A Natural Field Experiment Comparing Duluth and Cognitive Behavioral Therapy Treatment Options for Intimate Partner Violence Diversion Programs: The Winnebago County Experience*

Chad Cotti   Joshua Foster[†]

M. Ryan Haley   Shannon L. Rawski

University of Wisconsin - Oshkosh

November 16, 2017

Abstract

We use data from a three-year natural field experiment to study rates of recidivism in two types of diversion programs designed to reduce Intimate Partner Violence (IPV) among heterosexual partners. In one program (*Duluth*), efforts are focused on protecting women from male aggression through a psycho-educational program, regardless of the offender's sex. In the other program (*CBT*), efforts are focused on improving intra-household behaviors and communication skills through cognitive-behavioral therapy counseling. Using frame analysis and identity theory, we hypothesize that *Duluth*'s negative framing of a male offender's role in IPV will lead to higher rates of recidivism, as compared to *CBT*, which instead frames the offender's role in such a way that minimizes identity threat in the process of adapting behavior. In support of this hypothesis, we find that the IPV recidivism rate, measured as reconvictions for IPV, is 11 percentage points higher for offenders randomly assigned to a *Duluth* treatment program (14 percentage points higher among males). These experimental results are statistically and practically significant, indicating that the *Duluth* approach corresponds to meaningfully higher recidivism rates when compared to *CBT*, in a direct comparison of the two programs. To assess the robustness of our findings we extend the statistical inquiry to include several covariate-based analyses of the experimental data, additional recidivism measures, and survival analysis; each of these corroborates the primary experimental findings.

Keywords: Batterer Intervention; Domestic Violence; Experiments; Hazard Rate; Recidivism

1 Introduction

Intimate partner violence (hereafter, IPV) involves assualtive behaviors occurring between adults who are currently or were formerly engaged in an intimate relationship (Goolkasian, 1986). The occurance of IPV is a social problem affecting the health and well-being of community members across the country. Results from the Center for Disease Control's National Intimate Partner and Sexual Violence Survey (Black et al. 2011) suggest that 1 in 3 women and 1 in 4 men in the US have experienced IVP in their lifetime, often occurring before they were 25 years old. Additionally, about 1 in 4 women and 1 in 7 men report experiencing severe forms of physical violence at the hands of intimate partners including being hit with a fist/something hard or being slammed against something (Black et al. 2011). The U.S. Department of Justice reported that between the years of 2003 and 2012, IPV accounted for 15 percent of all violent crimes (Truman and Morgan, 2014). Max, Rice, Finkelstein, Bardwell, and Leadbetter (2004) suggested that the economic cost (e.g., physical and mental healthcare, lost productivity due to injury or premature death) of IPV against women in the US was $8.3 billion in the year 2003. This is likely an underestimate of the total economic cost of IPV, considering only female victims were included in the analysis.

Despite several decades of particularly heavy focus, IPV remains a challenging social problem. Courts, therapists, and other researchers routinely grapple with how to best address IPV. One potential intervention is the use of IPV offender treatment programs, aimed at preventing convicted IPV offenders from continuing to perpetrate violent acts against their intimate partners. The literature is thick with empirical studies from the many and varied IPV offender treatment programs applied in recent decades. Still, the debate over a "best" approach persists. Comparing the results between programs is often encumbered by different program designs and participant-screening processes, different outcome measures related to recidivism, different post-treatment observation durations, and varying levels of statistical sophistication (e.g., Feder and Dugan, 2002). As a result, estimates of IPV treatment efficacy range widely, sometimes suggesting that treatment is effective and other times suggesting that it is not (e.g., Davis et al., 2000; Feder and Wilson, 2005).

Amidst these statistical challenges and program-specific differentials is the debate over two specific treatment types: the *Duluth* model and the *Cognitive Behavioral Therapy (CBT)* model; see, for example, the series of papers between Dutton and Corvo (2006, 2007) and Gondolf (2007). Many important questions abound. Specifically, is one program better, and if so, *why* and *by what metric(s)*? Is there a theoretical explanation for why one treatment program is more effective than the other? Are the analyses complete enough to pass sound judgment on one paradigm versus the other, or do the many small program-specific design differentials inhibit a clean conclusion one way or the other?

To address the theoretical gap in the literature, our study draws on frame analysis (Goffman, 1974) and identity theory (Tajfel and Turner, 1985; Petriglieri, 2011). Specifically, we suggest the way IPV is framed in treatment programs will affect the potential roles offenders see for themselves in relationship conflict situations. If a training program creates only negative roles for offenders, then that treatment is likely to threaten the offender's identity and subsequently be rejected by the offender. However, if a treatment program offers offenders positve roles to serve, then treatment may be more readily accepted by the participants, and consequently more effective. While a variety of different paradigmatic views on IPV are beneficial to researchers, counselors, and law enforcement officials for the purposes of understanding IPV, we contend that only certain IPV paradigms will be effective for the specific purpose of preventing offenders' future use of IPV behaviors. Our theory offers insight into why one training program may work better than another, and how future training programs may be improved upon.

In addition to our theoretical contribution, our study also addresses several methodological and analytical concerns of prior research. Statistically navigating comparisons between IPV treatment programs is a nontrivial matter, and many approaches have been undertaken. We were particularly interested in the literature's criticisms of prior program evaluation and paradigm comparison efforts. What were the various statistical complaints, confusions, shortcomings, constraints, etc.? After collecting these issues, we distilled from them a set of ten desiderata that attempt to capture the literature's preferences (not ours) regarding the design and evaluation of IPV diversion programs. Specifically, a rigorous IPV treatment analysis study should:

D1: Describe the IPV charge.

D2: Describe all screening done to determine eligibility for IPV treatment; describe each treatment option and the defendant's latitude (if any) to opt-in or opt-out of the programmatic options; describe the legal ramifications of completed treatment, incomplete treatment, and refusal of treatment.

D3: Statistically assess outcome differences between completers, non-completers, and refusers of treatment.

D4: Use multiple measures of recidivism above and beyond self-reported revictimization.

D5: Employ a suitably representative sample of *at least* 100 offenders.

D6: Observe offenders, treated, untreated, and incompletely treated, for *at least* one year after treatment ends.

D7: Use clearly described statistical methods that are sophisticated enough to avoid fallacious inference; experimental or covariate-rich quasi-experimental methods are typically the most promising. This includes clearly defining the treatment and the control groups.

D8: Assert a manner in which the expected results can be logically generalized.

D9: Even-handedly apply treatments via competent therapists using consistent methods.

D10: Clearly articulate self-perceived empirical and methodological limitations.

We aspired to meet all ten desiderata (with imperfect success) in the experimental analysis of two recent, contemporaneous, co-located, and hitherto un-analyzed IPV offender treatment programs, one a *Duluth* model and one a *CBT* model. We augment our primary experimental findings by applying several covariate-centric methods to the experimental data.

The next section of the paper contains background information about the recent history of IPV criminological inquiry. Section three proposes our theory and develops the study hypothesis. Section four contains a detailed description of our local IPV diversion programs and the data we were able to obtain for their evaluation. Our findings are presented and discussed in section five. Section six contains conclusions, empirical and methodological limitations of our inquiry, and some thoughts regarding future research.

2 Background

Sherman (1992) provides an engaging recount of the sociological, political, and criminological history of domestic violence since the 1960s. In the 1960s, domestic violence—what would later be more generally referred to as IPV—was policed lightly, with arrest rarely resulting from minor domestic violence disturbances. In the 1970s, women's rights advocates began to exert significant legal and political pressure on police and lawmakers to more consistently and more forcefully address domestic violence, often advocating for "preferred arrest" or "mandatory arrest" in domestic violence cases.

Sherman's influence abruptly spiked in the 1980s pursuant to his implementation of the Minnesota Domestic Violence Experiment (MDVE), which was a novel attempt to apply the tenets of experimental design to the matter of police response to domestic violence calls. The MDVE results were often interpreted to mean that arrest was a superior form of intervention (i.e., a deterrent effect), which resulted in lower recidivism than the two competing treatments. While this interpretation skims over several important conditioning factors, it was nonetheless used to further advance a stricter approach in the policing of domestic violence disturbances. By 1992, many states had passed mandatory arrest laws for domestic violence calls, a remarkable departure from the standard practice in the 1960s. The MDVE became a landmark inquiry; praised, criticized, and replicated at great length. Its use of an experimental design was a particularly interesting development because it highlighted the evaluation challenges inherent to many human-centric issues that do not lend themselves to the traditional model of clinical experimentation espoused in pure statistics and often applied in the medical community.

As mandatory arrests laws burgeoned, it predictably increased workload demands on police, courts, and the penal system (e.g., Feder and Dugan, 2002), and also incited the inescapable question of what to do with all those arrested (e.g., Feder, 1997). As costs mounted, pressure naturally arose to study the efficacy of treatment options that might improve IPV recidivism outcomes and thereby reduce pressure on said institutions, and, of course, reduce the prevalence and severity of IPV assault. This naturally led to an ongoing debate about optimal program design. Is there a "best" IPV treatment regimen, or are there many equally effective approaches? The passion for finding a good outcome for all IPV-involved parties is clearly on display in the literature. However, finding a best solution has proved illusive, and continues to be debated. The reason for this lack of convergence stems from many of the issues captured in the ten desiderata set forth above; if these items differ meaningfully across programs, it is probable that the results from the program will likely differ as well. For example, desideratum D7 asserts the importance of suitably sophisticated statistical inference techniques, without which it is difficult or perhaps impossible to cleanly identify an average IPV treatment effect, much less make meaningful comparisons across studies.

In the end, the MDVE had enough gravity to help alter the trajectory of how Society interacted with IPV, and it helped propel Society into the churning discussion of how to effectively treat IPV offenders. We now turn our attention to a theoretical explanation for *why* any given IPV offender treatment program may be more or less effective. Using frame analysis and identity theory, we develop our study hypothesis, predicting which of the two treatment programs investigated in our study (i.e., The Duluth model vs. the CBT model) will be more effective at reducing IPV recidivism.

3   Theory & Hypothesis Development

IPV is a negative, yet complex social interaction between individuals who are involved in an intimate relationship that involves violent episodes. The relationship partners likely serve many reciprocal social roles in each other's lives, such as romantic partners, provider-dependent dyads (or co-dependent dyads), possibly parenting partners, and abuser-victim dyads. IPV is also not necessarily unidirectional, as either partner may take on the role of abuser at different points in the relationship (Larsen and Hamberger, 2015). Given the complexity of relationships involving IPV, it is no surprise that individuals involved in IPV often have trouble making sense out of the dynamics and violent episodes involved in their relationships (Catlett et al., 2010).

For convicted IPV offenders, court ordered treatment programs help offenders understand the complexity of IPV and their role in it. Past research based on Goffman's (1974) frame

analysis has indicated that training or treatment programs can be conceptualized as a sense-giving mechanism, whereby a program offers participants a frame for making sense out of a socially complex interaction, and through that frame defines the social roles of those involved in that interaction (Rawski, 2017). This sense-giving initiative may be necessary considering evidence suggesting that offenders and victims of IPV often disagree about the rates of occurrence of IPV in their relationships (Gangopadhyay et al., 2002). One reason for this is that offenders may frame IPV differently than victims or the legal system do. If treatment programs are intended to decrease IPV behaviors, it is essential that offenders agree with the treatment program's definition of IPV behaviors and learn to recognize these behaviors in their own actions.

One typically unanticipated problem with IPV interventions is that participants often come into treatment programs with their own frames for understanding IPV and their own role definitions. For instance, Catlett et al. (2010) found that offenders minimized their acts of violence by contrasting their behaviors with what they framed as "real" violence. For instance, one of their participants stated, " 'I've never really hit her, like punching or something like that. Yea, I slapped her before.' " (Catlett et al., 2010: 113). Offenders' framing of violent behaviors as not constituting violence is then coupled with denial of responsibility, rationalization, and justification for their violent actions, often involving victim blaming (Catlett et al., 2010). Further, some theories suggest that male violence against women may be viewed by some men as a functional method of conflict resolution (Dobash and Dobash, 1998), whereby the offender is not an "abuser" for using IPV to resolve a conflict, but rather is an effective "problem solver". This framing is further justified, in the male offenders mind, by their entitlement to power, status, and dominance over their female partners (Ptacek, 1988; Catlett et al., 2010). Jennings and Murphy (2000) corroborate this narrative with their theory, explaining that the social humiliation men experience from other men for failing to uphold hegemonic masculine ideals drives their violence toward women by entrenching men in the philosophy that women (as the less powerful partner) should not challenge them and deserve violence if they do not back down from a conflict. Thus, if offenders' framings of IPV contradict that of the treatment program, then many offenders will likely be confused and angry about why they were court ordered to attend treatment in the first place. Catlett et al. (2010) found evidence of these reactions from offenders in treatment and showed that offenders' inability to self-reflect on how they have framed their violent actions often led to attrition from IPV treatment programs. Indeed, many IPV offenders enter treatment in a state of denial, whereby they perceive the change encouraged by the treatment program as unnecessary (Gondolf, 1987). If treatment programs are unable to advance offenders past this initial denial stage, then those programs will be largely ineffective at changing offenders' use of IPV, empathy, and communication (Scott and Wolfe, 2003).

The idea that a convicted IPV offender may frame IPV and their role in it in a way that puts them in the best possible light and denies any harm they have caused is consistent with identity theory. Individuals have a core psychological motive to create and maintain a positive sense of self (Tajfel and Turner, 1985; Baumeister, 1998; Gecas, 1982; Hogg and Terry, 2000). This means that in social interactions, individuals seek to enact their positive identities, and have those positive identities affirmed by others. Conversely, individuals will also be motivated to avoid taking on negative roles in social interactions that may threaten their already established positive sense of self (Turner, 1982). So for IPV offenders, it should be expected that they will frame IPV as a conflict resolution strategy (Dobash and Dobash, 1998) and themselves as effective problem solvers, while at the same time denying the possibility that their actions constitute violence and that they are abusers.

When negative social roles are proscribed for an individual or past actions are reframed as evidence of negative role enactment, the focal individual will experience an identity threat (Petriglieri, 2011). Identity threats are "experiences appraised as indicating potential harm to the value, meanings, or enactment of an identity" (Petriglieri, 2011: 644). Offenders are likely to experience identity threats in reaction to treatment programs that present a negative frame for IPV and the offender's role (e.g., IPV is criminal behavior exhibited by violent abusers). This negative framing will threaten the offender's positive sense of self if the offender enters treatment with a positive framing of IPV and a perception that the role they play is also positive (e.g., IPV is a method of conflict resolution exhibited by effective problem solvers). This situation is likely to occur based on evidence suggesting that most offenders enter treatment in a state of denial, whereby they do not perceive their IPV behaviors as violence (Gondolf, 1987; Catlett et al. 2010). So, a negative frame offered in treatment would lead to the offender's perception that the treatment program is threatening the value, meaning, and enactment of the offenders "problem solver" identity, which would could potentially damage the offender's overall positive sense of self.

Identity threat may seem like a mildly negative psychological consequence of IPV treatment programs for offenders, necessary to put an end to IPV. After all, it would seem that in order to change the offender's behavior, the treatment must conflict with positive framings of IPV and threaten positive role definitions for abusers. However, research on sexual harassment training effectiveness, for example, has shown that identity threat reactions can often undermine treatment/training programs by motivating participants to backlash against the program and purposefully engage in the negative behaviors that the program is meant to prevent (Rawski, 2017). These findings are based on identity threat research that suggests a common method of coping with identity threat is to derogate the source of the threat. "Derogation of the source of the threat protects an individual's threatened identity by discrediting the source's validity, thus rendering irrelevant any potential harm"

(Petriglieri, 2011: 647). van Wormer and Bednar (2002) expressed concern for offenders' rejection of IPV treatment program narratives centered on feminist principals that define IPV as abusive men's way to control and dominate women. These authors recommended combining restorative justice (Carey, 1996) and strengths perspective (van Wormer, 1999) to highlight the positive strengths of offenders and how those strengths can be used to make amends. If this framing was offered in IPV treatment, offenders would more likely see a positive role for themselves and be less likely to experience identity threat and resist the treatment offered. This same logic has since been applied theoretically to diversity and sexual harassment training programs offered in policing organizations with toxic masculinity contest cultures (Rawski and Workman-Stark, 2017) and empirically to sexual harassment training in educational organizations (Rawski, 2017). These types of training too often frame trainees as "racists" or "sexists," rather than in positive roles such as potential "heroes" who are capable of reducing discrimination and harassment, once properly educated. Similarly, in IPV treatment programs, offenders will resist the role of abuser that is offered to them by a treatment program's framing of IPV, but may be more open to treatment that offers the offenders a positive role to play in reducing IPV.

If we apply this logic to the common treatment programs assigned to IPV offenders, we can determine if one treatment program is more likely to threaten participants' identities, and subsequently be rejected by the treatment participants. Two of the most popularly assigned IPV treatment programs are the *Duluth* model and the *Cognitive Behavioral Therapy (CBT)* model. The key to identifying which of these training programs is more likely to threaten offenders' identities lies in how each program frames IPV and the offender's role in it.

The *Duluth* model was one of the first treatment programs for IPV offenders and originated in women's shelters (Feder and Dugan, 2002). This treatment model is rooted in a feminist psychoeducational approach that frames IPV as a byproduct of a patriarchal society, which allows and encourages men to control and dominate women even with the use of violence (Scott, 2004; Pence and Paymar, 1993). This model makes use of the "power and control wheel," which depicts multiple ways that men exert abuse onto women; e.g., physical abuse, sexual abuse, economic abuse, or emotional abuse (van Wormer and Bednar, 2002). The power and control wheel is then juxtaposed with an "equality wheel" that depicts respectful and nonviolent behaviors aimed at challenging men's assumption that they are allowed to dominate and control their relationship partners with violence and abuse (Pence and Paymar, 1993; van Wormer and Bednar, 2002). The *Duluth* model is the predominate treatment program for IPV offenders in most US and Canadian communities, with some states even requiring that treatment programs adhere to *Duluth* principles in order to be certified (Babcock et al., 2004; Corvo et al., 2009). However, the *Duluth* model has been criticized for its lack of scientific underpinnings, its lack of involvement of qualified health

professionals, and its inability to address female abusers and IPV in same-sex couples (Pence and Paymar, 1993; Bohall et al., 2016).

The *Duluth* model frames IPV as a problem solely caused by men who take on the role of a violent abuser in order to control and dominate their partner (Pence and Paymar, 1993; Feder and Dugan, 2002). This frame stems from the feminist interpretation of IPV at a sociological level that focuses on institutionalized norms that promote patriarchy. While this frame is attractive in the sense that it can explain why IPV occurs without victim blaming, when this sociological frame is applied at the individual level it may result in IPV offenders' experiencing identity threat. The *Duluth* model frame for IPV results in offering primarily negative roles to IPV offenders during the treatment program. By labeling offenders as "abusers" and the sole cause of dysfunction and violence in their relationships, treatment programs based on the *Duluth* model do not provide positive roles for offenders to enact. Rather, the focus of the framing is on what offenders do wrong, a narrative that is likely to be rejected by many offenders, especially if they are in the denial phase of change (Catlett et al., 2010). Even though the equality wheel is offered as an alternative to abusive behavior, this information is still presented within the *Duluth* model frame that IPV is always and only the fault of men who are labeled as abusers. This negative role offering is even more likely to lead to identity threats because it occurs in a group setting where offenders may feel especially vulnerable and socially humiliated, which could in turn motivate more abusive behaviors (Jennings and Murphy, 2000). Indeed, in past research on identity threat reactions to training programs, identity threats motivated the behaviors that the training program was trying to prevent (e.g., sexual harassment; Rawski, 2017). Given the markedly negative roles offered to men in the *Duluth* model treatment for IPV, a similar result is expected, whereby this model may fail to successfully remediate, potentially leading to more IPV behaviors in the future. This assertion is already supported by results suggesting that one third of IPV offenders who participate in treatment based on the *Duluth* model reoffend (Herman et al., 2014).

Alternatively, the *CBT* model for IPV treatment is rooted in social learning theory (Bandura, 1973), which suggests that violence and abuse in intimate relationships are learned behaviors that have developed over time through observation and reinforcement (Fagen et al., 1983; Mihalic and Elliott, 1997). So, for instance, individuals who observe or experience abuse as children tend to perpetuate those behaviors onto their partners in adulthood (Oliver, 1993). Additionally, offenders who are rewarded for their abusive behaviors (e.g., obtaining compliance from their partner during a conflict) will be more likely to repeat those behaviors in the future, especially if they do not result in frequent punishments (e.g., arrests). Based on this theory of IPV as a learned behavior, the *CBT* treatment model focuses on developing offenders' skills for resolving conflict in nonviolent ways (Beck, 1999). This

model highlights how IPV is a suboptimal way of resolving conflict based on the negative consequences of IPV behaviors (e.g., an unhealthy relationship, partner injury, arrest) and promotes healthy conflict resolution techniques that result in more positive consequences.

The *CBT* treatment model frames IPV as a dysfunctional method of conflict resolution often triggered by psychological issues (e.g. stress, anxiety, depression), inadequate communication skills, and behavior modeling of mistreatment or violence observed in the past (Beck, 1976). By discussing IPV as a sub-optimal way of resolving conflict in relationships and introducing new skills (e.g., communication, anger management) that produce healthier forms of conflict resolution, the *CBT* model will likely tap into participants positive role identities, such as the problem solver role (Dobash and Dobash, 1998). Rather than threatening this positive role, the *CBT* model provides participants with tools to better and more effectively perform the role without resorting to IPV. As such, the *CBT* model is less likely to threaten offenders' identities. It is possible that *CBT*-based treatment may even affirm offenders' positive identities as problem solvers and motivate them to engage more with the treatment program to improve their conflict resolution skills. While the effects of *CBT* treatment for IPV have been modest, the *CBT* model has been shown to be the most effective to date (Babcock et al., 2004). Yet, Babcock et al. (2004) do caution that their findings should not be used to promote the *CBT* model over the *Duluth* model, highlighting the need for a direct experimental test of these two treatment types, a need that we specifically address in our experimental study.

Given the two framings of IPV offered by the *Duluth* model and the *CBT* model, we suggest that the *CBT* model will be more effective at reducing post-treatment IPV recidivism rates when directly compared with the *Duluth* model. This result is predicted to occur because the *Duluth* model frames IPV in a way that provides only negative roles for participants that are likely to threaten their identities and motivate more IPV in the future. In comparison, the *CBT* model frames IPV in a way that taps into participants' already established positive role identities, and trains participants to better perform these positive social roles without enacting IPV toward their relationship partners. Therefore:

Study Hypothesis: *The CBT Model treatment program will be more effective at reducing post-treatment recidivism of IPV than the Duluth Model treatment program.*

4   Study Design and Data

What seems to be infrequent in the literature is a direct comparison of a *Duluth* model to a *CBT* model in the same locale and during the same time frame. Thus, our study

presents an interesting inferential opportunity; specifically, it circumvents several conflating factors (e.g., differing political and macroeconomic conditions) that might otherwise arise when comparing a program of one type in one locale/time-frame to a program of the other type in a different locale/time-frame. Additionally, unlike in Dunford (2000), the population does not share dominant covariates (e.g., Navy-only, employed-only), but rather contains a broader representation of Society, as advocated for by Feder and Dugan (2002, pg. 351).

Unlike many prior studies, ours does not have an untreated group, but rather outcomes of those treated with the *Duluth* program are compared directly to the those treated with the *CBT* program. This head-to-head comparison permits an especially direct inquiry into the relative performance of each. Importantly, offenders in the two treatment groups were assigned by the local District Attorney's office to either the *Duluth* or *CBT* program in an automated, alternating fashion. No "tweaking" of the random assignment process was permitted.[1] We study this natural field experiment with the goal of understanding the relative effectiveness of the two distinct IPV diversion programs.[2]

Eligibility for treatment in either program was exactly the same for all participants; specifically they were all misdemeanor offenders, either first-time or with limited and dated priors. All those treated were initially arrested, served a criminal complaint, charged, and then offered treatment (governed by the aforementioned random assignment process). Upon completion of the treatment program, the charges were dismissed.

The study was conducted by the Winnebago County District Attorney's office in Oshkosh, Wisconsin, and the concomitant data were provided to us in a de-identified data set; additionally, we did not interact with any of the program participants in any way.[3] From January 2012 to January 2015, a total of 154 charged IPV offenders (male and female) were randomly assigned, as noted above, to one of the two IPV treatment programs. All of the subsequent criminal activity of these offenders, including IPV recidivism and non-IPV recidivism, was recorded by the District Attorney.

Unlike most prior studies, the specific experimental design permits *causal* inference regarding the effectiveness of *Duluth* versus *CBT* treatment; i.e., in theory the two treatment groups should differ in expectation only by the average differential treatment effect. However, even though the treatment assignment follows a completely randomized design, a single implementation of the experiment can result, by luck of the draw, in covariate imbalances

---

[1]The ethical implications of this assignment process were debated, but the churn in the literature ultimately led us to conclude, at the time the study commenced, that one treatment paradigm was not necessarily better than the other. Nonetheless, the final decision was not ours, but rather belonged to the District Attorney and his staff, which had been running both types of programs in parallel prior to our assessment.

[2]We use the phrase "natural field experiment" following the definition set forth in List (2008, pg. 206).

[3]Oshkosh, WI is part of the Fox River Valley area, which includes several adjacent cities, ranging from Oshkosh north to Green Bay; the aggregate population exceeds 650,000.

between the two treated groups; e.g., random assignment may, by chance, result in a disproportionate share of males or first-time offenders in one treatment or the other. Thus, we constructed our analysis in tiers. First, we give the direct results of the experiment, assessed by a simple t-test. Second, to address possible covariate imbalances and recidivism measured by duration, we apply logit and Cox proportional hazard analysis to the experimental data.

While the experiment conducted is not ideal, there are several reasons to believe this study is generalizable to other subjects in other locales in other time frames. First, it incorporates the criteria of a natural field experiment as outlined by Harrison and List (2004); i.e., the study is conducted with a naturally-occurring subject pool that participates in actual *Duluth* and *CBT* treatment programs and faces the natural consequences of recidivism. Second, since participants were unaware they were randomly assigned to one of two programs, there is little reason to believe the results were influenced by the study itself. Third, the two treatments were implemented at distinct facilities, which eases SUTVA concerns that treatment of one type was migrating from one treatment group to another.[4] Fourth, the metropolitan area from which the offenders arrived was large and diverse enough to reduce SUTVA non-interference concerns. Fifth, the *Duluth* program was administered by the same state-certified therapist and the *CBT* program was administered by the same advanced-degreed, licensed therapist. While the lack of variation across therapists introduces a potential confound in the interpretation of the experimental results, one benefit is that it reduces concerns about SUTVA consistency violations within each treatment. Sixth, both programs operated with similar group sizes.

In the following subsections, additional details of both models and summary-level statistics of individual characteristics are presented.

4.1 Local Duluth Model

In the *Duluth* program run by Winnebago County, the intervention program varied by sex. Men were presented with a program called "Domestic Violence Intervention: A Tiered Program for Men." This program began with an offender assessment in which consideration of risk factors put the participant into one of three tiers. The first tier was a 10-week group program designed to create awareness among men to their negative thinking and behaviors that generate IPV. The second tier was a 24-week group program that included the first tier plus a course on developing behavioral skills. The third tier was a 44-week group program that included both tier one and tier two plus self-analysis of male power, control and gender biases. All male individuals who participated in the *Duluth* program followed in this study were placed in the second tier. All enrollees were treated by the same state-certified therapist

---

[4]See the appendix for a broader discussion of SUTVA considerations in IPV studies to date.

in a group setting.[5] The local *Duluth* model was certified by the Wisconsin Batterer's Treatment Providers Association (WBTPA), which developed state-wide standards approved by the Wisconsin Governor's Council on Domestic Violence. Full curricular details are available from the authors by request.

The intervention program provided for females was called "Choices and Consequences." This was a 16-week group program designed to examine belief systems shaping the offender's domestic relationship. These individuals were encouraged to develop self-care safety planning strategies from their partner. Topics in this course included the cycle of violence, the effects of abuse, anger management, conflict management, and the power and control dynamics in relationships. Participants were encouraged to set goals for improving the equality in their domestic relationship and provided contact information on the resources available to them should IPV against them occur. Full curricular details are available from the authors by request.

4.2 Local CBT Model

The male and female batterers who participated in the *CBT* model program began by completing a comprehensive needs assessment similar to that of the men's *Duluth* model program. The program consisted of 16 weeks of *CBT* group counseling in which topics such as stress, anxiety, and anger management were discussed.[6] Solutions to conflict management were offered through examples of healthy relationships and effective communication skills. Specific issues that partners may need to address jointly, such as parenting and addiction, were discussed. Personal issues including grief, self-esteem, and decision making abilities were explored in the context of IPV so as to reduce the likelihood of reoffense. All enrollees were treated by the same advanced-degreed, licensed therapist in a group setting. Full curricular details are available from the authors by request.

4.3 Summary of Data

Table 1 presents the summary statistics of individual- and situation-specific factors. Importantly, we find that the randomization of offender assignment to the two programs led to a

---

[5]The lack of variation over therapists across both treatments is a shortcoming of this study. Ideally, enrollees would be randomly assigned across a set of therapists large enough to allow for standard errors to be clustered at the therapist level. However, this preferred alternative was infeasible. Despite this confound, we suggest the observed differences between treatment programs is, at least in part, reflective of the programs' conflict resolution techniques.

[6]Ideally, both treatment programs would have had the same number of sessions (e.g. both 16 weeks or both 24 weeks). The fact that those in the *CBT* program had fewer sessions would suggest, all else being equal, this treatment would be less effective at reducing IPV recidivism. However, given the results of this study, we must conclude the difference between the programs' effectiveness is potentially attenuated by *CBT*'s number of sessions. As a result the conclusions we draw between the programs is likely a conservative estimate of their relative effectiveness.

balanced representation across many of the individual characteristics. With 77 individuals participating in the *Duluth* program and the remaining 77 participating in the *CBT* program, there are no differences in race and average age, relatively small differences in prior convictions, and program duration. Propensity score matching identify the covariates in Table 1 are balanced across treatments, however some covariates are not ideally balanced; most notably there are more men in the *CBT* program; we specifically address this issue in a later subsection. While the *Duluth* program is significantly longer (for men) than the *CBT* program, participants were given some limited latitude as to when they attended sessions; but all sessions were required. Diversion programs were offered for a period of nine months and participants were allowed to postpone a week so long as it was completed within the stated time frame.

Table 1: Summary Statistics on Individual Characteristics

Offender Characteristics		
	Duluth Program	*CBT Program*
Observations	77	77
% Male	53.2%	67.5%
% White	89.6%	89.6%
Mean Age (at start)	32.5	32.5
% 1st time IPV Offender	74.0%	78.9%
% Completed Program	51.9%	67.5%
Mean Time Observed in Study (Days)	1248.6	1445.8

The preponderance of the data involve IPV between a spouse or a live-in partner ($\sim$75% for both programs). In descending order by frequency, the remaining $\sim$25% is represented by an ex-spouse/live-in, a family member, or a roommate. Offenders were 53.2% male in the *Duluth* program and 67.5% male in the *CBT* program. The *CBT* program has a higher completion rate of 67.5%, whereas the *Duluth* program has a 51.9% completion rate; this is likely related to the relative brevity of the former. The average number of days an offender is observed in our data set is approximately 3.5 years in the *Duluth* program and slightly under 4 years in the *CBT* program, but nonetheless quite similar.

5 Findings

Our primary and secondary findings appear in the following subsections. Following the theoretical motivations regarding frame analysis, we predict the relatively negative role offering given to male IPV offenders in the *Duluth* treatment program reduces its effectiveness, relative to the *CBT* treatment program. Overall, during the observation period, 8.4% of of-

fenders were convicted of a post-treatment IPV reoffense.[7] Some individuals committed a reoffense while still in the program, while others only after the program concluded. Our primary analysis focuses on post-treatment recidivism rates. One limitation we face when focusing on post-program reoffenses is a lack of variability among female offenders. In our data set, all instances of female recidivism in the *CBT* program were intra-program (i.e., prior to treatment completion). Sensitivity analyses involving intra-program recidivism were explored to augment the primary post-treatment analysis, the results of which are noted below where applicable.

5.1 Experimental Analysis

We begin by reporting the post-treatment probability of recidivism by program; see Figure 1. 13.0% of *Duluth*-treated offenders were convicted of a subsequent IPV act and 3.9% of *CBT*-treated offenders were convicted of a subsequent IPV act. A *t*-test reveals the estimated 9.1 percentage point difference in the probability of recidivism is statistically significantly lower in the *CBT* program at the five percent level ($p < 0.043$). This experimental outcome provides evidence in support of the *Study Hypothesis* and leads to Experimental Result 1.

Experimental Result 1: *The CBT treatment program was more effective than the Duluth treatment program at reducing post-treatment IPV recidivism.*

Because this analysis directly compares the *Duluth* and *CBT* models, we have no formal untreated control group. Thus, we can only determine if one treatment was more effective than the other, but not necessarily if either treatment was effective relative to no treatment. However, given the many results in the literature that conclude the *Duluth* model is, at worst, ineffective, we can plausibly deduce that the *CBT* model is, in fact, an effective IPV treatment option. By presuming that the *Duluth* model is effective to some degree (e.g., Davis et al., 2000), we only increase the plausibility of this deduction.

While many of the observable characteristics in Table 1 are closely balanced, several covariates, such as the percentage of male offenders, differ somewhat across treatments. As is frequently noted, recidivism is more likely among males. Thus, a disproportionate number of men in the *CBT* program may bias the results against the *CBT* treatment. To address this concern, we perform a logistic regression using the experimental data; the results appear in Table 2. Two separate estimates are reported, with coefficients on the left and the corresponding marginal effects on the right.

In the first panel, the regression model includes all offenders and accounts for observable

---

[7]We also looked at non-IPV recidivism, but no statistical differences were found between programs for this type of reoffense.

Figure 1: Probability of Recidivism
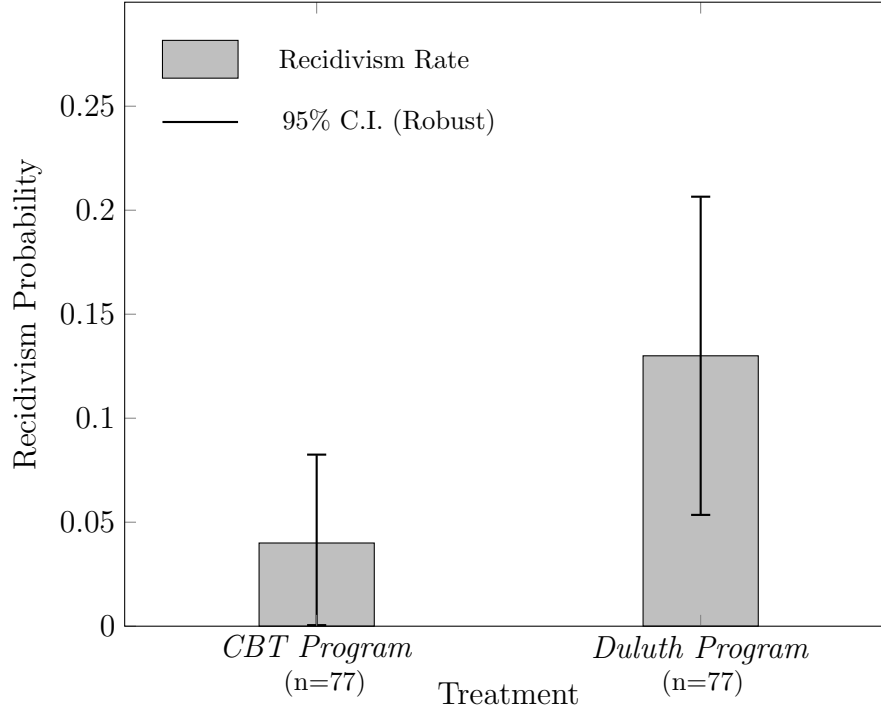
covariates; the dependent variable is post-treatment IPV conviction (recidivism). By this covariate-based estimate, the probability of recidivism is 11 percentage points higher among *Duluth* participants than *CBT* participants. This result is significant at the 5% level. The logit results also indicate that men are 13 percentage points more likely than women to be convicted of a post-treatment IPV act. We find no statistical evidence that age, prior conviction, or race affect IPV reoffense when looking at male and female offenders together (i.e., all offenders).

The estimates presented in the second panel of Table 2 are similar to the first, but focus only on male offenders. Given the theoretical motivation regarding frame analysis, it is predicted that male offenders are more likely to reoffend in the *Duluth* treatment program due to the negative framing of their role in IPV. The results suggest the *Duluth* model is statistically less effective (at the 5% level) at reducing IPV recidivism; specifically, participants in the *Duluth* model are 14 percentage points more likely to be convicted of a post-treatment IPV act than participants in the *CBT* program. This leads to Experimental Result 2.

Experimental Result 2: *Among male IPV offenders, the Duluth treatment program is less effective than the CBT treatment program at reducing post-treatment IPV recidivism.*

As a robustness check, the logit model was also estimated where recidivism included any

Table 2: Logit Estimation on Recidivism

Dependent Variable: Post-Treatment IPV Conviction

	All Offenders		Male Offenders	
	Coefficients	Marginal Effects	Coefficients	Marginal Effects
Duluth Model	1.59**	0.11**	1.48**	0.14**
	(0.68)	(0.05)	(0.74)	(0.07)
Age	-0.04	-0.003	-0.06*	-0.01*
	(0.03)	(0.002)	(0.04)	(0.003)
Male	1.89**	0.13**	–	–
	(0.91)	(0.06)		
1st Offense	0.93	0.07	0.77	0.07
	(0.77)	(0.06)	(0.86)	(0.08)
White	-0.22	-0.02	0.88	0.08
	(1.15)	(0.08)	(1.27)	(0.12)
Constant	-4.12***	–	-2.17	–
	(1.38)		(1.69)	
Observations	154		93	

Notes: ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively. Standard errors are reported in parentheses. Individuals in the *Duluth* program were observed for an average of 1249 days, while individuals in the *CBT* program were observed for an average of 1446 days.

IPV conviction during or after treatment (as opposed to only after treatment, as above).[8] The estimates and conclusions are qualitatively similar to those presented for the post-treatment-only recidivism measure. Finally, though the results are not included in Table 2, we find no interaction effect between diversion programs and the presence of a prior IPV conviction. This result suggests there is no statistical reason to favor one program over the other for repeat offenders. [9]

5.2 Hazard Analysis

In addition to the simple *t*-test and the subsequent logit analysis presented above, we also present a hazard analysis of recidivism, again using the experimental data. While this method does appear in criminological studies, it is somewhat less common than logit estimation. Here, it allows us to estimate how much time, on average, will pass prior to reoffense among those that do, indeed, reoffend after treatment. Specifically, the value of the hazard model is that it allows for the duration of time between reoffense to act as the

---

[8]These results are available from the authors upon request.
[9]These results are available from the authors upon request.

outcome of interest rather than simply defining the outcome as reoffense at any time. This is a potentially important distinction to investigate because it is possible that one treatment may result in a higher probability of reoffense (as identified in the above analysis), but the comparison treatment may result in shorter time to reoffense, albeit less likely overall. If this example were the case, it would mitigate the conclusions about superior performance between treatment programs on the whole. Nevertheless, understanding the relative impact of the treatment programs on time-to-reoffense, not simply the probability of reoffense, is also important in this context.[10]

5.2.1 IPV Hazard Results

The IPV hazard results, which appear in Table 3, can be interpreted using the logic displayed in the medical example (see Appendix). Here, we use Cox proportional hazard model, which assumes that the treatment and control hazard rates occur in the same proportion throughout the time line; this is a commonly applied way to facilitate the interpretation of covariate-based duration results. The results suggest an individual who completed the *Duluth* program, and who has not yet reoffended, is 4.19 times more likely to reoffend in the next day than someone who completed the *CBT* program. While this result is statistically significant at the 5% level, other considerations such as age, race, and whether or not they are a first time offender are not. However, the largest effect on reoffense is sex, as men are approximately 6.1 times more likely to commit a subsequent IPV offense, in a subsequent day, than women. These results suggest that all offenders (i.e., men and women) are less likely to reoffend on any given day after being randomly assigned to the *CBT* program versus the *Duluth* program.

Similar to those results from the logit model in Table 2, when the data are restricted to male offenders we find that men are 3.65 times more likely to reoffend, in a subsequent day, after completing the *Duluth* program (relative to *CBT* program completion), and adding an additional year to a offender's age reduces his likelihood of reoffending in a subsequent day by approximately 5%.

6   Discussion and Conclusions

This paper analyzes a natural field experiment to better understand the relative effectiveness of two competing models of IPV diversion. Individuals convicted of IPV were randomly assigned to either a stylized *Duluth* or *CBT* IPV treatment program. We study the behavior of those involved with these programs for over three years on average, and compare the rates of

---

[10]For readers unfamiliar with survival models, such as the Cox proportional hazard, we present a basic medical example of its functionality in the Appendix.

Table 3: Cox Proportional Hazards Estimation on Recidivism

Dependent Variable: Post-program IPV Offense

	All Offenders		Male Offenders	
	Hazard Ratio	95% C. I.	Hazard Ratio	95% C. I.
Duluth Model	4.19** (0.03)	[1.19, 14.79]	3.65** (0.05)	[0.99, 13.34]
Age	0.96 (0.22)	[0.91, 1.02]	0.95* (0.08)	[0.89, 1.01]
Male	6.13** (0.03)	[1.20, 31.44]	–	–
1st Offense	1.89 (0.43)	[0.39, 9.10]	2.05 (0.38)	[0.42, 10.13]
White	0.80 (0.82)	[0.11, 5.83]	2.10 (0.52)	[0.21, 20.48]
Observations	151		90	

Notes: ***, **, and * indicate significance at the 1%, 5% and 10% level,
respectively. P-values are reported in parentheses. Two individuals
(i.e. observations) were dropped due to missing program data.

IPV recidivism. Using frame analysis, we hypothesize the effectiveness of each program will be determined, in part, by their methods of assigning roles to offenders and victims of IPV. In the *Duluth* program, male offenders are provided with a relatively negative role offering in which they are charged with being the sole cause of dysfunction and violence in their relationship. This perceived identity threat may cause male offenders to respond adversely to the treatment program, which then would lead to an increased risk of recidivism. Alternatively, the *CBT* program relies on the already established positive role identities among offenders to address some of the causes of IPV, such as poor anger management or communication skills. These program differences motivate our hypothesis that the *CBT* treatment program will be more effective than the *Duluth* program at reducing IPV recidivism. Analyses using *t*-tests, non-linear probability models, and Cox proportional hazard models support this hypothesis, suggesting the *Duluth* program is less effective at curbing rates of recidivism overall, with stronger effects among male offenders.

Empirical and methodological limitations to our study include the following:

- Our sample size of 154 satisfied desiderata D5, however a larger sample size would have allowed for a more nuanced analysis of relevant factors.

- Each treatment was administered by a single therapist, which introduces a confound in measuring the relative effectiveness of each treatment program that cannot be properly

controlled for.

- Because we directly compared the *Duluth* and *CBT* models, we have no formal untreated control group. Thus, we can only determine if one treatment was more effective than the other, but not necessarily if the treatment was effective relative to no treatment.

- All treatment options entailed group therapy of various types and, as such, may expose the study to a SUTVA violation; however, it is worth noting that group therapy is the norm in these type of programs, and is not unique to our circumstance, nor to how these programs are often applied.

Like others that have utilized experimental design to evaluate IPV treatment, we advocate for additional future usage of this methodology, as it is the most likely to induce reliable inferential results. Future research might also make more extensive use of survival analysis as it is helpful at capturing the impact of treatment on the average time-to-reoffense, which can have important cost and workload implications for police, courts, and the penal system. Another matter for further inquiry concerns the tailoring of IPV treatment based on covariates of the offender, infraction, or victim.

References

[1] Babcock, J. C., Green, C. E., and C. Robie. 2004. "Does batters treatment work? A meta-analytic review of domestic violence treatment." *Clinical Psychology Review*, 23: 1023–1053.

[2] Bandura, A. 1973. "Aggression: A social learning analysis." Englewood Cliffs, NJ: Prentice Hall.

[3] Baumeister, R. F. 1998. "The self." In D. T. Gilbert, S. T. Fiske, G. Lindzey (Eds.) , *The handbook of social psychology*, Vols. 1 and 2 (4th ed.) (pp. 680-740). New York, NY, US: McGraw-Hill.

[4] Beck, A. 1976. *Cognitive therapy and the emotional disorders.* New York: International University Press.

[5] Beck, A. T. 1999. "Prisoners of hate: The cognitive basis of anger, hostility, and violence." New York, NY: HarperCollins.

[6] Berk, R. 2005. "Randomized experiments as the bronze standard." *Journal of Experimental Criminology*, 1: 416–433.

[7] Black, M.C., Basile, K.C., Breiding, M.J., Smith, S.G., Walters, M.L., Merrick, M.T., Chen, J., and M.R. Stevens. 2011. "The National Intimate Partner and Sexual Violence Survey (NISVS): 2010 Summary Report." Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention. Retrieved from: https://www.cdc.gov/violenceprevention/pdf/nisvs_report2010-a.pdf.

[8] Bohall, G., Bautista, M., and S. Musson. 2016. "Intimate partner violence and the Duluth model: An examination of the model and recommendations for future research and practice." *Journal of Family Violence*, 31: 1029–1033.

[9] Carey, M. 1996. "Restorative justice in community corrections." *Corrections Today*, 58(5): 152–156.

[10] Catlett, B. S., Toews, M. L., and V. Waliko. 2010. "Men's gendered constructions of intimate partner violence as predictors of court-mandated batterer treatment drop out." *American Journal of Community Psychology*, 45: 107–123.

[11] Corvo, K., Dutton, D., and W. Chen. 2009. "Do Duluth model interventions with perpetrators of domestic violence violate mental health professional ethics?" *Ethics & Behavior*, 19(4): 323–340.

[12] Davis, R., Taylor, B., and C. Maxwell. 2000. "Does batterer treatment reduce violence? A randomized experiment in Brooklyn - executive summary included." National Institute of Justice, Washington, DC.

[13] Dobash, R. P. and R. E. Dobash. 1998. "Violent men in violent contexts." In R. E. Dobash and R. P. Dobash (Eds.), *Rethinking violence against women* (pp. 141 168). Thousand Oaks, CA: Sage.

[14] Dunford, F. 2000. "The San Diego Navy Experiment: An assessment of interventions for men who assault their wives." *Journal of Consulting and Clinical Psychology*, 68: 468–476.

[15] Dutton, D. and K. Corvo. 2006. "Transforming a flawed policy: A call to revive psychology and science in domestic violence research and practice." *Aggression and Violent Behavior*, 11: 457–483.

[16] Dutton, D. and K. Corvo. 2007. "The Duluth model: A data-impervious paradigm and a failed strategy." *Aggression and Violent Behavior*, 12: 658–667.

[17] Fagen, J. A., Stewart, D. K., and K. V. Hansen. 1983. "Violent men or violent husbands? Background factors and situational correlates." In D. Finkelhor, R. J. Gelles, G. T. Hotaling, and M. A. Straus (Eds.), *The Dark Side of Families*. Beverly Hills, CA: Sage.

[18] Feder, L. 1997. "Domestic violence and police response in a pro-arrest jurisdiction." *Women and Criminal Justice*, 8: 79–98.

[19] Feder, L. and L. Dugan. 2002. "A test of the efficacy of court-mandated counseling for domestic violence offenders" The broward experiment." *Justice Quarterly*, 19: 343–375.

[20] Feder, L. and D. Wilson. 2005. "A meta-analytic review of court-mandated batterer intervention programs: Can courts affect abusers' behavior?" *Journal of Experimental Criminology*, 1: 239–262.

[21] Gangopadhyay Armstrong, T., Wernke, J. Y., Lisdahl Medina, K., and J. Schafer. 2002. "Do partners agree about the occurrence of intimate partner violence? A review of the current literature." *Trauma, Violence, & Abuse*, 3(3): 181–193.

[22] Gecas, V. 1982. "The self concept." *Annual Review of Sociology*, 81–33.

[23] Goffman E. 1974. "Frame Analysis: An essay on the organization of experience." Boston, MA: Northeastern University Press.

[24] Gondolf, E. 1987. "Changing men who batter: A developmental model for integrated interventions." *Journal of Family Violence*, 2: 335–349.

[25] Gondolf, E. 2007. "Theoretical and research support for the Duluth model: A reply to Dutton and Corvo." *Aggression and Violent Behavior*, 12: 644–657.

[26] Goolkasian, G. 1986. "Confronting domestic violence: The role of criminal court judges." Washington, DC: National Institute of Justice.

[27] Harrison, G. and J. List. 2004. "Field experiments." *Journal of Economic Literature*, 42: 1009–1055.

[28] Herman, K., Rotunda, R., Williamson, G., and S. Vodanovich. 2014. "Outcomes from a Duluth model batterer intervention program at completion and long-term follow-up." *Journal of Offender Rehabilitation*, 53: 1–18.

[29] Hogg, M. A., and D. J. Terry. 2000. "Social identity and self-categorization processes in organizational contexts." *The Academy Of Management Review*, 25(1), 121–140.

[30] Imbens, G. and T. Lemieux. 2008. "Regression discontinuity designs: A guide to practice." *Journal of Econometrics*, 142: 615–635.

[31] Jennings, J. L. and C. M. Murphy. 2000. "Male-male dimensions of male-female battering: A new look at domestic violence." *Psychology of Men and Masculinity*, 1(1): 21–29.

[32] Larsen, S. E. and L. K. Hamberger. 2015. "Mens and womens experience of IPV part II: A review of new developments in comparative studies in clinical populations." *Journal of Family Violence*, 30: 1007–1030.

[33] List, J. 2008. "Introduction to field experiment in economics with application to economics of charity." *Experimental Economics*, 11: 203–212.

[34] Max, W., Rice, D. P., Finkelstein, E., Bardwell, R. A., and S. Leadbetter. 2004. "The economic toll of intimate partner violence against women in the United States." *Violence and Victims*, 19(3), 259–272.

[35] Mihalic, S. W. and D. Elliott. 1997. "A social learning theory model of marital violence." *Journal of Family Violence*, 12: 21–47.

[36] Oliver, J. E. 1993. "Intergenerational transmissions of child abuse: Rates, research, and clinical implications." *American Journal of Psychiatry*, 150, 1315–1324.

[37] Pence, E., and M. Paymar. 1993. *Education groups for men who batter: The Duluth model.* New York: Springer.

[38] Petriglieri, J. L. 2011. "Under threat: responses to and the consequences of threats to individuals' identities." *Academy of Management Review*, 36(4), 641–662.

[39] Ptacek, J. T. 1988. "How men who batter rationalize their behavior." In A. L. Horton 7 J. A. Williamson (Eds.), *Abuse and religion: When praying isnt enough* (pp. 247–257). Lexington, MA: Lexington Press.

[40] Rawski, S. L. 2017. "The effects of identity threat reactions to sexual harassment training on training outcomes." Accepted for Presentation at the 77th Academy of Management Annual Meeting, Atlanta, Georgia.

[41] Rawski, S. L. and A. Workman-Stark. 2017. "Sexual harassment training in extreme masculinity Contests: Questions raised by a masculinity contest in policing." Accepted for Presentation at the 77th Academy of Management Annual Meeting, Atlanta, Georgia.

[42] Rubin, D. 1980. "Randomization analysis of experimental data: The fisher randomization test comment." *Journal of the American Statistical Association*, 75: 591–593.

[43] Scott, K. L. and D. A. Wolfe. 2003. "Readiness to change as a predictor of outcome in batterer treatment." *Journal of Consulting and Clinical Psychology*, 71(5): 879–889.

[44] Scott, K. L. 2004. "Predictors of change among male batterers: Application of theories and review of empirical findings." *Trauma, Violence, and Abuse*, 5(3): 260–284.

[45] Sherman, L. 1992. "The influence of criminology on criminal law: Evaluating arrests for misdemeanor domestic violence." *The Journal of Criminal Law and Criminology*, 83: 1–46.

[46] Tajfel, H. and J. C. Turner. 1985. "The social identity theory of intergroup behavior." In S. Worchel and W. G Austin (Eds.) Psychology of Intergroup Relations, 2nd ed. Chicago: Nelson Hall, 7–24.

[47] Truman, J. L. and R.E. Morgan. 2014. "Special Report: Nonfatal domestic violence 2003-2012." U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics. Retrieved from: https://www.bjs.gov/content/pub/pdf/ndv0312.pdf.

[48] Turner, J. C. 1982. "Towards a cognitive redenition of the social group." *Social identity and intergroup relations*, 15–40.

[49] van Wormer, K. 1999. "The strengths perspective: A paradigm for correctional counseling." *Federal Probation*, 63(1): 51–58.

[50] van Wormer, K. and S. G. Bednar. 2002. "Working with male batterers: A restorative-strengths perspective." *Families in Society*, 83(5/6): 557–565.

7 Appendix

7.1 SUTVA in IPV

The Stable Unit Treatment Value Assumption (SUTVA) is similar to the notion of statistical independence. The SUTVA has a "no interference" component and a "consistency" component.

For example, in an IPV treatment assessment setting, if some of the individuals randomly assigned to the treatment group happen to have negative pre-existing relationships, this would likely violate the SUTVA, and we could reasonably expect the apparent effectiveness of the treatment to be understated. As a second example, if treatment somehow clandestinely passes from a treated offender to an untreated offender, the non-interference component would be violated. As a third example, if treatment effort is uneven across treated offenders, this would violate the consistency component.

SUTVA violations are particularly likely in cases where the treatment and control regimens are applied to groups, members of which can then interact, or in any cases wherein members of a treatment group are able to receive heterogeneous treatments. Both increase the chance of "cross contamination," which in turn can cloud the results. There is no clean, general way to circumvent violations of the SUTVA; often the best remedy is careful construction and implementation of the inference process. SUTVA mindfulness is subsumed in desideratum D7, and we will revisit this topic when discussing our local IPV treatment programs. Fortunately, according to Berk (2005, pg. 421), minor and/or infrequent violations of the SUTVA may not meaningfully undermine the inference process; however, vigilance is unequivocally warranted regardless of this potential "breathing room." See Rubin (1980) and subsequent research for extended discussions of the SUTVA.

Both of the SUTVA components are often tacitly assumed to be satisfied in many inference settings (inside and outside of Criminology), but a study designer can benefit from explicitly considering the SUTVA components when planning an IPV treatment program evaluation.

7.2 Hazard Model Discussion and Example

In investigating determinates of duration before reoffense, a Cox proportional hazard model relates the time that passes before a "failure" occurs (in our case an IPV reoffense) to one or more characteristics about an individual that may also be associated with the failure (in our case treatment program assignment). Therefore, the proportional hazard approach provides a determination whether a participants hazard rate of failure is impacted by some treatment.

In particular, the Cox proportional hazard regression model produces an adjusted hazard ratio that accounts for baseline differences between offenders that may influence the choice to reoffend that are not already captured in the predictor variables. Within this hazard analysis

framework, the goal is to capture the factors that lead to reoffense as best as possible.

As an example, imagine a medical trial with 1000 patients, 500 of which receive a new medication thought to reduce heart attacks and 500 of which do not receive the new medication. Assume the trial lasts 60 months, and assume patients are allocated to treatment and control using a completely randomized design. Within this context, there will likely be heart attacks during the trial, some in the treatment group and some in the control group. The hazard rate can measure this. Specifically, the hazard rate, often denoted as $h(t)$, is the probability that a subject has a heart attack at time t assuming the subject has survived to month t.

Further suppose that 8 treatment and 10 control subjects suffer heart attacks in the first month. The hazard rate for the treatment group, $h(t = 1)$, would be $8/500 = 0.016$, while the hazard rate for the control group would be $10/500 = 0.02$. If, in the second month, the treatment group suffers 6 more heart attacks and the control suffers 9 more, the respective hazard rates $h(t = 2)$ would be $6/492 = 0.012$ and $9/490 = 0.018$. Upon completing the study, we would have 60 hazard rates for the treatment group and 60 hazard rates for the control group; the differences between them capture the treatment effect. One way to summarize this difference is by creating a ratio—a hazard ratio—of the treatment hazard rate at a given time divided by the control hazard rate at the same time. If the treatment and control groups showed no differences in heart attack frequency during the trial, the hazard ratio would equal one; i.e., the two hazard rates would be equal to each other, indicating that the medication appears ineffective in reducing heart attacks. Alternatively, a hazard ratio differing significantly from one indicates a meaningful difference between the treatment and control.