers are married, and 84% are the mother of the child. Further, 14% of the sample is white, 47% is Hispanic, 37% is black, and 3% is classified as "other" ethnicity. Regarding public assistance, 29% of these children and their families have received it in the past year. The percentage of the full sample that had experienced each act described by the corporal punishment scale is almost identical to those found in the smaller sample. For the first question (number of times thrown something at subject), 11% of the full sample had experienced this at least

once, and 12% of the smaller sample had experienced this at least once. The second question asked how often the subject was pushed or grabbed; 39% of the larger sample and 41% of the reduced sample experienced this at least once. The last question asked how often the subject was slapped, 60% of the larger sample and 55% of the smaller sample experienced this at least once. The fact that the original sample and analysis sample are so similar on important variables lends more support that the findings are representative of the larger sample. Although there is a large amount of missing data, the sample that remains does not differ drastically from the sample it was drawn from. Implications and consequences of this missing data will be discussed further in the last chapter.

Dependent Variables. Several dependent variables are used for the current study, including externalizing and internalizing behavior scales, as well as their subscales. Internalizing and externalizing behaviors at wave 2 are measured using the Reduced Child-Behavioral Checklist (CBCL) (Achenbach, 1991). The externalizing behavior scale is made up of two subscales, one on delinquent behavior and one on aggressive behavior. The internalizing behaviors scale is made up of three subscales: withdrawal, somatic complaints, and anxious/depressed behaviors (Crijnen, Achenbach, & Verhulst, 1999). To create the instrument, past research was investigated to identify common behavioral problems among children, and items were analyzed using principal components analysis to divide it into subscales. By doing this, the creation of these scales was empirically driven rather than by choosing items that were believed to belong together; they were named to identify the types of behavior represented by the items on each subscale. The externalizing subscales of delinquent and aggressive behavior assess

rule-breaking behaviors and other behaviors that are associated with disorders such as

conduct disorder (CD) and oppositional defiance disorder (ODD) (Sourander & Helstela,

2005). The aggressive behaviors subscale refers to actions termed hostile or destructive,

and the delinquent behaviors subscale is similar but includes more serious offenses such

as truancy. Internalizing behaviors are intended to assess internal stresses. The somatic

complaints scale asks about medical problems that have no known cause. Withdrawal

and anxious/depressed subscales have been indicative of disorders such as avoidance

disorder, overanxious disorder, and social phobia (Aschenbrand, Angelosante, & Kendall,

2005). Items on the withdrawal subscale refer more to symptoms of avoidance and

phobia, while items on the anxious/depressed scale refer more to symptoms of depression

and apprehensiveness about performance. The only difference between the reduced

version of this scale and the full scale is several questions from the externalizing scales

were eliminated because they were deemed to be unnecessary; other items in the scale

described these concepts adequately already.

All questions making up each scale have response categories of either "0," "1," or

"2," indicating the specified behavior is "not true," "sometimes true," or "often true" of

the child, respectively. These scales consist of summated responses to many questions.

Examples of internalizing behavior questions are if the child "sulks a lot," is

"withdrawn," "has headaches without medical cause," "complains of loneliness," and

"feels too guilty." Examples of externalizing behavior questions include "destroys own

things," "gets in many fights," "lies or cheats," "demands a lot of attention." The

internalizing behavior scale has 32 items and externalizing behavior scale has a total of

24 items. The withdrawal scale has 9 items, anxious/depressed has 14 items, and the

somatic complaints scale has 9 items. The aggressive behavior scale contains 13 items

and the delinquent behavior scale has 11 items. The internalizing behaviors scale has

been shown to have reliability coefficient of .89, and the externalizing behaviors scale

has been shown to have a reliability coefficient of .93 in past research (Achenbach,

1991). In the current study, the internalizing and externalizing scales for wave 2 both

have Cronbach's reliability coefficients of .89. Reliability coefficients for subscales are

as follows: withdrawal, .70, somatic complaints, .65, anxious/depressed, .76, aggressive

behavior, .87, and delinquent behavior, .68. On each of these scales, higher scores reflect

higher internalizing behavior, externalizing behavior, withdrawal, somatic complaints,

anxious/depressed behavior, aggressive behavior, or delinquent behavior. Pearson

correlations comparing externalizing and internalizing scales to others such as the Youth

Self Report (YSR) and Teacher's Report Form (TRF) that measure similar concepts have

shown it to have good construct validity (Achenbach, 1991).

Treatment Variable. The main independent variable, or treatment condition, was

measured using a portion of the Conflict-Tactics Scale for Parent-Child (CTSPC) (Kantor

& Jasinski, eds. 1997). The full scale measures how conflict is dealt with between

caregivers and their children, ranging from "talking things through calmly," to spanking,

to more severe abuse. Similar to past research (Straus et al, 1997; Kantor & Jasinski, eds.

1997), an index was created from three of the questions on the CTS concerning corporal

punishment. The questions ask primary caregivers how many times in the past year they

"threw something at him/her," "pushed, grabbed, or shoved him/her," and "slapped or

spanked him/her." The response categories are "0 = never," "1=once," "2=twice," "3=3-

5 times," "4=6-10 times," "5=11-20 times," and "6=more than 20 times." For each of the

three items responses were collapsed into a dichotomy, with "1" indicating at least one time in the past year and "0" indicating none in the last year, and then summed. Those that have a score of "1," "2," or "3" are collapsed into the treatment group, as they were exposed to corporal punishment, and those having a score of "0" are in the control group, indicating no exposure to corporal punishment.

Covariates. The covariates were grouped in several different categories. Individual characteristics of the child include age, race, sex, an impulsivity measure from the EASI temperament survey (Buss & Plomin, 1984), the number of times the child skipped school or class, whether or not the child has had special education for behavioral or emotional problems, and how many days the child has missed school or work due to mental or physical health. Sex of the subject is coded 0 = female, and 1 = male. Race was recoded from the categories Hispanic, Asian, Pacific Islander, black, white, Native American, and other into four dummy variables indicating white, black, Hispanic, or other. The white category was excluded from the analysis as a reference category. A dichotomous question asking whether or not the child has ever had special education for behavioral or emotional problems is coded 0 = no and 1 = yes.

The EASI temperament instrument was designed to measure several types of personality traits in children (Buss and Plomin, 1984). This instrument measures emotionality, activity, sociability, and impulsivity. The impulsivity measure is used in the current study, which is made up of four different subscales measuring inhibitory control, sensation seeking behavior, decision time, and persistence. Buss and Plomin (1984) refer to impulsivity as a child's tendency to resist urges rather than giving in to them quickly. Inhibitory control refers to the ability to delay an immediate response that

arises, or to resist temptations. Sensation seeking behavior is characterized by

unpredictability, and the desire to constantly try new things and have new experiences.

Decision time manifests itself in an impulsive person by quick actions without

contemplating consequences or other possible choices. Persistence refers to the desire to

seek novelty as soon as an activity becomes boring or dull. Examples of questions from

this scale include "I have trouble controlling my impulses" (inhibitory control), "I often

say the first thing that comes into my head" (decision time), "I'll try anything once"

(sensation seeking), and "I tend to give up easily" (persistence). There is a total of 20

questions, all have response categories of 1 = uncharacteristic of my child, 2 = somewhat

uncharacteristic, 3 = neither, 4 = somewhat characteristic, 5 = characteristic (Buss &

Plomin, 1984). Responses to items were summed for each individual, with higher scores

indicating more impulsivity. Responses to items from the four subscales were summed to

create a measure of impulsivity; this measure has a reliability of .76.

Internalizing and externalizing behavior scales from the CBCL at wave 1 were

also included in the list of covariates, to adequately control for previous behavior

problems. The internalizing behaviors scale has reliability coefficient of .85, and the

externalizing behaviors scale has a reliability coefficient of .89 for wave 1. As in the

wave 2 measure, higher scores on these scales indicate more internalizing and

externalizing behaviors. The items for each scale are the same as wave 2 with the

exception of several more items on the externalizing behavior scale. These differences

are noted in the Appendix.

Parental variables include age, sex, socioeconomic status, education, employment

status, and relationship status. The gender of the primary caregiver is coded as '0' female

or '1' male. The socioeconomic status (SES) of the family is coded on a continuous scale, with higher scores indicating higher family SES. This variable was created using the principal component of three variables; household income, maximum education level of primary caregiver and partner, and socioeconomic index (SEI) for primary caregiver's and partner's job ("PHDCN", n.d.). Education of the primary caregiver was recoded from five categories into a set of dummy variables, indicating whether or not the primary caregiver had less than a high school education, some high school, finished high school, or had education past high school. The reference group of "less than high school education" is excluded from the analysis for comparison purposes. Employment status of the primary caregiver was recoded from three categories (employed, unemployed for less than five years, and unemployed for more than five years) to a dichotomous variable indicating '0' unemployed or '1' employed. The relationship status of the primary caregiver to the subject was recoded from a variable containing 19 different categories into a series of four dummy variables. These categories were "mother is primary caregiver," "other female," "father," and "other male." For each variable, '1' indicates the subject's primary caregiver falls into the given category and '0' indicates they have a different relationship. The reference category of "mother is primary caregiver" was left out of this analysis for comparison purposes.

The Conflict Tactics Scale for Partner/Spouse (CTS) is also included. There could be behavioral consequences for a child who witnesses violence in their home, so it is important to control for this. There are several scales included in this instrument that measure ways of dealing with conflict between partners/spouses ranging from reasoning to violence. The scale for physical assault was used in this study. These questions

address violence between partners ranging from hitting to more serious violence such as firing a gun or using a knife. Questions include how many times have you…"threw something at him/her," "pushed or shoved him/her," and "beat him/her up." Each question has response categories the same as the Conflict Tactics Scale for parent and child described earlier, "0=never," "1=once," "2=twice," "3=3-5 times," "4=6-10 times," "5=11-20 times," "6 > 20 times" (Straus, 1979). Each question first asks the primary caregiver how many times they have exhibited a particular behavior in the past year, then asks how many times their spouse has. Two scales were created from these questions, one assessing the primary caregiver's behavior and the other assessing their spouse's. This scale has a reliability coefficient of .86 for the primary caregiver and .85 for items about their partner.

Family variables are also used as predictors, including measures of parental warmth, parental verbal skills, and supervision/monitoring from the Home Observation for Measurement of the Environment (HOME) instrument (Leventhal et al, 2004), whether or not the family is provided with public assistance, the size of household, number of siblings, and the amount of social support available.

The HOME instrument was created to assess aspects of developmental conditions and actions by caregivers that affect a child's well-being (Leventhal et al, 2004). Past research was assessed to create this list of questions, measuring aspects such as parental responsibility, emotional climate, learning materials, parental involvement, and family participation (Leventhal et al, 2004). Three subscales are used from this instrument. The parental warmth scale is made up of observational items coded dichotomously (0 if interviewer did not witness it, 1 if they did) such as "parent encourages child to

contribute" and "parent praises child twice during visit." Higher scores on this index indicate more warmth observed. The items on the parental warmth and parental verbal skills scales were not questions asked to primary caregivers, but rather were coded by observation of the interviewer. This scale has a reliability of .76. The parental verbal skills index is made up of questions coded the same as the warmth scale that assess the verbal and communication skills of the primary caregiver. Higher scores indicate more verbal skills. This scale has a reliability coefficient of .68. The scale measuring supervision/monitoring contains questions asked to the primary caregiver that relate to how well they supervise the subject, such as "subject has a set time (curfew) to be home on school nights" and "establishes rules for behavior with peers and asks questions to determine whether or not they are being followed." The reliability for this scale is .39. While this scale's internal consistency is low, the face validity of this scale indicates that all questions pertain to supervision and monitoring of the child, and the reliability coefficient is a conservative estimate. Based on this, the scale is still included in the model. For the supervision/monitoring scale, higher scores indicate more supervision.

There are several other covariates included that assess aspects of the family. Whether or not the subject has a family member with a criminal record, a family member with frequent legal problems, a family member with nerve problems, or a family member that has attempted suicide are all coded dichotomously, with a '1' indicating 'yes' and '0' indicating 'no'. Public assistance is measured dichotomously by a question asking the primary caregiver if the family has received public assistance in the past year. Both family size and number of siblings under 19 are included as covariates because of the potential for these two questions to measure different family members' presence. These

variables are measured continuously, with higher numbers indicating more siblings or more family members in the household.

The Provision of Social Relations Scale is used to assess the amount of social support available from family and friends (Turner, Frankel, & Levin, 1983). This instrument was designed to measure how much support one receives from others. There are two different subscales, one on family support and one on friend support. These are designed to be combined together to form a measure of overall social support. Respondents answer either 1 = very true, 2 = somewhat true, or 3 = not true to questions such as "I have at least one friend that I could tell anything to" and "No matter what happens, I know that my family will always be there for me should I need them." Responses to all questions on both subscales are summated to create a scale score for each individual, with higher scores indicating a lack of (or less) support (Turner et al, 1983). The reliability coefficient for this scale in the current study is .73.

Analytic Strategy

The first step in a propensity score matching analysis is to identify all important covariates. Based on previous research, many known indicators of whether or not a child will receive corporal punishment are included in the model. Even predictors that may not be statistically significant in a typical regression model are included in this type of analysis. As long as the variable is related in some way to the treatment condition it should be included; because unlike many other multivariate methods, PSM is not sensitive to the number of covariates included. A set of independent sample t-tests is performed for informative purposes only to explore preliminary differences in treatment

and control groups (Hahs-Vaughn & Onwuegbuzie, 2006), or who is exposed or not to corporal punishment.

Following preliminary analyses, the dichotomous treatment variable (corporal punishment) was modeled as an outcome of all the identified covariates using logistic regression. This model estimates a conditional probability, or propensity score, of exposure to corporal punishment for each child (Hahs-Vaughn & Onwuegbuzie, 2006). Stata 9.0 is used to carry out these analyses, specifically the PSMATCH2 program within Stata (Leuven & Sianesi, 2003). Once the propensity scores are created and matching occurs, another set of independent sample t-tests will be estimated, to ensure that a substantial reduction in selection bias is achieved before proceeding to other analyses.

In order to show robustness of the results, three different types of analytic techniques are used to ascertain the effects of corporal punishment on behavioral outcomes. Results from two different matching techniques are estimated, as well as OLS regression results. All of these analyses contain seven different models, each with a different dependent variable. These dependent variables are internalizing and externalizing behavior scales, and each subscale within these: somatic complaints, anxious/depressed, withdrawal, aggressive behavior, and delinquent behavior. The two types of matching are first discussed, then OLS regression, and last the estimation of the treatment effects.

The first type of matching used was nearest-neighbor, a one-to-one matching strategy. This technique takes the first treated subject on the list and then pairs them with the untreated subject with the closest propensity score. These two subjects are then removed from the list and the matching continues until the control units are all used or

there are no adequate matches. After the matching occurs, the matched subjects are compared on the outcome variables listed above.

The second type of matching used was kernel matching. This technique also begins with the first treated subject on the list and matches it with an untreated subject. However, matches are made based on a weighting process that takes into account the outcome measure for all subjects to find the best match(es). The weighted average of the outcomes for the untreated units is estimated and matches are chosen based on this and how close the untreated unit is to the treated unit (Sianesi, 2001). This technique allows more treated subjects to be included in the analysis than nearest-neighbor. After matching, the treatment and control group are compared on each outcome variable.

The first analysis is an Ordinary Least Squares (OLS) regression, with the propensity score and the treatment variable (corporal punishment) as independent variables. These models do not contain any other independent variables because the effect of all other covariates is contained within the propensity score. By including the propensity score as an independent variable, any significant effects that remain for corporal punishment represent its causal effect on the dependent variable. Any other independent variable that may have been included was already accounted for in the estimation of the propensity score.

Both types of matching will also provide an estimation of the "ATT", or average treatment effect. This will illustrate the differences between treatment and control groups both before and after matching. The ATT will be displayed both visually as figures and also numerically in a table. The difference between the treatment and control groups

after matching represents the causal effect of corporal punishment on internalizing and externalizing behaviors, including their respective subscales.

This chapter has presented information about the data, variables, and methods used for the current study. Each measure has been described in detail and the analytic strategy was then explained. The following chapters will present results from these analyses, as well as discussion and conclusions.

CHAPTER 5

RESULTS

This chapter will present results from all analyses used to test each of the research

hypotheses. First, independent samples t-tests will be reported to illustrate the pre-

matching differences between treatment and control groups. Second, results from a

logistic regression used to estimate propensity scores are then presented. Lastly, results

from a series of Ordinary Least Squares (OLS) regressions and the PSM models are

discussed. Each model is explained separately with corresponding tables or figures.

Throughout this chapter, results referred to as statistically significant are those with an

associated p-value of .05 or less, and/or a t-value of 2.0 or greater in absolute value.

T-tests and Matching Procedure

Results from pre-exposure/pre-matching independent samples t-tests are shown in

Table 2. This table shows that before matching occurs, there are numerous covariates

that exhibit statistically significant differences between treatment (exposed to corporal

punishment) and control (not exposed to corporal punishment) groups. The children

exposed to corporal punishment are significantly younger than those in the control group

(D=1.43; t-value=-7.86; p= .00). Children who experienced corporal punishment were

also more likely to be on public assistance (D=.05; t-value=2.00; p=.05) and have a male

primary caregiver (D=.05; t-value=-2.99; p=.00) than those not experiencing corporal

punishment. Children who experienced corporal punishment had primary caregivers with

slightly higher scores on verbal skills (D=.07; t-value=2.84; p=.01) and also on the

provision of social relations scale (D=.64; t-value=2.55; p=.01) than primary caregivers

that did not spank their children. Children exposed to corporal punishment were less

likely to have "other female" defined as their primary caregiver than those in the control group (D=.02; t-value=-2.01; p=.05). Children who experienced corporal punishment were also more likely to have received help in school for emotional problems (D=.03; t-value=2.96; p=.00). Children exposed to corporal punishment were more likely than those not exposed to have a family member with frequent legal problems (D=.05; t-value = 2.21; p=.03), a family member with nerve problems (D=.07; t-value=2.81; p=.01), a family member with a criminal record (D=.10; t-value=3.97; p=.00), and a family member who had attempted suicide (D=.04; t-value=1.98; p=.05).

A statistically significant difference was also observed for age of primary caregiver, the caregivers of children who were exposed to corporal punishment were approximately three years younger than the caregivers of those not spanked (D=2.78; t-value=6.80; p=.00). The mean scores on the conflict tactics scale for primary caregiver (D=1.18; t-value=3.83; p=.00), as well as the questions about their spouse's behavior (D=1.08; t-value=4.14; p=.00) were, on average, significantly higher for those children who were exposed to corporal punishment, indicating that children who were punished, on average, had more conflict in the home than their counterparts who were not punished.

The values for internalizing (D=2.26; t-value=5.92; p=.00) and externalizing (D=5.87; t-value=12.05; p=.00) behaviors scales were much higher for the treatment group as compared to the control group of children, indicating that children exposed to corporal punishment, on average, exhibited more externalizing and internalizing problems than those not exposed. Children in the treatment group also had higher scores for impulsivity than the control group (D=5.88; t-value=9.04; p=.00), indicating that children exposed to corporal punishment were, on average, more impulsive than those not

exposed. Children who experienced corporal punishment were less likely to have "dad" defined as their primary caregiver (D=.05; t-value=-3.12; p=.00).

The statistically significant results discussed above indicate that there is a large amount of bias between who is exposed to corporal punishment and who is not, a propensity score matching strategy is designed to eliminate this bias. On the other hand, some covariates did not show statistically significant differences between these two groups, including ethnicity, SES, family size, number of siblings, all education categories, "other male" defined as primary caregiver, primary caregiver employment and marital status, supervision/monitoring scale, parental warmth scale, truancy, and whether or not a family member had attempted suicide.

The next step in the analysis was to create the propensity score. Results from the logistic regression performed to do this can be found in Table 3. The logistic regression creates a conditional probability of experiencing corporal punishment, based on the list of covariates defined. As discussed in a previous chapter, research has shown that many of these covariates have an impact on whether or not a child experiences corporal punishment. Even if a covariate is only weakly related to exposure to corporal punishment, the weakness of this relationship is overcome by the reduction in bias it provides (Newgard et al, 2004).

As shown in Table 3, several covariates are statistically significant predictors of exposure to corporal punishment; again indicating that exposure to corporal punishment is not randomly experienced among these kids, but rather shows that selection bias is present. Selection bias was shown in the pre-matching, independent samples t-test reported earlier, but it is important to note that it still exists using a multivariate analysis

that controls for other covariates. Most notably the externalizing behaviors scale for

subjects taken at wave 1 shows a positive and statistically significant effect, indicating

that children who exhibit more externalizing behavior symptoms are more likely to

experience corporal punishment (b = .08; z = 6.22; p=.00), regardless of other covariates.

Age has a negative and statistically significant effect on exposure to corporal punishment,

indicating that younger children are more likely to experience corporal punishment (b = -

.13; z = -5.55; p=.00). Contrary to what was expected, children are significantly more

likely to experience corporal punishment if their primary caregivers have more verbal

skills (b =.33; z = 2.10; p=.02) and are male (b = -.57; z = -2.49; p=.01). The conflict

tactic scale for spouse has a positive and statistically significant effect on corporal

punishment, indicating that children are more likely to be exposed to corporal

punishment if their primary caregiver reports more conflict by their spouses (b = .05; z =

2.41; p=.01). Surprisingly, those families with fewer children are more likely to have a

child that experiences corporal punishment (b = -.14; z = -2.54; p=.01). Those children

who have a family member with a criminal record are also more likely to experience

corporal punishment (b =.49; z = 2.63; p=.01). Other covariates included did not show

statistically significant effects on corporal punishment, although as mentioned previously,

when using propensity score matching it is still important to keep them in the analysis if

there is adequate support for their affect on the treatment condition

     After estimating propensity scores, the first matching was done using nearest

neighbor one-to-one matching without replacement. As a result, there were 476 treated

observations dropped because they were above the maximum propensity score of the

control group. Those subjects that were dropped had a minimum propensity score of .76,

and the subjects that were used had a minimum of .05, indicating that no observations were dropped from the "bottom" of the sample. This resulted in a sample of N = 870 subjects that were matched pairs.

After matching, another set of independent sample t-tests was performed to ensure that selection bias was reduced; these results are in Table 4. As shown in Table 4, once children were matched on their conditional probability or propensity for being punished, statistically significant differences no longer existed between treatment and control groups for any of the covariates. Also included in this table is the percentage bias reduction that was achieved after matching. A large amount of the bias was eliminated for several variables, including whether or not family member has a criminal record (91%), nerve problems (82%), frequent legal problems (86%), attempted suicide (100%), impulsivity (89%), wave 1 internalizing (95%) and externalizing behavior (99%) scores, primary caregiver marital status (100%), subject age (73%), black ethnicity status (93%), primary caregiver had some high school (93%), conflict tactics scale (99%), conflict tactics scale for spouse's behavior (97%), and whether or not the child had school help for emotional problems (85%). There was at least some bias reduction for almost all other covariates, except SES, family size, "other male" defined as primary caregiver, primary caregiver employment status, parental supervision/monitoring, and parental warmth. It is possible that for some of these covariates, the sample was already adequately matched between treatment and control groups, resulting in less bias reduction for these covariates.

Treatment and Control Group Comparisons

Results are now presented for each type of analysis comparing treatment and control groups on behavioral outcomes at wave 2 assessment. These results directly test hypotheses outlined in chapter 2. As stated in chapter 2, research hypotheses predicted that there are direct causal effects of corporal punishment on both internalizing and externalizing behaviors, as well as their separate subscales. More specifically, it is expected that corporal punishment will increase both internalizing and externalizing behaviors at wave 2.

OLS Regression Results. Table 5 shows results from the OLS regression. The propensity score is used as an independent variable in this type of analysis. However, the only other variable included is the treatment indicator (dichotomous corporal punishment measure) because the propensity score contains the effect of all other covariates. The propensity score was calculated by taking into account the effect of the important covariates already defined. The coefficient or effect of this score cannot be interpreted in these analyses because it is not possible to disentangle each separate effect. The propensity score combines effects of all covariates into one conditional probability; therefore it is not possible to separate effects of one covariate from another. Given this information, a series of OLS models is another way to confirm results from other treatment and control group comparisons on behavioral outcomes at wave 2. Each outcome variable was modeled separately as a dependent variable, including the internalizing and externalizing scales as well as their separate subscales. As indicated in Table 5, corporal punishment does not have a statistically significant effect on any of the outcomes evaluated. In each model, the propensity score is statistically significant, but

again this effect cannot be directly interpreted, it only indicates that the combination of all covariates has an effect on the behavioral outcome. Ultimately, these analyses show that exposure to corporal punishment doesn't have a causal effect on internalizing and externalizing behavioral outcomes, once controlling for matching[1].

Nearest-Neighbor Matching Results. The second model used the one-to-one matches made previously and compared the means of treatment and control groups before and after matching. These matched pairs represent groups that are balanced on covariates, which provide a less biased estimate of the treatment effect or the average effects of corporal punishment. Results from these analyses are in Table 6. The first two columns indicate differences in the mean score on behavioral outcomes at wave 2 for those exposed to corporal punishment (treatment group) and those not exposed (control group). As illustrated, there are significant differences for all behavioral outcomes before nearest neighbor matching occurred, with t-values well above 2.0. For internalizing behaviors, children who experienced corporal punishment have higher scores than those who did not (D=2.14; t-value=4.78). Children who experienced corporal punishment also have higher scores on externalizing behaviors than those who have not (D=3.30; t-value=7.83). Those in the treatment group also have higher scores than the control group for withdrawal (D=.59, t value=3.69), somatic complaints (D=.43; t-value=3.22), anxious/depressed behavior (D=3.52; t-value=4.63), aggressive behavior (D=2.33; t-value=8.26), and delinquent behavior (D=.97; t-value=5.94).

The ATT refers to the "average treatment effect" of corporal punishment on each of the behavioral outcomes assessed at wave 2 after matching was performed, and shows

---

[1] In order to account for skewness present in the distribution of behavioral outcome variables, a series of negative binomial regression analyses was also conducted. Results confirmed the OLS regression output and therefore are not included.

the difference between treatment and control groups. For each outcome all significant differences have been eliminated, the mean scores are very similar and none are statistically significant. After the matching occurred, the difference between treatment and control group became very small; for internalizing behaviors the difference is .42, and for externalizing behaviors the difference is 04. Differences for each subscale are as follows: withdrawal, .19; somatic complaints, .20; anxious/depressed, .02; aggressive behavior, .16; delinquent behavior, .13. These results indicate that there is no significant effect of corporal punishment on the behavioral outcomes tested. Figure 1 indicates in graph form the reduced difference in externalizing behaviors before and after matching, and figure 2 shows these results in graph form for internalizing behaviors. Figures 3-7 illustrate these changes in mean differences for each subscale. These graphs are included to visually express what is contained in Table 6, and they were created using the mean scores from this table.

Kernel Matching Results. To assess the robustness of the results reported above, the final analyses were estimated using kernel matching, similar to caliper matching. This technique uses weighted averages of the outcomes to determine which non-treated unit(s) is the best match for treated subjects within a fixed area surrounding the treated subject. Caliper matching uses the same process but without weighting. As mentioned in the PSM chapter, the type of matching used normally does not have differing effects on results, as long as there is adequate overlap in the distributions of propensity scores between the treatment and control group. The weighting process used in kernel matching takes into account outcomes and therefore may be slightly better, but in general caliper and kernel matching will provide very similar results. Using this matching technique

allows a larger number of subjects to be included, there are only 41 observations eliminated, 5 control and 36 treated. The logistic regression for this analysis is not included, because results are very similar to Table 3.

Results from the comparison of matched pairs using this technique are in Table 7. In the case of each behavioral outcome, significant differences in mean scores for treatment and control groups existed before matching. Those children who had experienced corporal punishment had higher scores on both internalizing (D=2.14; t-value 4.78) and externalizing behaviors (D=3.30; t-value=7.83). Those in the treatment group also had higher scores on each subscale, withdrawal (D=.59; t-value 3.69), somatic complaints (D=.43; t-value=3.22), anxious/depressed behavior (D=1.13; t-value=4.63), aggressive behavior (D=2.33; t-value=8.26), and delinquent behavior (D=.97; t-value=5.94).

After matching, the ATT indicates that corporal punishment has no significant effect on internalizing and externalizing behaviors, or their subscales. As illustrated, there are no significant differences between treatment and control groups after matching. The differences between treatment and control groups after matching are no longer statistically significant. The mean values for the control group are slightly higher for all outcomes except somatic complaints in this model, even though these differences are not considered statistically significant. Mean values were higher for both treatment and control groups for all outcomes than those found with nearest-neighbor matching, probably due to the increased sample size with this type of matching. OLS regression models were also performed with each outcome as a dependent variable using the propensity scores from the kernel matching technique. Results were the same as those

presented in Table 5, with no significant effects of corporal punishment and identical

values for the coefficients of propensity scores in each model.

However, the results of the pre- and post- matching t-tests from the sample using

kernel matching were somewhat different. The pre-matching independent samples t-tests

are the same because the same sample is used (refer again to Table 2). After matching

occurred, statistically significant differences for several covariates between the treatment

and control groups were still present, results from these t-tests are in Table 8. Those

children who received corporal punishment were slightly younger than those who did not

(D=.02; t-value=6.50; p=.00). Children who experience corporal punishment were also

more likely to be black (D=.07; t-value=-2.11; p=.04) and also more likely to be "other"

ethnicity (D=.01; t-value=2.36; p=.02). Those children in the treatment group were more

likely to be on public assistance than those in the control group (D=.07; t-value=-3.49;

p=.00). The children who experienced corporal punishment had younger primary

caregivers than those who did not (D=.36; t-value=5.23; p=.00) and were more likely to

be male (D>.01; t-value=3.19; p=.00). Those children who were in the treatment group

were less likely to have "dad" defined as their primary caregiver (D=.01; t-value=3.18;

p=.00). Children who experienced corporal punishment also had primary caregivers with

higher scores on the conflict tactics scale for their behavior (D=.03; t-value=-3.16; p=.00)

and spouse's behavior (D=.28; t-value=-2.94; p=.00). The treatment group also had

higher scores on wave 1 internalizing (D=.36; t-value=-4.66; p=.00) and externalizing

behaviors (D=.73; t-value=-10.63; p=.00) than the control group. The children who

experienced corporal punishment had slightly lower scores for impulsivity than those

who had not experienced corporal punishment (D=.58; t-value=-7.85; p=.00). Children

who experienced corporal punishment were more likely to have been truant (D=.03; t-value=-3.11; p=.00). The children exposed to corporal punishment were also more likely to have a family member with a criminal record (D=.01; t-value=-4.07; p=.00), a family member with nerve problems (D=.01; t-value=-1.93; p=.05), and a family member with frequent legal trouble (D=.02; t-value=-2.85; p=.00). Even though these significant differences remain, it is likely due again to the increased sample size because as presented, the actual mean values for these two groups are very close on all covariates.

There was still a large amount of bias reduction for many of the covariates, including primary caregiver gender (100%), impulsivity (90%), supervision/monitoring (93%), conflict tactics scale (97%), conflict tactics scale for spouse's behavior (74%), both internalizing (93%) and externalizing (88%) behaviors, primary caregiver age (87%), "dad" defined as primary caregiver (92%), "other female" defined as primary caregiver (94%), provision of social relations (61%), subject age (99%), subject gender (82%), and whether or not a family member has criminal record (99%) and nerve problems (76%). Some amount of bias was reduced for all other covariates except black and Hispanic ethnic categories, family size, number of siblings, public assistance, primary caregiver graduated high school, "other male" defined as primary caregiver, primary caregiver marital status, parental warmth, and truancy. Caliper matching was also used to estimate treatment effects, but results are not presented here for the sake of redundancy. Results were essentially identical to those found with kernel matching.

This chapter has described all results from this study. Results from each step of the analytic strategy outlined in the previous chapter were presented. The next chapter

will provide discussion of these results and conclusions that were drawn from these

analyses.

CHAPTER 6

DISCUSSION & CONCLUSION

This final chapter will discuss the results presented previously and conclusions

that can be drawn from these results. Also presented in this chapter is a section

containing limitations of this study and future research directions given the findings of

the current study. The final section will conclude this study.

Discussion

This study assessed the causal effects of corporal punishment on behavioral

outcomes for children using one of the largest, most ambitious, data collection efforts in

the history of sociology, the Project on Human Development in Chicago Neighborhoods.

Past research has shown that corporal punishment is related to a host of negative

outcomes, including antisocial behavior (Grogan-Kaylor, 2005a; Grogan-Kaylor, 2004;

Straus & Mouradian, 1998), depressive symptoms (Turner & Muller, 2004), and

psychological distress (Turner & Finkelhor, 1996). Although a small body of research

shows otherwise (Larzelere, 2004), most research supports the argument that negative

behavioral outcomes are associated with experiencing corporal punishment. However, as

discussed in previous chapters, most of these studies have been unable to discern the

direction of causality and accurately estimate the causal effects of corporal punishment,

largely due to the fact that they did not approximate experimental conditions. The

current study has moved closer towards doing this by using a PSM method, which

matches children based on their probability of being punished before comparing them on

behavioral outcome measures.

The results from the current study are contradictory to past research. Most of this research shows that corporal punishment has negative consequences for children. First, a large amount of selection bias was present before the matching occurred which led to children who displayed certain behaviors and had certain characteristics to be much more likely to experience corporal punishment. Second, as a way to adjust for selection bias, the two groups of children (one having experienced corporal punishment and the other having not) were paired so that their probabilities of experiencing corporal punishment were very similar. As such, more accurate estimates of the causal effects of corporal punishment could be estimated, resulting in findings inconsistent with what many studies conclude regarding the negative consequences of corporal punishment.

As shown in the results chapter, the current study indicates that corporal punishment does not have an effect on behavioral outcomes. All hypotheses presented in Chapter 2 were not supported. After the matching occurred, there were no significant differences between the treatment and control groups of children, indicating that they did not differ on behavioral outcomes. This finding was present for internalizing behaviors and externalizing behaviors, as well as their subscales—withdrawal, anxious/depressed behavior, somatic complaints, delinquent behavior, and aggressive behavior. In each case, the treatment and control differences disappeared after the subjects were adequately matched.

A series of multivariate regression analyses upheld these results. For each behavioral outcome, the dichotomous corporal punishment variable did not have a statistically significant effect on any of the outcomes of interest. The propensity score affected these outcomes significantly and had large coefficients, but as mentioned

previously this number cannot be directly interpreted. It only indicates that the

combination of all covariates affects the behavioral outcome. By using a propensity

score approach, the results of the current study are more robust than those from prior

research. The findings across two different types of matching were virtually identical,

lending support to their accuracy.

Only one part of the analyses showed that there were remaining differences

between treatment and control groups after matching. The results from the second set of

independent sample t-tests after kernel matching indicated that there were still significant

differences between some of the covariates, but this is likely due to the increased sample

size. With a sample of over 1,300 children, even very small differences become

significant. The actual mean values for treatment and control groups after matching are

very close to one another. For example, the mean score on internalizing behaviors for the

treatment group is 8.86, and for the control group it is 9.16. The scores for externalizing

behaviors are 9.03 for the treatment group and 9.62 for the control group. The model

containing somatic complaints as the dependent variable showed identical values of 1.77

as the mean for both treatment and control groups, but the p value for this model was .00,

suggesting there were significant differences between these two groups.

Both types of matching provided substantial bias reduction for many of the wave

1 covariates. Importantly, they both resulted in greater than 85% reduction for the wave

1 measures of the outcome variables of internalizing and externalizing behaviors. This

bias reduction was necessary in order to show that corporal punishment was the cause (or

not the cause) behind behavior problems measured at wave 2. One of the central

arguments of those against corporal punishment is that it causes behavior problems such

as the ones measured in this study. However, if it is true that these behavior problems existed prior to experiencing that punishment, this argument may not carry much weight. The bias reduction achieved through PSM ensures that this effect is more accurately estimated. Children were only compared with others exhibiting similar behavioral, family, and individual characteristics measured at wave 1.

For example, imagine a hypothetical child, "Chris," who has numerous behavior problems and acts out at school and at home. Now imagine another child, "Mark," who also displays similar behaviors as Chris. Their externalizing behavior scores are comparable when taken at a specific time period. Mark and Chris are very alike in other ways, such as their age, ethnicity, and family characteristics. The main difference between the two is that Chris' parents spank him for bad behavior and Mark's parents do not. When Mark's and Chris' externalizing behavior scores are compared at a later time period, possibly a year later, they are still very similar. If this is true, then it is difficult to argue that the corporal punishment experienced by Chris is the cause of his behavior. Propensity score matching does exactly what this example describes, but with a large group of children and a large number of covariates. This study replicated on a larger scale this example, indicating that children who exhibit internalizing and externalizing behavior problems do not experience an increase in these behaviors as a result of corporal punishment.

The results of this study are opposite of what was predicted, and opposite of findings from a large amount of research on corporal punishment, thus indicating that more research is needed. This relationship has not been fully investigated, as evidenced by the current study's discrepant findings. Although the current study makes several

improvements over past research on corporal punishment, it is not without limitations.

The next section in this chapter will discuss these limitations and make suggestions for

future research.

Limitations and Future Research

Although this study was designed to better assess the causal effects of corporal

punishment, it does have limitations. These limitations point to directions for future

research to better understand the effects corporal punishment has on childhood outcomes.

Future attempts to disentangle these effects will hopefully utilize what was learned in this

study and incorporate ways of overcoming the obstacles presented below.

First, the corporal punishment measure used for the treatment indicator was

condensed to a dichotomous variable, children experiencing one or more of three types of

corporal punishment were classified as being in the same group. It did not contain

information about frequency or severity of the punishment used. Had either of these

factors been taken into account, results may have been different. It is possible that those

children who are spanked very frequently by their parents have more behavioral problems

than others. The same could be true for more severe punishment. Further research

should investigate the differential effects of corporal punishment based on severity and

frequency. Wissow (2001) found that parents who spank more frequently also showed

less nurturing than other parents and also were more likely to use other types of discipline

practices. This more frequent spanking combined with less nurturance is an important

finding that suggests the need for further investigation. Larzelere and Kuhn (2005) also

found in a meta-analysis that the effect sizes for physical punishment were worse than

other discipline practices when it was very severe or used as the predominant disciplinary

method. Frequency and severity are important in assessing the way a child is affected by exposure to corporal punishment.

The questions in the PHDCN data provide more detailed information than what was used in this study, and using this more detailed data could provide important results. One possible way to investigate this issue is by estimating a series of propensity scores using several dichotomous variables that each indicates one distinct frequency of punishment. For example, four different propensity scores would be generated, the first using a dichotomous variable where "0" = other number of punishment, and "1" = at least one instance of punishment. The second propensity score is generated using a dichotomous treatment condition where "0" = other number of punishment, and "1" = two instances of punishment; the creation of propensity scores would continue in this way until all available frequency categories had been used. The next step would be to estimate an ordinary regression model with a set of dummy variables indicating each frequency category. This type of analysis would allow the effects of each frequency category to be compared to one another.

Second, the current study explores the direct causal effect of corporal punishment, but it is possible that this is not the most relevant nor important relationship. Simons et al (2000) found that a lack of warmth from parents increased the negative effects of corporal punishment. The current study did not investigate any interaction effects between corporal punishment and other factors such as parenting. Perhaps corporal punishment does have an effect on behavioral outcomes of children in the PHDCN study, but the effects might only be observed under certain parenting conditions. It has also been suggested that when corporal punishment is administered impulsively by parents the

effects are stronger (Straus & Mouradian, 1998). These interactions between corporal

punishment and parenting styles should be investigated more closely. The PHNCN data

contains information about parental warmth (which was used as a covariate in this study),

future research could assess the interaction between these values and corporal

punishment.

Third, previous research has indicated that younger children (especially toddlers)

experience more corporal punishment (Dietz, 2000; Giles-Sims et al, 1995; Grogan-

Kaylor & Otis, 2007). Dietz (2000) found that children less than 6 years old were more

likely to be spanked, and Giles-Sims et al (1995) showed that children between 3 and 5

years old were spanked more frequently than other age groups. Because of differences in

instruments given to each cohort of children, the three-year old cohort was not included

in this study. Several important covariates were measured differently with this cohort. It

is possible that this group of children could affect the results of the study had they been

used because younger children are more likely to be spanked. This also relates to the first

limitation presented in this section; if the three year old cohort received more frequent or

more severe corporal punishment their effects could be different than other cohorts of

children. If this cohort had been included there may have been effects present that were

not seen in the results of this study. The three year old children in the PHDCN data had

the highest percentage of experiencing corporal punishment out of all cohorts. Future

research with this data set should take into account this cohort of children.

Fourth, another limitation of this study was the large amount of missing data. As

mentioned previously, the large number of covariates used contributed to this because

only cases that had valid answers to all questions were included. The missing data issue

could have affected results significantly if these data were not were not missing at random. The large number of cases missing could indicate there is a systematic reason as to why they are missing, in which case results could be biased. However, Allison (2001) recommends listwise deletion as an appropriate way to handle missing data when using regression if the value of the missing variable does not depend on the dependent variable. As mentioned previously the characteristics, including distributions of corporal punishment and behavioral outcome variables, were very similar for the analysis sample and original larger sample. For regression analysis, he suggests this type of strategy is better than multiple imputation or maximum likelihood, two ways to infer missing values. In the future, one of these methods could be employed to increase sample size. Though, as mentioned in the previous chapter, when the original sample was analyzed it did not differ greatly from the smaller sample on important demographics. This provides some support that the results are still representative.

Fifth, although several behavioral outcomes were assessed in the current study, other outcomes could not be investigated because of the lack of consistent outcome measures across cohorts. For instance, cohorts 9, 12, and 15 contain a measure of self-reported offending which asks questions about crime ranging from minor offenses to violent ones. Since the six-year old cohort did not answer this portion of the study, it could not be included as an outcome measure. While empirical research investigating this relationship is not common, Straus (1996) argues that corporal punishment is one way in which children learn violence is an acceptable way to resolve conflict, and this belief is internalized at a young age. He argues that physical punishment contributes to a more violent society. Future research may provide important information about the

relationship between corporal punishment and offending behavior. There are also other behavior outcomes (such as oppositional defiance disorder) and cognitive outcomes that were not investigated in this study.

Sixth, this study used the additive effects of a set of covariates to predict the likelihood of experiencing corporal punishment, but did not address the differential impact of these covariates, such as race. As pointed out in the results chapter, the propensity score takes into account all of these covariates together. However, there are some covariates (such as race and gender) that may be more important than others. A propensity score matching analysis could be used to investigate this further by splitting the sample by race or gender before estimating differences between treated and control groups, as corporal punishment may have a causal effect for males, but not females or blacks, but not whites. If some covariates have more impact than others on whether or not a child is physically punished these aspects should be concentrated on in research. Past research has shown that the demographic characteristics of race and gender are influential in who is physically punished, which could in turn effect behavioral outcomes. Several studies have indicated that both African-American and male children are more likely to be spanked (Day et al, 1998; Giles-Sims et al, 1995; Dietz, 2000). Day et al (1998) used NSFH data to divide parents into groups based on race and marital status, results showed across most groups that boys were spanked more often than girls. Giles-Sims et al (1995) also showed that boys were more likely to be spanked and black parents were more likely to spank their children, although neither of these groups showed significant differences in how often they were spanked.

Seventh, although the PSM analytic strategy is an improvement over other methods that attempt to estimate causal effects, the propensity score is still based only on observable characteristics that are used as covariates. It is possible that there are covariates either not observed or that were not included in this analysis that could have important effects on the likelihood of a child experiencing corporal punishment, and/or behavioral outcomes. Specific information about both child abuses that may occur in the home as well as whether or not the parent has a history of child abuse could potentially impact these findings. It is also possible that genetic factors influence results. For example, parents who have short tempers and/or are impatient and quick to use physical punishment may have children who also impatient and impulsively behave badly. These similarities may not be accurately accounted for using an observable measure. Future research should investigate possible important covariates that were not included such as these.

Eighth, the results could be affected by the types of matching techniques used in this study. When nearest-neighbor matching was used, there was a large number of treated subjects (470) eliminated from the analysis because they fell outside the distribution of propensity scores for the control group. These children were those that had high propensity scores, indicating a high likelihood of receiving punishment. If these children had been included in the analysis and in fact, had a match in the control group, results may have been different. However, treatment and control groups were also compared on outcome variables using both kernel and caliper matching and there were still no statistically significant effects of corporal punishment. Both of these techniques only eliminated a small number of observations, proving more support results.

Ninth, the generalizability of results from this study is an important consideration. Because the data were collected in Chicago and contain only information about children and caregivers living in that particular study, it is not possible to assert these findings would hold across all locations. It is possible that there are specific conditions or characteristics unique to the residents of Chicago that produced the findings of the current study. In the future, more research in other locations should be done in order to learn more about the generalizability of these findings.

Lastly, the analyses in this study investigated differences between children measured for two waves in an interval of 2.5 years apart as opposed to assessing intra-individual change and continuity. It did not assess trajectories of internalizing and externalizing behaviors of these children over a longer developmental time period, as this would require at least three different measurement times. Perhaps children exposed to corporal punishment may have quite different longitudinal patterns of internalizing and externalizing behaviors compared to their counterparts not exposed. Growth curve analyses could perhaps shed light on how children's behavioral trajectories are influenced by corporal punishment. Straus and Kantor (1994) found that the more corporal punishment a teenager experienced, the more likely they were to be abusive as adults. Grogan-Kaylor (2005a) also found that corporal punishment had more of an effect on antisocial behavior when it was used on teenagers. This research supports differential effects of physical punishment based on the age of the child. An analysis of growth curves could better discern such patterns of behavioral development as it relates to corporal punishment experienced in childhood.

Conclusion

Although much of the past research indicates there is a relationship between corporal punishment and behavioral outcomes, this study does not support this finding. The direct causal effect of corporal punishment was estimated by using a more appropriate statistical method, propensity score matching. This method allowed children to be matched based on pre-existing covariates that affect their likelihood of experiencing corporal punishment or not. After these matches were made, differences on several behavioral outcomes were estimated. Results indicate that there is no causal effect of corporal punishment. However, this study contains several limitations and future research is needed to fully understand this relationship.

In conclusion, this study has added significant information to the body of research on corporal punishment. The findings should be interpreted with the limitations presented above in mind, but nonetheless as important. Rather than punishment received, it appears that behavioral characteristics and other covariates already observed account for differences in later externalizing and internalizing behaviors. There is still sufficient room for further research on this topic however, due to the strong associations present in past research that supports effects of corporal punishment. In closing, this study has laid the foundation for a modest research agenda on corporal punishment, with hopes that answers to some of these questions will be ascertained in the near future.

REFERENCES:

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Burlington, VT: University of Vermont Department of Psychiatry.

Afifi, T. O., Brownbridge, D.A., Cox, B.J., & Sareen, J. (2006). Physical punishment, childhood abuse, and psychiatric disorders. *Child Abuse and Neglect 30*, 1093-1103.

Allison, P.D. (2001). *Missing Data*. Sage University Papers Series on Quantitative Applications in the Social Sciences, 07-136. Thousand Oaks, CA: Sage.

Aschenbrand, S.G., Angelosante, A.G., & Kendall, P.C. (2005). Discriminant Validity and Clinical Utility of the CBCL With Anxiety-Disordered Youth. *Journal of Clinical Child & Adolescent Psychology 34*, 734-746.

Benjet, C. & Kazdin, A.E. (2003). Spanking children: the controversies, findings, and new directions. *Clinical Psychology Review 23*, 197-224.

Biller, H. B. & Soloman, R. S. (1986). *Child Maltreatment and Paternal Deprivation*. Lexington, MA: Lexington Books.

Bingenheimer, J. B., Brennan, R.T. & Earls, F.J. (2005). Firearm Violence Exposure and Serious Violent Behavior. *Science 308*, 1323-1326.

Bitensky, S. H. (1998). Spare the Rod, Embrace our Humanity: Toward a New Legal Regime Prohibiting Corporal Punishment of Children. *Michigan Journal of Law Reform 31*, 353.

Brennar, V. & Fox, R.A. (1998). Parental Discipline and Behavior Problems in Young Children. *The Journal of Genetic Psychology 159*, 251-256.

Buss, A. & Plomin, R. (1984). *Temperament: Early Developing Personality Traits*. Hillsdale, NJ: Erlbaum.

Campbell, D.T. & Stanley, J.C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally College Pub. Co.

Couture, L.A. (2006). Child Advocate.org website, sections concerning laws on corporal punishment. Retrieved February 13, 2007. (www.childadvocate.org)

Crijnen, A.A.M., Achenbach, T.M. & Verhulst, F.C. (1999). Problems Reported by Children of Parents in Multiple Cultures: The Child Behavior Checklist Syndrome Constructs. *American Journal of Psychiatry 156*, 567-574.

Day, R. D., Peterson, G.W., & McCracken, C. (1998). Predicting Spanking of Younger and Older Children by Mothers and Fathers. *Journal of Marriage and Family 60*, 79-94.

Dehejia, R. H. & Wahba, S. (2002). Propensity Score-Matching Methods for Nonexperimental Causal Studies. *The Review of Economics and Statistics 84*, 151-161.

Dietz, T. L. (2000). Disciplining Children: Characteristics Associated with the Use of Corporal Punishment. *Child Abuse & Neglect 24*, 1529-1542.

Donnelly, M. (2005). Putting Corporal Punishment of Children in Historical Perspective. In M. Donnelly & M.A. Straus (Eds.), *Corporal Punishment of Children in Theoretical Perspective* (pp. 41-54). New Haven, CT: Yale University Press.

El-Bassel, N., Gilbert, L. Wu, E. Go, H., & Hill, J. (2005). Relationship Between Drug Abuse and Intimate Partner Violence: A Longitudinal Study Among Women Receiving Methadone. *American Journal of Public Health 95*, 465-470.

Ford, R.G. (1950). Population Succession in Chicago. *American Journal of Sociology 56*, 156-160.

Foster, E. M. (2003). Propensity Score Matching: An Illustrative Analysis of Dose Response. *Medical Care 41*, 1183-1192.

Gangl, M. (2006). Scar Effects of Unemployment: An Assessment of Institutional Complementarities. *American Sociological Review 71*, 986-1013.

Gershoff, E. T. (2002). Corporal Punishment by Parents and Associated Child Behavior Experiences: A Meta-Analytic and Theoretical Review. *Psychological Bulletin 128*,539-579.

Giles-Sims, J., Straus, M.A., & Sugarman, D.B. (1995). Child, Maternal, and Family Characteristics Associated with Spanking. *Family Relations 44*, 170-176.

Guo, S., Barth, R., & Gibbons, C. (2004). *Introduction to Propensity Score Matching: A New Device for Program Evaluation*. Presented at the Annual Conference of the Society for Social Work Research, New Orleans.

Grogan-Kaylor, A. (2004). The Effect of Corporal Punishment on Antisocial Behavior in Children. *Social Work Research 28*, 153-162.

Grogan-Kaylor, A. (2005a). Corporal Punishment and the Growth Trajectory of Children's Antisocial Behavior. *Child Maltreatment 10*, 283-292.

Grogan-Kaylor, A. (2005b). Relationship of Corporal Punishment and Antisocial Behavior by Neighborhood. *Archives of Pediatric and Adolescent Medicine 159*, 938-942.

Grogan-Kaylor, A. & Otis, M.D. (2007). The Predictors of Parental Use of Corporal Punishment. *Family Relations 56*, 80-91.

Hahs-Vaughn, D.L. & Onwuegbuzie, A.J. (2006). Estimating and Using Propensity Score

Analysis With Complex Samples. *The Journal of Experimental Education 75*, 31-

65.

Hill, J.L., Waldfogel, J., Brooks-Gunn, J., & Han, W.J. (2005). Maternal Employment

and Child Development: A Fresh Look Using Newer Methods. *Developmental

Psychology 41*, 833-850.

Inter-university Consortium for Political and Social Research. Sections of website

concerning Project for Human Development in Chicago Neighborhoods

("PHDCN"). Retrieved between November 1, 2005 and November 30, 2005.

(http://www.icpsr.umich.edu/PHDCN/)

Kantor, G.K. & Jasinski, J.L. (Eds.). (1997). *Out of the Darkness: Contemporary

Perspectives on Family Violence*. Thousand Oaks, CA: Sage.

Larzelere, R.E. (2000). Child Outcomes of Nonabusive and Customary Physical

Punishment by Parents: An Updated Literature Review. *Clinical Child and

Family Psychology Review 3*, 199-221.

Larzelere, R.E. & Kuhn, B.R. (2005). Comparing Child Outcomes of Physical

Punishment and Alternative Disciplinary Tactics: A Meta-Analysis. *Clinical

Child and family Psychology Review 8*, 1-37.

Leuven, E. & Sianesi, B. (2003). PSMATCH2: Stata module to perform full Mahalanobis

and propensity score matching, common support graphing, and covariate

imbalance testing. Version 1.2.3. Retrieved November 15, 2006.

(http://ideas.repec.org/c/boc/bocode/s432001.html)

Levethal, T., Selner-O'Hagan, M.B., Brooks-Gunn, J., Bingenheimer, J.B., & Earls, F.E. (2004). The Homelife Interview from the Project on Human Development in Chicago Neighborhoods: Assessment of Parenting and Home Environment for 3- to 15-Year-Olds. *Parenting: Science and Practice 4*, 211-241.

Lewis, D.A. & Maxfield, M.G. (1980). Fear in the Neighborhoods: An Investigation of the Impact of Crime. *Journal of Research in Crime & Delinquency 17*, 160-190.

Lopata, H.Z. (1978). Contributions of Extended Family to the Support Systems of Metropolitan Area Widows: Limitations of the Modified Kin Network. *Journal of Marriage and the Family 40*, 355-364.

Lundquist, J. H. (2004). When Race Makes No Difference: Marriage and the Military. *Social Forces 83*, 731-757.

Magnuson, K. A. & Waldfogel, J. (2005). Preschool Child Care and Parents' Use of Physical Discipline. *Infant and Child Development 14*, 177-198.

McCord, J. (2005). Unintended Consequences of Punishment. In M. Donnelly & M.A. Straus (Eds.), *Corporal Punishment of Children in Theoretical Perspective* (pp. 165-169). New Haven, CT: Yale University Press.

McLoyd, V.C. & Smith, J. (2002). Physical Discipline and Behavior Problems in African American, European American, and Hispanic Children: Emotional Support as a Moderator. *Journal of Marriage and the Family 64*, 40-53.

Molnar, B. E., Buka, S.L., Brennan, R.T., Holton, J.K., & Earls, F.E. (2003). A Multilevel Study of Neighborhoods and Parent-to-Child Physical Aggression: Results from the Project on Human Development in Chicago Neighborhoods. *Child Maltreatment 8*, 84-97.

Mowrer, E.R. (1950). Social crises and social disorganization. *American Sociological Review 15*, 60-66.

Newgard, C.D., Hedges, J.R., Arthur, M., & Mullins, R.J. (2004). Advanced Statistics: The Propensity Score—A Method for Estimating Treatment Effect in Observational Research. *Academy of Emergency Medicine 11*, 953-961.

Paolucci, E.O. & Violato, C. (2004). A Meta-Analysis of the Published Research on the Affective, Cognitive, and Behavioral Affects of Corporal Punishment. *The Journal of Psychology 138*, 197-221.

Penrod, J.D., Deb, P., Luhrs, C., Dellenbaugh, C., Zhu, C.W., Hochman, T., Macijewski, M., Granieri, E., & Morrison, R.S. (2006). Cost and Utilization Outcomes of Patients Receiving Hospital-Based Palliative Care Consultation. *Journal of Palliative Medicine 9*, 855-860.

Ringel, J.S., Ellickson, P.L., & Collins, R.L. (2007). High school drug use predicts job-related outcomes at age 29. *Addictive Behaviors 32*, 576-589.

Rodriguez, C.M. (2003). Parental Discipline and Abuse Potential Affects on Child Depression, Anxiety, and Attributions. *Journal of Marriage and the Family 65*, 809-817.

Rosenbaum, P.R. & Rubin, D.B. (1984). Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association 79*, 516-524.

Rubin, D.B. (1997). Estimating Causal Effects from Large Data Sets Using Propensity Scores. *Annals of Internal Medicine 127*, 757-763.

Sampson, R.J. (1997). The Embeddedness of Child and Adolescent Development: A Community-Level Perspective on Urban Violence. In J. McCord (Ed.) *Violence and Childhood in the Inner City* (pp. 31-77). Cambridge, MA: Cambridge University Press.

Sampson, R.J., Laub, J.H., & Wimer, C. (2006). Does Marriage Reduce Crime? A Counterfactual Approach to Within-Individual Causal Effects. *Criminology 44*, 465-508.

Simons, R.L., Wu, C., Lin, K., Gordon, L., & Conger, R.D. (2000). A Cross-Cultural Examination of the Link Between Corporal Punishment and Adolescent Antisocial Behavior. *Criminology 38*, 47-79.

Shaw, C.R. & McKay, H.D. (1942). *Juvenile Delinquency and Urban Areas*. Chicago: University of Chicago Press.

Sianesi, B. (2001). *Implementing Propensity Score Matching Estimators with STATA*. Presented at the UK Stata Users Group, VII Meeting, London.

Slade, E.P. & Wissow, L.S. (2004). Spanking in Early Childhood and Later Behavior Problems: A Prospective Study of Infants and Young Toddlers. *Pediatrics 113*, 1321-1330.

Sourander, A. & Helstela, L. (2005). Childhood predictors of externalizing and internalizing problems in adolescence: A prospective follow-up study from age 8 to 16. *European Child & Adolescent Psychiatry 14*, 415-423.

Straus, M.A. (1979). Measuring Intrafamily Conflict and Violence: The Conflict Tactics (CT) Scales. *Journal of Marriage and the Family 41*, 45-88.

Straus, M.A. (1994). *Beating the Devil out of Them: Physical Punishment in American Families*. San Francisco: Jossey-Bass/Lexington Books.

Straus, M.A. (1996). Spanking and the Making of a Violent Society. *Pediatrics 98*, 837-842.

Straus, M.A. (2001). New Evidence for the Benefits of Never Spanking. *Society 38*, 52-60.

Straus, M.A. & Kantor, G.K. (1994). Corporal Punishment of Adolescents by Parents: A Risk Factor in the Epidemiology of Depression, Suicide, Alcohol Abuse, Child Abuse, and Wife Beating. *Adolescence 29*, 543-563.

Straus, M.A. & Gimpel, H. (1994). Corporal Punishment and Economic Achievement: A Theoretical Model and Some Preliminary Data. In M.A. Straus *Beating the Devil out of Them: Physical Punishment in American Families* (pp. 137-146). San Francisco: Jossey-Bass/Lexington Books.

Straus, M.A., Sugarman, D., & Giles-Sims, J. (1997). Spanking by Parents and Subsequent Antisocial Behavior of Children. *Archives of Pediatrics and Adolescent Medicine 151*, 761-767.

Straus, M.A, Hamby, S.L., Finkelhor, D., Moore, D.W., & Runyan, D. (1998). Identification of Child Maltreatment with the Parent-Child Conflict Tactics Scales: Development and Psychometric Data for a National Sample of American Parents. *Child Abuse and Neglect 22*, 249-270.

Straus, M.A. & Mouradian, V.E. (1998). Impulsive Corporal Punishment by Mothers and Antisocial Behavior and Impulsiveness of Children. *Behavioral Sciences and the Law 16*, 353-374.

Straus, M.A. & Stewart, J.H. (1999). Corporal Punishment by American Parents: National Data on Prevalence, Chronicity, Severity, and Duration, in Relation to Child and Family Characteristics. *Clinical Child and Family Psychology Review 2*, 55-70.

Turner, H.A. & Finkelhor, D. (1996). Corporal Punishment as a Stressor Among Youth. *Journal of Marriage and the Family 58*, 155-166.

Turner, H.A. & Muller, P.A. (2004). Long-Term Effects of Child Corporal Punishment on Depressive Symptoms in Young Adults. *Journal of Family Issues 25*, 761-782.

Turner, R.J., Frankel, B., & Levin, D. (1983). Social support: Conceptualization, measurement, and implications for mental health. In J.R. Greenley (Ed.), *Research in Community and Mental Health* (pp. 67-111), vol. III, Greenwich: JAI Press.

Wissow, L. (2001). Ethnicity, Income, and Parenting Contexts of Physical Punishment in a National Sample of Families with Young Children. *Child Maltreatment 6*, 118-129.

APPENDIX A

ITEMS FOR SCALED VARIABLES

Child Behavioral Checklist Scales

Withdrawal

Would rather be alone than with others

Refuses to talk

Is secretive, keeps things to self

Shy or timid

Stares blankly

Sulks a lot

Underactive, slow moving, or lacks energy

Unhappy, sad, or depressed

Withdrawn, doesn't get involved with others

Somatic Complaints

Feels dizzy

Is overtired

Aches and pains, not including headaches

Headaches

Nausea, feels sick

Problems with the eyes

Rashes or other skin problems

Stomach aches or cramps

Vomiting, throwing up

Anxious/Depressed

Complains of loneliness

Cries alot

Fears he/she might do something bad

Feels he/she has to be perfect

Feels or complains that no one loves him/her

Feels others are out to get him/her

Feels worthless or inferior

Is nervous, highstrung, or tense

Too fearful or anxious

Feels too guilty

Self-conscious or easily embarrassed

Suspicious

Unhappy, sad, or depressed

Worries

Delinquent Behavior

Doesn't seem to feel guilty after misbehaving

Hangs around with others who get in trouble

Lies or cheats

Prefers being with older kids

Runs away from home

Sets fires

Disobedient at home

Disobedient at school

Swears or uses obscene language

Thinks about sex*

Truant, skips school

Uses alcohol or drugs*

Destroys things belonging to his/her family or others*

Aggressive Behavior

Argues a lot

Brags*

Cruelty, bullying or meanness to others

Demands a lot of attention

Destroys own things

Destroys other's things

Disobedient at home

Disobedient at school

Jealous*

Gets in many fights

Attacks people*

Screams a lot

Shows off*

Stubborn, sullen, or irritable

Sudden changes in mood or feeling

Talks too much*

Has temper tantrums hot temper

Threatens people

Loud*

*item not included in wave 2 measure

(Achenbach, 1991)

Items for HOME scales

Parental Warmth Scale

> Parent talks with child twice during visit

> Parent answers child's questions orally

> Parent encourages child to contribute

> Parent mentions skill of child

> Parent praises child twice during visit

> Parent voices positive feelings to child

> Parent caresses, kisses, or hugs child

> Parent responds positively to praise of child

Primary Caregiver Verbal Skills

> Primary caregiver's speech is distinct, clear, and audible to the interviewer

> Primary caregiver initiates verbal interchanges with the interviewer, asks
> questions, makes spontaneous comments

> Primary caregiver expresses ideas freely and easily and uses statements of
> appropriate length for conversation

> Primary caregiver appears to readily understand the interviewer's questions

(Leventhal et al, 2004)

Supervision/Monitoring Scale

> Subject has a set time (curfew) to be home on school nights

> Subject has a set time (curfew) to be home on weekend nights

> Requires subject to sleep at home on school nights

When primary caregiver is not at home, reasonable procedures are established for

subject to check in with primary caregiver or other designee on weekends or after

school

After school subject goes somewhere that adult supervision is provided

Establishes rules for behavior with peers and asks questions to determine whether

they are being followed

Subject is not allowed to wander in public places without adult supervision for

more than three hours

Has had contact with two of the subject's friends in the last two weeks

Has visited with school or talked to the teacher or counselor within the last three

months

Conflict Tactics Scale for Partner/Spouse

When you had an argument in the past year…

How many times have you thrown, smashed, hit, or kicked something?

How many times have you thrown something at him/her?

How many times have you pushed, grabbed, or shoved him/her?

How many times have you slapped him/her?

How many times have you kicked, bit, or hit him/her with a fist?

How many times have you tried to or hit him/her with something?

How many times have you beat him/her up?

How many times have you threatened him/her with a knife or gun?

How many times have you used a knife or fired a gun?

(Straus, 1979)

Provision of Social Relations

Family Support Scale

No matter what happens, I know that my family will always be there for me
should I need them.

Sometimes I'm not sure if I can completely rely on my family.

My family lets me know they think I'm a worthwhile person.

People in my family have confidence in me.

People in my family help me find solutions to my problems.

I know my family will always stand by me.

Friend Support Scale

When I'm with my friends I feel completely able to relax and be myself.

I share the same approach to life that many of my friends do.

People who know me trust me and respect me.

When I want to go out and do things, I know that many of my friends would enjoy
doing these things with me.

I have at least one friend that I could tell anything to.

I feel very close to some of my friends.

People who know me think I am good at what I do.

My friends would take the time to talk about my problems, should I ever want to.

Even when I am with my friends, I feel alone.

(Turner et al, 1983)

EASI Temperament Survey

Impulsivity Measure

Inhibitory Control

      I have trouble controlling my impulses

      Usually I can't stand waiting

      I can tolerate frustration better than most

      I have trouble resisting my cravings (for food, cigarettes, etc.)

      I like to spend my money right away rather than save it for long-range goals

Decision Time

      I often say the first thing that comes into my head

      I often have trouble making up my mind

      I like to plan things way ahead of time

      I often act on the spur of the moment

      I like to make detailed plans before I do something

Sensation Seeking

      I generally seek new and exciting experiences and sensations

      I'll try anything once

      I sometimes to "crazy" things just to be different

      I'm happiest in familiar surroundings

      I get bored easily

Persistence

      I generally like to see things through to the end

      I tend to hop from interest to interest quickly

I tend to give up easily

Unfinished tasks really bother me

Once I get going on something I hate to stop

(Buss & Plomin, 1984)

APPENDIX B

IRB APPROVAL

GEORGIA SOUTHERN UNIVERSITY

OFFICE OF RESEARCH SERVICES & SPONSORED PROGRAMS
POST OFFICE BOX 8005
STATESBORO, GEORGIA 30460-8005
TELEPHONE: (912) 681-5465

March 8, 2007

Sara Z. Morris
P.O. Box-8051

Dear Sara Morris,

After a review of your proposed research project numbered: **H7169**, and titled **"The Causal Effect of Corporal Punishment on Behavioral Outcomes: A Propensity Score Analysis"**, it appears that your research involves activities that do not require approval by the Institutional Review Board according to federal guidelines.

According to the Code of Federal Regulations Title 45 Part 46, your research protocol is determined to be exempt under the following exemption category(s):

> Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

Therefore, as authorized in the Federal Policy for the Protection of Human Subjects, I am pleased to notify you that your research is exempt from IRB approval. You may proceed with the proposed research.

Sincerely,

N. Scott Pierce
Director of Research Services and Sponsored Programs

A UNIT OF THE UNIVERSITY SYSTEM OF GEORGIA • AFFIRMATIVE ACTION/EQUAL OPPORTUNITY INSTITUTION

Table 1

Descriptive statistics (N = 1346)

	Mean	Std. deviation	Min.	Max.
Outcomes				
Externalizing	8.40	7.40	0.00	41.00
Internalizing	8.51	7.74	0.00	48.00
Aggressive	5.65	4.97	0.00	26.00
Delinquent	2.75	2.84	0.00	17.00
Withdrawal	2.56	2.73	0.00	15.00
Somatic Complaints	1.67	2.27	0.00	16.00
Depressive/Anxious	4.29	4.21	0.00	24.00
Treatment				
CP dichotomous	0.68	0.47	0.00	1.00
Covariates				
Demographics				
Subject Age	10.04	3.19	5.04	16.37
Subject Gender (1=male)	0.50	0.50	0.00	1.00
White	0.18	0.38	0.00	1.00
Black	0.31	0.46	0.00	1.00
Hispanic	0.48	0.50	0.00	1.00
Other	0.03	0.18	0.00	1.00
SES	0.08	1.44	-2.92	3.52
Family Size	5.33	1.87	2.00	14.00
Number of Siblings < 19	2.18	1.56	0.00	10.00
Public assistance	0.26	0.44	0.00	1.00
Primary Caregiver Characteristics				
PC Age	36.36	7.12	20.70	67.95
PC Gender (1=male)	0.09	0.29	0.00	1.00
PC less than high school	0.21	0.41	0.00	1.00
PC some high school	0.21	0.41	0.00	1.00
PC graduated high school	0.13	0.33	0.00	1.00
PC education past high school	0.46	0.50	0.00	1.00
PC Mom	0.86	0.35	0.00	1.00
PC Dad	0.09	0.28	0.00	1.00
PC other female	0.05	0.21	0.00	1.00
PC other male	0.01	0.07	0.00	1.00
PC married (1=yes)	0.66	0.48	0.00	1.00
PC employed (1=yes)	0.61	0.49	0.00	1.00
Supervision/Monitoring	8.38	0.89	3.00	9.00
Warmth	6.95	1.92	0.00	9.00
PC Verbal Skills	3.89	0.44	0.00	4.00
Provision of Social Relations	20.55	4.36	15.00	38.00
Conflict Tactics Scale	2.24	5.30	0.00	46.00
Conflict Tactics Scale--Spouse	1.94	4.51	0.00	48.00
Child Characteristics				

	Mean	Std. deviation	Min.	Max.
Wave 1 Internalizing	7.36	6.64	0.00	45.00
Wave 1 Externalizing	11.16	8.78	0.00	54.00
Impulsivity	54.50	11.48	25.00	93.00
School help for emotional problems	0.03	0.18	0.00	1.00
Truant past year	0.04	0.20	0.00	1.00
Family Characteristics				
Family member w/ criminal record	0.30	0.46	0.00	1.00
Family member w/ nerve problems	0.19	0.39	0.00	1.00
Family member w/ legal problems	0.17	0.38	0.00	1.00
Family member attempt suicide	0.13	0.34	0.00	1.00

* p < .05

Table 2

*Pre-matching t-tests using nearest
neighbor matching (N = 1346)*

	Mean			
	Treated	Control	t value	p value
Demographics				
Subject Age	9.58*	11.01*	-7.86	0.00
Subject Gender (1=male)	0.51	0.46	1.73	0.08
Black	0.32	0.29	1.19	0.23
Hispanic	0.48	0.49	-0.53	0.60
Other	0.03	0.04	-1.44	0.15
SES	0.09	0.06	0.27	0.79
Family Size	5.33	5.33	-0.04	0.97
Number of Siblings < 19	2.14	2.25	-1.12	0.26
Public assistance	0.28*	0.23*	2.00	0.05
Primary Caregiver Characteristics				
PC Age	35.47*	38.25*	-6.80	0.00
PC Gender (1=male)	0.08*	0.13*	-2.99	0.00
PC some high school	0.22	0.19	1.41	0.17
PC graduated high school	0.12	0.14	-0.60	0.55
PC education past high school	0.47	0.44	0.91	0.37
PC Dad	0.07*	0.12*	-3.12	0.00
PC other female	0.04*	0.06*	-2.01	0.05
PC other male	0.01	0.00	0.21	0.83
PC married (1=yes)	0.64	0.68	-1.22	0.22
PC employed (1=yes)	0.62	0.60	0.56	0.58
Supervision/Monitoring	8.40	8.34	1.12	0.26
Warmth	6.95	6.94	0.13	0.90
PC Verbal Skills	3.91*	3.84*	2.84	0.01
Provision of Social Relations	20.76*	20.12*	2.55	0.01
Conflict Tactics Scale	2.62*	1.44*	3.83	0.00
Conflict Tactics Scale--Spouse	2.29*	1.21*	4.14	0.00
Child Characteristics				
Wave 1 Internalizing	8.08*	5.82*	5.92	0.00
Wave 1 Externalizing	13.05*	7.18*	12.05	0.00
Impulsivity	56.40*	50.52*	9.04	0.00
School help for emotional problems	0.04*	0.01*	2.96	0.00
Truant past year	0.05*	0.03*	1.57	0.11
Family Characteristics				
Family member w/ criminal record	0.33*	0.23*	3.97	0.00
Family member w/ nerve problems	0.21*	0.14*	2.81	0.01
Family member w/ legal problems	0.19*	0.14*	2.21	0.03
Family member attempt suicide	0.14*	0.10*	1.98	0.05

* p < .05

Table 3

Logistic regression results (N = 1346)

	Coefficient	Std. error	z	p value	95 % conf. interval	
Demographics						
Subject Age	-0.13*	0.02	-5.55	0.00	-0.18	-0.09
Subject Gender (1=male)	0.17	0.13	1.28	0.20	-0.09	0.43
Black	0.04	0.22	0.20	0.84	-0.38	0.47
Hispanic	0.00	0.21	0.02	0.99	-0.41	0.42
Other	-0.39	0.38	-1.03	0.30	-1.14	0.36
SES	0.03	0.08	0.41	0.68	-0.12	0.19
Family Size	0.06	0.05	1.28	0.20	-0.03	0.16
Number of Siblings < 19	-0.14*	0.06	-2.54	0.01	-0.26	-0.03
Public assistance	0.02	0.20	0.08	0.94	-0.38	0.41
Primary Caregiver Characteristics						
PC Age	-0.02	0.01	-1.82	0.07	-0.04	0.00
PC Gender (1=male)	-0.57*	0.23	-2.49	0.01	-1.02	-0.12
PC some high school	0.15	0.23	0.65	0.52	-0.30	0.60
PC graduated high school	0.07	0.26	0.28	0.78	-0.44	0.58
PC education past high school	-0.02	0.27	-0.06	0.95	-0.55	0.52
PC Dad[a]						
PC other female	-0.51	0.33	-1.53	0.13	-1.16	0.14
PC other male	0.71	1.03	0.68	0.49	-1.32	2.73
PC married (1=yes)	0.31	0.17	1.79	0.07	-0.03	0.66
PC employed (1=yes)	0.26	0.16	1.69	0.09	-0.04	0.57
Supervision/Monitoring	0.03	0.07	0.38	0.71	-0.12	0.17
Warmth	-0.02	0.04	-0.51	0.61	-0.09	0.05
PC Verbal Skills	0.33*	0.16	2.10	0.04	0.02	0.64
Provision of Social Relations	0.01	0.02	0.47	0.64	-0.02	0.04
Conflict Tactics Scale	0.01	0.02	0.71	0.48	-0.02	0.05
Conflict Tactics Scale--Spouse	0.05*	0.02	2.41	0.02	0.01	0.09
Child Characteristics						
Wave 1 Internalizing	0.00	0.01	0.18	0.86	-0.02	0.03
Wave 1 Externalizing	0.08*	0.01	6.22	0.00	0.06	0.11
Impulsivity	0.01	0.01	1.54	0.12	0.00	0.03
School help for emotional problems	0.75	0.53	1.41	0.16	-0.29	1.79
Truant past year	0.30	0.36	0.84	0.40	-0.40	1.00
Family Characteristics						
Family member w/ criminal record	0.49*	0.19	2.63	0.01	0.12	0.86
Family member w/ nerve problems	0.21	0.19	1.11	0.27	-0.16	0.58
Family member w/ legal problems	-0.28	0.23	-1.24	0.22	-0.72	0.16
Family member attempt suicide	-0.05	0.22	-0.23	0.82	-0.48	0.38
Constant	-0.74	1.14	-0.65	0.52	-2.98	1.50

* p < .05

[a]this variable was kicked out of analysis due to collinearity.

Table 4

Post- matching t-tests using nearest-neighbor matching (N = 870)

Covariate	Mean		% Bias reduction	t value	p value
	Treated	Control			
Demographics					
Subject Age	10.63	11.01	73.30	-0.89	0.37
Subject Gender (1=male)	0.50	0.46	36.20	0.44	0.66
Black	0.29	0.29	92.80	0.24	0.81
Hispanic	0.48	0.49	40.90	-0.72	0.47
Other	0.34	0.44	39.30	-0.24	0.81
SES	0.13	0.06	-207.30	0.87	0.39
Family Size	5.20	5.34	-3322.40	-0.92	0.36
Number of Siblings < 19	2.20	2.25	57.30	-0.20	0.84
Public assistance	0.24	0.23	68.60	0.37	0.71
Primary Caregiver Characteristics					
PC Age	37.78	38.25	83.20	-0.15	0.88
PC Gender (1=male)	0.10	0.13	50.20	-0.70	0.48
PC some high school	0.19	0.19	93.10	0.05	0.96
PC graduated high school	0.14	0.14	40.50	0.16	0.88
PC education past high school	0.46	0.44	21.50	0.55	0.58
PC Dad	0.10	0.12	55.50	-0.50	0.61
PC other female	0.06	0.06	63.00	-0.32	0.75
PC other male	0.00	0.00	-158.10	-1.31	0.20
PC married (1=yes)	0.68	0.68	100.00	0.25	0.80
PC employed (1=yes)	0.58	0.60	-30.90	0.09	0.92
Supervision/Monitoring	8.41	8.34	-7.00	1.14	0.25
Warmth	7.02	6.94	-444.90	0.28	0.78
PC Verbal Skills	3.88	3.84	42.50	0.90	0.37
Provision of Social Relations	20.23	20.12	81.90	0.15	0.89
Conflict Tactics Scale	1.46	1.44	99.00	-0.32	0.75
Conflict Tactics Scale--Spouse	1.12	1.21	97.20	-0.62	0.53
Child Characteristics					
Wave 1 Internalizing	5.93	5.82	95.10	-0.56	0.56
Wave 1 Externalizing	7.22	7.18	99.30	-1.18	0.24
Impulsivity	51.14	50.52	89.40	0.16	0.87
School help for emotional problems	0.01	0.01	84.80	-0.94	0.35
Truant past year	0.04	0.03	25.10	0.99	0.32
Family Characteristics					
Family member w/ criminal record	0.22	0.23	91.30	-0.28	0.78
Family member w/ nerve problems	0.16	0.14	82.00	-0.11	0.91
Family member w/ legal problems	0.15	0.14	85.80	0.29	0.77
Family member attempt suicide	0.11	0.11	100.00	-0.43	0.67

* p < .05

Table 5

OLS regression results (N = 870)

		b	SE	p-value
Internalizing				
	CP measure	0.19	0.50	0.70
	P score	10.64*	1.11	0.00
	constant	1.19	0.71	0.10
Externalizing				
	CP Measure	0.03	0.42	0.94
	P score	17.85*	1.00	0.00
	constant	-3.69*	0.62	0.00
Anxious/Depression				
	CP Measure	-0.10	0.26	0.70
	P score	6.70*	0.60	0.00
	constant	-0.18	0.38	0.64
Somatic Complaints				
	CP Measure	0.20	0.15	0.17
	P score	1.22*	0.34	0.00
	constant	0.71*	0.22	0.00
Withdrawal				
	CP Measure	0.09	0.17	0.62
	P score	2.72*	0.40	0.00
	constant	0.66*	0.26	0.01
Delinquent Behavior				
	CP Measure	-0.12	0.17	0.46
	P score	5.97*	0.39	0.00
	constant	-1.20*	0.25	0.00
Aggressive Behavior				
	CP Measure	0.15	0.30	0.58
	P score	11.87*	0.65	0.00
	constant	-2.49*	0.42	0.00

* $p < .05$

Table 6

Comparison results from nearest-neighbor matching (N = 870)

	Treated	Control	Difference	Std. Error	t value
Internalizing					
Unmatched	9.20	7.06	2.14*	0.45	4.78
ATT	7.49	7.06	0.42	0.48	0.88
Externalizing					
Unmatched	9.48	6.17	3.30*	0.42	7.83
ATT	6.21	6.17	0.04	0.42	0.10
Withdrawal					
Unmatched	2.75	2.17	0.59*	0.16	3.69
ATT	2.36	2.17	0.19	0.18	1.10
Somatic Complaints					
Unmatched	1.81	1.38	0.43*	0.13	3.22
ATT	1.58	1.38	0.20	0.15	1.37
Anxious/Depressed					
Unmatched	4.65	3.52	1.13*	0.24	4.63
ATT	3.54	3.52	0.02	0.25	0.09
Aggressive					
Unmatched	6.41	4.08	2.33*	0.29	8.26
ATT	4.24	4.08	0.17	0.28	0.61
Delinquent					
Unmatched	3.07	2.10	0.97*	0.16	5.94
ATT	1.97	2.10	-0.13	0.16	-0.79

* p < .05

Table 7

Comparison results from kernel matching (N = 1305)

	Treated	Control	Difference	Std. Error	t value
Internalizing					
Unmatched	9.20	7.06	2.14*	0.45	4.78
ATT	8.86	9.16	-0.30	0.61	-0.48
Externalizing					
Unmatched	9.48	6.17	3.30*	0.42	7.83
ATT	9.03	9.62	-0.58	0.57	-1.03
Withdrawal					
Unmatched	2.75	2.17	0.59*	0.16	3.69
ATT	2.65	2.70	-0.06	0.22	-0.25
Somatic Complaints					
Unmatched	1.81	1.38	0.43*	0.13	3.22
ATT	1.77	1.77	0.00	0.17	0.00
Anxious/Depressed					
Unmatched	4.65	3.52	1.13*	0.24	4.63
ATT	4.45	4.70	-0.24	0.33	-0.73
Aggressive					
Unmatched	6.41	4.08	2.33*	0.29	8.26
ATT	6.11	6.30	-0.19	0.38	-0.49
Delinquent					
Unmatched	3.07	2.10	0.97*	0.16	5.94
ATT	2.92	3.32	-0.40	0.22	-1.80

* p < .05

Table 8

Post-matching t-tests using kernel matching (N = 1305)

	Mean		% Bias		
	Treated	Control	reduction	t value	p value
Demographics					
Subject Age	9.65*	9.67*	98.90	6.50	0.00
Subject Gender (1=male)	0.51	0.52	81.80	-1.54	0.12
Black	0.32*	0.38*	-108.20	-2.11	0.04
Hispanic	0.48	0.41	-313.80	1.69	0.09
Other	0.03*	0.02*	59.40	2.36	0.02
SES	0.10	0.11	45.20	0.22	0.83
Family Size	5.31	5.11	-4398.80	1.50	0.14
Number of Siblings < 19	2.14	2.25	-9.70	-0.10	0.92
Public assistance	0.27*	0.34*	-44.30	-3.49	0.00
Primary Caregiver Characteristics					
PC Age	35.71*	36.07*	87.00	5.23	0.00
PC Gender (1=male)	0.08*	0.08*	99.50	3.19	0.00
PC some high school	0.21	0.19	14.50	-0.54	0.59
PC graduated high school	0.12	0.17	-315.40	-0.86	0.39
PC education past high school	0.47	0.48	74.30	-0.18	0.86
PC Dad	0.07*	0.08*	92.40	3.18	0.00
PC other female	0.04	0.04	94.30	1.44	0.15
PC other male	0.01	0.00	-367.40	0.21	0.84
PC married (1=yes)	0.65	0.60	-39.40	1.87	0.06
PC employed (1=yes)	0.62	.61	30.50	-.15	0.88
Supervision/Monitoring	8.41	8.42	92.90	-1.85	0.06
Warmth	7.00	6.89	-737.60	0.38	0.70
PC Verbal Skills	3.91	3.87	34.80	-0.45	0.66
Provision of Social Relations	20.62	20.38	61.30	-0.96	0.34
Conflict Tactics Scale	2.43*	2.40*	96.70	-3.16	0.00
Conflict Tactics Scale--Spouse	2.05*	1.77*	73.70	-2.94	0.00
Child Characteristics					
Wave 1 Internalizing	7.56*	7.92*	93.00	-4.66	0.00
Wave 1 Externalizing	12.24*	12.97*	87.60	-10.63	0.00
Impulsivity	55.83*	56.41*	90.20	-7.85	0.00
School help for emotional problems	0.03	0.02	39.80	-0.67	0.50
Truant past year	0.05*	0.08*	-71.80	-3.11	0.00
Family Characteristics					
Family member w/ criminal record	0.33*	0.32*	99.00	-4.07	0.00
Family member w/ nerve problems	0.20*	0.19*	75.50	-1.93	0.05
Family member w/ legal problems	0.18*	0.20*	59.90	-2.85	0.00
Family member attempt suicide	0.14	0.12	59.30	-1.42	0.16

* p < .05

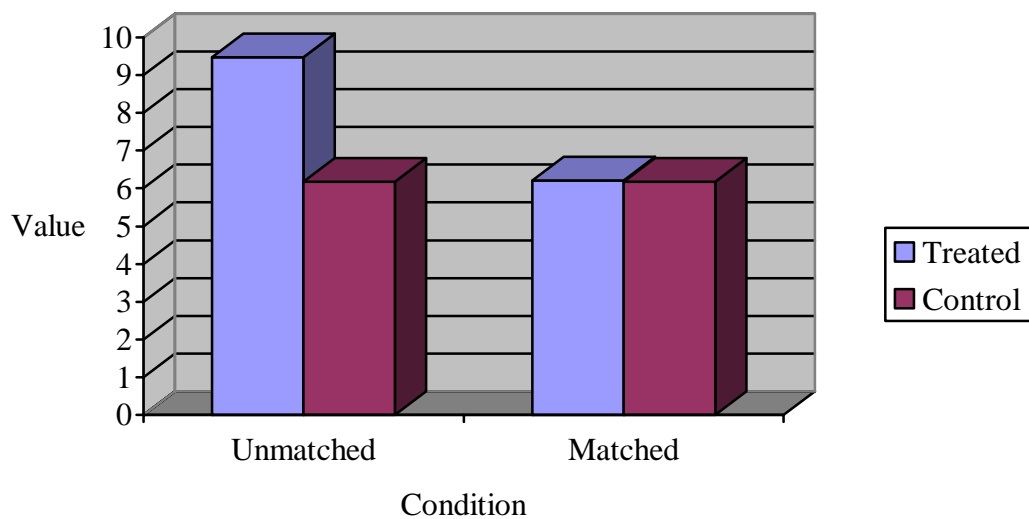*Figure 1*. Externalizing Behavior Results Before and After Matching

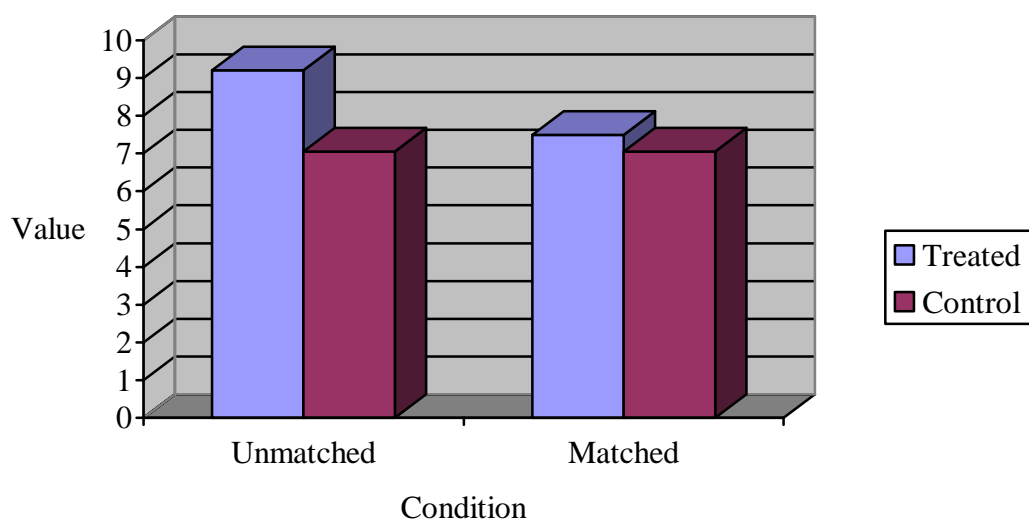*Figure 2*. Internalizing Behavior Results Before and After Matching

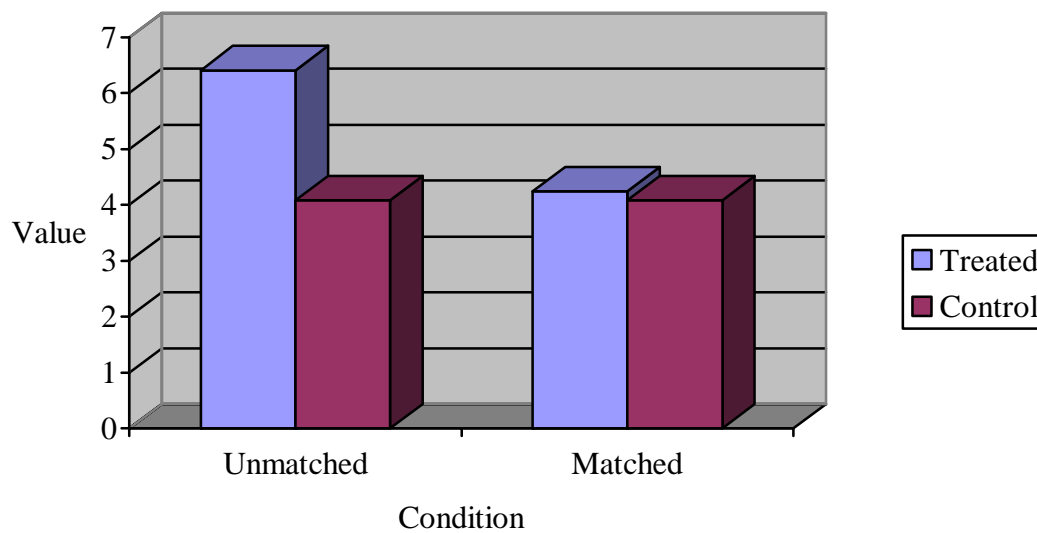*Figure 3*. Aggressive Behavior Results Before and After Matching

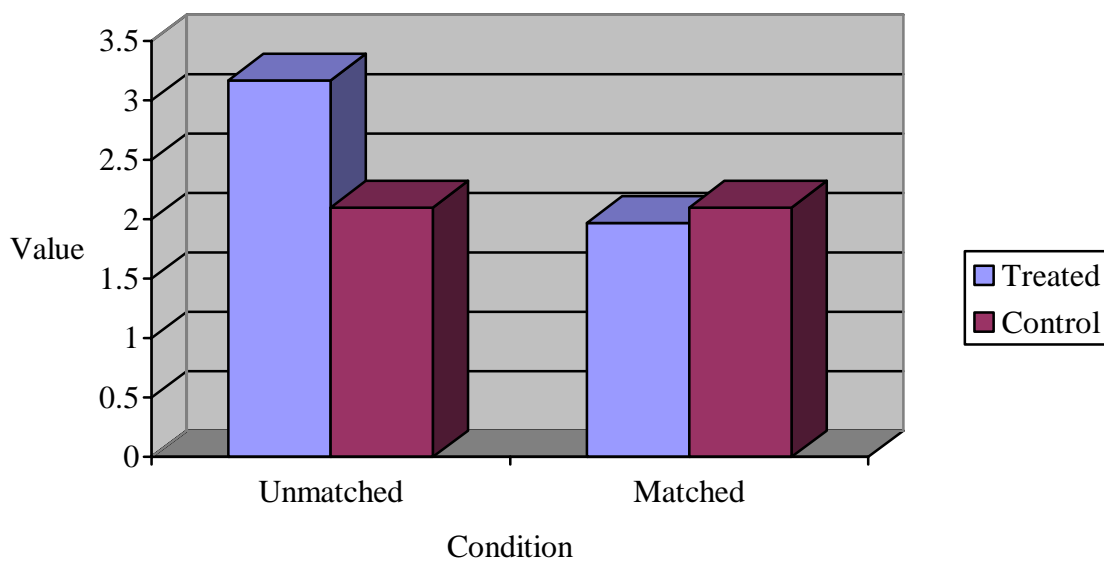*Figure 4*. Delinquent Behavior Results Before and After Matching

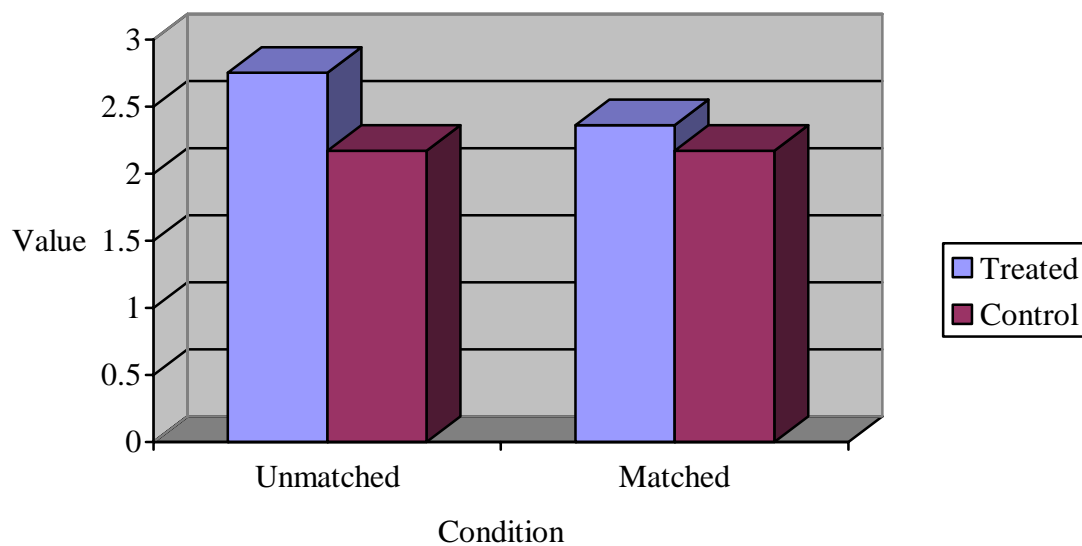*Figure 5.* Withdrawal Results Before and After Matching

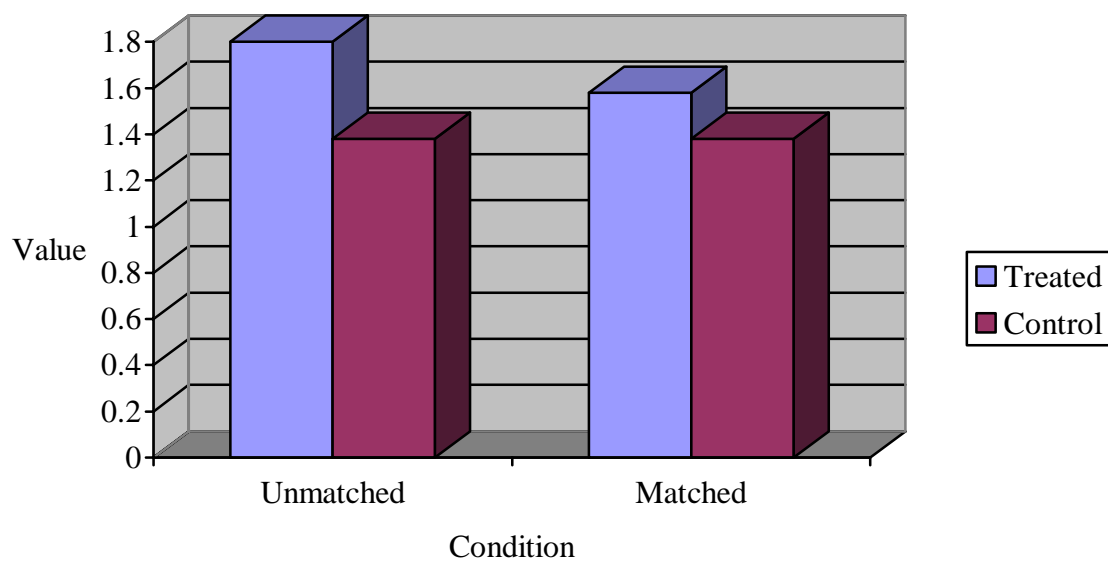*Figure 6.* Somatic Complaint Results Before and After Matching

Figure 7. Anxious/Depressed Results Before and After Matching.