P. S. HOYMAND

PHD

2003

This is to certify that the
dissertation entitled

MANAGING A PROSECUTOR'S OFFICE DOMESTIC
VIOLENCE CASELOAD TO INCREASE ASSAILANT
ACCOUNTABILITY: A SYSTEM DYNAMICS APPROACH

presented by

PETER SVEND HOVMAND

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Social Work

Diane G. Levande
Major Professor's Signature

5-6-03
Date

MANAGING A PROSECUTOR'S OFFICE DOMESTIC VIOLENCE CASELOAD TO INCREASE ASSAILANT ACCOUNTABILITY: A SYSTEM DYNAMICS APPROACH

By

Peter Svend Hovmand

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

School of Social Work

2003

ABSTRACT

MANAGING A PROSECUTOR'S OFFICE DOMESTIC VIOLENCE CASELOAD TO INCREASE ASSAILANT ACCOUNTABILITY: A SYSTEM DYNAMICS APPROACH

By

Peter Svend Hovmand

Domestic violence is a major social problem. Many communities have adopted a coordinated approach to increasing victim safety and assailant accountability. Most communities continue to rely on the criminal justice system as the primary means of holding assailants accountable. Criminal justice system reforms such as mandatory arrest policies have increased the number of arrests and increased the role of the prosecutor's office in holding assailants accountable for their actions. This research considered the question of how the domestic violence caseload of a prosecutor's office was dynamically related to holding assailants accountable.

The research used a single case design of a prosecutor's office from 1998 to 2001 with a system dynamics approach, which is a method that uses analysis of computer models to understand the relationships between the structure and behavior of complex systems in terms of nonlinear feedback loops. Data sources for the model building included numerical time series and key informant interviews. A major challenge for this study was the problem of modeling sparse time series associated with small populations. The problem was solved by evaluating the realism of the models' responses to a family of smoothed numerical time series. A series of models were built and tested. The resulting baseline model was formulated mainly in terms of caseflows, which included arrests, review of warrants, prosecution of cases, and

the allocation of prosecutor resources between review of warrants and prosecution of cases.

The results found general support for the hypothesis that case dispositions are affected by caseloads. The baseline model did a reasonable job of reproducing the general trends, but it did not capture the dynamics of dismissals of repeat arrests. Analysis of the baseline model showed that reductions in prosecutor resources to domestic violence caused warrant reviews to be dominated by caseload pressures as opposed to case attributes. The baseline model also showed that if there were sufficient resources for prosecuting domestic violence cases, then the prosecutor's office could handle large increases in new cases without affecting prosecution dynamics. Analysis of the feedback loops revealed that when caseloads were stable, accountability and caseloads were both controlled by the allocation of resources to reviewing warrants. Analysis of arrests led to the identification of a potential positive feedback loop from first time arrests of male assailants to subsequent arrest of female victims. That is, arresting some types of assailants led to an increased risk of victims being arrested as assailants learned police officers' criteria for a domestic assault and subsequently used that knowledge to get victims arrested. Separate analyses and key informant interviews supported this finding. This would represent an unintended consequence of mandatory arrest laws for domestic violence. Key informants suggested a time-critical counteracting feedback mechanism, namely, victim interviews within a week of the arrest. Other findings included indications that the prosecutor's office resource allocation problem was dynamic, and that trials play a significant role in determining the outcomes of all case dispositions. System dynamics showed great potential for studying the prosecution of assailants and community responses to domestic violence.

ACKNOWLEDGMENTS

This dissertation was by no means an individual achievement. Many people supported the effort in one or more ways. The faculty and staff of the Michigan State University School of Social Work provided a unique environment for pursuing doctoral research on domestic violence using a system dynamics approach. Diane Levande, my chair, enthusiastically supported the research and intellectual excursions from beginning to end. She shared a passion for understanding and confronting violence against women, systems thinking, and philosophy. Cynthia Jackson-Elmoore brought a keen objectivity through her focused questions and insights. Marilyn Frye helped me develop a more critical ear for the language, metaphors, and politics that have shaped the philosophy underlying this dissertation. Ralph Levine provided numerous opportunities for learning the basics of system dynamics in the classroom, gaining modeling experience on research projects, and challenged me with more advanced concepts. Cris Sullivan helped me put on the page what I had learned from working in the community.

Leslie Hagen provided early opportunities for me to see the potential of this research. Cressida Heyes provided a solid foundation in her political philosophy seminar for thinking about power and posed early questions about the nature of system dynamics for social science research. David Garvin introduced me to notion of obfuscation, Mike Jackson showed me the practice of accountability, and James Nelson helped me develop these ideas philosophically his Wittgenstein seminar. David Lounsbury introduced me to Ralph Levine and system dynamics. Jack Homer impressed upon me the importance of grounding reference modes in real numerical time series. David Ford provided an invaluable opportunity to learn how to model resource allocation problems and answered my questions about feedback loop dominance.

Lori Post, Chris Maxwell, and Ellen Hayes generously supported this work by allowing me to use their data and provided access to key informants. This research was partially supported through research on various grants by the Centers Disease Control and Michigan Department of Community Health.

MSU Safe Place funded the majority of my doctoral studies through graduate assistantships and kept me in touch with the immediate issues facing communities. Holly Rosen has been an inspiration in her work as a community activist, leader in the battered women's movement, and social worker.

Stephen Jefferson, Bev Henrichsen, Peg Whalen, and Melissa White were amazing throughout the process. They read countless drafts and heard numerous presentations at the earliest and most confusing stages, as did my parents, Beverly and Svend Hovmand, and brother, Lars Hovmand, who asked me questions along the way that made the rest easy. Charlene Warner and Eyvind Hovmand Warner were always there. They felt the ups and downs, the twist and turns, as much as me.

None of this would have been possible, however, without the generosity of the participants I interviewed and their dedication to help victims of domestic violence and hold assailants accountable. They took time out of their hectic schedules, even on a Friday evening, to talk about their experiences.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This dissertation is about improving the effectiveness of prosecution of batterers in cases of domestic violence. There are many ways that one might think about domestic violence, but it is foremost a facet of sexism–the oppression of women by men. Domestic violence names a category of practices used by men in an effort to domesticate women into compulsory heterosexual, monogamous, and patriarchal relationships through the use or threat of violence.[1] To say that these practices are used in an *effort* to domesticate women means that the effects on women are neither uniform, total, nor universal. Domestic violence is coercive at both the individual and macro level. And to say that these practices are used by men means that men as a social group[2] benefit from the individual batterer who abuses by drawing on the specter of the abusive relationship to structure the relationships of women to men.

[1]This does not presuppose that women cannot be abusive, men cannot be victims of domestic violence, or that same-sex domestic violence does not happen. Nor does it mean that "other" kinds of cases are uninformative in the analysis of sexism and the oppression of women.

[2]The notion of a social group is a macro or structural concept. Social groups are socially constructed categories, a part of social reality, and refer in the sense of Iris Young (1990, pp. 42-47) to the classes that structure the relationship between oppressor and oppressed. These classes are populated by individuals who's membership varies or might not even fit the category term. So theoretically, being male does not mean that one is a man. But, it raises the question about the extent that males might be able to meaningfully separate from membership or association with the men as a social category.

Domestic violence includes practices such as a batterer verbally abusing a woman, isolating her from friends and family, preventing her from getting or keeping a job, making threats of violence, and beating, raping, or killing her.

Legislators and local agencies have responded incrementally to increased public pressure by funding domestic violence domestic programs, providing civil remedies like personal protection orders, criminalizing batterer behavior, and reforming procedures for handling cases of domestic violence. In terms of holding assailants accountable, most reforms have focused narrowly on domestic violence as a criminal behavior. Some of the promoted changes to the criminal justice system have included adoption of pro or mandatory arrest policies, mandatory referral of domestic assault police reports to the prosecutor's office, implementation of special prosecution units, and court mandated referrals to batterer intervention programs. (Pro-arrest policies encourage police officers to make warrantless arrests in cases of misdemeanor domestic violence whereas mandatory arrest policies require police officers to make a warrantless arrest in cases of misdemeanor domestic violence.) However, jurisdictions have varied considerably in their actual implementation.

Reformers have tended to explain resistance to change in terms of attitudes since individuals working in the criminal justice system have frequently blamed the victim, minimized the abuse, or colluded with the assailant. Appointing a police chief or electing a prosecutor or judge who understands the dynamics of domestic violence does make a difference, but this approach is limited to the extent that reforms are not institutionalized. Moreover, institutionalizing reforms through changes in organizational policies does not by any means guarantee their effectiveness, especially when the changes ignore the real working conditions of practitioners. Some in the battered women's movement like Ellen Pence (1999) have similarly reconsidered their approach to understanding the criminal justice system:

> In the past, we might have asked, 'Why did this practitioner take this action?' Now we ask, 'How was this practitioner institutionally organized to take this action?' Instead of seeing actions as the result of what goes on in the head of a judge or probation officer or police officer, we see it as the result of what is going on in the work practices. (pp. 37-38)

Reforms that pay attention to how practices are structured are more likely to lead to organizational changes that are sustainable, and hence last long enough for the results to accumulate and impact assailant accountability.

1.1 Research question

One obvious way to consider how practitioners are institutionally organized in the criminal justice system is to look at the structure of their work-flows. This has become increasingly important as more communities have adopted mandatory arrest policies, which have effectively restricted police officers' discretion, moved more cases into the prosecutor's office, and thus displaced discretion in the criminal justice system from police officers to prosecutors. Hence, many communities are noticing increases in the number of domestic violence cases where warrants are denied or cases are dismissed, potentially undermining the benefits of mandatory arrest and prosecution. That is, there appears to be a relationship between the caseloads of prosecutor's offices and dispositions that affects the extent to which communities can hold assailants accountable via the criminal justice system. Thus, the general research question for this study is: how does a prosecutor's office dynamically manage its domestic violence caseload?

1.2 System dynamics

Understanding how a prosecutor's office manages its domestic violence caseload can be conceptualized in a variety of ways. In this study, it is understood in terms of feedback control mechanisms or feedback loops. A system dynamics approach is used to identify and understand the feedback loops in terms of patterns of feedback loop dominance. The notion of feedback loop dominance is relevant to complex systems where the dynamic behavior of the system is being understood in terms of the structure of feedback loops, where the set of feedback loops generating the system's behavior over time also changes over time.

System dynamics is a method for modeling and evaluating complex systems of feedback loops, which are formulated as systems of coupled nonlinear ordinary differential equations. Such systems of equations are generally impossible to solve analytically. So system dynamics modeling relies heavily on numerical methods and computer simulation to solve this system of differential equations. The purpose of model building in system dynamics is to understand and solve problems in terms of feedback loops. The interest in studying how a prosecutor's office manages its domestic violence caseload is to increase assailant accountability. So the question for this study in terms of system dynamics becomes one of understanding the relationship between the (a) mechanisms regulating the caseload in a prosecutor's offices, and (b) the mechanisms regulating the prosecutor's office holding batterers accountable. This can be understood as identifying the feedback loops (if any) that dominate both some measure of accountability and the prosecutor's office caseload during a particular phase of dynamic behavior.

1.3 Overview

Chapter 2 provides the broad theoretical framework for this study and a review of the existing literature on domestic violence arrests, prosecution research, and batterer intervention programs, reviewing several lines of domestic violence and prosecution related research. Included is a summary of the research on deterrence from arrest, evaluations of batterer intervention programs, and review of prosecution research including the prosecution of domestic violence cases and plea bargaining.

Chapter 3 covers the methods. One of the main goals for this study is to develop methods for studying community responses to domestic violence with particular attention paid to modeling small populations. The chapter describes the use of a single case study design, use of the main database for generating the numerical time series from the prosecutor's office violence against women (VAW) data base, methods for handling the resulting sparse time series associated with small populations, and the development of a finite state machine approach to studying the transition patterns of individuals within the database.

Chapter 4 describes the results, starting with descriptive statistics of the VAW database and comparisons against census demographic graphics. The results from testing three descriptive models are then discussed. These models simply test to what extent (a) simple systems of differential equations can reproduce the general behavior pattern of the numerical time series, and (b) to what extent dispositions can be formulated as functions of caseload relative to other formulations. This is followed by discussion of the development and testing of the baseline model, which begins with each submodel and concludes with an analysis of the integrated baseline model. A variety of tests are conducted on the baseline model, including sensitivity to various parameters, frequency response, and smoothing parameters used in handling the sparse time series. The relationship between the behavior of the baseline model and structure of the feedback loops is then analyzed in terms of identifying patterns of feedback loop dominance. The next major section of Chapter 4 covers follow-up analysis examining the disaggregation of arrests by type of relationship and sex of suspect. The chapter concludes with a summary of the key informant interviews, which were based on questions developed during the development and testing of the baseline model, and proposed inclusion of additional feedback loops for a revised model.

Chapter 5 is the conclusion. This chapter summarizes the major findings of the study, major limitations, policy implications, and directions for future research.

Chapter 2

Review of research literature

2.1 Overview

This chapter introduces the problem of domestic violence as a major social problem and argues that we need to better understand how prosecutors manage their domestic violence caseloads in order to improve and assess the limits of the criminal justice system response. The chapter summarizes the previous modeling and simulation work on the criminal justice system response to domestic violence, which points to a need to better understand the linkage between arrests and batterer intervention programs. This is followed with a review of the literature covering the three main aspects of the criminal justice system response: police and deterrence effects from arrest, batterer intervention programs and their treatment effects, and prosecutors and their decision making. The chapter concludes with a discussion of how prosecutor decision making can be understood as a dynamic problem of how prosecutors manage their domestic violence caseloads.

2.2 Domestic violence

With 25% of women reported having been raped or physically assaulted by an intimate partner during their lifetime (Tjaden & Thoennes, 1998), woman abuse continues to be a major social problem in the United States. Battering involves a constellation of tactics, including emotional abuse, isolating the victim from resources, preventing her from getting or keeping a job, coercion and threats of violence, rape, and murder. "Battering is far more than a single event, even for the woman who is hit once, because it teaches a profound lesson about who controls a relationship and how that control will be exercised" (Schechter, 1982, p. 17). In order to make the distinction between the single event and pattern within this dissertation, *abuse* will refer to a specific behavior including physical abuse, sexual abuse, emotional abuse, economic abuse, destruction of property, coercion, and threats; *battering* refers to a pattern of abusive behaviors; *domestic violence* includes both abuse and battering; and *men who batter* refers to men who use tactics of abuse and battering against their intimate partners (Pennsylvania Coalition Against Domestic Assault, 1992, Section K. Definitions). Hence, the terms 'male batterers', 'batterers', 'abusers', and 'assailant' will be used interchangeably to refer to adult men (ages 18 and over) who batter women.

Abusers use battering tactics to keep or delay a woman from leaving an abusive relationship, with violence continuing and often escalating after she leaves. In a Michigan study, 37.4% of the women with one or more violent partners said that the violence continued after separation and 45.6% indicated that the violence increased (Largo, Smith, Thrush, McCrohan, & Rafferty, 1999). Adding to the assailant's behaviors are economic, social, and institutional barriers such as no alternative housing, lack of support from friends or family, difficulty finding a job and coordinating childcare, and risk of losing child custody to the assailant.

Before the modern battered women's movement, the State colluded with assailants by failing to extend the criminal law to domestic violence (Fagen, 1996; Mullender, 1996; Schechter, 1982). Police used their discretion to avoid making arrests and prosecutors did not pursue cases where the victim-offender had an intimate relationship. With no legal protection and no shelters until the 1970's, many women who did leave had no option but to return to the abuser. Police, social workers, and other professionals judged a woman's decision to return to the abuser as an indication that she did not really want to leave. It was a "laissez-fair approach that dominated police and court response" (Pence, 1999, p. 14). Feminists pioneered the battered women shelters as a way to increase victim safety and initiated system reform efforts focused on the criminal justice system's response to domestic violence and holding the assailant accountable. Today, many communities have implemented a coordinated response to responding to domestic violence (Mullender, 1996; Pence, 1999). People working to end domestic violence often speak of reform in terms of enhancing victim safety and increasing assailant accountability. "Determining whether the coordinated community response enhances victim safety and increases offender accountability is central to the impact evaluation of a community intervention project" (Shepared, 1999, p. 183).

The long-term goal of this research is increasing assailant accountability. Communities rely primarily on the criminal justice system to hold assailants accountable through arrest, prosecution, probation monitoring, and referral to court mandated batterer intervention programs. "The main challenge appears to be in making the existing components of interventions work together more decisively and consistently. They need to hold men accountable for their behavior" (Gondolf, 2002, p. 218).

2.3 Operational definition of accountability

With respect to the criminal justice system and the prosecutor's office, the most visible aspect of a accountability[1] is the verdict of guilty, either through a trial or guilty plea. Perhaps the simplest operational definition for accountability as a practice for the prosecutor's office is then the proportion of cases that result in a verdict of guilty. While this simple definition is in many ways convenient, it is flawed in the sense that it depends on cases actually flowing through (or at least into) the criminal justice system. For example, a prosecutor's office might have a high proportion of its cases resulting in a guilty verdict, but this would not represent accountability as a practice if police officers were not arresting abusers because the prosecutor's office refused to prosecute most cases. And in contrast, a prosecutor's office might not have prosecuted any cases while still representing accountability as practice if, in some future history, the practice of battering was extinct. To get at the notion of accountability as practice, one must look not at the frequencies of dispositions, but at the structures that explain those frequencies. That is, accountability as practice is the structuring of dispositions. In terms of system dynamics, this can be defined as follows: accountability as practice is the structure of feedback loops that determine the proportion of dispositions resulting in a guilty verdict.

[1]The term 'accountability' has several conflicting uses in popular discourse, particularly when one is talking about patterns of behaviors in the context of oppression, where one social group tends to obfuscates abuses as a tactic for maintaining power and control over another social group. However, a thorough philosophical analysis of accountability is beyond the scope of this dissertation and a topic for future research.

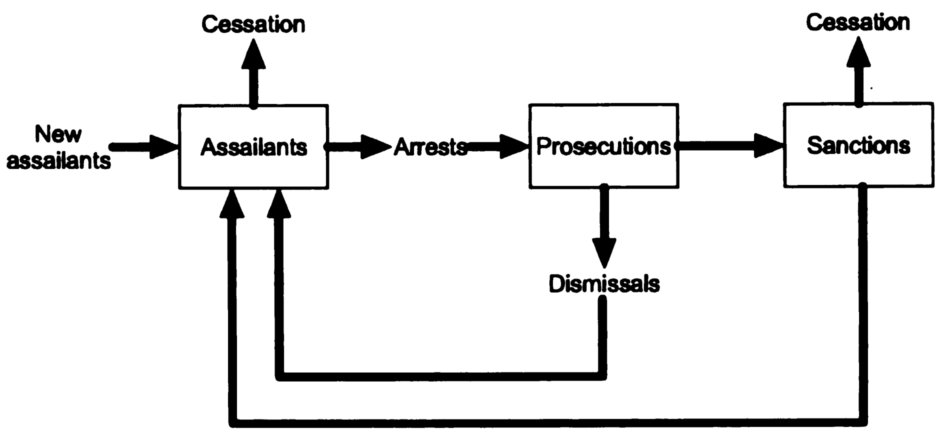
2.4 The criminal justice response

An idealized stock-and-flow picture of the criminal justice response to domestic violence is shown in Figure 2.1. Text within boxes represent stocks of individuals, whereas text with no surrounding box represent flows of individuals between stocks. Stocks represent the number of individuals in a given state. Individuals can, theoretically, remain in a given state indefinitely.

New assailants at risk of arrest come into the community's stock of *Assailants,* which represents individuals committing behaviors that fit the criteria for criminal domestic violence. Note that individuals move into and out of *Assailants* with changes in criminal law, that is, a change in the boundary of a category. For example, the criminalization of domestic violence will increase the number of individuals at risk of an arrest.

Some assailants might be arrested and move via *Arrests* into the prosecutor's office caseload, *Prosecutions,* others might stop committing criminal domestic violence and leave the community's stock of *Assailants* before they are arrested via *Cessation*, and the rest will remain in the stock of *Assailants* (perhaps indefinitely). *Cessation* includes assailants that no longer abuse their intimate partners (a change in behavior or death) and assailants who moved to non-criminal battering tactics (either by changing their own behavior or by changes in the enforcement of criminal law). Some of the cases being prosecuted will be denied or dismissed, and return to the stock of assailants via *Dismissals*, while others will move into the stock of cases with *Sanctions*, either through informal agreements or formal guilty pleas. Of those that are sanctioned, some will discontinue committing criminal domestic violence and leave the system of assailants via *Cessation*, while others will return to *Assailants.*

Figure 2.1: Criminal Justice Response to Domestic Violence



Understanding how this simple model of the criminal justice response functions to hold assailants accountable is essential for both improving the current system and understanding its limits. The criminal justice system is, however, a complex system of interacting subsystems. Developing a scientific understanding of how such systems work is challenging because they entail sets of feedback loops that are inherently difficult to identify and assess using traditional statistical methods that rely on the classic linear cause-effect relationship. Research into complex systems has therefore been mostly theoretical in the sense of social theory or mathematical modeling and computer simulation. Computer models and simulations have been used to study a variety of social phenomena, but only recently has the approach been applied to the problem of domestic violence. The major challenge for these early efforts has been conceptual, that is, how to represent the problem and prevailing theories in terms of quantified variables and their relationships.

Previous modeling work on the criminal justice response to domestic violence (see Figure 2.1) focused on understanding the specific problem of batterer intervention programs not appearing to have any measurable effect at the community level despite increases in court mandated referrals (Hovmand, 2000). One explanation by service providers for the lack of impact was that as referrals increased, more agencies entered the batterer intervention market, which increased competition and lowered standards of accountability. Specifically, programs with large groups could afford to hold participants accountable and risk losing a group member whereas small groups could not. The model assumed that all batterer intervention programs had a positive effect size and that every arrest resulted in some type of intervention. Analysis of the model indicated that the problem was not caused by market competition or deteriorating standards, but policies where the chances of being referred to a batterer intervention program after an arrest were proportional to the severity of criminal behavior. Specifically, the intervention programs were most effective when the least criminally severe assailants had the highest chances of being arrested and referred to the batterer intervention program, for two reasons. First, removing assailants at the lowest levels of criminal behavior reduced the number who would escalate to higher levels of criminal behavior. Second, assailants at the lowest level of criminal behavior were also closest to the desired state of not battering and hence had a better chance of stopping violence, versus just reducing the severity of violence. The key finding was that the effectiveness of the criminal justice system response to assailants was sensitive to how the probability of an arrest and mandated referral varied by the severity of criminal behavior, that is, the link between *Assailants* and *Sanctions* in Figure 2.1, the prosecutor's office.

While the police response and batterer intervention programs have been extensively researched, relatively few studies have looked at the prosecution of domestic violence cases. The next three sections review the research relevant to the criminal justice system response to domestic violence: police and deterrence effects from arrest, batterer intervention programs and their treatment effects, and prosecutors and their decision making. The basic argument is that while there is little evidence of deterrence effects from arrests, batterer intervention programs are showing some tentative indications of being able to reduce recidivism. Hence, the limitation of community responses is currently the link between arrests and batterer intervention programs, that is, the prosecutor's office.

Prosecutors play a major role in the flow of misdemeanor criminal cases, from authorizing the arrest warrants to plea bargaining. Prosecutors decide whether there was enough evidence to establish probable cause for the domestic assault arrest. Contrary to popular perception, most criminal cases resulting in guilty dispositions are settled through plea agreements. This is true for both felonies and misdemeanors. In plea agreements, prosecutors and defense attorneys might negotiate the charge, the sanction, or both. When an agreement has been reached between the prosecutor and defendant, the agreement is presented in front of the judge for final approval. While judges have the final say on sentences, their options are ultimately limited to what the prosecutor presents in court.

2.4.1 Police and deterrence effects from arrest

The police response to domestic violence has several facets, including reporting domestic violence (Coulter, 1999) and decisions to arrest (Rigakos, 1997; Robinson, 2000). In the discourse on reducing domestic violence, the major focus of police response has been whether there is a deterrence effect from arrest on the future behavior of assailants. Interestingly, most of the domestic violence legal reforms have not questioned the deterrence theory behind legal intervention (Fagen, 1996).

Deterrence theory presupposes a rational choice model of behavior where the "pain" of sanction is weighed against the "pleasure" of illegal behaviors and individuals refrain when the pain outweighs the pleasure (Maxwell, 2000). There are several ways in which this rational choice model is wrong. Rational decisions about what to do are often framed by our imagination and expectations in a given situation. Some behaviors "make sense" when one considers a perspective on a situation that has no other viable alternatives. Moreover, we rarely have complete information, but have to try and balance various degrees of uncertainty. Nonetheless, deterrence theory has played a major role in research on domestic violence arrests and their effects.

Sherman and Berk (1984) found support for a specific deterrence effect from arrest with batterers in the widely publicized Minneapolis Experiment. The study provided empirical support for the benefits of arresting assailants in domestic violence cases, which together with several large lawsuits and pressure from advocacy groups, led states to pass pro or mandatory arrest statues (Davis, Smith, & Nickles, 1998). Michigan, for example, passed legislation taking effect on July 1, 1994 (PA 62) that required all police departments to implement pro or mandatory arrest policies. Pro-arrest policies allow police officers to make a warrantless arrest without having actually witnessed the assault. Mandatory arrest policies require police officers to make the arrest in such circumstances.

As a follow-up to the original Minneapolis Experiment, the Spouse Abuse Replication Project (SARP) tried to replicate the original study. The researchers were mixed in their analysis about whether arrest deterred assaults (Zorza & Woods, 1994). Analyzing data from Milwaukee, Sherman et al. (1992) did not find support for an overall deterrence effect, but they did find support for a deterrence effect for white, Hispanic, and employed abusers and an escalation effect for abusers who were Black or unemployed. Sherman et al. (1992) interpreted the result as supporting the claim that there is a deterrence effect with abusers who have a stake in conformity while those who do not have a stake in conformity retaliate against their victims because they have less to lose.

The replication studies contained some serious flaws and the results should be read as inconclusive (Garner, Fagan, & Maxwell, 1995). Zorza and Woods (1994, p. 66) point to some of the issues, including not considering the escalation of domestic violence, officers' statements, misclassifying assailants and their risk, including self-defense cases, problems with censored data, and failing to consider the impact of the courts and prosecutors' responses to domestic violence. While some have called for more studies, others have raised concerns about the continuance of mandatory arrest laws, given the absence of an overall effect, since they might disempower victims or result in less reporting and seeking of medical help (Smith, 2000). There are, however, a number of other reasons for favoring mandatory arrest. Arrest temporarily separates the assailant from the victim, providing a window of opportunity for her to seek shelter or make other arrangements, and sends a message to her family and community that domestic violence is illegal, which might ultimately reduce the intergenerational transmission of domestic violence (Zorza & Woods, 1994). Arrest also provides the principal pathway of moving assailants from the community into the batterer intervention program.

2.4.2 Batterer intervention and treatment effects

While there is no conclusive evidence that mandatory arrest or prosecution have an overall deterrence effect on abusers, there is evidence that supports a cautious optimism that batterer treatment programs do lower recidivism rates for at least some assailants. There have always been some "therapeutic" interventions with men who batter, but the 1980's represented a period of development and growth for batterer intervention programs. Policy changes like pro or mandatory arrest increased the referrals and demand for court mandated batterer intervention services (Sonkin, 1988). By the late 90's, approximately 80% of the batterer intervention programs' referrals were court-mandated (Healey, Smith, & O'Sullivan, 1998).

A group approach for men who batter is now the preferred method over individual, couples, or family treatment (Tolman & Edleson, 1989). In couples and family therapy, the clinician is to remain neutral with respect to each individual and avoid taking sides. Such a stance risks the victim feeling blamed for the abuse and increases the likelihood that victim will be silenced in the sessions by the batterer's threats outside the office (Kaufman, 1992; Tolman & Edleson, 1989). In addition, Tolman points out that many batterers will seek treatment only when the woman is about to leave. After being assured that she is not going to leave, he continues abusing. This deepens the cycle of abuse, leaving the victim feeling betrayed by the therapists and professionals, and making subsequent efforts to leave even more difficult. Despite these observations and concerns by advocates, many couples and family therapists feel they are successful in treating family violence.

From a national survey, Gondolf (1990) identified three different treatment theories or modalities for court-mandated batterer programs: therapeutic, psycho-educational, and didactic-confrontational. Therapeutic approaches focus on treating one or more emotional problems. Psycho-educational approaches stress cognitive restructuring and development of social skills. Didactic-confrontational modalities emphasize consciousness raising and taking responsibility for abuse. Programs differ not only theoretically, but also in their implementation. As a result of the wide variation in programs, many states have now implemented minimum standards for court mandated batterer intervention programs (Bennett & Piet, 1999). It remains to be seen whether having the standards improves the quality of the interventions and if that improvement leads to a decline in recidivism. But without the standards, nearly any program would qualify as a batterer intervention program, and this could place the lives of victims at even greater risk.

Problems with methodology in evaluating batterer intervention programs have often hindered firm conclusions from being made about reductions of battering (Eisikovits & Edleson, 1989; Fagen, 1996; Petrik, Gildersleeve-High, McEllistrem, & Subotnik, 1994). Issues include: high attrition rates, variability in consequences for not attending or dropping out, delays between the original incident and the first session, availability of culturally specific groups, no standard follow-up measures or intervals, few control groups, and low survey response rates (Davis & Taylor, 1999; DeHart, Kennerly, Burke, & Follingstad, 1999). However, more recent research is beginning to suggest that some programs do contribute to decrease in rearrest, at least with some men (Davis & Taylor, 1999; Gondolf, 1995, 1999; Taylor, Davis, & Maxwell, 2001).

2.4.3 Prosecution

The increase in arrests has moved more assailants into the prosecution caseflow, but relatively few studies have focused on the prosecution of domestic violence (Cahn, 1992). This should be alarming for two reasons. First, the prosecution caseflow has consistently been found to have a large attrition rate or "sieve effect" attributed to the exercise of prosecutor discretion (Moyer, 1982). Prosecutors use their discretion to make decisions about the number and kinds of charges, whether to prosecute, what to offer and negotiate in plea bargaining, and what to recommend for a sentence. The second reason is that it is unclear what happens when there is an increase in domestic violence cases referred to the prosecutor's office. Presumably, there is an increase in convictions and referrals, but what about the number of cases denied for prosecution, dismissed, or diverted to something other than batterer intervention programs?

The next four sections review some of the existing research on prosecutor decision-making: charging decisions, decision to prosecute, plea bargaining, and effect of case deposition on future behavior. While discussed separately, these decisions and their effects are not necessarily distinct. For example, plea bargaining can easily include decisions to reduce charges. The decision to prosecute might rely on information about the effectiveness of prosecution on a given type of case. And charging decisions might be made expecting a subsequent reduction of charges during the plea bargaining.

2.4.3.1 Charging decision

The first decision prosecutors normally make is whether or not to charge the defendant and on which charge(s). Prosecutors can decide to not charge a defendant at all, nolle prosequi, or add to, remove, or change the charge at arrest. Charging decisions typically determine the court that the case will be heard in and prosecutors might lower the charges in order to get a more favorable judge or avoid constraints on sentencing or plea agreements.

Schmidt and Steury(1989) found that prosecutors were more likely to issue criminal charges when the victim was injured, a weapon was used, defendant failed to appear at the charging conference or had a prior record. But, prosecutors were less likely to file charges when the victim and defendant continued to have a sexual relationship, were living together, the defendant was employed, and there were no indications of alcohol or drugs at the incident.

There are major problems in terms of both the reliability and validity of domestic violence suspects' prior records. First, indicators of whether or not a suspect has prior domestic violence offenses are notoriously unreliable. Criminal history databases are often incomplete with respect to misdemeanor arrests and convictions, and assailants will often move from one jurisdiction to another to avoid accumulating a record. Second, domestic violence related offenses can include a wide range of criminal behaviors, from stalking to destruction of property and disturbing the peace. Most criminal records do not, at this point, record whether or not a crime is related to domestic violence. Some record keeping systems now make it possible to indicate whether or not any crime is domestic violence, but their implementation is slow.

While Schmidt and Steury's (1989) study has been the only one focusing on charging decisions in domestic violence cases, other researchers have looked at prosecutorial discretion for other types of crime. Albonetti (1992), for example, considered prosecutorial discretion to reduce burglary and robbery charges, finding that seriousness of the case and indicators of the defendant's character (having a prior record, young, habitual offender) lowered the chances of charges being reduced. Albonetti tested and did not find race or gender to be factors in the decision to reduce charges, which might be due to reliance on regression models. Logistic regression models estimate the probability of an outcome from a set of explanatory variables. One major limitation of such approaches is the risk of oversimplifying into cause-effect models complex decisions processes that involve feedback loops.

For example, prosecutors often negotiate the charge with the possible sentences in mind. Hence, the charging and sentencing decision are not independent, but interact across time. Prosecutors might also charge cases as misdemeanors or felonies as a way to get a judge considered more favorable for the prosecution. Similarly, mandatory sentencing guidelines and other constraints on criminal behavior for certain types of crimes may influence prosecutors' charging decisions. For example, McCoy (1984) argued that California's Determinant Sentencing Law, which was intended to reduce judges' sentencing discretion, resulted in prosecutors reducing their initial charges in order to avoid the constraints on plea bargaining in felony cases. And in a more recent analysis, Taha (2001) found that the impact of Sentencing Reform Act of 1984 on federal prosecutors was an increase in more precharge bargaining which resulted in defendants pleading guilty faster and to lesser charges.

There has also been some suggestion that organizational factors between police and prosecutors play a role in charge decisions. For example, police agencies monitor prosecutor's charging decisions as a guide for determining standards for a viable criminal case (Schmidt & Steury, 1989). And McCoy (1984) described other effects in the context of a hydraulic theory of discretion in the criminal justice system, whereby efforts to limit discretion in one part of the system resulted in increases in another part. That is, instead of limiting discretion, discretion was simply displaced from one part of the criminal justice system's response to another. Her point is important in the context of domestic violence because as communities limit police discretion to arrest and restrict judicial discretion, an unintended consequence might be the displacement and concentration of discretion in the prosecutor's office.

2.4.3.2 Decision to prosecute

Once charges have been filed, the prosecutor can decide whether to go forward or delay prosecution, usually for up to a year with misdemeanor cases. Albonetti (1987) considered the effects of uncertainty on prosecutor discretion, looking at variables such as the types of evidence, defendant-victim relationship, defendant arrested at scene, gender, race, defendant's prior record, offense type, use of weapon, type of victim, victim provocation, and statutory severity. Albonetti hypothesized that a prosecutor's uncertainty about winning a case would decrease the likelihood of prosecuting the case. Albonetti found support for and concluded that uncertainty about winning a case at initial stages of prosecution guided the decision to prosecute. Albonetti also noted that warrantless arrests were less likely to be prosecuted. Most misdemeanor domestic violence cases involve warrantless arrests. In warrantless arrest, police officers make the initial decision, and prosecutors review that decision with the result of either authorizing or denying the arrest warrant. In contrast, prosecutors are more likely to prosecute cases when they make the initial screening decision.

Prosecutors do make decisions based on the strength of a case. And, stronger cases are more winnable by definition. But, it is unclear how Albonetti's sense of uncertainty differs from the notion of the strength of a case. If one wants to say that the strength of the case is the same thing as the uncertainty of winning a case, then Albonetti's findings are much less interesting. One might be tempted to simply say that uncertainty of winning and the strength of the case are the same thing. But there are differences between the two concepts, especially in cases involving violence between intimate partners.

A prosecutor can have a strong case in terms of evidence, victim credibility, use of a weapon, and so on and still be uncertain as to whether or not a jury would find the batterer guilty if the case went to trial. The prosecutor and defense attorney's assessments of what would happen if a case went to trial, and their certainty in that assessment, is likely to have a significant influence over the plea bargaining process and the decision to prosecute. Thus, the uncertainty lies not in the strength of the case, but in assessing the "going rate" for a case with a given set of attributes. The going rate is a state variable and subject to both mis-perception and adjustment as the outcomes of similar cases change over time. For example, if the prosecutor offers pleas below the going rate, the going rate will eventually drop. Likewise, if the prosecutor manages to secure convictions above the going rate, the going rate will gradually increase. But what the real going rate is at any point in time is uncertain. If there is an argument to be made about the relationship between uncertainty and prosecutor decision making, then it is in the prosecutor's estimation of the going rate for a particular case.

Martin (1994) looked specifically at the decision to prosecute domestic violence cases and found that prosecution was more likely when defendants committed more serious offenses, had prior arrests for domestic violence, if the charge was criminal trespassing or harassment, or either victim or defendant was using alcohol or other drugs during the incident. Martin did not find that victim injury predicted prosecution. In contrast to Martin (1994), Hirschel and Hutchinson (2001) could only find two explanatory variables that were significant in their model of the decision to prosecute: victim injury and the victim arguing with police against a citation/arrest. And, Dawson and Dinovitzer's (2001) model of the decision to prosecute contained only one statistically significant explanatory variable, namely, a victim's willingness to cooperate.

Albonetti and Hepburn (1996) looked at the decision to defer prosecution in felony drug possession cases and found that older males with prior arrest records or multiple charges were less likely to have prosecution deferred in lieu of drug treatment, while defendants arrested for marijuana possession were more likely. Albonetti and Hepburn did not find support for a direct relationship between the defendant's minority status and the prosecutor's diversion decision, but they did find support for a more complicated interaction between prior record, age, and minority status. Specifically, Albonetti and Hepburn (1996) argued that (a) defendants with prior contact with the criminal justice system were more likely to know how to negotiate a deferred prosecution, and (b) defendants with minority status were more likely to have increased contact with the criminal justice system, which would be the direct consequence of such practices as racial profiling.

This last argument is significant because it points to an unintended consequence of a suspect's contact with the criminal justice system that has an impact on future dispositions. Specifically, one usually thinks of criminal history as something that accumulates. As suspects accumulate prior convictions, subsequent sentences are enhanced. But Albonetti and Hepburn's (1996) argument points to a second, compensating mechanism, whereby increased contact with the criminal justice system contributes to the accumulation of experience manipulating the system, and might as Albonetti and Hepburn point out, actually lead to lessening of sanctions relative to the arrest charge. The possibility of this second mechanism has major implications in the context of domestic violence and arrest policies. If there is such a mechanism, then it means that an unintended consequence of pro and mandatory arrest will be batterers who become more sophisticated in manipulating the criminal justice system and controlling victims. This is similar to the concern that some have with batterer intervention programs, namely, that these groups effectively provide batterers with an opportunity to share battering tactics and become more effective at manipulating victims and services providers.

2.4.3.3 Plea-bargaining decisions

Plea bargaining is a loose term for negotiations between the prosecutor and defense attorney, and can include everything from how the defendant will be charged to sentencing, pretrial release, and victim restitution (Dawson, Smith, & DeFrances, 1993). For the prosecutor, the main advantage of plea agreements is having the guilty verdict without having to invest the resources needed to prepare a case for trial. For the defense attorney, plea agreements can obviously result in more favorable sanctions, but not always. Hollander-Blumoff (1997) argues, for example, that the private defense attorney is likely to pursue the interests of their client, but a public defender or court appointed attorney might have other interests in mind since they are repeat players in their negotiations with the prosecutor. Consequently, public defenders might encourage their client to accept a guilty plea at least in part as a way to build credit with the prosecutor for another case.

Plea agreements are controversial from a victim's rights perspective because advocates and the public tend to see the results as the criminal justice system being soft on crime (McCoy, 1993). Independent of how one feels about plea agreements they have become an essential part of the business of disposing cases in a timely manner with limited resources. Even small reductions in the number of plea agreements can result in large increases in the number of cases going to trial. This has led to the popular hypothesis that prosecutors seek guilty pleas in an effort to reduce the caseload pressure, i.e. their decisions are governed by the economics of a disposition assembly line (for a discussion of this, see McCoy, 1993). The caseload pressure hypothesis is unappealing in the sense that justice might vary for reasons outside the principles of justice. In contrast, the professional norms hypothesis states that decisions to seek a guilty plea are better understood as a function of shared professional norms about fairness, legal standards, and punishment (McCoy, 1993). McCoy (1993) argues that prosecutors and defense attorneys share common views about the "going rate" for a "normal crime" and plea bargaining becomes simply an efficient way of handling cases with predictable outcomes. Cases are expedited without compromising the principles of justice.

In one of the earliest studies, Heuman (1975) considered the impact of caseload pressure on plea bargaining by studying trial dispositions from 1880 through 1954 in Connecticut Superior Courts. Heuman was responding to criticisms of plea bargaining that argued that its use was a recent phenomenon brought on by expanding caseloads. Heuman found that trial verdicts represented between 10 and 20% of all dispositions from 1880 through 1954, and hence plea bargains had always been a part of the criminal justice process and remained relatively stable.[2] Heuman (1975) then compared dispositions and caseloads between low and high volume superior courts and found no evidence for a direct relationship between caseload pressure and plea bargaining, but qualified the results, saying that caseload pressure might have indirect effects, namely:

> The prosecutor may nolle[3] the marginal case which he might have pursued for a plea earlier. He may offer to reduce more charges and recommend lighter sentences, or he may simply demand more severe sentences after trial. (p. 527)

Rhodes (1976) considered the impact of caseload pressure in terms of delays and applied William Landes's model of plea bargaining as a market transaction. In this model,

> the plea bargaining process, which is of central importance to modern jurisprudence, can be characterized as a market transaction in which the prosecutor "buys" guilty pleas in exchange for promises of sentence leniency (Rhodes, 1976, p. 311).

Rhodes (1976) assumed that prosecutors would be interested in maximizing the total mean number of years in punishment and derived an expression that optimized

[2]Heuman's conclusion about the existence and stability of plea bargaining in the criminal justice system is popular, but not without disagreement. For example, Alschuler (1979) argues that plea bargaining was infrequent and discouraged, and only started during the 1920's with the expansion of prosecutor offices and bureaucratization, reaching its present form during the due process revolution of the 1950's and 1960's.

[3]For a prosecutor's office to nolle or *nolle prosequi* a case is to abandon the lawsuit.

years in punishment with prosecutor resources. His results suggested that (a) the rate that prosecutors use plea bargaining could be understood as a combination of legal strengths of the case and court delay, and (b) plea bargaining was essentially a market clearing mechanism.

McDonald (1985) took a more comprehensive view with interviews and hypothetical cases, and found caseloads to be general determinants of the need to plea bargain, but did not actually predict which cases were plead or the terms of the plea agreement. McDonald did find that with greater caseloads, less attention was given to less serious crimes, increasing the likelihood that plea-bargains for less serious cases would become more generous. In ranking prosecutors' information cues to a range of hypothetical cases, McDonald observed that the participants did not pay much attention to court caseload in their plea bargaining decision. Instead, they focused on the strength of the case, seriousness of the offender, and seriousness of the offense.

These studies relating caseload to prosecutor decision making (Heuman, 1975; McDonald, 1985; Rhodes, 1976) all suffer from two problems. First, they focused on felony cases. Most criminal domestic violence cases are misdemeanors, and misdemeanors are generally considered less serious offenses. The second major problem with these studies is that they focused on court caseloads, not prosecutor caseloads. Court caseloads and prosecutor caseloads are arguably related, but prosecutor caseloads are more likely to have a direct impact on the prosecutors day-to-day decision making. This impact could either be direct if prosecutors decided to deny warrants or dismiss cases based on their current caseload. Or there could be indirect effects as higher caseloads might mean less time to investigate and prepare cases, which might lead to more dismissals.

McAllister and Bregman (1986) also studied plea bargaining and found that as the probability of conviction and length of sentence increased, prosecutors were less willing to offer plea agreements while defense attorneys were more willing to accept plea agreements. McAllister and Bregman concluded that there was no evidence of a bias by prosecutors and defense attorneys toward plea bargains. However, there were two major problems in McAllister and Bregman's experimental design. First, they essentially assumed that prosecutors and defense attorneys would be trying to optimize some utility function, specifically, the difference between (a) current plea offer, and (b) the product of the likelihood of conviction and sentence if convicted at trial. If the chances of a conviction were 50% and the average sentence 2 years, defense attorneys would be looking for a plea offer of 1 year or less while prosecutors would be looking for a plea offer of 1 year or more. But, this is essentially a rational model of decision making with full information. One might object to this criticism by pointing out that there are still guesses involved. For example, that participants were only provided with the probability of conviction or the average sentence. But, *knowing* the probability of conviction or average sentence is to know a real parameter value, not an estimate of it. In reality, we can only statistically estimate parameter values. The real decisions that prosecutors face is one with uncertain information, which is what Albonetti (1987) was trying to get at.

The second major problem in McAllister and Bregman's (1986) study is their use of hypothetical situations to draw inferences about real-world decision making. Such studies have been successful at showing that decision making frequently does not conform to the classical model of rational choice. But, these types of experiments simply need to demonstrate that people do not follow the rational choice model in some situations. If, for example, people generally followed the rational choice model in making decisions, then one would expect people to use that model in both real situations and hypothetical situations. And a corollary, if people do not follow the rational choice model in hypothetical situations, then it is unlikely that they follow the rational choice model in real situations. There has been much experimental research showing that people do not follow the rational choice model in hypothetical situations. But having shown that people do not follow a particular decision model in a hypothetical situation is not the same as showing what they actually do in real situations. Specifically, generalizing claims about prosecutor and attorney decision making from hypothetical situations into the real world introduces a major external validity problem, especially if one considers the role of heuristics in the theory of bounded rationality. Ebbesen and Konečni (1980) conducted an experiment illustrating this external validity problem. They compared the results of judges' sentencing decisions in hypothetical cases and real court cases. While they were able to identify a model of the information cues used in judges' decisions in both situations, they found that the models differed. That is, the associations between information cues and sentencing decisions from hypothetical decisions did not correspond to what they actually observed the same judges to be doing in the courtroom with real cases. Ebbesen and Konečni's (1980) point is that one would not know this unless one observed and tested decision models in real situations.

2.4.3.4 Effects of case disposition decision

One reason for even wanting to understand prosecutor decision making is the perceived impact of these decisions on defendants' future behavior. Two studies have considered the effect of prosecuting domestic violence cases on defendants' rearrest rates. Ford and Regoli (1992) evaluated no-drop versus drop permitted prosecution policies. In no-drop policies, victims are not allowed to drop the charges and the prosecutor pursues the case seeking a finding of guilty, often using other forms of evidence such as photos, medical reports, police reports, 911 transcripts, and other witnesses. Ford and Regoli did not find an overall difference between no-drop versus drop permitted prosecution in terms of subsequent reports of abuse. However, they did find that women in the drop-permitted group who elected to proceed with prosecution were less likely to report subsequent abuse than women assigned no-drop prosecution, whereas women in the drop-permitted group who dropped charges were the most likely to report subsequent violence. Ford and Regoli called for more studies to understand these results, but speculated that one explanation might be that victims were empowered by a combination of their ability to drop charges and alliance with the criminal justice agencies. Davis et al. (1998) also studied the effect of court disposition (dismissal, probation with treatment, and jail) on six-month rearrest rates. Davis et al. did not find evidence of a deterrence effect from prosecutor outcomes.

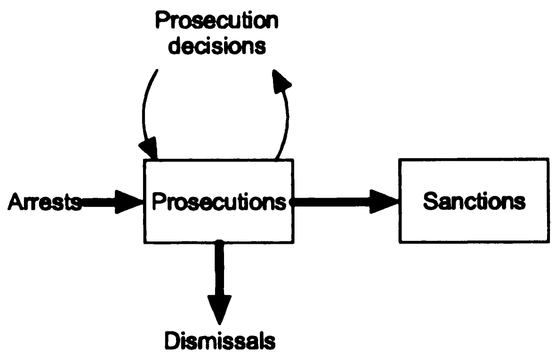
2.5 Discussion

The current research on prosecutor decision-making does not address the general issue of how prosecutor discretion with domestic violence cases is affected by the prosecutor's caseload size. The published studies each have one or more of the following shortcomings. First, many studies on prosecutor's decision making simply did not consider the prosecutor's caseload size as a factor. Second, the studies that did consider the relationship between caseload size and decision making focused on felony cases and studied the court's caseload, not the prosecutor's. The vast majority of criminal domestic violence cases are charged as misdemeanors, not felonies. A large portion of these cases are simply not prosecuted and handled informally outside the court's docket. So it is likely that the caseloads of prosecutors and courts differ in size and that court caseloads are not a good proxy for prosecutor caseloads when it comes to studying decision behavior. Third, while a number of studies suggested that there might be a more complicated and indirect relationship between prosecutor caseload size and decision making, most simply treated each decision as statistically independent of previous decisions. However, both the caseload pressure hypothesis and professional norms hypothesis suggest feedback mechanisms that involve information about previous decisions. In the caseload pressure hypothesis, for example, previous decisions to file criminal charges would increase the caseload and in turn discourage charges being filed in new cases. And in the professional norms hypothesis, defense attorneys and prosecutors would be using previous case dispositions to determine the likely outcome or going rate for a current case. In both hypotheses, the determinants of the current decision facing a prosecutor would include state variables indicating outcomes of previous decisions. Both caseloads and the going rate are likely to vary over time, making the determinants of prosecutor decision making dynamic and not static.

2.5.1 Caseload pressure and caseload management

The question about whether caseload size affects prosecutor decision making is essentially equivalent to asking whether prosecutors have ways of managing the size of their caseload. Prosecutor decisions affect their caseload. Arrests contribute to the caseload while dispositions remove cases from the caseload. The net effect of referrals and dispositions depends on how fast prosecutors dispose of cases relative to the arrest rate. Denying prosecutions will have an immediate effect on the caseload, whereas taking a case to trial will delay the disposition. If the prosecutor's caseload size affects prosecutor disposition decisions, then caseload size and decision making will exist within a feedback loop, specifically, a caseload adjustment mechanism. Conversely, if prosecutors try to manage their caseload, they will have to monitor the size of their caseload and use that information in the decisions that affect the caseload. Thus the prosecutions depicted in Figure 2.1 on page 12 should be extended to include the relationships between the prosecutor's caseloads and decisions that affect case dispositions (see Figure 2.2)

Figure 2.2: Prosecution of Domestic Violence Cases



The major consequence of this extension is that considering caseload as a potential determinant of prosecutor decision making implies the existence of one or more feedback loops. Feedback loops are generally difficult to identify and understand using linear methods like regression. The relationship between *Prosecutions* and *Prosecution decisions* in Figure 2.2 is reciprocal in that one variable affects the other and visa versa. Linear methods estimate parameters by fitting some observed data to a hypothesized model, which is generally assumed to be open loop in nature in order to satisfy the mathematics for finding a solution. Reciprocal relationships violate this assumption and result in misleading estimates. Methods like regression and structural equation modeling do have ways of representing time variant, nonlinear variables, usually through the creation of dummy variables, but they simply represent descriptions of the time variant behavior, not the underlying mechanisms generating that behavior (Hedström & Swedberg, 1998). By analogy, it would be like modeling a plant's growth over time curve instead of the things that affect its growth. While potentially helpful, one is not using a quantitative model to specify and test hypotheses about growth, but instead using a model to simply describe the observations. What is needed is an account of how the underlying feedback mechanisms generate the observed behavior. The problem of understanding the link between *Arrests* and *Sanctions* in Figure 2.1 on page 12 posed at the beginning of this chapter can thus be re-conceptualized in terms of identifying the underlying mechanisms regulating and adjusting *Prosecutions*.

2.5.2 Structure of prosecutor decision making

Until now, most of the discussion on prosecutors and their caseloads has been at the individual level. Individuals do display patterns in decision making, but what is often of more interest is the structuring of that decision making. That is, there is clearly some variation in how individual prosecutors make decisions. For example, some prosecutors refuse to prosecute cases unless the victim cooperates while others take a much more aggressive stance against domestic violence and proceed with evidence based prosecutions that do not rely on the victim's cooperation in the process. But, understanding the structure that constrains all prosecutors in their decision making, whether they take domestic violence seriously or not, is just as important because it ultimately represents the main limiting condition on developing an effective criminal justice system response for holding assailants accountable. The structure of individual prosecutor decision making is arguably at the level the prosecutor's office, where one or more assistant prosecuting attorneys are involved in criminal domestic violence cases. So the question becomes, how does a prosecutor's office manage its domestic violence caseloads?

Thinking of the prosecutor's office response to domestic violence in terms of a caseload management problem has several advantages. First, it has immediate relevance to prosecutor's office in terms of day-to-day operations, which both motivates individual participation in a study and is likely lead to real-world application. Second, understanding the problem in terms of caseload management helps us understand not only how to improve the criminal justice response, but also understand its limits, i.e. what are the theoretical potentials of various caseflow management policies? It might turn out, for example, that there is only a marginal potential for improving the criminal justice response to domestic violence in terms of holding assailants accountable. This would suggest (1) that communities should find alternative ways of holding the majority of assailants accountable, and (2) resources for improving the criminal justice system should be focused more on improving victim safety as opposed to increasing assailant accountability.

Finally, viewing the prosecutor's role in terms of caseload management helps relate the issue to a class of more general problems in social science. For example, caseload management problems are really a subset of resource allocation problems. In the context of service delivery systems, a resource allocation problem concerns how one allocates resources to provide services. When resources are fixed, the resource allocation problem reduces to a caseload management problem because caseload is the only variable that can be manipulated to balance resources with services. But this poses a question: What if prosecutor's office resources are increased, for example, through a grant to fund a domestic violence prosecution unit? Will the increase improve the response? The answer depends on the nature of the resource allocation problem. If services are proportional to resources, then the answer is probably yes. But if the resource allocation problem is dynamic, more resources might actually create new demands that cancel any potential benefits.

Highway construction in a congested metropolitan area is a classic example of a dynamic resource allocation problem. Roads are built to alleviate traffic congestion. But, reduced traffic congestion in a busy metropolitan area has the unintended consequence of increasing construction of commercial and housing real-estate, which in turn increases traffic congestion to the original or higher levels. Understanding the nature of the resource allocation problem (and hence the caseload management problem) helps identify and explain the unintended consequences of various policies that a prosecutor's office might have with respect to prosecuting domestic violence cases.

2.6 Conclusion

This chapter reviewed the social science research on the criminal justice systems, covering deterrence effects from arresting abusers, batterer intervention programs, and prosecutions. It was then argued that changing practices of holding abusers accountable involved, at least as presently implemented in the criminal justice system response, a better understanding of the flow of cases from the arrest of abusers under pro or mandatory arrest laws to the prosecution and referral of abusers to batterer intervention programs. This led to the main question for this study: how do prosecutors dynamically manage their misdemeanor domestic violence caseloads?

Chapter 3

Methods

3.1 Overview

The previous chapter argued that to better understand the link between arrests and referrals to batterer intervention programs, we need a better understanding of how prosecutors manage their domestic violence caseloads. Previous research on prosecutor decision making relied on methods that are best suited for studying linear cause-effect relationships. However, the effects of caseloads on case dispositions and referrals arguably entails one or more adjustment mechanisms, that is feedback loops. This chapter describes the main methods that were used to generate and analyze the numeric time series and study the feedback loops that structure the complex relationships between how a prosecutor's office manages its domestic violence caseload and case dispositions.

3.2 Single case study design

Though most prosecutor's offices operate under the same legal system within a state, the offices are largely autonomous with respect to state oversight and the demands of their jurisdictions vary significantly from small rural counties to large metropolitan areas. Thus it is likely that prosecutor's offices vary in how they balance the demands of their communities with their available resources and priorities. Mechanisms that prosecutors use to dynamically manage their caseloads are therefore likely to vary significantly from office to office. The only way that one might assess the extent of variation is to consider each office as an independent case, from which one can later abstract commonalties and try to identify generic structures. Prior to making any comparisons, however, one must have a reliable and generalizable method for identifying and assessing hypotheses about how a prosecutor's office manages its domestic violence caseload. Therefore, this study focused on developing a robust method for identifying and assessing hypotheses on caseload dynamics using a single case study research design (Yin, 1994) of a single prosecutor's office.

3.3 Studying dynamic systems

There are any number of ways that one might go about studying and analyzing data on the dynamics of a prosecutor's office domestic violence caseload, from qualitative studies to statistical analysis of numerical time series. There are, however, major epistemological challenges to consider when studying the dynamics of complex social systems.

The first epistemological challenge is the misperception of feedback in complex dynamic systems. People often invoke event-based, open-loop views of causality that ignore feedback mechanisms and fail to appreciate the significance of delays, nonlinearities, and the distinction between stocks and flows (Sterman, 2000, p. 27). The tendency is to think of causation in terms of physical action involving human agency that stems from our prototypical metaphor of causation:

> At the heart of causation is its most fundamental case: the manipulation of objects by force, the volitional use of bodily force to change something physically by direct contact in one's immediate environment. It is conscious volitional human agency acting via direct physical force that is at the center of our concept of causation. (Lakoff & Johnson, 1999, p. 177)

The further one moves away from the prototypical of manipulating objects, the more figurative the application of causation becomes. The effect of this tendency to see physical action as prototypical causation is to grossly misperceive the relationship between structures of feedback loops and dynamic behavior, even when people have full and complete information about the underlying structure of feedback loops (Moxnes, 2000). That is, one might have a good idea of what the structure is, but not why or how it generates the dynamic behavior of interest.

One might try and meet this challenge by more formally specifying the relationships between variables in terms of quantitative relationships. But such an approach comes with its own problems, which raises the second major epistemological challenge when studying complex social systems: quantitative models of complex systems are inherently under-determined, meaning that they rarely have a solution to a given set of equations and data. Or put differently, there are too many variables for too few independent equations.

> In the quantitative realm, engineers and econometricians have long struggled with the problem of uniquely identifying the structure and parameters of a system from its observed behavior. Elegant and sophisticated

> theory exists to delimit the conditions in which one can identify a system from its behavior alone. In practice the data are too scarce and the plausible alternative specifications are too numerous for statistical methods to discriminate among competing theories. The same data often support wildly divergent models equally well, and the conclusions based on such models are not robust. (Sterman, 2000, p. 26)

This emphasis on uniquely identifying the structure and parameters dominates statistical thinking[1]. It is arguably driven by what Lakoff (1987) calls the traditional view of objectivism, part of which is the idea that words, numbers, mental representations and so on, get their meaning from a correspondence with the external world. Wittgenstein discusses this view as the picture theory of meaning (1921/1974). But, this emphasis on correspondence is impossible to satisfy when it comes to complex dynamic systems because such systems are generally under-determined.

3.4 System dynamics

System dynamics (Forrester, 1961/1999, 1971; Richardson, 1991; Richardson & Pugh, 1986) takes an approach to studying complex system that differs from the emphasis on identifying the structure and parameters of a system from behavior alone. Part of the system dynamics paradigm is that numerical data represent only a small fraction of our knowledge of complex systems (Forrester, 1980). People experienced with some aspect of a system can often have accurate mental models of specific parts of a system, usually in terms of simple linear cause-effect relationships. When their understanding of that part of the system is realistic, it is because they are close to that part of the system. System dynamics uses that "local" informa-

[1]For example, see Suppes and Zinnes' (1963) work on measurement theory.

tion to simplify and develop the structure of feedback loops that might account for dynamic behavior along with available numerical data, and then tests that hypothesis that a given structure can account for the observed dynamic behavior using computer simulation. Knowledge about complex systems is not acquired through trying to identify the structure or estimate parameters, but understanding which structures are capable of explaining the dynamic behavior and why (in terms of explicitly formulated feedback mechanisms).

System dynamics modeling has been successfully used to study complex systems in case study designs. The term 'system dynamics' often refers to a variety of theories and approaches to research on systems, but in this study, it means the method pioneered by Jay W. Forrester in the 1960's as industrial dynamics (Forrester, 1961/1999, 1971; Richardson, 1991; Richardson & Pugh, 1986). System dynamics is a method "for understanding certain kinds of complex problems," specifically dynamic problems where quantities change over time and involve the notion of feedback (Richardson & Pugh, 1986, p. 1). Within system dynamics itself, approaches vary in their emphasis on quantification of variables and reliance on computer simulation for model analysis (Lane, 2001). The approach used in this study emphasizes the role of computer simulation in the building and testing of a system dynamics model.

The main purpose in system dynamics is to understand the relationships between the structure of feedback loops and behavior over time. That is, *how* does a particular structure of feedback loops reproduce a specific pattern of behavior–only being able to reproduce the behavior is not enough. In system dynamics, the relationship between the structure of feedback loops and behavior can be formally describe in terms of feedback loop dominance (Ford, 1999). The general notion is that the behavior of a variable is controlled by one or more feedback loops. The feedback loops that dominate a given variable at a given point in time are called the dominant feedback loops. Complex behavior patterns emerge as the dominance of feedback loops with respect to a given variable change over time. That is, feedback dominance over a variable often shifts from one feedback loop to another. These shifts in feedback loop dominance are largely insensitive to variations in parameters (Richardson, 1991). The dominance of a feedback loop can also be hidden by other feedback loops. That is, the control that a feedback loop might exert on variable might be masked by another feedback loop, called a shadow feedback loop. And to complicate matters some more, shadow feedback loops can be nested. Intuitively, the effect is that one might not notice the effects of a dominant feedback loop upon a variable until one actively identifies and disables the shadow feedback loops. Thus, system dynamics is usually used to understand the relationship between structure and behavior in terms of feedback loop dominance as opposed to the more conventional approach in statistics of parameter estimation.

3.5 Case selection

The prosecutor's office selected for this study had several attractive features. First and foremost, the office was participating in a larger multi year evaluation study of coordinated community responses to domestic violence. Researchers at Michigan State University had already established strong collaborative relationships with the coordinating council, local shelter, prosecutor's office, and law enforcement. Strong collaborative relationships with stakeholders would be critical to gaining access to key informants and numerical data. Second, researchers were already collecting, cleaning, and analyzing substantial amounts of secondary data from various agencies, including shelters, law enforcement, and the prosecutor's office. That meant potentially having several measures of the same variable from which one could assess the reliability. Having access to multiple data sets also meant having more information to use for calibration and testing of models. This would lessen the amount of speculation required for certain trends and model parameters.

A third feature of this prosecutor's office was that it served a rural county with approximately 60,000 people. That might at first seem an undesirable trait in a study, but it provided an opportunity to solve the main barrier to conducting future multiple case study designs using system dynamics and involving small populations. In the context of time series analysis, the small sample sizes of rural counties and other small population phenomena (e.g. rare diseases) imposes a major limitation on the number of data points available for statistical analysis. Specifically, in small populations events happen much less frequently than with larger populations. So one generally has to aggregate over larger periods of time than larger populations in order to attain adequate sample sizes for statistical analysis. One might, for example, be forced to aggregate the arrests over years instead of months or weeks. Aggregating over larger periods of time reduces the number of data points one has across time. This is not a problem if the phenomenon under study is stable over time, but if one is interested in the dynamics over time, then aggregating over larger time intervals smooths out changes. And if the time intervals are too long relative to how fast a quantity is changing in time, one might not see indications of the phenomenon at all. One might try to avoid this problem by aggregating data from several small communities. But such an approach would essentially presuppose that their aggregated representation reflected a common underlying structure. There are many ways that small populations can vary, from economic conditions to culture and geography. This problem of small populations is major methodological challenge facing any future multiple case study design that tries to look at structural differences of responding to domestic violence between various communities and populations. Solving this problem would make it much easier to demonstrate the feasibility of such a multiple case study design and work with more diverse communities, including both rural and metropolitan communities, as well as smaller subpopulations.

3.6 Violence against women database

The prosecutor's office violence against women (VAW) database was the main source of case data for demographic comparisons and construction of numerical time series. The prosecutor's office maintained the VAW database in Microsoft Access, which was imported into R for data manipulation and statistical analysis. Text variables were coded and dates were converted to POSIXlt representations. Some variables, like the reason for denial or dismissal, contained text descriptions along with dates of when cases were denied or dismissed. Such text descriptions were parsed for date formats, and converted into POSIXlt representations.[2] POSIXlt dates had to be corrected for Y2K. Ages of victims and suspects were calculated as the difference between the date of offense and date of birth.

The VAW database (n=6358) contained a subset of the prosecutor's office cases, including records of personal protection orders (n=1029), 911 calls involving violence against women for the county (n=2821), warrant requests (n=1993), and 48-hour reports by police departments (n=474) covering offenses from approximately January 1, 1994 through June 10, 2002. However frequency distributions of types of cases by year suggested a pattern of inconsistent data entry before January 1, 1998.

[2]POSIXlt is a specific way of representing dates and times as a long named list of vectors representing seconds, minutes, hours, days of month, months of the year, years, day of week, day of year, and daylight savings flag.

The primary focus on this research was on the prosecution of assailants, all of which started as warrant requests. Thus the initial subset (n=1457) included warrant requests with valid dates of offense on or between January 1, 1998 and December 31, 2001.

Table 3.1: Distribution of Records by Year and Type of File in Violence Against Women (VAW) Database

Year	48HR	911	PPO[1]	WARRANT
< 1995	0	0	0	2
1995	0	444	0	168
1996	0	535	0	46
1997	5	257	2	33
1998	113	0	152	**229**
1999	85	994	193	**235**
2000	123	588	162	**240**
2001	95	0	121	**200**
2002[2]	49	0	63	82

[1]Based on date that PPO was authorized.
[2]Limited to records through June 10, 2002.

The VAW database contained first names, last names, and middle initials along with sex, date of birth, and address for both suspects and victims. A unique list of individuals was generated by first generating a unique sorted list of concatenated string of the name and date of birth for both suspects and victims with each row being assigned a unique personal identification (PID) number. Since the list was sorted by string names and dates of birth, it was relatively easy to identify multiple rows that appeared to refer to the same unique individual. In such cases, string representations of the individual were assigned the first PID number of the matching string representation. Victim and suspect names were then replaced with corresponding PID numbers (n=2050 for 1457 records of warrant requests). Two records were eliminated because names were missing (leaving a total of 1455 records of warrants). Another 17 records were dropped from the analysis because their dates were obviously invalid, for example, a record having a dismissal date that came after the date that a plea offer was accepted. The main variables of the resulting individual data matrix, $IDMAT$ (n=1438), is shown in Table 3.2. Basic descriptive statistics were run on $IDMAT$ for the all warrants and selected subsets.

3.7 Estimating rates and times for state transitions

Each record in $IDMAT$ represented one warrant request. An individual who was arrested for a probation violation would appear again as another warrant request. Using the PID numbers, it was possible to estimate the number of times individuals appeared in the database as well as the number of days between their appearances. Such statistics could be used to estimate recidivism rates, re-victimization rates, commutative hazard functions, and mean times to failure. For example, some individuals were recorded as victims in some records and suspects in others. Rearrest rates might also depend on the type of sanction. At first, each set of transition statistics was calculated manually. Each set of transition statistics usually led to more questions about alternative explanations for the observed frequencies and transition times, which in turn led to writing another calculation. It quickly became apparent that transition patterns of individuals through the prosecutor's office could be fairly complex in their own right. To facilitate a exploration of the possible patterns, a more general approach was required.

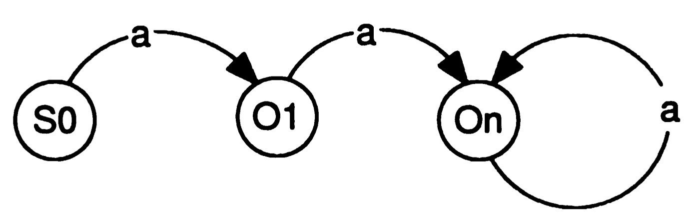
Table 3.2: Variables in the Individual Data Matrix ($IDMAT$)

Variable	Description
File.Type	Type of file
DateOFF	Date of offense
Charge	Arrest charge
PolDept	Police department
Prosecutor	Prosecutor
SPID	Suspect's PID number
SDOB	Suspect's date of birth
SAGE	Suspect's age
SRace	Suspect's race/ethnicity
SSex	Suspect's sex
RelType	Victim-suspect relationship
VPID	Victim's PID number
VDOB	Victim's date of birth
VAGE	Victim's age
VRace	Victim's race/ethnicity
VSex	Victim's sex
A.BType	Type of assault
DatePOA	Date plea offer accepted
TrialTyp	Type of trial
Judge	Judge
Dispo	Disposition
Dismissal.Denial.Reason	Reason for denial/dismissal
Probmnth	Months of probation
DOSmnth	Months of delay of sentence
CGroup	Counseling group referral
DateTRL	Trial date
DateDenied	Date denied
DateDismissed	Date dismissed
DateAUTH	Date warrant authorized
Date.VMeet	Date met with victim
Jaildys	Jail days

3.7.1 Modeling state transitions

The solution was to model individuals' state transitions with a finite state machine (FSM), which is essentially an abstract way to describe a system with a finite number of states and finite number of transitions from one state to another state. Figure 3.1 shows a diagram of a simple FSM with three states (S_0, O_1, and O_n) and one condition (a) that triggers transitions from the initial state S_0 to O_1, and from $O_1$to O_n, and then from O_n to O_n.

Figure 3.1: Finite State Machine



Finite state machines, automata, and Turing machines are typically used in computation theory and such applications as computer language compiler design (Lewis & Papdimitriou, 1981). Finite state machines have also been developed and recently applied in various engineering applications for modeling and diagnosing system failures (Özveren & Willsky, 1990; Sampath, Sengupta, Lafortune, Sinnamohideen, & Teneketzis, 1995; Sampath, Raja, Lafortune, Sinnamohideen, & Teneketzis, 1996; Rozé & Cordier, 2002). Table 3.3 shows a simple description of a finite state machine corresponding to Figure 3.1 that was used for calculating the state transitions associated with assailants' rearrest. Three states are specified: an initial state S_0 that all individuals start out in by default, the state O_1that individuals end up in after their first appearance as a suspect, and the state $O_n$which move into after subsequent offenses. Transitions from one state to another state are triggered by a condition being satisfied. In Table 3.3, individuals move from one state to another state when the individual's PID number is equal to the suspect's PID number for a given record.

3.7.2 S-language implementation of finite state machines

Functions were then written in R, a variant of the S programming language (Becker, Chambers, & Wilks, 1988; Venables & Ripley, 2000), which took a description such as that in Table 3.3 and applied it to a subset of *IDMAT* to generate a table of transitions along with their transition times. The output could be used to calculate statistics on transition times and construct transition matrices. FSMs were used for calculations of first offense-repeat offense transitions, which could be further analyzed by the type of sanction imposed and type of referral to a counseling group. The FSM depicted in Figure 3.1 and Table 3.3 was used to identify cases that were first offenses and repeat offenses *IDMAT*. FSMs were also used to calculate the transitions from one type of victim-offender relationship to another type of relationship, transitions from one type of offense to another type, and test for the reliability of demographic variables such as sex and race/ethnicity.

3.8 Battering tactics

The *A.BType* variable in *IDMAT* contained a text description of the types of battering tactics used during the assault. Initial inspection of frequency distributions suggested that there might be differences in the types of tactics reported. Many descriptions included multiple tactics, typically separated by a forward slash ("/"). These tactics were parsed into a sorted list of tactics, with each row corresponding

Table 3.3: Finite State Machine Example

Element	Example		
States	S_0, O_1, O_n		
Conditions	Condition	Expression	
	a	PID==SPID	
Transition table	S_i	Condition	S_{i+1}
	S_0	a	O_1
	O_1	a	O_n
	O_n	a	O_n

to one tactic. This list was then cleaned to remove obvious duplicates. For example, "Threw Object" and "threw object" were both recoded as "THREW OBJECT". The result was a list of 279 unique descriptions, 15 of which appeared to name more general categories (see Table 3.4).

Several methods were considered for coding the battering tactics in hope of finding some differences between either warrants involving female suspects and male suspects or differences between first offense suspects and repeat offense suspects. One approach would be to manually recode each description in terms of an existing scale or published frequency distribution, but this would arguably have made it difficult to assess what the resulting distributions represented since many records contained descriptions that did not easily fit into existing coding schemes. Instead, a single large indicator matrix of tactics was constructed with each row corresponding to a record in $IDMAT$ and each column corresponding to one of the 279 unique descriptions resulting in $TACMAT$ (6358 rows by 279 columns). One could then

Table 3.4: Major Categories of Battering Tactics

Category	% of all cases
ARGUED	5
BIT	22
CHOKED	12
SEXUAL ASSAULT	9
GRABBED	10
KICKED	16
PULLED	2
PUNCHED	19
PUSHED	34
SLAPPED	27
STALKED	4
THREATENED	16
THREW DOWN	1
THREW OBJECT	5
USED A WEAPON	1

use $TACMAT$ in a cluster analysis to generate an atheoretical set of classifications that might identify combinations of tactics that corresponded to different types of suspects (e.g. victims versus assailants, first offenders versus repeat offenders). Or, one could use an existing theory and recode records from $TACMAT$. While simulations demonstrated the computational feasibility of running a cluster analysis on $TACMAT$, the results were disappointing because police officers had already coded their police reports in terms of legal criteria for an arrest. That is, instead of describing what happened in terms of the specific battering tactics, the fields appeared to contain labels that corresponded to the legal criteria that police officers might have used to justify an arrest decision. For example, while there were some descriptions of battering tactics that made distinctions about the type of object thrown, most simply labeled such behavior as "threw object."

3.9 Numerical time series

IDMAT was a case or individual level data matrix that included dates of certain critical events. These dates could be aggregated over time intervals to produce numerical time series of rates (e.g. number of events per unit time). Preliminary inspection of the *IDMAT* suggested that aggregating events (arrests, warrant authorizations, denials, dismissals, etc.) over months would be sufficient for identifying and representing the time variant behaviors involved with how a prosecutor's office manages its domestic violence caseload. This was based on the assumption that the shortest time constants of interest would be on the order of 30 days since the mean number of days from the date of the offense for a warrant being approved was initially estimated to be 31 days with the mean number of days from authorization to plea offer being accepted and dismissed being 53 days and 96 days respectively. Aggregating events on months would have allowed a sufficient sample size to avoid the problems associated with sparse time series. However, using the median as a more robust estimate of the central tendency and plotting the cumulative distributions over days from offense led to a starkly different and more realistic picture. Over 50% of the warrants were being authorized within a day, 50% of cases were being plead in 29 days or less after the arrest, and 50% of dismissals were happening in 70 days or less after arrest. With less than one warrant request per day, aggregating on days would mean having to deal with sparse time series (Figure 3.3c is a simulated example of sparse time series) as the number of cases per day jumped randomly between values of 0 and 8. Aggregating on weeks would avoid the problem of having zeros as discontinuities, but still suffer from not having enough cases. Specifically, as the rates (number of events per unit time) that one would be aggregating on became smaller, the totals per unit time would become increasingly sensitive to both the size of the aggregating bins and their origin of the aggregating bins (Härdle, 1990).

3.9.1 Illustration of sparse time series

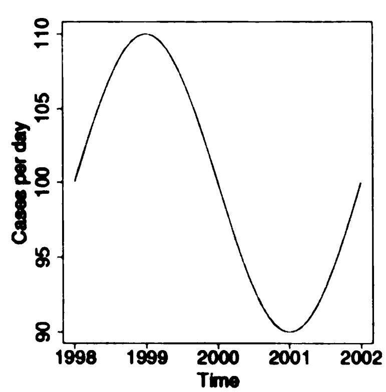
The impact of smaller rates on the sensitivity of aggregations to bin size and their origin can be seen through a series of simulations where the expected number of events per unit time decreases. Figure 3.2 shows three simulated time series of the rate or expected number of cases per day with respective overall means of 100, 10, and 1 (note that, going from left to right, the vertical scaling decreases by a factor of ten for each graph). In the case of IDMAT, the actual mean number of arrests was on the order of one case per day.

The number of events per unit time is always an integer value equal to or greater than zero. This is typically modeled as a Poisson distribution. One can generate a simulated time series of observations by randomly sampling one case per unit time from a Poisson distribution with the expected rates corresponding to the time series shown in Figure 3.2. Specifically, $O_m(t) = P(\lambda_m(t))$ where, $O_m(t)$ is the observed value at time t, $P(\lambda_m(t))$ is a random value sampled from the Poisson distribution with a rate of $\lambda_m(t)$. The rate $\lambda_m(t)$ is a function of time and the mean expected value m such that $\lambda_m(t) = m + 0.1 \cdot m \cdot \sin(t \cdot 2\pi/1461)$, where m is the overall mean (i.e., 100, 10, or 1)

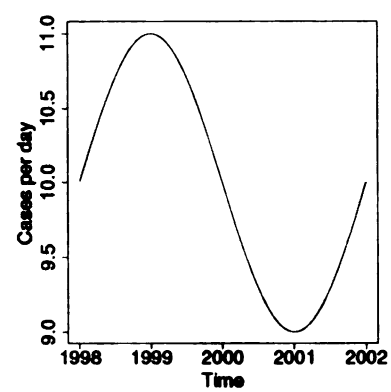
The resulting noisy time series are shown in Figure 3.3. As m gets smaller, the effects of the distribution being bounded at zero become more pronounced. Figure 3.3a would typically be seen as a noisy but continuous time series. But as m gets smaller, the result is something that looks much more discrete with quite a few days with zero cases. Figure 3.3c would be a good example of a sparse time series.

When the number of events per unit time is relatively high, aggregating the noisy time series over quarters results in a reasonable approximation of the original

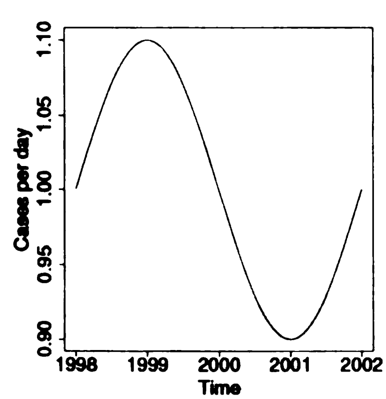
Figure 3.2: Original Expected Rate of Cases per Day



(a) Mean = 100

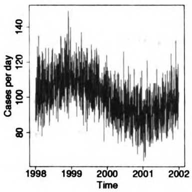


(b) Mean = 10

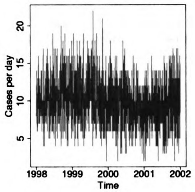


(c) Mean = 1

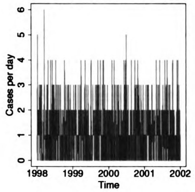
Figure 3.3: Noisy Time Series



(a) Mean = 100



(b) Mean = 10



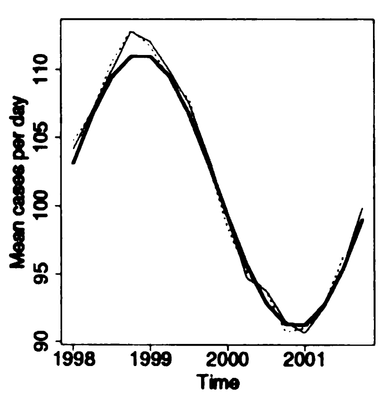
(c) Mean = 1

expected values, but not so for sparse time series. Figure 3.4 shows the result of aggregating the counts over quarters. The thick solid line shows the original expected values, the thin solid line shows the aggregated counts over quarters (divided by the number of days per quarter to standardize the values), and the thin dashed line shows the effect of offsetting the origin of the bins by as little as nine days.[3] Figure 3.4a illustrates how aggregating the number of cases per day on quarters results in a reasonable approximation of the original expected values and is insensitive to minor variations in the origin of the aggregating bins. However, as m gets smaller, the aggregated time series does a worse job of approximating the original time series and becomes more sensitive to variations in the origin of the bins. Figure3.4b might still be a reasonable approximation, but one would be hard pressed to identify the original time series in Figure 3.4c. Figure 3.4c also shows how even a small change in the origin of the bins of only nine days can affect quarterly totals, especially the first, second, and third quarters of 2000.

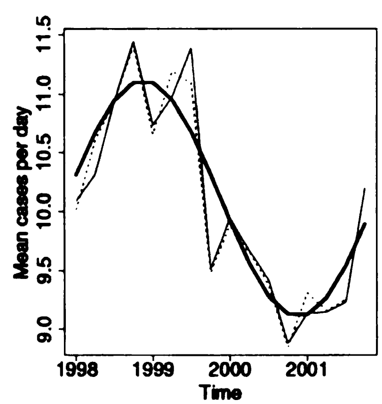
The effects of small time series on aggregated values get worse as the size of the aggregating bins get smaller. Figure 3.5 shows the results of aggregating the number of cases over months which are, again, divided by the width of the aggregating bin to standardize the values. While Figure 3.5a still appears to be a reasonable approximation of the original time series, Figure 3.5b is starting to become questionable and showing more effects from varying the origin of the bin size. Figure 3.5c is unrecognizable.

[3]The choice of offsetting the origin by nine days is arbitrary. The main point is that the offset is small relative to the width of the bins.

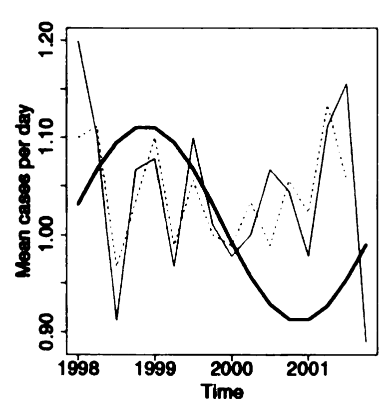
Figure 3.4: Quarterly Aggregated Time Series



(a) Mean = 100

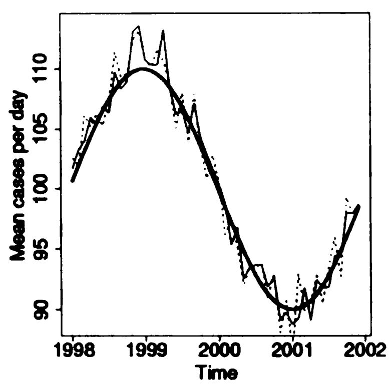


(b) Mean = 10

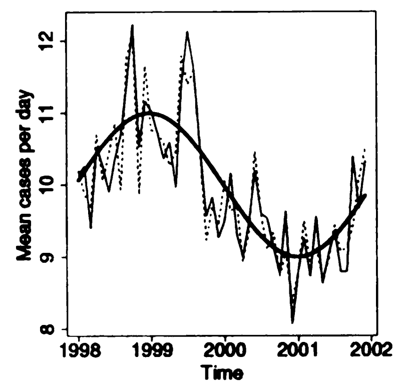


(c) Mean = 1

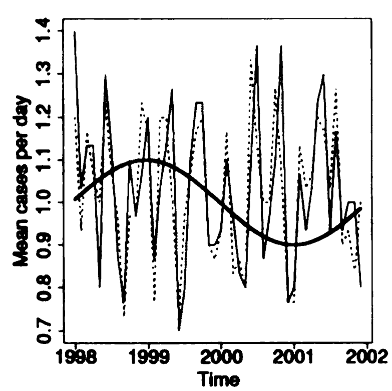
Figure 3.5: Monthly Aggregated Time Series



(a) Mean = 100



(b) Mean = 10



(c) Mean = 1

3.9.2 Conventional approaches to handling noisy data

Approaches to handling noisy time series fall into three general classifications: filtering, smoothing, and prediction (Anderson & Moore, 1979). Filtering works in real time, trying to recover the original value from the noisy signal $O(t)$ at time t. The tuner on a radio or television is a classic example of a filter. Smoothing uses a range of values, $(t - \Delta t,\ t + \Delta t)$, to estimate $O(t)$. There are many approaches to smoothing, but perhaps the most familiar is a moving average. Prediction uses a set of past values, $(t - \Delta t,\ t)$, to estimate $O(t)$ for a set of future values, $(t,\ t + \Delta t)$: for example, by fitting an equation to past values in order to predict future values.

When generating reference modes from real data, one is generally concerned with identifying trends and working with time series spanning the entire time horizon. That is, for each point in time t, one can use information coming both before and after t to estimate some value at t. Thus, one is generally looking to select a method for smoothing noisy time series. If one knows a priori the underlying probability distribution of a given time series, then one can apply regression techniques to estimate the parameters of the underlying distribution. Some variables in this study might fit one of the many available parametric probability distributions. But this cannot, in general, be assumed because of the nature of feedback.

One possibility is the use of non-parametric techniques to estimate the density distribution (Härdle, 1990; Scott, 1992). Such techniques do not assume a specific underlying distribution, but rather, are based on the empirical distribution of a given time series. However the application of smoothing techniques introduces some problems. Smoothing always distorts oscillations in time series in one or more ways. Frequencies may be entirely eliminated, peaks flattened out, and delays introduced. Smoothing also introduces the problem of truncating the first and last parts of a time series since most procedures use an interval of values to estimate a given point. Thus, one is often not sure whether or not some important features have been removed by smoothing the time series, or whether the features that one is trying to model are in fact, not just artifacts of the smoothing algorithm.

3.9.3 Solution to the sparse time series problem

In system dynamics modeling, one is generally concerned with identifying and studying the underlying structure generating a particular pattern of behavior over time. One uses numerical time series, not to identify the system of equations as one might in traditional time series modeling, but to test the model's behavior against real data. The purpose of such a test is generally not to see whether or not the model yielded the precise values of the observed data, but whether the overall behavior pattern is realistic. That is, one can make a distinction in system dynamics between the *real numeric values* and *realistic numeric values* of a variable. The real numeric value of a variable is the actual value that the variable takes on at a given point in time. If the variable is the number of warrant requests on a given day, then there is a real theoretical numeric value for the expected number of warrant requests on that day. The problem of estimating the density distribution of the number of warrant requests per day is concerned with estimating that numeric value, and it is this distribution that is sensitive to the particular technique used to smooth the time series. A realistic numeric value, however, is something more general. For any point in time, there can be many values that would be realistic, only one of which is the real numeric value of that variable at that time.

A good robust model should be able to describe the dynamics over a range of situations. Given a variety of inputs, the model should be able to reproduce corresponding outputs that are realistic. In system dynamics, researchers often test their models using a variety of inputs to explore the model's behavior over extreme conditions, oscillations, and random perturbations (Forrester, 1961/1999, 1971; Forrester & Senge, 1980; Richardson & Pugh, 1986; Sterman, 2000). Careful inspection of the outputs and feedback loop behavior often reveals structural flaws or important insights into the model's behavior. These inputs are typically idealized in some form as step functions, pulses, sine waves, or random samples from a parametric distribution. But there is no reason why a robust model should not also be able to handle a much wider range of inputs, including the results of not just a specific smoothing algorithm, but the results from an entire collection of smoothing algorithms.

More specifically, a robust system dynamics model should be able to produce reasonable approximations of smoothed numeric time series outputs given corresponding smoothed numeric time series for an input, provided that both the input and output numeric time series have been smoothed using the same algorithm. That is, a robust system dynamics model should be able to reproduce the qualitative behavior of two different time series, each smoothed using a different algorithm or with different smoothing parameters, provided that the model's input (if there is any) was also smoothed using the same algorithm and smoothing parameters.

3.9.4 Rates and stocks numeric time series

Rates were calculated by aggregating cases by dates of key events from $IDMAT$. To generate time series for first offenses versus repeat offenses, or for each type of relationship, the time series were calculated from subsets of $IDMAT$. The resulting rates were stored in a matrix of time series, $TSMAT$, with various suffixes to indicate the subset (e.g. $TSMATall$, $TSMATfirst$, $TSMATrepeat$). The generation of time series for stock variables was more complicated. Unlike rates, time series of stock variables had to be calculated by first specifying the date variables that indicated entering and exiting of the stock and then counting the number of cases that were in that stock for each day (see Tables 3.5 and 3.6).

Table 3.5: Times Used to Generate Caseload Statistics

Variable	Description
t_0	Date of offense
t_1	Date warrant was denied or authorized
t_2	Date warrant was authorized
t_3	Date of plea offer was accepted, case being dismissed, or trial
t_4	Date of plea offer was accepted, case dismissed, case denied, or trial

Table 3.6: Entry and Exit Times for Caseload or Stock Variables

Range	Stock variable
$[t_0, t_1]$	Warrants
$[t_2, t_3]$	Cases
$[t_0, t_4]$	Total

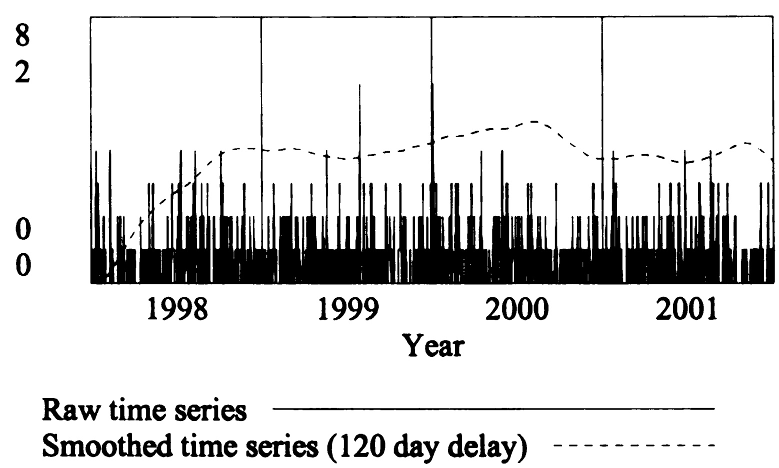
All time series generated in R were then exported to a Vensim compatible text file. Smoothing algorithms were implemented as separate Vensim models, which read the original input and saved the output to data files that could be read by the Vensim system dynamics models. Smoothed data sets were generated and compared for first and third order exponential smooths with 30, 60, 120, and 180 day delays. The mean values of the sparse time series were at first used as initial values in the smoothing models in order to minimize differences between the cumulative totals of the original and smoothed data. This approach was subsequently abandoned as (a) starting with initial values of zero in the smoothing algorithm corresponded to stocks in the system dynamics model variables being initially empty, and (b) the resulting step input response of the system dynamics model corresponded to a realistic step input associated with the prosecutor's office establishing a specialized domestic violence prosecution unit. Figure 3.6a shows a comparison between the raw time series for warrant requests and the results of smoothing the times series using a third-order 120 day delay exponential filter. The solid (vertical) lines show that the warrant request time series is a sparse time series, while the dashed line shows the smoothed version of the warrant request time series. This smoothed time series can be considered an estimate of the expected value of the number of warrant requests per day. The mean of the sparse time series and smoothed time series should be equal (the extreme of a smoothed time series would simply be a constant value equal to the mean). Figure 3.6b shows the effects of varying a smoothing parameter, specifically the delay, from 30 to 180 days. Notice that as one increases the delay, the time series appears smoother with fewer high frequency oscillation.

3.10 Building system dynamics models

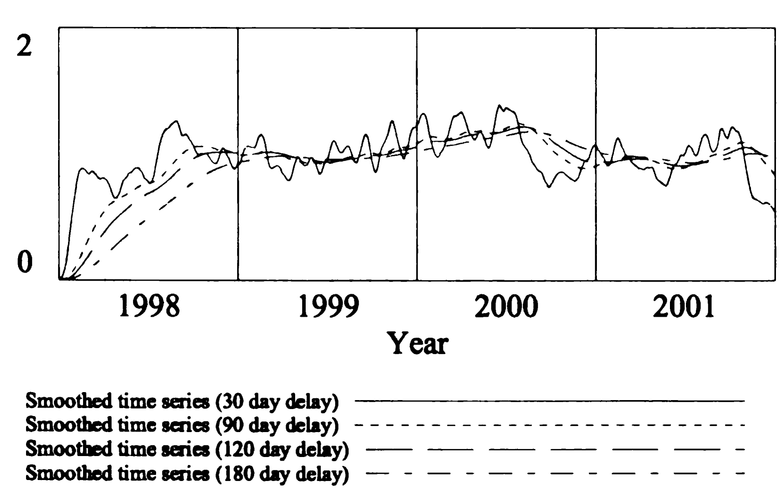
There were three major phases of constructing the system dynamics models. The goal of the first phase was to build the smallest models that could reproduce the total caseload over time. These started out as single-stock models that represented dispositions as (a) strictly a function of caseload, (b) strictly a function of case attributes, and (c) a mixture of caseload and case attributes. These models were only descriptive in nature and did not provide any account of the underlying mechanisms. The second phase of modeling focused on building the initial baseline model, which essentially provided the first full dynamic hypothesis of how the prosecutor's office managed its domestic violence caseload. This model was used to identify key stock variables, refine the organization of data by subgroups, and develop questions for the initial interviews with key informants. The third phase of model building used data from key informant interviews to revise the model and provide additional infor-

Figure 3.6: Comparisons of Warrant Request Time Series

(in cases per day)



(a) raw versus smoothed time series with 120-day delay



(b) Smoothed time series with 30, 90, 120, and 180 day delays mation on various functional relationships between variables and model parameters.

A modular approach was taken in the development of the baseline and revised models. That is, submodels were developed and tested separately, i.e. partial-model testing (Homer, 1983). That is, submodels or basic units of structure were developed, each entailing a dynamic hypothesis that the structure was capable of reproducing the observed time series behavior. Where inputs like arrests or warrant authorizations were required from earlier stages of the prosecutor's caseflow, real inputs were used to drive the submodel and the simulated outputs were compared against the real data. Thus, for example, real arrest rates were used to drive the warrant processing submodel and the simulated results of authorization and denial rates compared against the real data. Such tests frequently revealed flaws in the submodel's formulation, which were subsequently revised and tested again. When the submodel was structurally capable of generating a reasonable reproduction of the real data, uncertain model parameters were calibrated against the real data. Calibrated parameters that were unrealistic indicated additional problems with the submodel's structure. After a submodel finally succeeded in reproducing the reference mode, additional models were built, often based on the first submodel. These variations of the same submodel could then be compared in order to assess the impact of various modifications on the submodel's ability to reproduce the reference mode. In many cases, simplifying models had a negligible impact on the submodel's performance.

The use of partial-model testing has two main advantages over whole model-testing during the formulation of a model. First, it minimizes the amount of information used to calibrate the model. As models get larger, more parameters are introduced, which need to be calibrated and tested. In large models, parameters in one part of the model can be adjusted to compensate for behaviors in another part of the model. This is what one would expect in large complex systems. The problem in model building, however, is that with the available software, it is relatively easy to fit large models. The price for fitting large models is that one does not know to what extent the parameters in one part of the model are compensating for other parts of the model. That is, the larger the model, the more opportunity one has to adjust constants in such a way that the simulated behavior more accurately reassembles the real behavior. Partial-model testing avoids this by focusing on small submodels with a few parameters and thereby limiting the number of parameters and interactions. Thus, partial-model testing uses less information during model development and allows whole-model testing to reveal new structural flaws. The second main advantage of partial-model testing is that it generally leads to a much easier development path and understanding of why a model behaves as it does. Careful study of a single submodel, including dimensional consistency tests, testing and calibrating against real data, and sensitivity analysis will lead to a more thorough understanding of each component. It also is generally much easier to isolate the problem in a submodel than trying to trace the problem down at the whole-model level.

Whether working on a model or submodel, the basic approach is the same. Initial work focused on laying out the basic stock-and-flow diagram along with formulating the basic relationships between variables. The dimensional consistency of units in the equations frequently gives some important clues as to how an expression might be formulated and provides one way to structurally check the model (Richardson & Pugh, 1986). Thus all formulations included units and dimensional consistency checks. Simulations were then used to explore and test the behavior of the model, which often led to more modifications. Some models were unable to reproduce the intended behavior after several iterations, which often indicated after further analysis reasons for ruling out that particular structure. Models that passed this stage of development and appeared capable of reproducing the qualitative behavior over time were then calibrated to smoothed data for comparisons. Models were calibrated in Vensim by specifying the one or more output variables to match and the model parameters to vary. Since output variables have different magnitudes, for example warrant requests are on the order of 1 case per day while caseloads are on the order of 100 cases, the output variables to match on were initially weighed by dividing the time series values by their standard deviation. Better results seemed to be achieved, however, by dividing each time series by their average value over time.

3.11 Key informant interviews

The second phase of model building generated many questions about realistic values for certain parameters and critical relationships between relationships. It was also important to establish whether or not patterns identified in the numerical time series were valid from the perspective of stakeholders.

A snowball sampling procedure was used to identify key community informants, who were selected on the basis of their expertise with domestic violence and prosecution of domestic violence cases. Two initial contacts were made to a domestic violence coalition coordinator and assistant prosecuting attorney. These key informants suggested other potential participants. A total of five key informants from the community were interviewed, and included victim advocates, domestic violence coalition coordinators, assistant prosecuting attorneys, and legal assistants. All five key informants had experience working with victims of domestic violence and had attended training workshops specific to working with domestic violence.

A list of questions from the baseline model was prepared prior to the initial contact with a key informant. The key informant was then contacted via telephone or email to schedule an initial telephone call. The first part of the initial telephone call included the reading of the consent script. Participants provided verbal consent by continuing with the interview. The date of their consent was recorded in a log, along with their contact information, participant number, agency, and total number of interview hours. Interviews were unstructured with the list of questions serving as a general guide. Handwritten notes were taken during the interview and then typed to disk after the interview. Analysis of the interview notes was informal, and involved reading the handwritten and typed notes, and identifying key informants' descriptions of trends and feedback mechanisms in a laboratory notebook.

3.12 Conclusion

This chapter provided an overview of the methods used to generate and analyze the numeric time series along with summary of the system dynamics approach taken in this study. Specific attention was paid the construction of smoothed time series from sparse time series that could be used to generate the reference modes for the various system dynamics models. The chapter also discussed the use of finite state machines as a way to model the transitions of individuals within a social service delivery system.

Chapter 4

Results

4.1 Overview

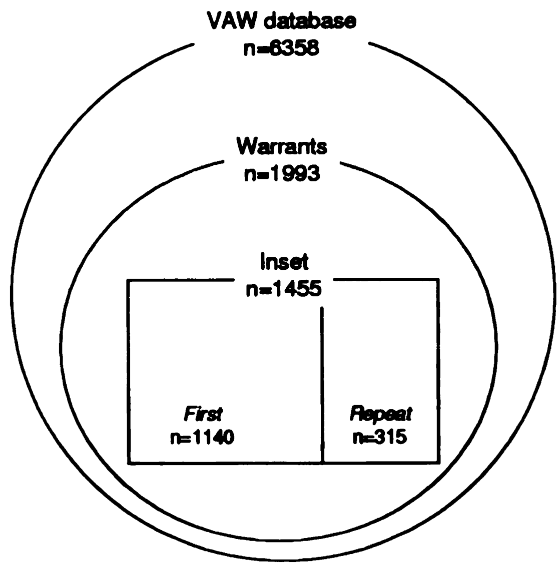
This chapter presents and discusses the results of this study. The first part of the chapter focuses on documenting the violence against women (VAW) database. Demographics, case attributes, and dispositions are considered with respect to all the warrants in the VAW database, those selected for analysis, and first and repeat arrest subsets. The second part of the chapter briefly describes three descriptive models and compares how well they are able to reproduce the behaviors of the real data. The third part focuses on the development of the baseline model, going through each submodel, its variations, and performance in response to real inputs. The chapter then continues by evaluating the performance of the baseline model and presents the results from sensitivity analyzes and analysis of patterns in feedback loop dominance. The fourth section diagnoses some of the problems in the model's behavior in terms of identifying an additional positive feedback loop between first time arrests and new cases. This is followed by a brief summary of the key informant interviews and presentation of the revised model.

4.2 Demographics

Based on 2000 U.S. Census statistics, the jurisdiction of the prosecutor's office covered a county of approximately 70,000 people. About 55,000 persons were 15 years of age or older. Although criminal justice cases and individuals represent different units of analysis, comparisons between the distributions of census demographics and case demographics indicate differences in arrests rates. That is, if the arrest rates were the same over all groupings of individuals, then one would expect the demographic distributions of rates to be proportional to the demographic distributions of census data. One can make similar comparisons between different subsets of the prosecutor's office's data, e.g. between first arrests and repeat arrests. Table 4.1 through Table 4.3 show the results of comparing the distribution of census demographics against the demographics for all warrants, those selected in the analysis set, first arrests, and repeat arrests, which are respectively selected with the following binary variables: *Warrants*, *Inset*, *First*, and *Repeats*. These variables take on the value of true if the case is in that specific set and otherwise false. Thus, all the cases within the VAW that start out as warrant requests have a true value for the variable *Warrants*. *Inset* indicates all the cases that are in the analysis set, which is a subset of all the cases selected by *Warrants*. *Firsts* and *Repeats* are variables that respectively indicate cases that are first domestic violence arrests and repeat domestic violence arrests within the subset of cases indicated by *Inset*. Since a case that is *Inset* is either a first domestic violence arrest or a repeat domestic violence arrest, the union of all cases indicated by *Firsts* and *Repeats* should be identical with *Inset*. The set relationships between cases indicated by *Warrants*, *Inset*, *Firsts*, and *Repeats* is shown in Figure 4.1.

Differences between the demographic distribution of warrants and those selected

Figure 4.1: Subsets of Violence Against Women Database



for the analysis would suggest problems with (a) reliability of coding and data entry, or (b) missing data mechanisms that correlate with demographics. Comparisons between all warrants and those in the selected set for analysis ("inset") in Tables 4.1-4.3 reveal that any such discrepancies can be treated as negligible. Since first time arrests are a unique one-time attribute of individuals, the demographic distribution of suspects in first arrest cases is identical to the demographic distribution of suspects at the individual level. That is, the census distributions are directly comparable to the first arrest suspect distributions.

Overall, there were more male suspects than female suspects, and more female victims than male victims (see Table 4.1), which would be consistent with the general understanding of domestic violence that it is primarily committed by men against women. Note that female suspects include battered women who were arrested by the police and women who were abusing other women in same sex relationships, and male victims includes cases where the victim was arrested and men who were being abused by other men in same sex relationships. There was little variation between the distributions of sex across the different subsets. However, repeat arrests had a higher proportion of male suspects and female victims than warrants, those in the set, and first arrests.

Table 4.1: Distribution of Sex by Suspects and Victims Across Subsets of Violence Against Women Database

	Census	*Warrants*		*Inset*		*First*		*Repeats*	
		Suspects	Victims	Suspects	Victims	Suspects	Victims	Suspects	Victims
Female	49%	22.9%	74.4%	23.2%	73.9%	25.2%	72.2%	16.1%	80.3%
Male	51%	77.1%	25.6%	76.8%	26.1%	74.8%	27.8%	83.9%	19.7%
Total	100%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Warrants included information about race or ethnicity of the victim and suspect. Table 4.2 compares the distributions of suspects and victims by all warrants, those in the analysis set, first arrests, and repeat arrests. Whereas census figures treat Hispanics as an attribute separate from race, warrants made no such distinction. Thus the census race category of white is not directly comparable to warrant race category of white non-Hispanic. Furthermore, with a small population that is 96% white, small variations in the frequencies in other racial and ethnic categories will have a disproportionate effect on percentage comparisons. Thus several repeat arrests of the same individual can severely skew the percentages. Nonetheless, such differences cannot be ignored or go undocumented because they may point to important biases in the criminal justice system's response to racial and ethnic minorities. And one might expect such biases to be more accentuated, not less, in a county where the population is mostly white.

The vast majority of both victims and suspects were recorded as white non-Hispanic. In comparison to 2000 county census demographics for persons 15 years

Table 4.2: Distribution of Race/Ethnicity by Suspects and Victims Across Subsets of Violence Against Women Database

	Census		*Warrants*		*Inset*	
	Male	Female	Suspects	Victims	Suspects	Victims
African American	3%	1%	4.4%	1.5%	5.1%	1.6%
Asian American	1%	1%	0.3%	0.2%	0.4%	0.2%
Hispanic	1%	1%	0.2%	0.2%	0.2%	0.2%
Native American	2%	2%	1.3%	1.0%	1.2%	1.2%
Other	na	na	0.1%	0.1%	0.0%	0.0%
Unknown	na	na	0.1%	0.1%	0.1%	0.1%
White non-Hispanic	95%	96%	93.7%	96.9%	93.0%	96.7%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

	Census		*First*		*Repeats*	
	Male	Female	Suspects	Victims	Suspects	Victims
African American	3%	1%	4.1%	1.7%	8.6%	1.3%
Asian American	1%	1%	0.2%	0.2%	1.3%	0.3%
Hispanic	1%	1%	0.3%	0.2%	0.0%	0.3%
Native American	2%	2%	1.3%	1.0%	1.0%	1.7%
Other	na	na	0.0%	0.0%	0.0%	0.0%
Unknown	na	na	0.1%	0.2%	0.0%	0.0%
White non-Hispanic	95%	96%	94.0%	96.8%	89.1%	96.3%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

of age or older, the census distribution of white females is consistent with the warrant distribution of white non-Hispanic females. In contrast, the distribution of warrants with African American female victims is higher than the corresponding census distribution (1.5% versus 1%), but lower for Asian American female victims (0.2% versus 1%), Hispanic female victims (0.2% versus 1%), and Native American female victims (1.0% versus 2%). Relative to the census distributions, white non-Hispanic males appear at nearly the same rate (94.0% versus 95%), but are under-represented in repeat arrests (89.1% versus 94.0%). In contrast, cases involving African American suspects appear more frequently than the corresponding census figures would suggest (4.1% versus 3%), and at more than twice the rate for repeat offenses (8.6% versus 4.1%). While Asian American suspects were less likely to appear in comparison to the census statistics (0.2% versus 1%), they appear more six times more frequently in repeat arrests (1.3% versus 0.2%). Hispanic and Native American suspects are both under-represented in first arrests and repeat arrests.

Table 4.3 summarizes the age distributions. The age distributions of both suspects and victims are concentrated in the range of 15 and 44 years, which is consistent with general demographic patterns in the criminal justice system. This does not mean that persons under 15 or over 40 are not being victimized or perpetrating violence against women and children. The age distribution of suspects and victims reflect both age distributions in incidence rates and reporting rates.

4.3 Case attributes

Table 4.4 summarizes the distribution of charges at the time of arrest. Sexual assault is typically a felony arrest, but the remaining charges are, for the most part, misdemeanors. There are two major distinctions between domestic violence arrests

Table 4.3: Distribution of Age by Suspects and Victims Across Subsets of Violence Against Women Database

	Census		*Warrants*		*Inset*		*First*		*Repeats*	
	Male	Female	Suspects	Victims	Suspects	Victims	Suspects	Victims	Suspects	Victims
Under 5	7%	7%	0.0%	1.2%	0.0%	0.9%	0.0%	0.7%	0.0%	1.6%
5-9	7%	7%	1.2%	1.6%	0.0%	1.3%	0.0%	1.3%	0.0%	1.3%
10-14	6%	6%	0.3%	6.3%	0.0%	6.2%	0.0%	6.1%	0.0%	6.5%
15-19	9%	8%	9.4%	16.4%	8.8%	17.0%	9.0%	16.0%	8.1%	20.6%
20-24	9%	8%	19.6%	18.2%	20.0%	18.4%	20.2%	18.7%	19.0%	17.3%
25-29	8%	8%	16.3%	14.3%	16.0%	14.5%	15.2%	15.3%	19.0%	11.8%
30-34	8%	8%	15.7%	12.9%	15.9%	12.6%	16.2%	12.6%	14.8%	12.4%
35-39	7%	7%	16.0%	12.2%	17.1%	12.6%	16.7%	12.4%	18.4%	13.4%
40-44	7%	8%	8.7%	9.0%	9.2%	8.7%	9.5%	8.7%	8.1%	8.8%
45-49	8%	8%	6.6%	3.6%	6.8%	3.4%	6.5%	3.6%	7.7%	2.6%
50-54	6%	6%	3.3%	2.2%	3.3%	2.3%	3.6%	2.2%	2.3%	2.6%
55-59	4%	4%	1.4%	1.1%	1.3%	1.2%	1.4%	1.4%	1.0%	0.7%
60-64	3%	4%	0.7%	0.7%	0.7%	0.8%	0.8%	1.0%	0.3%	0.3%
65-69	3%	3%	0.5%	0.1%	0.6%	0.1%	0.5%	0.1%	0.6%	0.0%
70-74	3%	3%	0.2%	0.0%	0.2%	0.0%	0.2%	0.0%	0.3%	0.0%
75-79	2%	3%	0.1%	0.1%	0.1%	0.0%	0.1%	0.0%	0.0%	0.0%
80-84	1%	2%	0.1%	0.0%	0.1%	0.0%	0.0%	0.0%	0.3%	0.0%
Over 85	1%	2%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.0%	0.0%
Total	100%	100%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

and assault and battery arrests. First, police can make a warrantless domestic assault arrest in Michigan if there is evidence of violence, and the victim and suspect (a) are a current or former married couple, (b) share a residence, or (c) have a child in common.[1] A warrantless arrest means that the police officers do not have to directly witness the crime being committed in order to make the arrest. One legal implication of this is that the arrest warrant is actually issued after the arrest has been made. The second major distinction between a misdemeanor domestic violence arrest and misdemeanor assault and battery is that misdemeanor domestic violence arrests carry a maximum penalty of 93 days in jail in contrast to assault and battery, which carries a maximum penalty of 90 days in jail. This means that police officers are required to finger print suspects in misdemeanor domestic violence arrests, which essentially provides a state wide mechanism to track domestic violence suspects. Police departments are not required to finger print and report misdemeanors with a maximum jail sentence of 90 days or less. This means that a criminal background check through Michigan State Police would not reveal a previous arrest for assault and battery. Prior to 1994, assailants arrested on assault and battery could evade the system by moving to a different county. Each time they were arrested would then essentially be treated as the first offense. So a misdemeanor domestic assault having a maximum sentence of 93 days in jail means, with one important exception of the "first" offense, that domestic assault arrests will appear in a criminal history background check. The exception is that Michigan law provides a mechanism (769.4a) whereby first time convictions for domestic assault can receive a deferred sentence, and upon completion of probation requirements, essentially have no criminal record of the arrest and conviction. A second arrest and conviction would then essentially appear as the first official arrest, conviction, and sentence.

[1]The law was amended in 2002 to include dating relationships.

Most cases in the prosecutor's violence against women database involved misdemeanor domestic violence arrests (79.8%), followed by sexual assault (12%), stalking (5.5%), assault and battery (2.2%), and personal protection order (PPO) violations (0.5%). These distributions are consistent between all warrants, those in the analysis set, and first arrests. Repeat arrests, however, are less likely to involve domestic violence (77.1% versus 79.5%) and sexual assault (8.7% versus 12.0%), but more likely to involve stalking (10.3% versus 5.5%).

Table 4.4: Arrest Charges Across Subsets of Violence Against Women Database

		Warrants		*Inset*		*First*		*Repeat*
Assault and battery	42	2.2%	38	2.6%	27	2.4%	11	3.5%
Sexual assault	228	12.0%	162	11.3%	135	12.0%	27	8.7%
Domestic violence	1516	79.8%	1136	79.0%	897	79.5%	239	77.1%
PPO violation	9	0.5%	8	0.6%	7	0.6%	1	0.3%
Stalking	104	5.5%	94	6.5%	62	5.5%	32	10.3%
Total	1899	100.0%	1438	100.0%	1128	100.0%	310	100.0%

The VAW database included an open ended text field describing the nature of the relationship between the victim and suspect. Cases that could be clearly identified as involving intimate partner violence were coded as domestic violence, cases that could otherwise be clearly identified as involving two people living together where coded as cohabiting, and the remaining cases were coded as other. Table 4.5 summarizes the frequency distributions of the type of relationship.

4.4 Case dispositions

While a common perception of the criminal justice system is that most cases are resolved in court trials, very few actually are. Table 4.6 summarizes the dispositions. About half the cases either do not have the warrant authorized by the prosecutor's

Table 4.5: Type of Relationship Between Victim and Suspect Across Subsets of Violence Against Women Database

	Warrants		*Inset*		*First*		*Repeat*	
Domestic violence	745	37.7%	549	38.2%	438	38.8%	111	35.8%
Cohabiting	572	28.9%	342	23.8%	266	23.6%	76	24.5%
Other	659	33.4%	547	38.0%	424	37.6%	123	39.7%
Total	1976	100.0%	1438	100.0%	1128	100.0%	310	100.0%

office (39.2%) or end up eventually being dismissed (10.8%). Pleading no contest is essentially the same thing as pleading guilty. Cases that are nolled or *nolle prosequi* are cases where the prosecutor's office dismisses the case because the lawsuit has been abandoned. The remaining cases are resolved mostly through plea agreements (40.5%). Only 1.4 percent of the cases are resolved through a court trial. Table 4.7 summarizes the reasons given by prosecutors for denying or dismissing a case, the main one being lack of evidence.

Table 4.6: Dispositions Across Subsets of Violence Against Women Database

	Warrants		*Inset*		*First*		*Repeat*	
Denied	695	39.2%	594	42.7%	490	44.5%	104	35.9%
Dismissed	192	10.8%	131	9.4%	94	8.5%	37	12.8%
Guilty at trial	11	0.6%	6	0.4%	5	0.5%	1	0.3%
Nolled	6	0.3%	6	0.4%	5	0.5%	1	0.3%
Not guilty at trial	15	0.8%	10	0.7%	8	0.7%	2	0.7%
Plead guilty	719	40.5%	541	38.9%	413	37.5%	128	44.1%
Plead no contest	137	7.7%	102	7.3%	85	7.7%	17	5.9%
Total	1775	100.0%	1390	100.0%	1100	100.0%	290	100.0%

Guilty dispositions, either through trial convictions or plea agreements, can result in a variety of sanctions. Table 4.8 summarizes the various combinations of sanctions along with their distribution. Percentages are in terms of all cases within the set, not just those cases where a finding of guilt was determined. Thirty-seven

Table 4.7: Distribution of Reasons Given by Prosecutor for Denial or Dismissal Across Subsets of Violence Against Women Database

	Warrants		*Inset*		*First*		*Repeat*	
Civil	6	0.9%	6	1.0%	5	1.1%	1	0.8%
Apologized	2	0.3%	2	0.3%	2	0.4%	0	0.0%
Complied with treatment	7	1.1%	7	1.2%	7	1.5%	0	0.0%
Plead to other charge	20	3.1%	18	3.0%	6	1.3%	12	9.7%
Lack of evidence	510	78.7%	473	79.2%	384	81.2%	89	71.8%
Dismissed by judge	2	0.3%	2	0.3%	0	0.0%	2	1.6%
Not in best interest of justice	3	0.5%	3	0.5%	3	0.6%	0	0.0%
Other	46	7.1%	40	6.7%	30	6.3%	10	8.1%
Police request	3	0.5%	3	0.5%	2	0.4%	1	0.8%
Probate	12	1.9%	10	1.7%	8	1.7%	2	1.6%
Victim	37	5.7%	33	5.5%	26	5.5%	7	5.6%
Total	648	100.0%	597	100.0%	473	100.0%	124	100.0%

percent of cases that were in the analysis set (*Inset*) did not receive any type of sanction. Approximately 11% that were in the analysis set (*Inset*) were referred to some type of batterer intervention program (BIP). Delay of sentence (DOS) was agreed to in approximately 14% of the cases in the analysis set (*Inset*).

The county has two batterer intervention programs, although neither meets the minimum state standards for batterer intervention services and key informants generally felt that this was one of the major areas lacking in their community response. Table 4.9 summarizes the referrals for cases receiving a guilty disposition. The 10-week batterer intervention program (BIP) falls far short of the 26-week minimum required by state standards. The 26-week BIP follows more of an anger management model, but does not follow an accountability model like the Duluth model (Pence & Paymar, 1993). Anger management for batterers is contraindicated (Gondolf, 2002; Healey et al., 1998; Pence & Paymar, 1993). The "other" referrals are to substance abuse assessment and treatment programs, individual psychological or ministerial

Table 4.8: Distribution of Consequences from Prosecution Across Subsets of Violence Against Women Database

	Warrants		*Inset*		*First*		*Repeat*	
None	1223	61.4%	846	37.0%	668	37.2%	178	36.5%
BIP only	52	2.6%	25	1.1%	20	1.1%	5	1.0%
BIP and DOS	130	6.5%	100	4.4%	93	5.2%	7	1.4%
BIP and Jail	132	6.6%	95	6.6%	69	6.1%	26	8.4%
DOS only	196	9.8%	152	10.6%	128	11.3%	24	7.7%
Jail only	259	13.0%	220	15.3%	150	13.3%	70	22.6%
Total	1992	100.0%	1438	100.0%	1128	100.0%	310	100.0%

BIP=Batterer intervention program
DOS=Delay of sentence

counseling.

Table 4.9: Distribution of Counseling Referrals Across Subsets of Violence Against Women Database

	Warrants		*Inset*		*First*		*Repeat*	
10 week BIP	277	31.4%	186	28.6%	160	31.7%	26	17.8%
26 week BIP	38	4.3%	34	5.2%	22	4.4%	12	8.2%
Anger management	48	5.4%	9	1.4%	7	1.4%	2	1.4%
None	170	19.3%	125	19.2%	85	16.8%	40	27.4%
Other	348	39.5%	297	45.6%	231	45.7%	66	45.2%
Total	881	100.0%	651	100.0%	505	100.0%	146	100.0%

4.5 Time series statistics

The VAW database contained a number of date variables, which were used to calculate time series of the number of events per day and estimate caseload statistics. Table 4.10 summarizes the means and standard deviations of raw observed time series of rate variables in units of cases per day for all cases in the analysis set, first arrests, and repeat arrests. That is, there are on average 0.9843 warrant requests per day. Along the same lines, Table 4.11 provides a summary of the means and standard deviations for the raw estimates of the stock variables.

Table 4.10: Statistics for Raw Time Series of Rate Variables

(in units of cases per day)

	Inset		*First*		*Repeat*	
	Mean	SD	Mean	SD	Mean	SD
Warrant requests	0.9843	1.1647	0.7721	0.9990	0.2122	0.5090
Denials	0.3970	0.7736	0.3265	0.6648	0.0705	0.2911
Warrant request authorized	0.6037	0.8591	0.4606	0.7066	0.1431	0.4245
Meeting victim	0.3463	0.7007	0.2642	0.5589	0.0821	0.3207
Plea agreement accepted	0.4285	0.8528	0.3326	0.6926	0.0958	0.3594
Dismissed	0.0869	0.3247	0.0630	0.2672	0.0240	0.1739
Trial dates	0.0315	0.2290	0.0274	0.2076	0.0041	0.0640
Referred to 10 week BIS	0.1218	0.4074	0.1054	0.3662	0.0164	0.1424
Referred to 26 week BIS	0.0233	0.1596	0.0151	0.1218	0.0082	0.0903
Referred to anger mgt.	0.0062	0.0866	0.0048	0.0691	0.0014	0.0370
Referred to other counseling	0.1916	0.5615	0.1492	0.4422	0.0424	0.2446
Guilty but no referral	0.0787	0.2960	0.0541	0.2292	0.0246	0.1718
Received delay of sentence	0.1636	0.4937	0.1444	0.4273	0.0192	0.1602
Received jail	0.2047	0.5311	0.1410	0.4079	0.0637	0.2733
Received probation	0.3717	0.7865	0.2957	0.6528	0.0760	0.3190

4.6 Descriptive models

The next step was to develop descriptive models capable of reproducing the general trend in the prosecutor's office domestic violence caseload. That is, they simply attempted to reproduce the behavior pattern as opposed to describing the underlying mechanisms generating the behavior pattern. This was essentially analogous to a large family of statistical modeling approaches where the goal is to develop a mathematical representation of the data. All of the descriptive models were driven

Table 4.11: Statistics for Raw Time Series of Stock Variables

(in units of cases)

	Inset		*First*		*Repeat*	
	Mean	SD	Mean	SD	Mean	SD
Warrants	18.3292	5.9436	13.9432	4.6960	4.3860	2.0514
Cases	72.4887	22.1962	52.0678	14.7212	20.4209	8.4000

with real time series data of warrant requests that had been smoothed using a 120-day delay third order exponential smoothing algorithm.[2]

Each model was then calibrated to *Caseload*. There were several reasons for only calibrating on *Caseload* (as opposed to also including *Plea offers accepted* and *Dismissals*). First, such simple models will be unable to replicate higher order dynamics such as oscillations that are endogenous to the system. This means that there will be a significant amount of error between the actual and simulated time series. So one can either (a) try to optimize the fit on one observed variable and distribute the error to the other observed variables, or (b) optimize the fit on several variables and spread the error across them. The second reason for only calibrating on *Caseloads* is that it is a stock variable and the hypothesized determinant of the dispositions *Plea offers accepted* and *Dismissals,* which are rates. That is, the question in this study is focused on understanding whether or to what extent caseloads affect dispositions. If one has distributed the error between both caseloads and dispositions, then one is likely to underestimate the possible effects that caseload has on dismissals and

[2]A third order exponential smoothing algorithm was used because it is the lowest order smoothing algorithm where peaks are approximately distributed symmetrically (as opposed to asymmetrically) and neighboring points around a peak are weighed more heavily than distant points. The decision to use a 120-day delay was chosen by comparing various delays and looking for shortest delay that clearly identified the general trends. Ultimately, however, the question is not whether the optimal smoothing algorithm was selected, but whether the model reproduces realistic time series behavior over a wide range of inputs generated from a variety of smoothing algorithms and hypothetical scenarios.

plea offers being accepted. Moreover stocks are generally more reliable than rates as observable measures. Thus, calibration was restricted to optimizing the fit to *Caseloads*.

Calibration was achieved by optimizing the fit between the actual smoothed time series of *Caseload* and each descriptive model's simulated time series of *Caseload* using the Powell optimization procedure in Vensim with the original values as the starting point. This procedure tries to optimize the fit between two or more time series by varying the constants over a given range. The degree of fit or payoff function, P, is calculated as the sum squared differences between actual and model values:

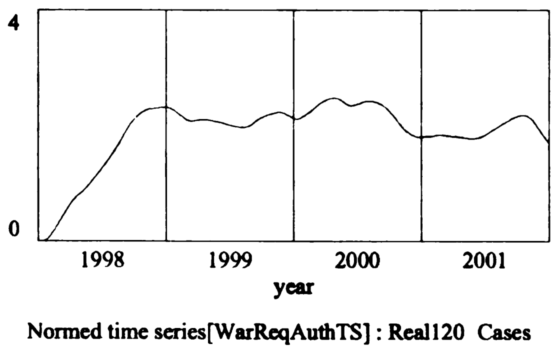
$$P = \sum_{i=1}^{n} -[W_i(M_i(t) - A_i(t))]^2,$$

where n is the number of variables to fit to, W_i is the weight for the i-th variable, $M_i(t)$ is the models value for the i-th variable at time t, and $A_i(t)$ is the actual value for the i-th variable at time t. Weighting the differences between the model and actual values has no impact on the fit when there is only one variable in the payoff function P (i.e., when $n = 1$). However, when the payoff function P is written in terms of more than one variable ($n > 1$), one might decide to assign different weights to each difference because one either wants to (a) weight more reliable measurements of the actual values more heavily than less reliable measurements, or (b) adjust for differences in the scaling of variables.

The constants available for modification will vary from model to model. Constants were added incrementally to the calibration procedure if they improved fit and dropped in favor of estimates from real data in the VAW database if they did not improve fit. The results from calibration were a set of values for each model that optimized the fit between the actual and simulated *Caseload*.

Figure 4.2 shows the smoothed and normed warrant requests data used to drive the descriptive models. *Normed time series* is a subscripted variable and *[WarReqAuthTS]* designates the subscript being referenced, which is in this case warrant requests authorized. *Real120* refers to the name of the real data set used to drive the specific simulation.

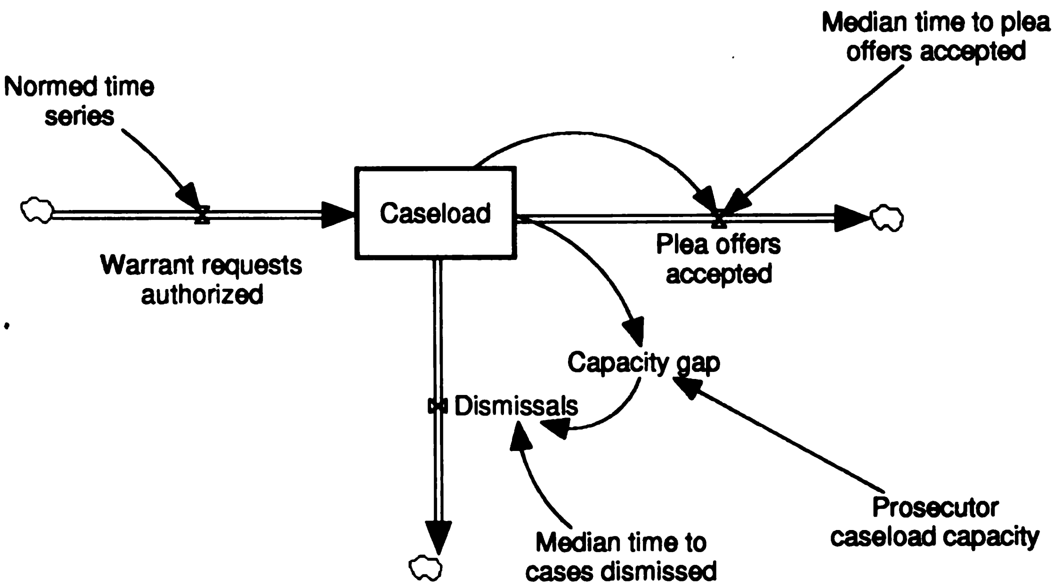
Figure 4.2: Smoothed Time Series of Warrant Requests Generated from Violence Against Women Database



The first model, PACM120 in Figure 4.3, modeled *Plea offers accepted* and *Dismissals* as functions of *Caseload*. *Dismissals* was defined in terms of *Capacity gap*, the difference between some theoretical *Prosecutor caseload capacity* and the actual *Caseload*. Adjustment times were defined in terms of *Median time to plea offers accepted* and *Median time to cases dismissed*. Medians were used at this stage as more robust estimators of central tendency, which was an assumption that would turn out to be problematic in later models.

The results after calibrating PACM120 are shown in Figure 4.4. PACM120 was calibrated by modifying *Prosecutor caseload capacity*. Observed variables (solid line) are shown with PACM120's simulated results (dashed line) for *Caseload*, *Plea*

Figure 4.3: Descriptive Model with Dispositions as a Function of Caseload Pressure, PACM120



offers accepted, and *Dismissals.* The best way to evaluate how well the model performs behaviorally is to visually compare the simulated data against the real data (Sterman, 2000). Traditional fit statistics alone do not say very much about how well the model reproduces the observed behavior because they focus on point-by-point fit, which can be very misleading. For example, all three of the following descriptive models correlated with $r > 0.8$, yet they departed from the observed behavior in noticeable ways. What one wants to do in assessing fit is not just know how much error there is, but the nature of the error. Theil inequality statistics (Sterman, 2000; Theil, 1966) provide one way to assess where the errors are distributed by dividing the error into three components: bias (U^M), unequal variation (U^S), and unequal covariation (U^C). Specifically,

$$
\begin{aligned}
U^M &= \frac{(\overline{M}-\overline{A})^2}{MSE} \\
U^S &= \frac{(S_M - S_A)^2}{MSE} \\
U^C &= \frac{2(1-r)S_M S_A}{MSE}
\end{aligned}
$$

where $\overline{M}$ is the mean of the model's time series, $\overline{A}$ is the mean of the actual time series, S_M is the standard deviation of the model's time series, S_A is the standard deviation of the actual time series, MSE is the mean squared error or $\frac{1}{n}\Sigma(M_i - A_i)^2$, and r is the correlation between the model and actual time series (Theil, 1966). The three components sum to one, that is, $U^M + U^S + U^C = 1$.

Generally speaking, one is most concerned with systematic errors as opposed to random or unsystematic errors. A high U^M statistic indicates bias in a model and systematic error, usually from poorly estimated parameters. A high U^S statistic can indicate systematic error in terms of either the model not reproducing the general trend or having the wrong amplitudes in oscillations. And, a high U^C indicates random or unsystematic error. When $U^S + U^C = 1$, the model has the same mean and trend and is unsystematic unless one is interested in studying the oscillations. The mean absolute error divided by the mean of the observed variable ($MAE/mean$) was used to assess the size of the error. $MAE/mean$ is dimensionless and reported as percentages. A $MAE/mean = 9\%$ indicates that the model had an average error between simulated and actual values of nine percent from the actual data.

PACM120 did very well for a single stock model in terms of reproducing general behavior pattern. Single stock models are inherently limited in their ability to generate more complicated behavior modes. Specifically, single stock variables are unable to generate oscillating behaviors. The size of the error was relatively low, only 9%, most of which was unsystematic, $U^C = 0.97$. PACM120 did not, however, do so well with reproducing the behavior of the rates, *Plea offers accepted* and *Dismissals*. *Plea offers accepted* had a moderately sized error of 16%, but half of that was systematic, $U^M = 0.49$. *Dismissals* had an even larger error of 31%. However, $U^S + U^C = 0.92 \approx 1$. Thus, if one is not looking to study the oscillations (which single stock systems cannot reproduce anyway), then one can consider the error unsystematic even though the size of the error is large.

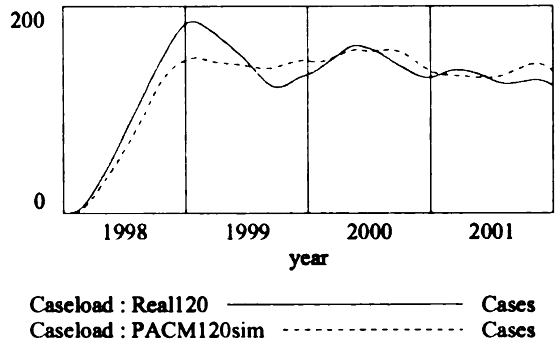
PACM131 in Figure 4.5 modeled *Dismissals* as a function of caseload, the same as PACM120, but *Plea offers accepted* was written as a function of *warrant requests authorized*. This essentially describes the situation where plea offers are accepted on the basis of the attributes of a criminal case, which might be assumed to be stochastic.

PACM131 was calibrated by varying *Proportion of warrants POA* and *Prosecutor caseload capacity*. PACM131 did about the same as PACM120 in reproducing the observed caseload (see Figure 4.6). But PACM131 performed dramatically worse with modeling *Plea offers accepted* and *Dismissals* than PACM120. The amount of error in *Plea offers accepted* increased from 17% to 21% with even more bias than PACM120. The amount of error in *Dismissals* was more than 170% with a most of that error now concentrated as bias and systematic ($U^M = 0.81$).

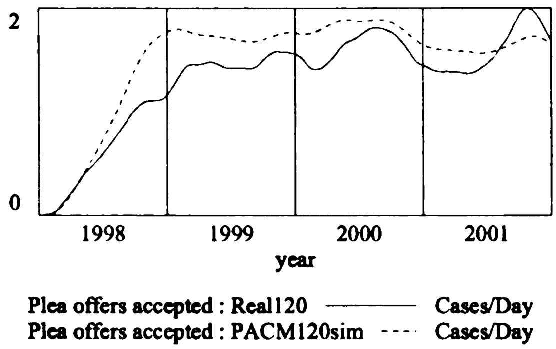
PACM142 made *Dismissals* a function of *Warrant requests authorized* (see Figure 4.7. This would describe the situation where all dispositions from prosecution (plea offers and dismissals) could be described strictly in terms of the attributes of cases. PACM142 was calibrated by varying *Proportion of warrants POA*, *Mean time to plea offers accepted*, *Mean time to cases dismissed*, and *Proportion of warrants dismissed*. PACM142 did about as well as reproducing *Caseload* as PACM120 and PACM130. PACM142 did a much better job of reproducing *Plea offers accepted*, with an average

Figure 4.4: Simulated Results from PACM120

(a) Caseload ($MAE/mean$ = 9%, $U^M = 0.02$, $U^S = 0.00$, $U^C = 0.97$)



(b) Plea offers accepted ($MAE/mean$ = 16%, $U^M = 0.49$, $U^S = 0.01$, $U^C = 0.49$)



(c) Dismissals ($MAE/mean$ = 31%, $U^M = 0.08$, $U^S = 0.16$, $U^C = 0.76$)

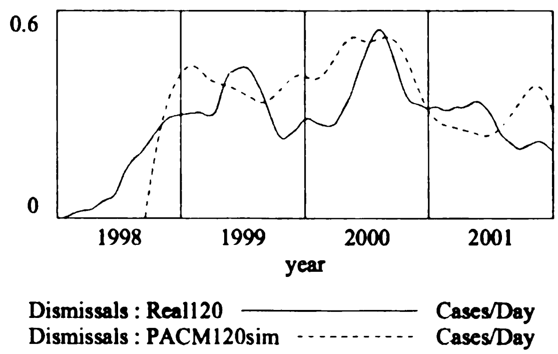
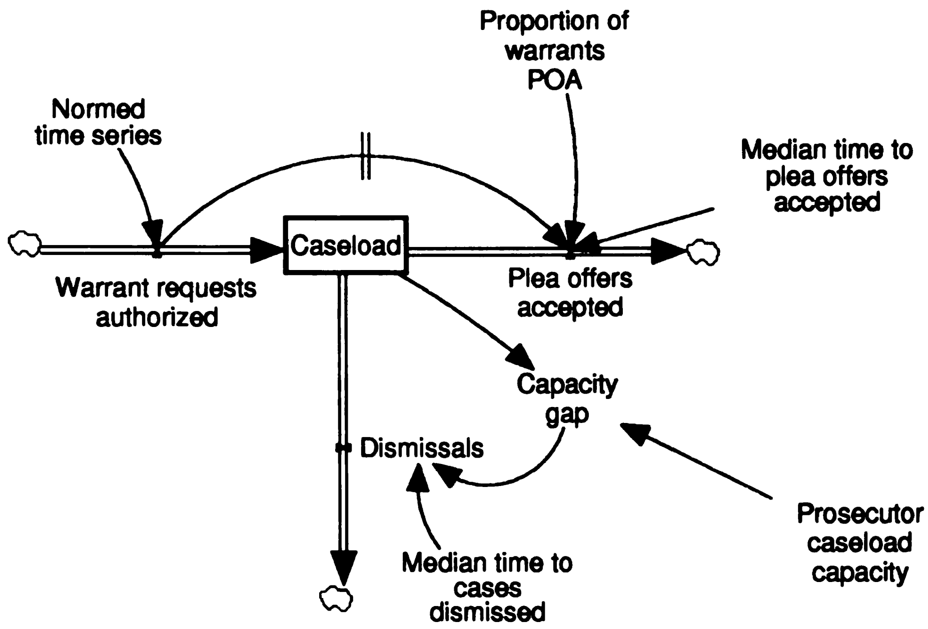


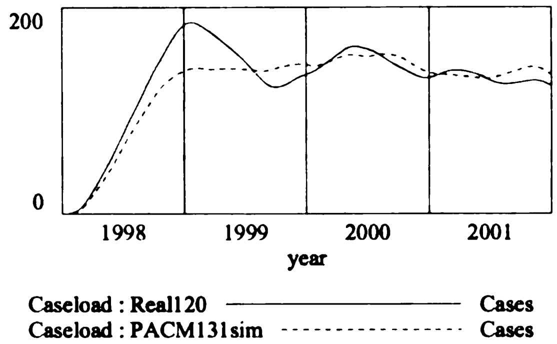
Figure 4.5: Descriptive Model with Dispositions as a Function of Case Attributes and Caseload Pressure, PACM131



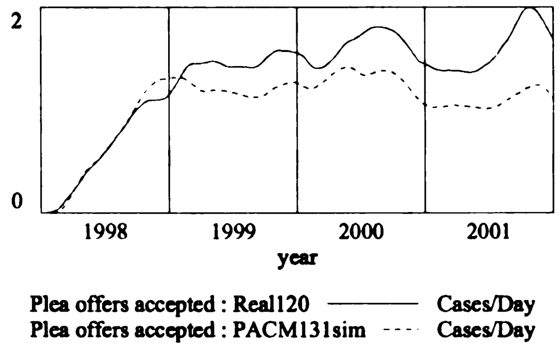
error of only 9% (down from 17% in PACM120 and 21% in PACM130), which was mostly unsystematic. With exception to 2001, PACM142 did an excellent job of reproducing *Plea offers accepted*. It did not, however, do a good job of reproducing *Dismissals* with a mean error of 98% (versus 30% for PACM120 and 171% for PACM130), which was mostly systematic in terms of bias. That PACM142 did better than PACM130 at reproducing *Plea offers accepted* when both formulated *Plea offers accepted* as a function of *Warrant requests authorized* might seem at first surprising. But all three of these models were calibrated against the observed data and errors redistributed to optimize the fit of *Caseload*. The effect of this procedure was to trade error in one variable for another. And, more importantly, making *Dismissals* a function of *Warrant requests authorized* essentially decoupled *Dismissals* from *Caseloads*.

Figure 4.6: Simulated Results from PACM131

(a) Caseload ($MAE/mean = 10\%$, $U^M = 0.08$, $U^S = 0.01$, $U^C = 0.90$)



(b) Plea offers accepted ($MAE/mean = 21\%$, $U^M = 0.56$, $U^S = 0.16$, $U^C = 0.28$)



(c) Dismissals ($MAE/mean = 171\%$, $U^M = 0.81$, $U^S = 0.14$, $U^C = 0.05$)

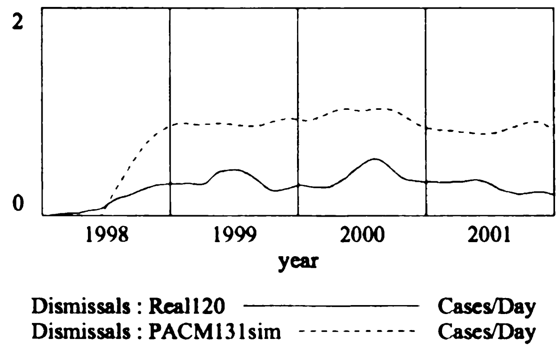
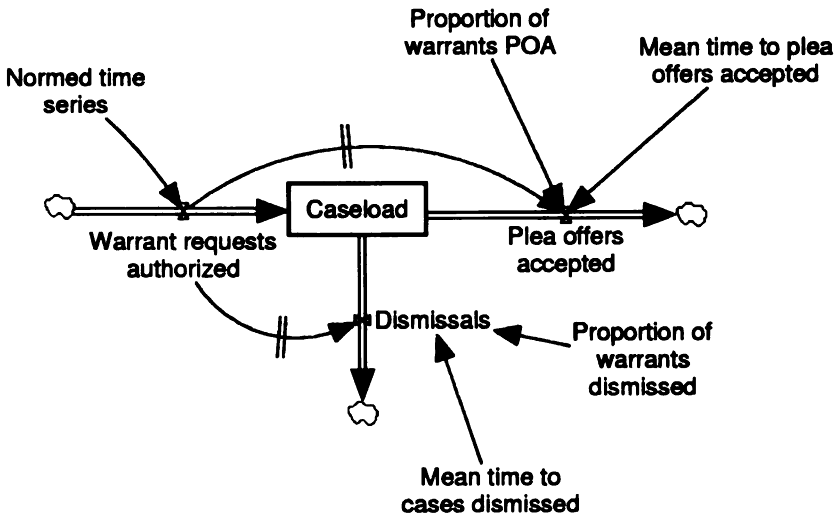


Figure 4.7: Descriptive Model with Dispositions as a Function of Case Attributes, PACM141



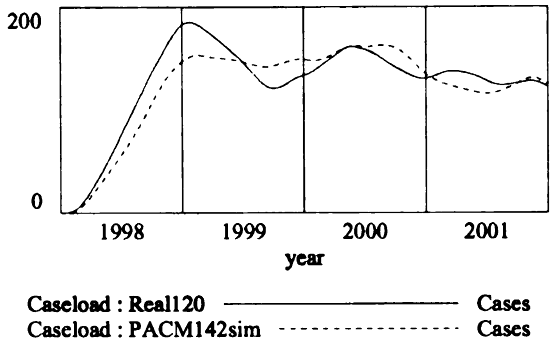
While all three descriptive models performed reasonable well at reproducing *Caseload,* PACM120 did a generally better job of reproducing *Plea offers accepted* and *Dismissals* than PACM130 and PACM142. The implication is that caseloads provide a better *descriptive* account of dispositions than case attributes. However, none of these models provided an account as to *how or why* that might be the case. While system dynamics models can include such highly aggregated relationships, the goal of system dynamics is generally to understand how the structure of the feedback loops might generate the dynamic behavior pattern of interest.

4.7 Baseline model

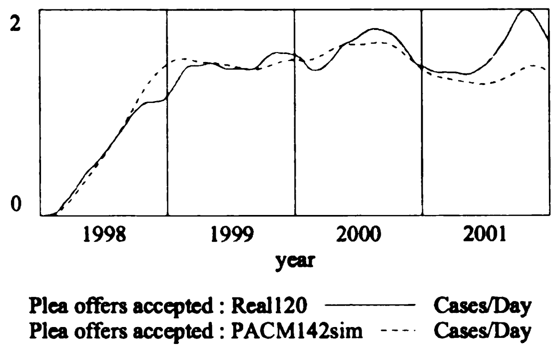
The primary purpose of the baseline model was to develop an initial structure of feedback loops capable of reproducing the primary reference mode. This model was

Figure 4.8: Simulated Results from PACM142

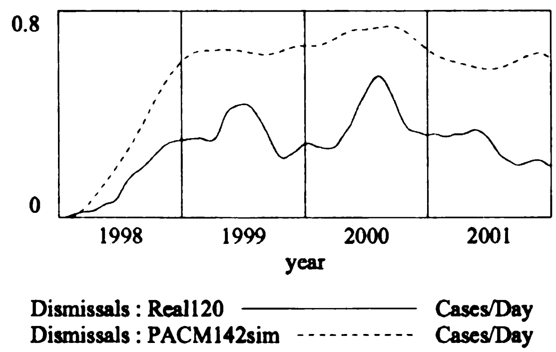
(a) Caseload ($MAE/mean = 10\%$, $U^M = 0.08$, $U^S = 0.00$, $U^C = 0.92$)



(b) Plea offers accepted ($MAE/mean = 9\%$, $U^M = 0.09$, $U^S = 0.05$, $U^C = 0.87$)



(c) Dismissals ($MAE/mean = 98\%$, $U^M = 0.85$, $U^S = 0.07$, $U^C = 0.08$)



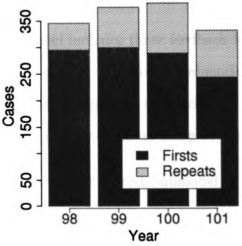
also used to focus questions for interviews with key informants and to further refine the reference modes. The baseline model was developed using a modular approach, starting with the assailant submodel, then the warrant processing submodel, and finally the the case processing submodel. The warrant processing and case processing submodels were initially developed using the real data as exogenous inputs as opposed to the simulated values from the preceding submodel. Thus, the warrant processing submodel was developed using the real arrest data for first time and repeat offenses for inputs, while the case processing submodel was developed using the real authorization rates for first time and repeat offenses for inputs.

4.7.1 Arrest submodel

After examining the caseload by first and repeat arrests, it became evident that repeat arrests represented a significant portion of the prosecutor's office caseload. Figure 4.9 shows the overall trends of first time arrests and repeat arrests. The total height of the bars represents the total number of arrests for that year. The black portion of the bar represents the first time arrests for that year, whereas the gray portion represents the repeat arrests for that year. First time arrests were individuals who did not have a prior record of arrest in the VAW database, while repeat arrests had a prior record of arrest in the VAW database. Since both first arrests and repeat arrests contribute to the caseload, repeat arrests could potentially consume a significant portion of prosecutor's office resources. Moreover, if holding assailants accountable impacts assailants' behaviors, then there would be a feedback loop linking case dispositions with rearrest and cessation rates. Thus, it became clear that one would eventually want to include an arrest submodel as part of the baseline model and disaggregate cases by first time versus repeat arrests or cases.

Arrests are a rate (cases per unit time), while the "supply" of cases that are at

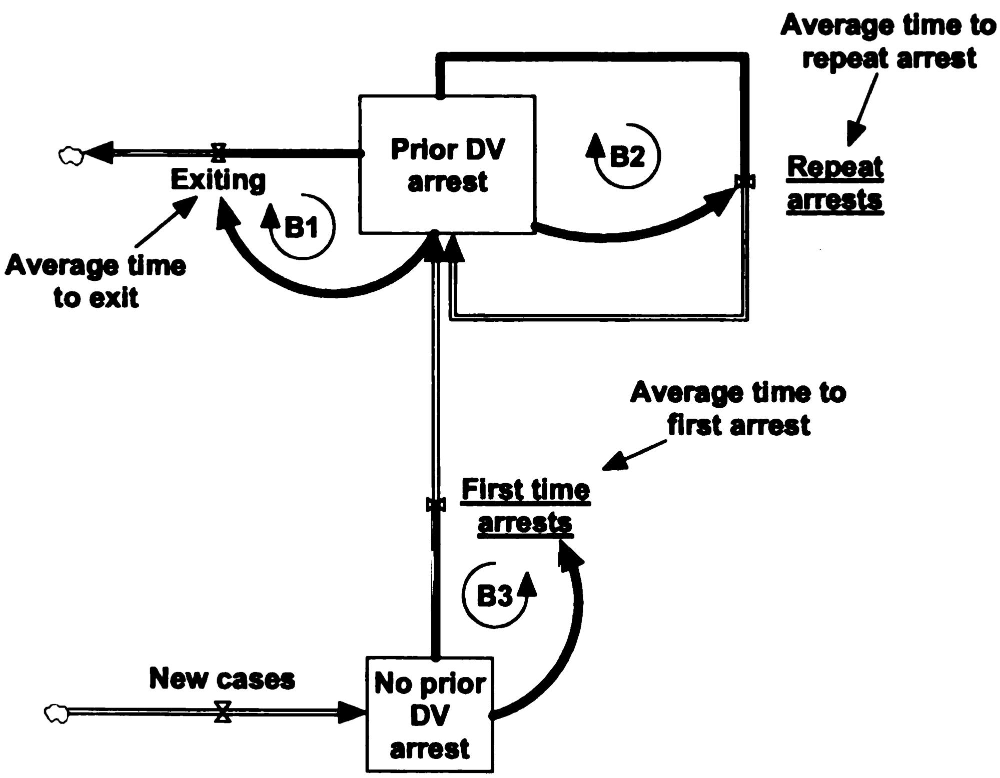
Figure 4.9: Trends in First Time Arrests Versus Repeat Arrests



risk of an arrest is generally limited. That is, arrest rates are constrained by the number of cases committing criminal behaviors. Most arrests in the VAW database are first time arrests (see Figure 4.9). If the number of domestic violence cases at risk of arrest remain constant, then one would expect the sum of first time and repeat arrests to be constant over time. This could be represented as a transfer of cases from first time arrests to repeat arrests, with no net change of cases entering or leaving the system. The basic arrest submodel representing this logic is shown in Figure 4.10, which has one stock of individuals with no prior domestic violence arrest, *No prior DV arrest*, and one stock of assailants with priors, *Prior DV arrest*. *No prior DV arrest* represents cases with no prior domestic violence arrest and at risk of an arrest. *Prior DV arrest* represents cases with prior domestic violence arrests and at risk of repeat arrest. With a January 1, 1998 through December 31, 2001 time horizon, *First time arrests* refers to arrests where the case had no prior domestic violence arrest on or after January 1, 1998, and *Repeat arrests* refers to arrests where the case had a prior domestic violence arrest on or after January 1, 1998.

Although the arrest submodel includes three feedback loops (B1, B2, and B3), it is essentially descriptive in nature. The arrest submodel does not provide any explanation as to how or why individuals are arrested, how their behaviors increase the risk of arrest, or why they stop or continue being at risk of arrest.

Figure 4.10: Arrest Submodel



The arrival or creation of individuals at risk of arrest is formulated as constant flow of cases per day, *New cases*, into *No prior DV arrest*. This represents the creation of cases that are at risk of being arrested but have not yet been arrested.

It is critical to note that the flow and accumulation of cases in the baseline model includes both assailants and victims that were arrested. That is, the baseline model does not disaggregate arrests by assailants and victims of domestic violence.

The time to being arrested after onset of violence is modeled as an exponential distribution, some cases being arrested quickly while others taking much longer. The rate *First time arrests* going out of the stock *No Prior DV Arrests* is written as a function of *No Prior DV Arrests,* forming the balancing feedback loop that moves cases from *No Prior DV Arrests* to *Prior DV Arrest* (B3). Specifically,

$$First\,time\,arrests = \frac{No\,Prior\,DV\,arrest}{Average\,time\,to\,first\,arrest}. \tag{4.1}$$

Some of these cases in the stock *Prior DV Arrest* will be re-arrested at the rate of *Repeat arrests*, which is determined by the the constant *Average time to repeat arrest* and the stock *Prior DV arrest*, forming another balancing feedback loop (B2). The specific equation for *Repeat arrests* is,

$$Repeat\,arrests = \frac{Prior\,DV\,arrest}{Average\,time\,to\,repeat\,arrest}. \tag{4.2}$$

Since cases in *Prior DV arrest* do not change states when there is a repeat arrest, cases moving "out of" *Prior DV arrest* via *Repeat arrests* do not actually leave *Prior DV arrest.* Mathematically, the flow *Repeat arrests* has no net effect on the stock *Prior DV arrest.* While this representation might be confusing from a system dynamics perspective, it is clearer from a non-system dynamics view because both *First time arrests* and *Repeat arrests* are, in the real system, rates of cases moving through the system in units of cases per day. A more typical system dynamics oriented representation of *Repeat arrests* might be to disaggregate *Prior DV arrest* into several stocks, each representing the accumulation of another prior. This is certainly possible, but such a disaggregation would (a) increase the complexity of the model and number of constants that would need to be identified, while also (b) reducing the number of cases available for estimating the constants for the additional stocks. Such a modification might work if one had a longer time horizon. But with only a four year time horizon, the opportunity to observe multiple rearrests is limited. Technically speaking, the cases become heavily censored to the right as one tries to distinguish the number of rearrests. Thus, the representation of *Repeat arrests* as a flow from *Prior DV arrest* to *Prior DV arrest* is the result of a trade-off between the time horizon and level of aggregation.

Some cases will cease to be at risk of another arrest. Some cases will migrate to avoid further prosecution, some assailants might displace their battering tactics from criminal to non-criminal behaviors, a few might die, and others might stop abusing their intimate partners. These cases leave *Prior DV arrest* through the rate *Exiting* that is a function of the *Prior DV arrest* and the *Average time to exit*, which forms another balancing feedback loop (B1). The expression for *Exiting* is similar to the previous equations,

$$Exiting = \frac{Prior\,DV\,arrest}{Average\,time\,to\,exit}. \tag{4.3}$$

Tables 4.12 and 4.13 summarize the variables in the basic assailant submodel.

Table 4.12: Variables in Arrest Submodel

Name	Units	Mean	SD
First time arrests	Cases/day	0.72	0.15
No prior DV	Cases	92	19
Repeat arrests	Cases/day	0.19	0.08
Prior DV arrest	Cases	148	60
Exiting	Cases/day	0.59	0.24

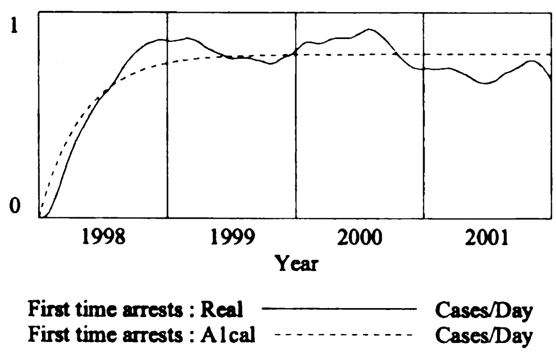
Overall, the behavior of the arrest submodel performs reasonably well in terms of

Table 4.13: Constants in Arrest Submodel

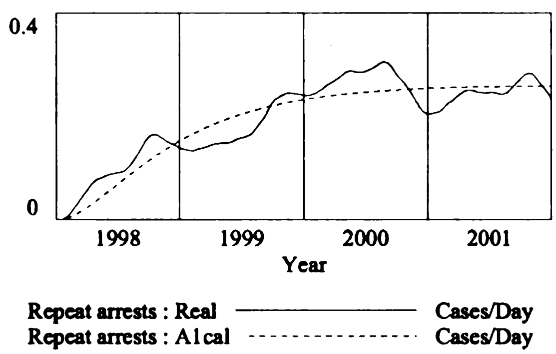
Name	Units	Value
New assailants	Cases/day	0.797
Average time to first arrest	Days	127
Average time to repeat arrest	Days	765
Average time to exit	Days	250

reproducing the general trends in *First time arrests* and *Repeat arrests,* as depicted in Figure 4.11. Table 4.14 summarizes the fit statistics. However, the real data show indications of second or higher order behavior that the arrest submodel is unable to replicate.

Figure 4.11: Simulated Results from Arrest Submodel



(a) First offense arrests



(b) Repeat offense arrests

4.7.2 Warrant processing submodel

The development of the warrant processing submodel started with developing a basic descriptive model and then proceeded with the development of underlying mechanisms for actually processing the warrants. While warrant requests do accumulate,

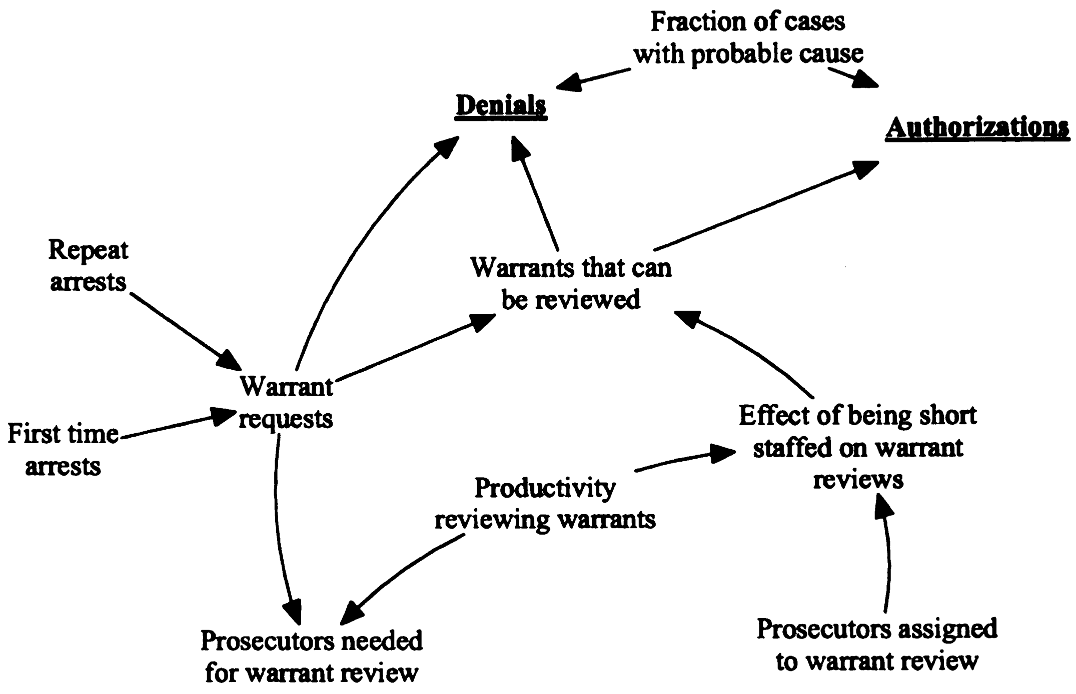
Table 4.14: Summary of Arrest Submodel Fit

Variable	r	$\frac{MAE}{Mean}$	U^M	U^S	U^C	Source(s) of error
First time arrests	0.91	9%	0.00	0.24	0.76	Unsystematic
Repeat arrests	0.93	13%	0.01	0.00	0.99	Unsystematic

decisions on whether to deny or authorize a warrant happen on the order of hours to days, which is very short relative to the four year time horizon of the baseline model. Having stock variables change quickly relative to the time horizon appears to introduce a computational problem for the optimization algorithm in Vensim. Specifically, calibrating a warrant processing submodel formulated as stocks took approximately 20 minutes on a AMD Durron 600 MHz system with 196MB in contrast to less than 1 minute for all the other submodel structures in the baseline model. This could become prohibitive in a revised model with additional structures. Thus a decision was made to treat the variables in the warrant processing submodel as auxiliaries instead of stocks. The resulting submodel is shown in Figure 4.12. There are no stocks in the warrant processing submodel, thus there are no feedback loops within the warrant processing submodel itself.

Warrant requests arrive from *First time arrests* and *Repeat arrests.* Although first time and repeat arrests are likely to differ in any number of dimensions (hence the reason for disaggregating arrests into first time and repeat offenses), all warrant requests are basically processed the same way whether or not the warrant request involves a first or repeat offense. Thus, the basic structure of processing warrants is the same for both first time arrests and repeat arrests. When the basic structure of a model is repeated, it is often easier to understand the model by subscripting variables. Thus, *Warrant requests* is simply a subscripted variable with two elements, *first* and *repeat,* where $Warrant\,requests[first] = First\,time\,arrests$ while

Figure 4.12: Warrant Processing Submodel



$Warrant\,requests[repeat] = Repeat\,arrests$. Real numerical time series for *First time arrests* and *Repeat arrests* were used for the testing and calibration of the warrant processing submodel.

Warrant requests generate a demand for prosecutors to review warrants, *Prosecutors needed for warrant review*, which is formulated as,[3]

$$Prosecutors\,needed\,for\,warrant\,review_i = Warrant\,requests_i / Productivity\,reviewing\,warrants_i.$$

The actual number of warrant requests that can be reviewed is constrained by the number of warrant requests available for review and the resources available for reviewing warrants. This is formulated as,

$$\begin{aligned} Warrants\,that\,can\,be\,reviewed_i = MIN\; & (Warrants\,requests_i,\\ & ZIDZ(Warrant\,requests_i,\\ & SUM_{j \in S}(Warrant\,requests_j)) * \\ & Effect\,of\,being\,short\,staffed\,on\,warrant\,reviews). \end{aligned}$$

Since warrant requests are not accumulated in this submodel, all requests have to be decided upon as they come in. To handle this, *Denials* is formulated as the sum of the cases that do not have probable cause and cases that cannot be reviewed because there are not enough prosecutors. Specifically,

[3]The function *ZIDZ(x,y)* is used instead of standard division where *x/0* would result in a discontinuity. When y is not zero, *ZIDZ(x,y)* equals x/y. When y is zero, *ZIDZ(x,y)* equals zero. The subscript i is an element of the set $S = first, repeat$. The expression $SUM_{j \in S}(\cdot)$ means sum the elements within the brackets over the subscript j.

$$\begin{aligned} \mathit{Denials}_i = \; & (1 - \mathit{Fraction\,of\,cases\,with\,probable\,cause}_i) * \\ & \mathit{Warrants\,that\,can\,be\,reviewed}_i + \\ & (\mathit{Warrant\,requests}_i - \mathit{Warrants\,that\,can\,be\,reviewed}_i). \end{aligned}$$

Authorizations is simply the product of *Fraction of cases with probable cause* and *Warrants that can be reviewed,*

$$\begin{aligned} \mathit{Authorizations}_i = \; & \mathit{Fraction\,of\,cases\,with\,probable\,cause}_i * \\ & \mathit{Warrants\,that\,can\,be\,reviewed}_i. \end{aligned}$$

The *Effect of being short staffed on warrant reviews* is simply the number of cases per day that the *Prosecutors assigned to warrant review* can process, that is,

$$\begin{aligned} \mathit{Effect\,of\,being\,short\,staffed\,on\,warrant\,reviews} = \; & \mathit{Productivity\,reviewing\,warrants} * \\ & \mathit{Prosecutors\,assigned\,to\,warrant\,review}. \end{aligned}$$

Table 4.15 lists the variables while Table 4.16 lists the constants. The resulting behavior of the initial warrant processing submodel is shown in Figure 4.13, where the solid lines show the actual values from the smoothed reference mode and the dashed lines show the simulated values. Figures 4.13a-d show how the calibrated model appears to do an excellent job of replicating the rates variables. And Table 4.20 summarizes the fit statistics, where *First time authorizations* and *Repeat authorizations* have no error since the submodel is being driven by the real data.

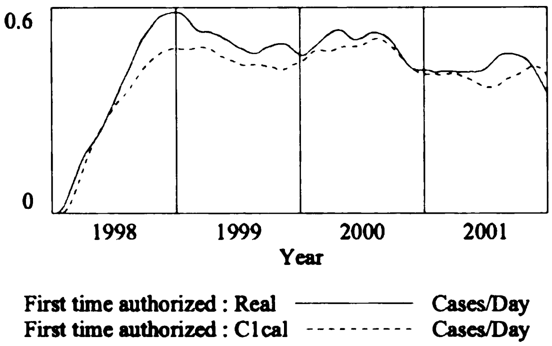
Table 4.15: Variables in Warrant Processing Submodel

Name	Units	Mean	SD
Warrants that can be reviewed[first]	Cases/day	0.73	0.15
Warrants that can be reviewed[repeat]	Cases/day	0.19	0.08
Authorizations[first]	Cases/day	0.40	0.08
Authorizations[repeat]	Cases/day	0.12	0.05
Denials[first]	Cases/day	0.32	0.07
Denials[repeat]	Cases/day	0.07	0.03
Prosecutors assigned to warrant review	Persons	0.52	0.24
Effect of being short staffed on warrant reviews	Cases/day	4.1	1.9
Prosecutors needed for warrant review[first]	Persons	0.09	0.02
Prosecutors needed for warrant review[repeat]	Persons	0.02	0.01

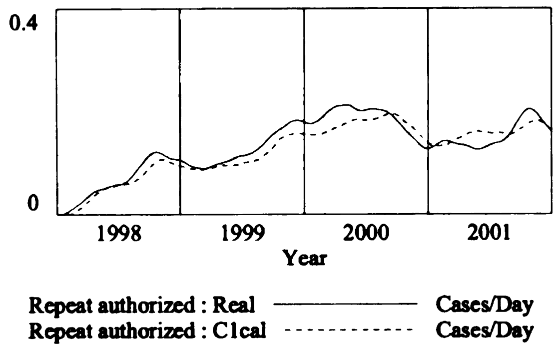
Table 4.16: Constants in Warrant Processing Submodels

Name	Units	Value
Productivity reviewing warrants	Cases/person/day	8
Number of prosecutors	Persons	2
Fraction of cases with probable cause[first]	Dimensionless	0.555
Fraction of cases with probable cause[repeat]	Dimensionless	0.641

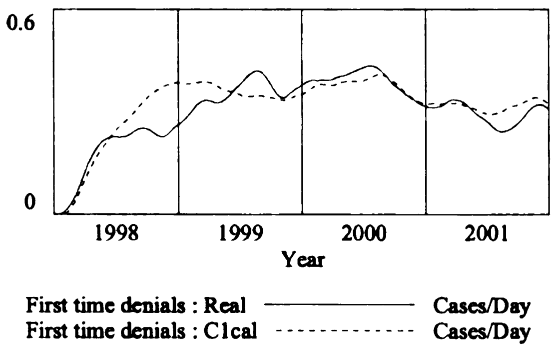
Figure 4.13: Simulated Results from Warrant Processing Submodel



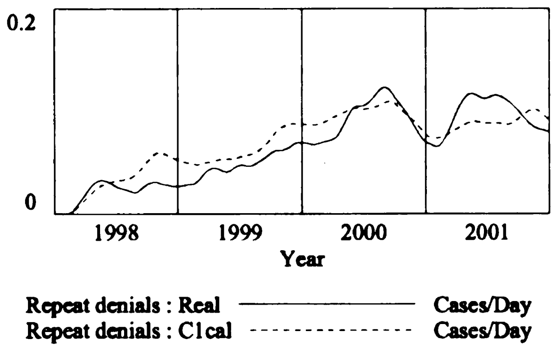
(a) First offense authorizations



(b) Repeat offense authorizations



(c) First offense denials



(d) Repeat offense denials

Table 4.17: Summary of Warrant Submodel Fit

Variable	r	$\frac{MAE}{Mean}$	U^M	U^S	U^C	Source(s) of error
First time arrests	1.00	0%	-	-	-	No error
Repeat arrests	1.00	0%	-	-	-	No error
First time authorized	0.97	9%	0.50	0.18	0.31	Systematic, bias
Repeat authorized	0.96	9%	0.13	0.09	0.78	Unsystematic
First time denials	0.86	12%	0.11	0.05	0.84	Unsystematic
Repeat denials	0.90	22%	0.14	0.18	0.68	Unsystematic

4.7.3 Case processing submodel

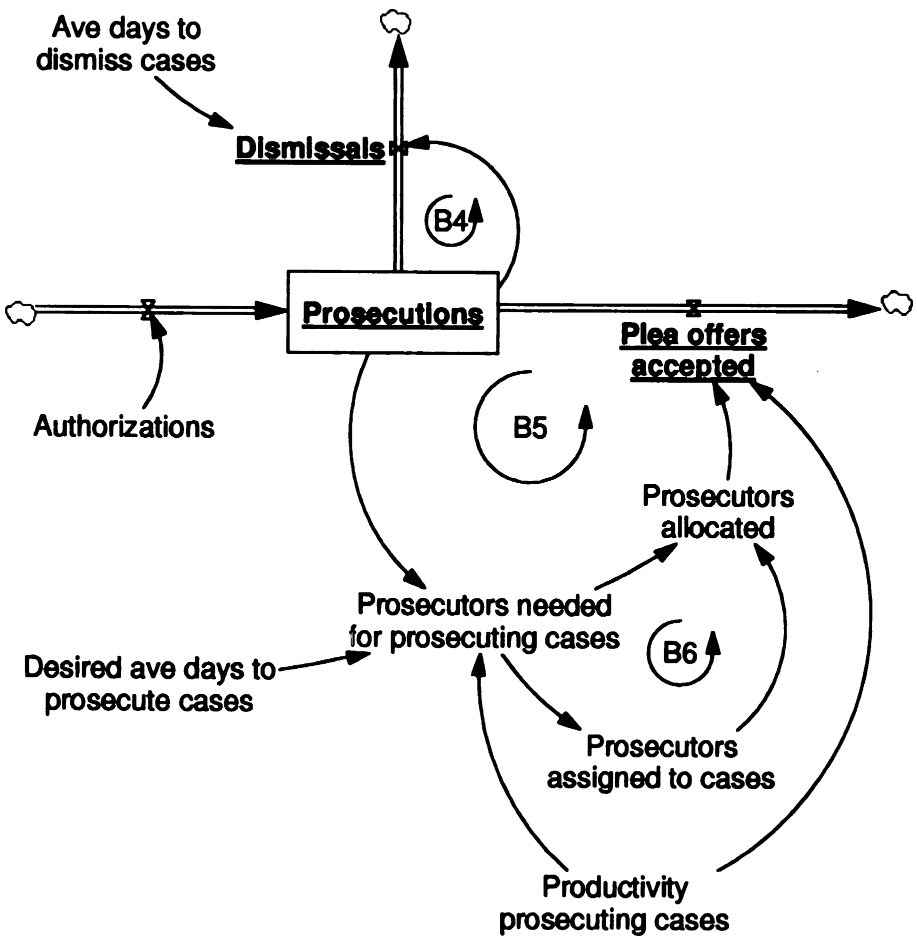
The vast majority of warrant requests authorized result in either a dismissal or plea agreement. Less than 2 percent of authorized warrants actually go to trial (see Table 4.6 on page 77). Thus dispositions in the case processing submodel were limited to plea agreements and dismissals. Prosecuting cases requires resources over time, which must be allocated between the different types of cases. The simplest model of this is to consider the rate that plea offers can be accepted as proportional to the available resources, which is inversely proportional to the number of cases being actively prosecuted. Hence, the time that it takes to negotiate a plea agreement is, unlike the time it takes to review warrants, variable. The basic structure of the case processing submodel is shown in Figure 4.14.

Again, the basic structure is considered the same for first and repeat offenses. Cases that have been authorized enter the stock of cases being actively prosecuted, *Prosecutions.*

Cases that are being actively prosecuted, *Prosecutions*, create a demand on prosecutor's office resources, *Prosecutors needed for prosecuting cases.* The number of *Prosecutors needed for prosecuting cases* is a function of prosecutor's office caseload, *Prosecutors* (cases), divided by the number of cases that can be prosecuted per person per day (cases/person/day), *Productivity prosecuting cases*, and the desired time to successfully prosecute a case (days):

$$Prosecutors\,needed\,for\,prosecuting\,cases = Prosecutions/\left(\frac{Productivity\,prosecuting\,cases}{Desired\,ave\,days\,to\,prosecute\,cases}\right)$$

Figure 4.14: Case Processing Submodel



How that demand is met depends on whether or not there are enough resources. If there are enough resources, then the number of *Prosecutors allocated* applied to *Plea offers accepted* equals the number of *Prosecutors needed for prosecuting cases,* which forms the the balancing feedback loop B5. However, if there are not enough resources, then the number of *Prosecutors allocated* to *Plea offers accepted* is a function of the number of *Prosecutors assigned to cases,* which forms the balancing feedback loop B6.

The actual number of prosecutors allocated, *Prosecutors allocated,* depends on the number needed, *Prosecutors needed for prosecuting cases,* and the number of prosecutors assigned to prosecute cases, *Prosecutors assigned to prosecute cases* (as opposed to reviewing warrants). This is written as,

$$\begin{aligned}
Prosecutors\,allocated_i = MIN\ & (Prosecutors\,assigned\,to\,cases * \\
& ZIDZ(Prosecutors\,needed\,for\,prosecuting\,cases_i, \\
& SUM_{j\in S}(Prosecutors\,needed\,for\,prosecuting\,cases_j)), \\
& Prosecutors\,needed\,for\,prosecuting\,cases_i).
\end{aligned}$$

The actual rate the plea offers are accepted is function of the number of prosecutors allocated to the securing plea agreements and the productivity of prosecutors prosecuting cases. Thus, *Plea offers accepted,* is written as a product of the *Prosecutors allocated* and *Productivity prosecuting cases.* Specifically,

$$\begin{aligned}
Plea\,offers\,accepted_i = \ & Prosecutors\,allocated_i * \\
& Productivity\,prosecuting\,cases.
\end{aligned}$$

Cases that do not result in a plea offer being accepted are dismissed, but up until that point, these cases are still being prosecuted, and hence contribute to *Prosecutors needed for prosecuting cases.* Cases are dismissed via the flow *Dismissals,* which is written as a function of *Prosecutions* and *Ave days to dismiss case*, and forms the feedback loop B4. Specifically,

$$Dismissals_i = \frac{Prosecutions_i}{Average\ days\ to\ dismiss\ case}.$$

Thus, *Dismissals* increases as the number of cases in *Prosecutions* increases and decreases as the *Average days to dismiss case* increases. Table 4.18 summarizes the variables in the case processing submodel and Table 4.19 summarizes the constants.

Table 4.18: Variables in Case Processing Submodel

Name	Units	Mean	SD
Prosecutions[first]	Cases	50	17
Prosecutions[repeat]	Cases	15	7.8
Dismissals[first]	Cases/day	0.03	0.02
Dismissals[repeat]	Cases/day	0.02	0.01
Prosecutors needed for prosecuting cases[first]	Persons	0.31	0.11
Prosecutors needed for prosecuting cases[repeat]	Persons	0.09	0.05
Prosecutors assigned to cases	Persons	1.48	0.25
Prosecutors allocated[first]	Persons	0.31	0.11
Prosecutors allocated[repeat]	Persons	0.09	0.05
Plea offers accepted[first]	Cases/day	0.30	0.10
Plea offers accepted[repeat]	Cases/day	0.09	0.05

Table 4.19: Constants in Case Processing Submodel

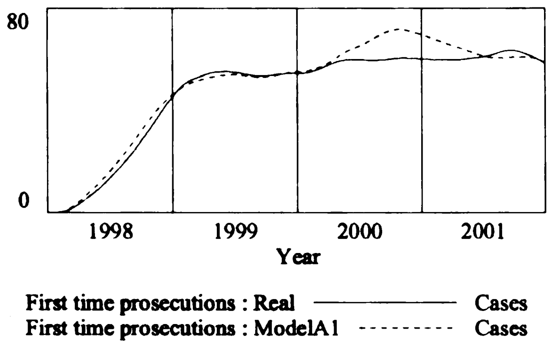
Name	Units	Value
Ave days to dismiss cases	Days	786
Desired ave days to prosecute cases	Days	167
Productivity prosecuting cases	Cases/person/day	0.96

A comparison of the case processing submodel after calibration against the real data is shown in Figure 4.15. The submodel did a better job of reproducing the behavior pattern for first offenses than repeat offenses. With exception of 2000, it did especially well at reproducing the behavior pattern for first time prosecutions (Figure 4.15a). Although it was able to reproduce the general behavior pattern for repeat prosecutions (Figure 4.15b), the simulated behavior pattern appears to lag the real behavior pattern, indicating a problem with a misspecified delay. And it is generally unable to reproduce suggested oscillations in the real data, especially Figure 4.15d through Figure 4.15f. These oscillations in Figures 4.15d through 4.15f do not appear in the real data of warrant authorizations (see Figure 4.19c and 4.19d). That is, the oscillations are not the result of an exogenous input. Nor could they be generated by a simple parametric variation in the case processing submodel since it is a first order system and first order systems are incapable of generating oscillations. Thus, the oscillations must be endogenous to a higher order system, either a more complicated version of the case processing submodel involving two or more stock variables or an interaction between various components of the prosecutor's office caseflow, resources, and information. Table 4.20 summarizes the fit of the case processing submodel. Like the warrant processing submodel, the case processing submodel was developed using real data for input (*Authorizations*, which are broken out as *First time authorized* and *Repeat authorized*).

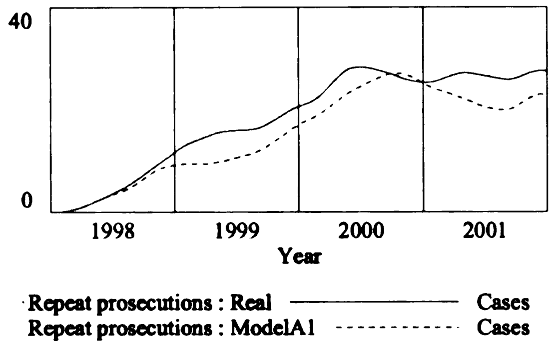
4.7.4 Allocation submodel

Prosecutor resources are distributed in proportion to various resource demands in the allocation submodel (shown in Figure 4.16). The allocation submodel has two main balancing loops, B7 and B8, which distribute prosecutor resources between warrant reviews and prosecuting cases. Loop B7 allocates prosecutors to reviewing

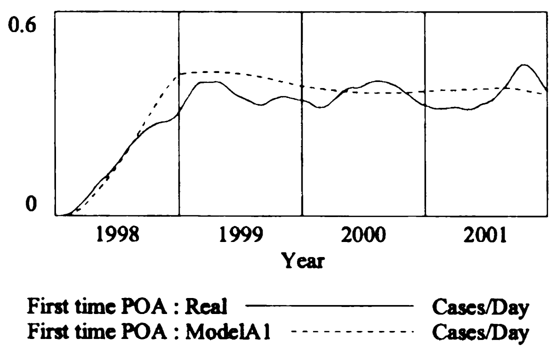
Figure 4.15: Simulated Results from Case Processing Submodel



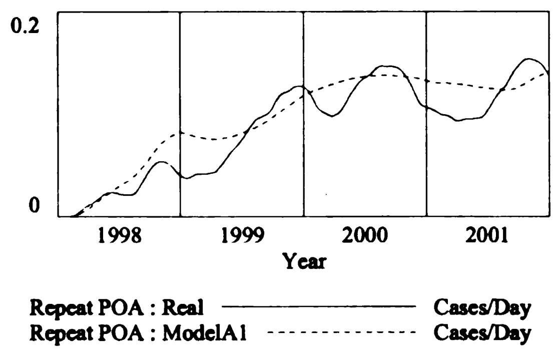
(a) First time cases



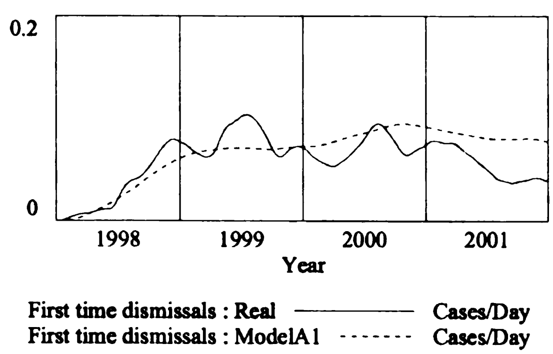
(b) Repeat offense cases



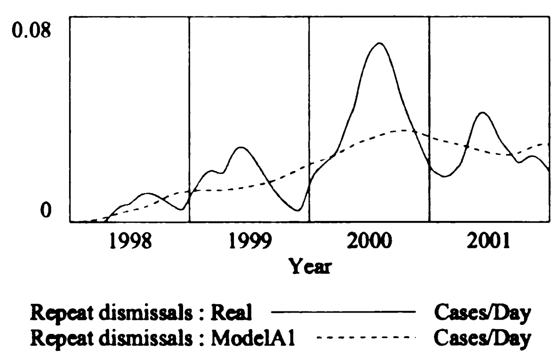
(c) First offense plea aggreements



(d) Repeat offense plea agreements



(e) First offense dismissals



(f) Repeat offense dismissals

Figure 4.16: Allocation Submodel

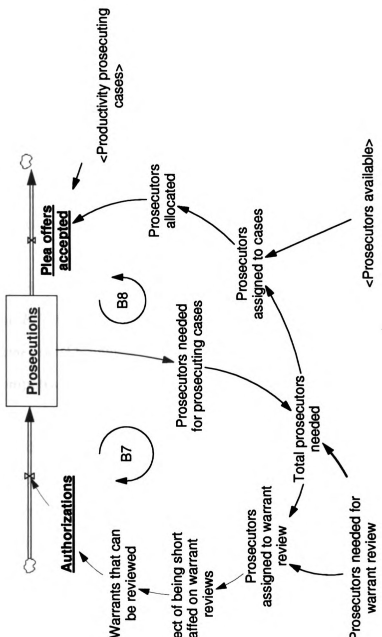


Table 4.20: Summary Case Processing Submodel Fit

Variable	r	$\frac{MAE}{Mean}$	U^M	U^S	U^C	Source(s) of error
First time authorized	1.00	0%	-	-	-	No error
Repeat authorized	1.00	0%	-	-	-	No error
First time prosecutions	0.98	9%	0.41	0.11	0.48	Systematic, bias
Repeat prosecutions	0.98	15%	0.57	0.14	0.42	Systematic, bias
First time POA	0.93	13%	0.13	0.03	0.84	Unsystematic
Repeat POA	0.92	18%	0.15	0.06	0.79	Unsystematic
First time dismissals	0.70	30%	0.14	0.01	0.85	Unsystematic
Repeat dismissals	0.74	37%	0.04	0.21	0.75	Unsystematic

warrants while loop B8 allocates prosecutors to prosecuting cases.

The mechanism described by loop B7 is as follows: As the number of cases increases in *Prosecutions,* the number of *Prosecutors needed for prosecuting cases* increases, which decreases the number of *Prosecutors assigned to warrant review,* lowers the *Effect of being short staffed on warrant reviews* and lowers *Warrants that can be reviewed.* This slows *Authorizations* and ultimately decreases *Prosecutions.*

B8 competes with loop B7 for resources. As there are more cases in *Prosecutions,* there are more *Prosecutors needed for prosecuting cases,* which increases *Total prosecutors* needed. This increases the *Prosecutors assigned to cases* and *Prosecutors allocated,* which increases the rate of *Plea offers accepted.* As *Plea offers accepted* increases, *Prosecutions* decreases.

The *Total prosecutors needed* is the sum of prosecutors needed for reviewing warrants and prosecutors needed for prosecuting cases, which is written as,

$$\begin{aligned} Totaprosecutors\,needed = \;& SUM_{j \in S}(Prosecutors\,needed\,for\,warrant\,review_j)+ \\ & SUM_{j \in S}(Prosecutors\,needed\,for\,prosecuting\,cases_j). \end{aligned}$$

Allocation of prosecutors is formulated to be directly proportional to demand. That is, the larger the demand for prosecutors for a particular activity relative to other activities, the greater the allocation. This allocation scheme does not reflect any prioritization. *Prosecutors assigned to warrant review* and *Prosecutors assigned to cases* written as,

$$\begin{aligned} \textit{Prosecutors assigned to warrant review} = \; & \textit{Prosecutors available} * ZIDZ(\\ & SUM_{j \in S}(\textit{Prosecutors needed for warrant review}_j), \\ & \textit{Total prosecutors needed}), \end{aligned}$$

and

$$\begin{aligned} \textit{Prosecutors assigned to cases} = \; & \textit{Prosecutors available} * \\ & ZIDZ(SUM_{j \in S}(\textit{Prosecutors needed for prosecuting cases}_j), \\ & \textit{Total prosecutors needed}). \end{aligned}$$

The equations for *Effect of being short staffed on warrant reviews* and *Warrants that can be reviewed* have already been covered in Section 4.7.2, and the equation for *Prosecutors allocated* was covered in Section 4.7.3. The only new variable in the allocation submodel is *Total prosecutors needed*, which is summarized in Table 4.21.

Table 4.21: Variables in Allocation Submodel

Name	Units	Mean	SD
Total prosecutors needed	Persons	0.52	0.18

4.7.5 Behavior of baseline model

Figure 4.17 shows a general overview of the resulting baseline model after combining the various submodels. Note that in the overview, feedback loops B5, B6, and B8 all appear to be within the same feedback loop. The formulation for the allocation of prosecutors under various conditions introduced three separate mechanisms, each of which was represented by a feedback loop. At the level of the overview, however, these distinctions disappeared.

The baseline model's behavior is strictly endogenous, meaning behavior is generated entirely from within the model's structure as opposed to being driving by external inputs from real data as was done during the development of the warrant processing and case processing submodels. Some minor recalibration was required to optimize the fit with the real data.

4.7.5.1 Behavior of simulated time series

Figure 4.18 shows the simulated first time and repeat arrest rates against the actual data.

Figure 4.19 shows comparisons of the simulated data for the warrant sector against the corresponding variables from the real data. Although this part of the model does a reasonable job reproducing the overall trend and means, it does not replicate the shorter variations in contrast to response of the warrant submodel's performance with real data for input (see Figure 4.13 on page 103), which is in large part due to using the output from the arrest submodel. Likewise, the simulated data from the prosecution sector reproduces the overall trend and means, but not the shorter variations (see Figure 4.20 on page 117).

With very few feedback loops between the various submodels, one should not expect this baseline model to do more than reproduce the general trends. Table 4.22 summarizes the fit statistics along with an evaluation of the type of error based on the Theil statistics. Overall, the model did a reasonable job of reproducing the means and general trends. There were some problems with *Repeat arrests, First time*

Figure 4.17: Overview of Baseline Model

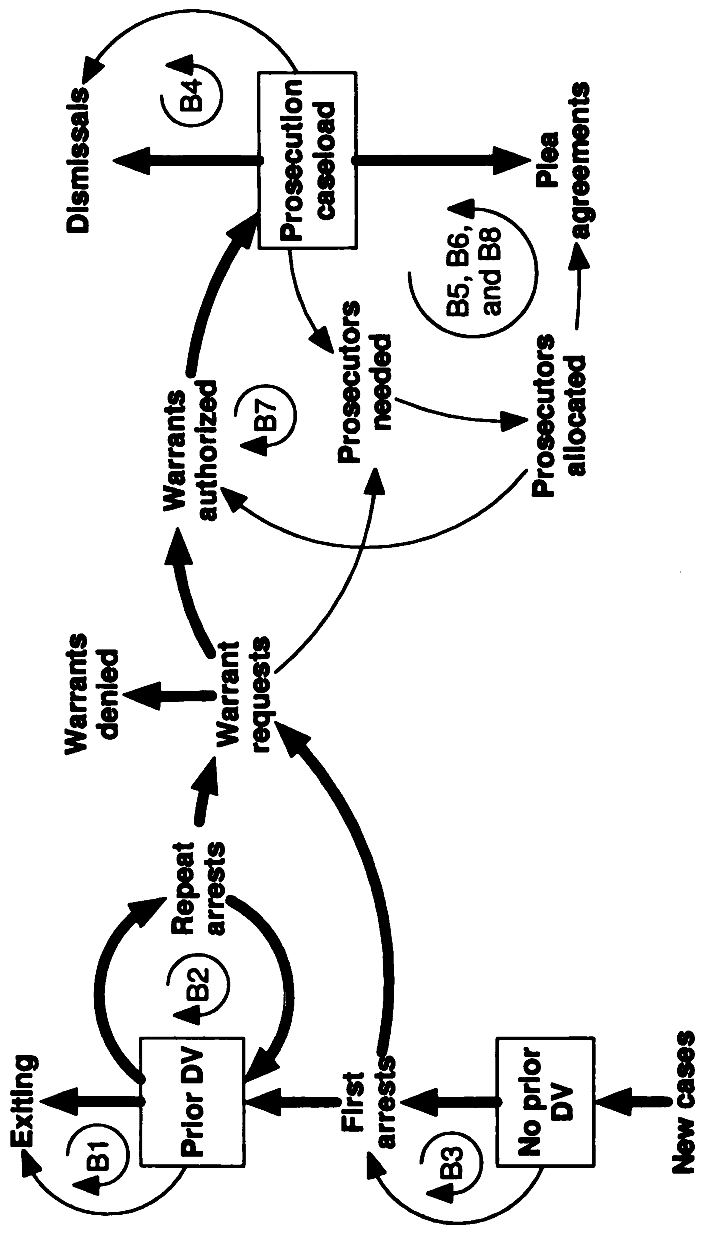
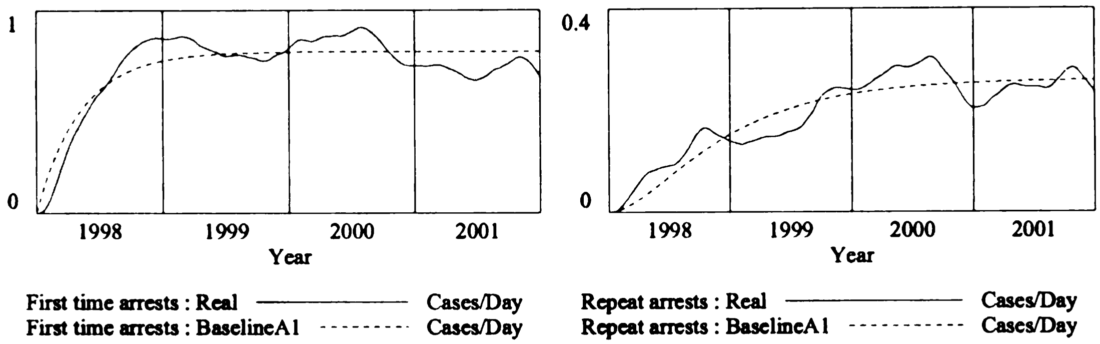


Figure 4.18: Simulated Results of Arrests from Baseline Model



(a) First time arrests

(b) Repeat offense arrests

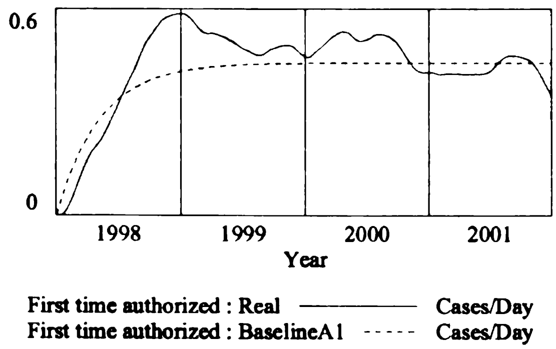
authorized, First time dismissals, and *Repeat dismissals* that need to be investigated further. *Repeat dismissals* is the most troubling because of the large oscillations.

4.7.6 Sensitivity analysis

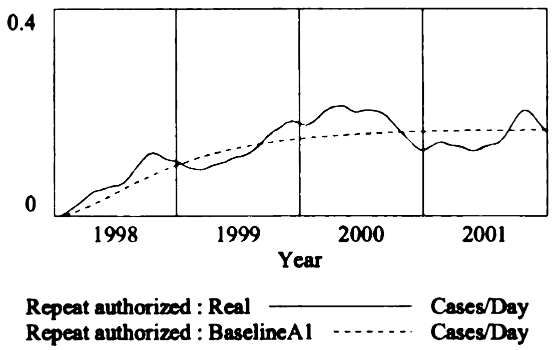
There are three types of sensitivity to variations in parameters and structure that one might consider: numerical, behavior mode, and policy (Richardson & Pugh, 1986; Sterman, 2000). Questions about numerical sensitivity are concerned with the extent specific values from the model vary in response to changes in model parameters or structure, behavior mode sensitivity focus on changes in behavior patterns, and policy sensitivity raise issues about the extent that policy recommendations vary in response to changes in a model.

The structure of the feedback loops in the baseline mode can be divided into two sets that are separated in terms of feedback. The first set contains the arrests, while the second set contains the warrants, prosecutions, and allocation. The two sets are isolated in terms of feedback in the sense that variables in the second set do

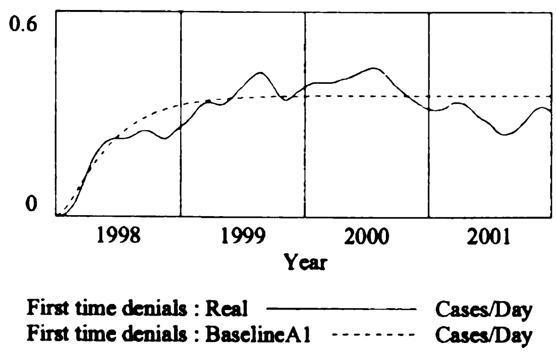
Figure 4.19: Simulated Results of Warrants from Baseline Model



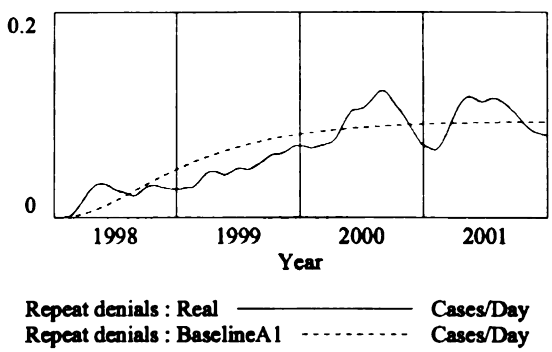
(a) First offense authorizations



(b) Repeat offense authorizations

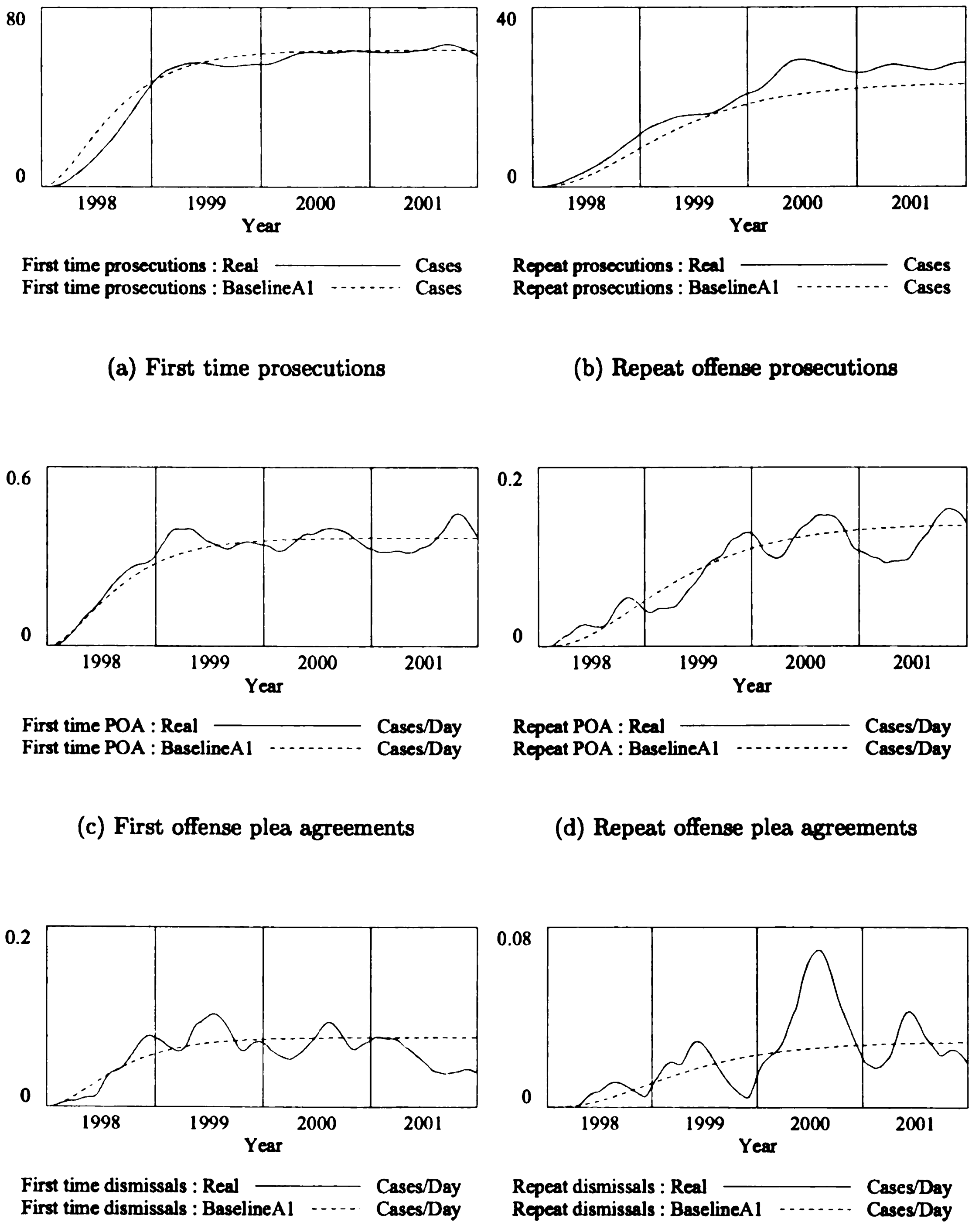


(c) First time denials



(d) Repeat denials

Figure 4.20: Simulated Results of Prosecutions from Baseline Model



(a) First time prosecutions

(b) Repeat offense prosecutions

(c) First offense plea agreements

(d) Repeat offense plea agreements

(e) First time dismissals

(f) Repeat dismissals

Table 4.22: Summary of Baseline Model Fit

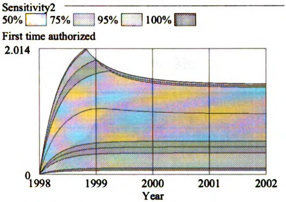
Variable	r	$\frac{MAE}{Mean}$	U^M	U^S	U^C	Source(s) of error
First time arrests	0.91	9%	0.00	0.26	0.74	Unsystematic
Repeat arrests	0.93	13%	0.10	0.00	0.99	Systematic, bias
First time authorized	0.88	14%	0.18	0.34	0.48	Systematic, bias and trend
Repeat authorized	0.87	19%	0.07	0.01	0.91	Unsystematic
First time denials	0.86	16%	0.12	0.28	0.60	Unsystematic
Repeat denials	0.85	26%	0.07	0.12	0.81	Unsystematic
First time prosecutions	0.99	6%	0.31	0.26	0.43	Systematic, bias and trend
Repeat prosecutions	0.99	17%	0.71	0.16	0.12	Systematic, bias
First time POA	0.94	9%	0.05	0.01	0.94	Unsystematic
Repeat POA	0.93	17%	0.03	0.00	0.97	Unsystematic
First time dismissals	0.74	23%	0.07	0.05	0.89	Unsystematic
Repeat dismissals	0.68	39%	0.09	0.29	0.62	Unsystematic

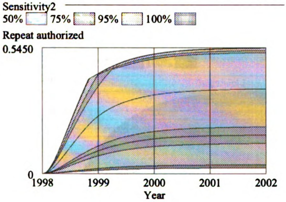
not feedback either information or material flows to the first set. Three sensitivity analyses were conducted. The first analysis considered the sensitivity of observed variables to increases in the number of new cases the system. The second analysis considered the impact of oscillations in the *Average time to first arrest.* And the third analysis examined the impact of varying the number of prosecutors.

4.7.6.1 Sensitivity to the number of new cases

The original value for *New cases* was 0.797 cases per day. Varying *New cases* along a random uniform distribution between ±50% (from 0.39 to 1.20 cases per day) showed no noticeable impact on the qualitative behavior of any of the observed variables. Increasing the range of the random uniform distribution did eventually reveal changes in the qualitative behavior of the baseline model (see Figures 4.21 through 4.25). At *New cases* $\approx$ 2.4, *First time authorized, Repeat authorized, First time denials, First time POA,* and *Repeat POA* all begin to show noticeable changes in their qualitative behavior over time. However, this is nearly three times the base rate of 0.797 cases per day. That is, according to the baseline model, the rate that new cases come into the system can increase by as much as 300% before the system begins to change behavior modes.

Figure 4.21: Sensitivity of Authorizations to Increase in Number of New Cases

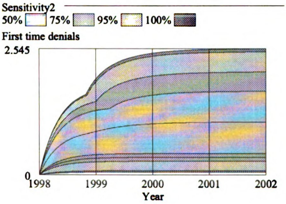




4.7.6.2 Frequency response

Making arrests endogenous to the model with a simple two-stock structure dramatically reduced the complexity of the inputs to the warrant and case processing sectors. One might therefore want to know how the system responds as a whole to

Figure 4.22: Sensitivity of Denials to Increase in Number of New Cases



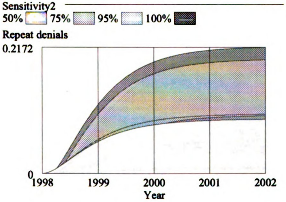
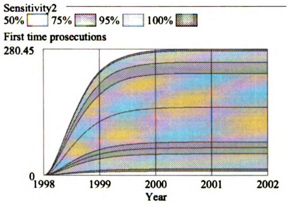


Figure 4.23: Sensitivity of Prosecution Caseload to Increase in Number of New Cases



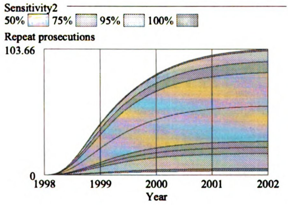


Figure 4.24: Sensitivity of Plea Offers Being Accepted to Increase in Number of New Cases

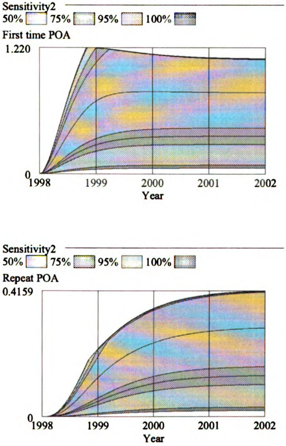
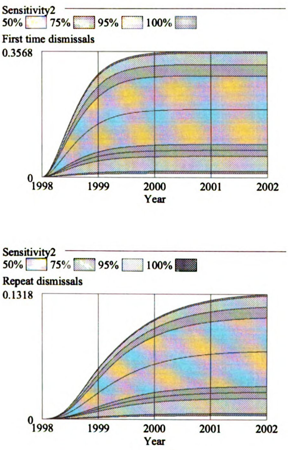
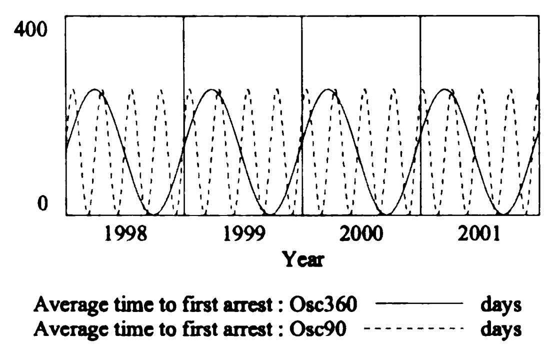


Figure 4.25: Sensitivity of Dismissals to Increase in Number of New Cases



higher frequency oscillations. This can be considered by varying a model parameters in the arrest submodel to simulate a higher frequency oscillation in arrest rates. A simple sinusoidal wave was used to drive the *Average time to first arrest* (see Figure 4.26).

Figure 4.26: Oscillations Used for Studying Frequency Sensitivity of Baseline Model



(a) First time prosecutions

$$\text{Osc360: } Average\,time\,to\,first\,arrest = 200 + sin(t \cdot 2\pi/360)$$
$$\text{Osc90: } Average\,time\,to\,first\,arrest = 200 + sin(t \cdot 2\pi/90)$$

The results of subjecting the baseline model to a 360-day oscillation and 90-day oscillation in the average time to first time arrests are shown in Figures 4.27 and 4.28. The 90-day oscillation (Osc90) had an significant impact on first time arrests resulting in non-differentiable peaks or spikes, which had effects that cascaded through the entire system. These spikes put the model under an extreme test. The baseline model started to show effects of discontinuities under the 90-day oscillation with first and repeat arrest plea offers accepted (Figures 4.28c and 4.28d). This was caused by the sharp peaks in the warrant requests placing a very brief but intense demand on prosecutor resources, temporarily creating a shortage of resources for prosecuting cases and negotiating plea agreements. While these effects were not evident in the effects from the 360-day oscillation on repeat arrests as shown in Figure 4.28c, the 360-day oscillation did create similar discontinuities as the 90-day oscillation for repeat arrests Figure 4.28d. The baseline model clearly has some difficulties handling the extreme conditions of oscillations on the order of 90 days or less.

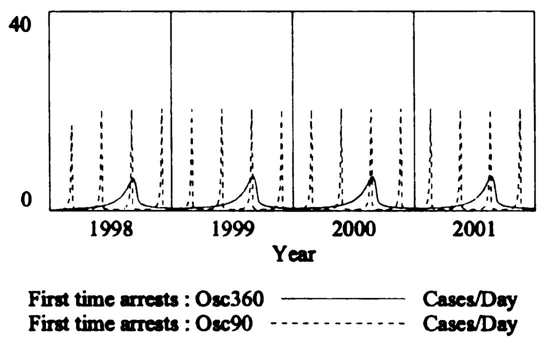
4.7.6.3 Sensitivity to number of prosecutors

The baseline model should show some sensitivity to the number of prosecutors available. To test this, the number of prosecutors available was varied along a uniform random distribution from 0.1 to 4.0 prosecutors. The results are shown in Figures 4.29 through 4.33. Figure 4.29 shows a bifurcation in *First time authorized*, that is, a change in the behavior mode that depends on the number of prosecutors. As the number of prosecutors drops below 0.6, behavior mode of *First time authorized* switches from a strictly increasing function to increasing and then decreasing function. The baseline model did not display any such pattern with *Repeat authorized*. The behavior modes for the other main variables appeared to be relatively robust in that they did not appear to change behavior modes with variation in the number of prosecutors.

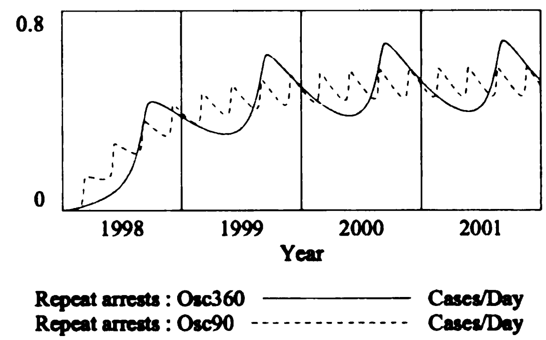
4.7.6.4 Sensitivity to smoothing parameter

The submodels and baseline model were developed and calibrated using time series that were exponentially smoothed with a third order, 120-day delay. To evaluate the sensitivity of the model to assumptions based on the smoothing parameters, five sets of smoothed time series were generated, each with a different delay (30, 60, 90, 120, and 180 days). All five sets were generated using the same basic third order exponential delay.

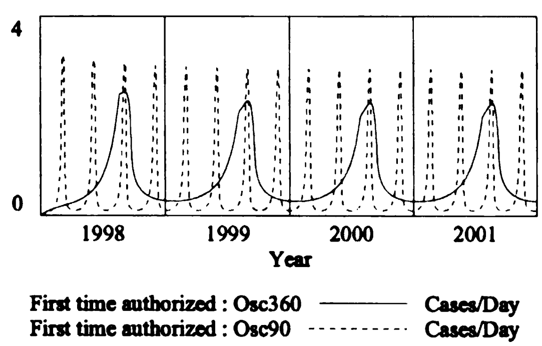
Figure 4.27: Sensitivity of Arrests and Warrants to Frequency



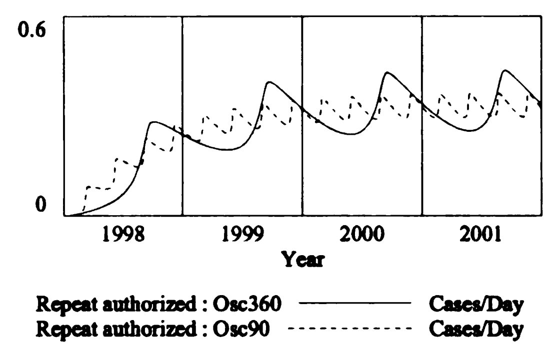
(a) First time arrests



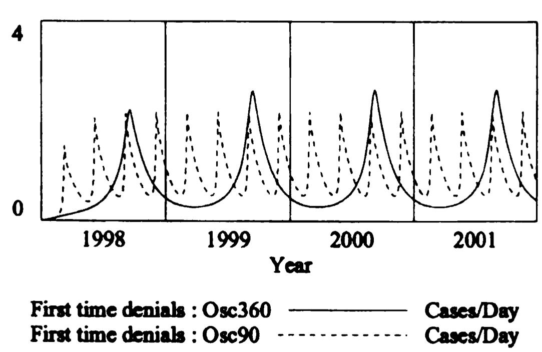
(b) Repeat offense arrests



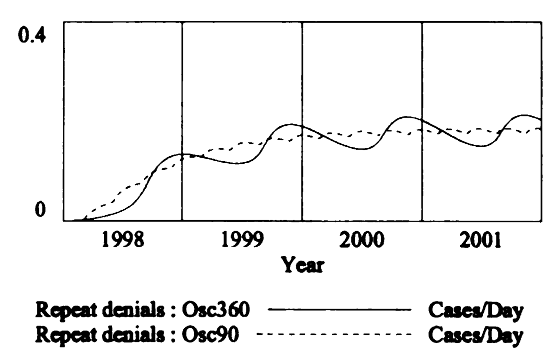
(c) First offense authorizations



(d) Repeat arrest authorizations

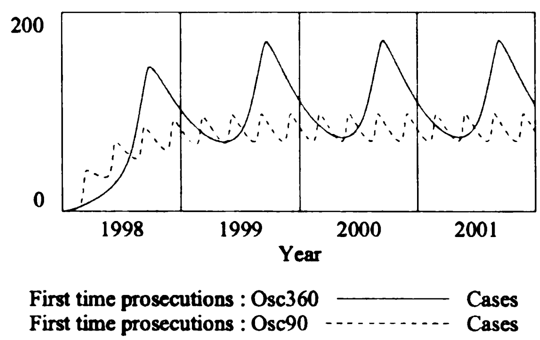


(e) First arrest denials

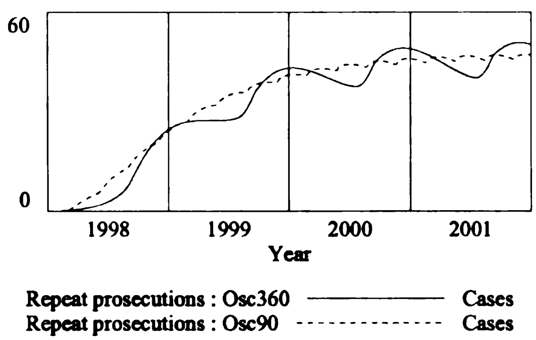


(f) Repeat arrest denials

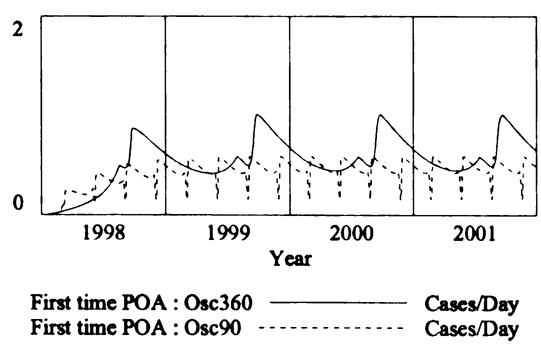
Figure 4.28: Sensitivity of Prosecutions to Frequency



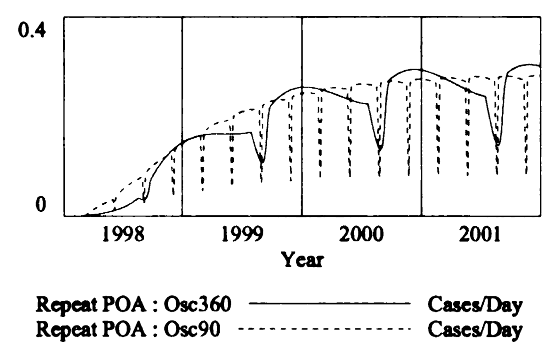
(a) First arrest prosecutions



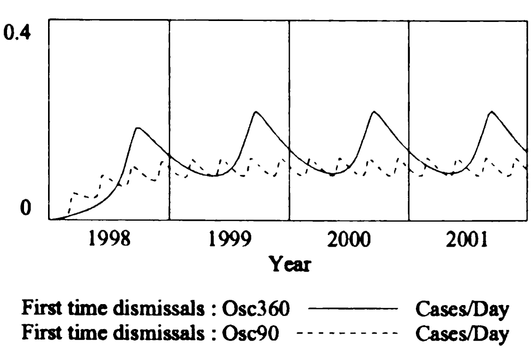
(b) Repeat arrest prosecutions



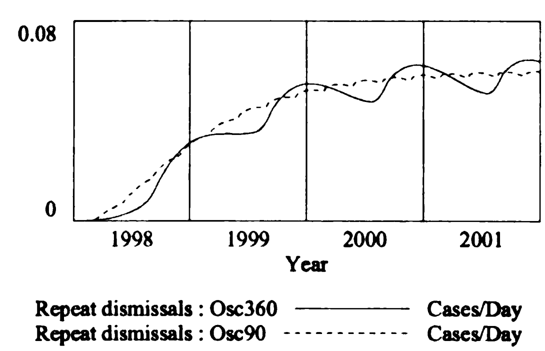
(c) First arrest plea offers accepted



(d) Repeat arrest plea offers accepted

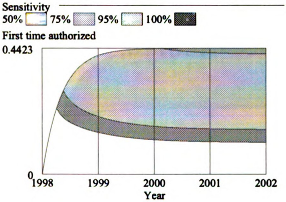


(e) First arrest dismissals



(f) Repeat arrest dismissals

Figure 4.29: Sensitivity of Authorizations to Decrease in Number of Prosecutors



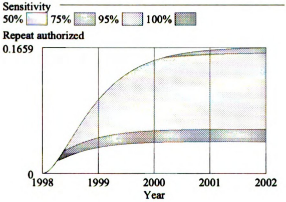


Figure 4.30: Sensitivity of Denials to Decrease in Number of Prosecutors

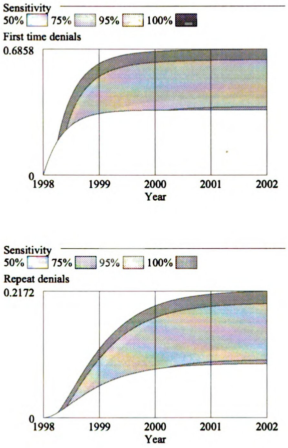
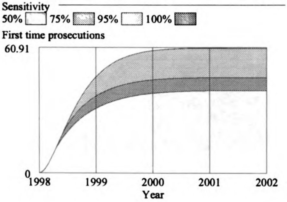


Figure 4.31: Sensitivity of Prosecution Caseload to Decrease in Number of Prosecutors



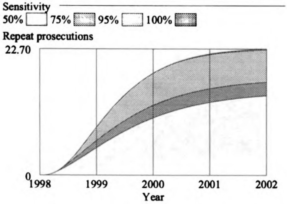
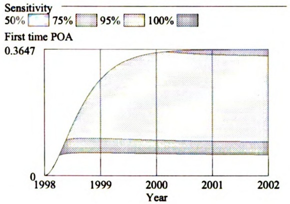


Figure 4.32: Sensitivity of Plea Offers Being Accepted to Decrease in Number of Prosecutors



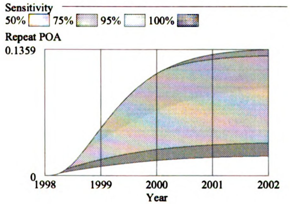
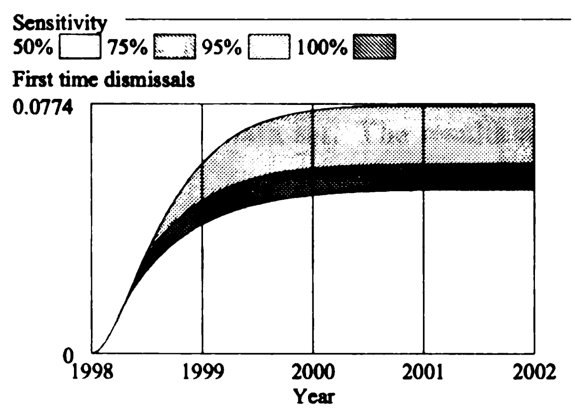
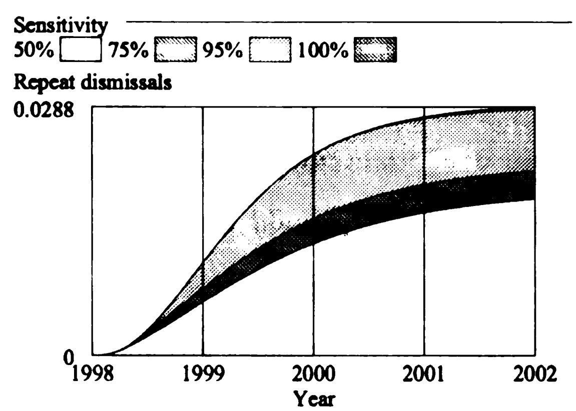


Figure 4.33: Sensitivity of Dismissals to Decrease in Number of Prosecutors





If the baseline model is robust with respect to assumptions about the smoothing parameters, then the baseline model's performance in reproducing the time series behavior would ideally remain constant as the smoothing parameters were varied. Specifically, if real input and output data are smoothed using two different parameters, τ_1 and τ_2, a model that is insensitive to assumptions about smoothing should be able to reproduce the outputs corresponding to the inputs for both τ_1 and τ_2 equally well. To test this, real data sets smoothed using different parameters were used as inputs to the baseline model, and the results were compared against the corresponding variables in the real data set. The results are summarized in Tables 4.23 through 4.27.

As the delay was shortened, the errors in *Authorizations* for both first and repeat arrests increased (7% for first time cases and 11% for repeat arrests), but the systematic component of the error tended to decrease. *Authorizations* was therefore sensitive to the delay in terms of both the size of the error and component of error. Likewise, *Denials* was sensitive to the delay, with relatively large errors for repeat arrests. *Dismissals* for first time arrests showed a pattern similar to *Denials*. *Dismissals* for repeat arrests indicated a dramatic increase in the error, going from 26% to 80%, although the systematic bias was relatively low for delays less than 180 days. Similar results were found for *Plea offers accepted*. However, *Prosecutions* for both first time and repeat offenses appeared to be relatively insensitive to changes in the smoothing parameter. First time arrests kept the error rate to 10% across all five sets, while repeat arrests kept the error rate to between 16 and 17%. Note, however, that *Prosecutions* is a stock and accumulates quantities over time whereas the remaining observed variables are all rates.

Table 4.23: Sensitivity of Authorizations to Smoothing Delays

(a) First Time arrests

Delay (days)	r	$\frac{MAE}{Mean}$	U^M	U^S	U^C
30	0.73	16%	0.12	0.16	0.72
60	0.88	11%	0.25	0.18	0.56
90	0.95	10%	0.40	0.18	0.42
120	0.97	9%	0.50	0.18	0.31
180	0.99	9%	0.62	0.19	0.20

(b) Repeat Arrests

Delay (days)	r	$\frac{MAE}{Mean}$	U^M	U^S	U^C
30	0.86	19%	0.03	0.14	0.82
60	0.93	12%	0.07	0.12	0.81
90	0.95	10%	0.11	0.10	0.79
120	0.96	9%	0.13	0.09	0.78
180	0.98	8%	0.18	0.07	0.75

Table 4.24: Sensitivity of Denials to Smoothing Delays

(a) First Time Arrests

Delay (days)	r	$\frac{MAE}{Mean}$	U^M	U^S	U^C
30	0.55	22%	0.03	0.16	0.81
60	0.69	16%	0.06	0.11	0.83
90	0.80	13%	0.09	0.07	0.84
120	0.86	12%	0.11	0.05	0.84
180	0.92	11%	0.14	0.04	0.82

(b) Repeat Arrests

Delay (days)	r	$\frac{MAE}{Mean}$	U^M	U^S	U^C
30	0.67	42%	0.03	0.22	0.75
60	0.81	29%	0.07	0.21	0.72
90	0.87	25%	0.11	0.20	0.69
120	0.90	22%	0.14	0.18	0.68
180	0.94	19%	0.19	0.14	0.67

Table 4.25: Sensitivity of Dismissals to Smoothing Delays

(a) First Time Arrests

Delay (days)	r	$\frac{MAE}{Mean}$	U^M	U^S	U^C
30	0.40	53%	0.03	0.35	0.62
60	0.61	35%	0.06	0.16	0.78
90	0.74	26%	0.08	0.06	0.86
120	0.82	20%	0.10	0.02	0.88
180	0.91	14%	0.11	0.00	0.89

(b) Repeat Arrests

Delay (days)	r	$\frac{MAE}{Mean}$	U^M	U^S	U^C
30	0.40	80%	0.01	0.54	0.46
60	0.54	57%	0.03	0.45	0.52
90	0.65	45%	0.05	0.40	0.55
120	0.74	37%	0.08	0.36	0.56
180	0.86	26%	0.17	0.33	0.51

Table 4.26: Sensitivity of Plea Offers Accepted to Smoothing Delays

(a) First Time Arrests

Delay (days)	r	$\frac{MAE}{Mean}$	U^M	U^S	U^C
30	0.68	19%	0.00	0.05	0.94
60	0.85	13%	0.02	0.00	0.97
90	0.91	11%	0.04	0.00	0.96
120	0.94	9%	0.04	0.02	0.95
180	0.97	8%	0.03	0.07	0.90

(b) Repeat Arrests

Delay (days)	r	$\frac{MAE}{Mean}$	U^M	U^S	U^C
30	0.73	32%	0.02	0.12	0.86
60	0.86	23%	0.03	0.04	0.93
90	0.90	19%	0.05	0.01	0.94
120	0.93	16%	0.08	0.00	0.92
180	0.97	12%	0.15	0.02	0.83

Table 4.27: Sensitivity of Prosecutions to Smoothing Delays

(a) First Time Arrests

Delay (days)	r	$\frac{MAE}{Mean}$	U^M	U^S	U^C
30	0.94	10%	0.16	0.06	0.78
60	0.95	10%	0.18	0.04	0.78
90	0.96	10%	0.20	0.03	0.76
120	0.97	10%	0.24	0.02	0.74
180	0.98	10%	0.35	0.01	0.65

(b) Repeat Arrests

Delay (days)	r	$\frac{MAE}{Mean}$	U^M	U^S	U^C
30	0.97	17%	0.61	0.18	0.20
60	0.98	17%	0.63	0.20	0.16
90	0.99	17%	0.65	0.22	0.12
120	0.99	17%	0.67	0.25	0.09
180	0.99	16%	0.67	0.29	0.05

4.7.7 Dynamic behavior of feedback loops

The main interest in this study is the dynamic relationship between caseload and accountability, as measured as a proportion of warrant requests that result in plea offers being accepted. The first part of this relationship is to understand how the pattern of feedback loop dominance changes over time with respect to prosecutor's office caseload, *Prosecutions*. This first result provides a dynamic account of how, according to the baseline model, the prosecutor's office manages its domestic violence caseload. The second part of this relationship is to identify the pattern of feedback loop dominance over time with respect to the proportion of cases that are held accountable, specifically:

$$Percent\,held\,accountable \;=\; \frac{\sum Plea\,offers\,accepted_i}{\sum Warrant\,requests_i}. \tag{4.4}$$

This second result shows how accountability is dynamically related to caseload.

Table 4.28 summarizes the major feedback loops in the baseline model. All eight feedback loops are balancing or negative feedback loops, as indicated by the "B" prefix. That is, there are no reinforcing or positive feedback loops (which would be denoted with a "R" prefix) in the baseline model. Feedback loops B4 through B8 are subscripted. For example, there is one B4 feedback loop for cases involving first time arrests and another feedback loop for cases involving repeat arrests. Identifying patterns of feedback loop dominance therefore means identifying which of these feedback loops (individually or in various combinations) dominate the behavior of a given variable of interest during a given time interval.

Analyzing a model in terms of feedback loop dominance has typically been done through various ad hoc testing procedures where the gains of loops are modified in order to identify changes in behavior patterns. Such approaches, while insightful, can miss more subtle dynamics, especially when it comes to identifying shadow feedback loops. A more formal approach involves systematically testing whether or not and in what combinations various feedback loops change the behavior pattern of a specific variable during a given interval of time.

Time series can be decomposed into three basic or atomic behavior patterns (Ford, 1999): linear, exponential, and logarithmic. A system in equilibrium is a special case of a linear behavior pattern. An exponential behavior pattern corresponds to a variable that is increasing or decreasing exponentially. That is, not only is the variable increasing or decreasing, but the *rate* that it is increasing or decreas-

Table 4.28: Major Feedback Loops in Baseline Model

Loop	Description	Subscripts (i)	Variables
B1	Cases exiting stock of cases with prior DV arrest	-	*Prior DV arrest, Exiting*
B2	Repeat arrests	-	*Prior DV arrest, Repeat arrests*
B3	First time arrests	-	*No prior DV arrest, First time arrests*
B4	Dismissals	first, repeat	*Prosecutions, Dismissals*
B5	Prosecutions with enough staff	first, repeat	*Prosecutions, Prosecutors needed for prosecuting cases, Prosecutors allocated, Plea offers accepted*
B6	Prosecutions when short staffed	first, repeat	*Prosecutions, Prosecutors needed for prosecuting cases, Prosecutors assigned to cases, Prosecutors allocated, Plea offers accepted*
B7	Allocation of prosecutors to authorizing warrants	first, repeat	*Prosecutions, Prosecutors needed for prosecuting cases, Total prosecutors needed, Prosecutors assigned to warrant review, Effect of being short staffed on warrant reviews, Warrants that can be reviewed, Authorizations*
B8	Allocation of prosecutors to prosecuting cases	first, repeat	*Prosecutions, Prosecutors needed for prosecuting cases, Total prosecutors needed, Prosecutors assigned to cases, Prosecutors allocated, Plea offers accepted*

ing is *increasing*. Finally, a logarithmic behavior pattern indicates a variable that is increasing or decreasing, but the *rate* that the variable is increasing or decreasing is *decreasing*. These atomic behavior patterns can be identified by looking at the sign of the atomic behavior pattern index (ABPI), $\partial(|\partial x/\partial t|)/\partial t$. Specifically,

$$\begin{array}{ll} \text{linear} & \partial(|\partial x/\partial t|)/\partial t = 0, \\ \text{exponential} & \partial(|\partial x/\partial t|)/\partial t > 0, \\ \text{logarithmic} & \partial(|\partial x/\partial t|)/\partial t < 0. \end{array}$$

Ford (1999) proposes a formal procedure of testing a model's feedback loops for identifying not only the set of dominant feedback loop, but also shadow feedback loops. The advantage of this procedure is that it leads to a complete and formally defensible understanding of the dynamic relationship between a specific variable during a given period of interest in terms of one or more feedback loops. The main disadvantage is that the the number of steps involved with the procedure grows "exponentially" with the number of feedback loops and variables to consider.

Since the differential equations for the baseline model are solved numerically as simulations, the derivatives (e.g. $\partial x/\partial t$) used for identifying the atomic behavior pattern need to be approximated. Specifically,

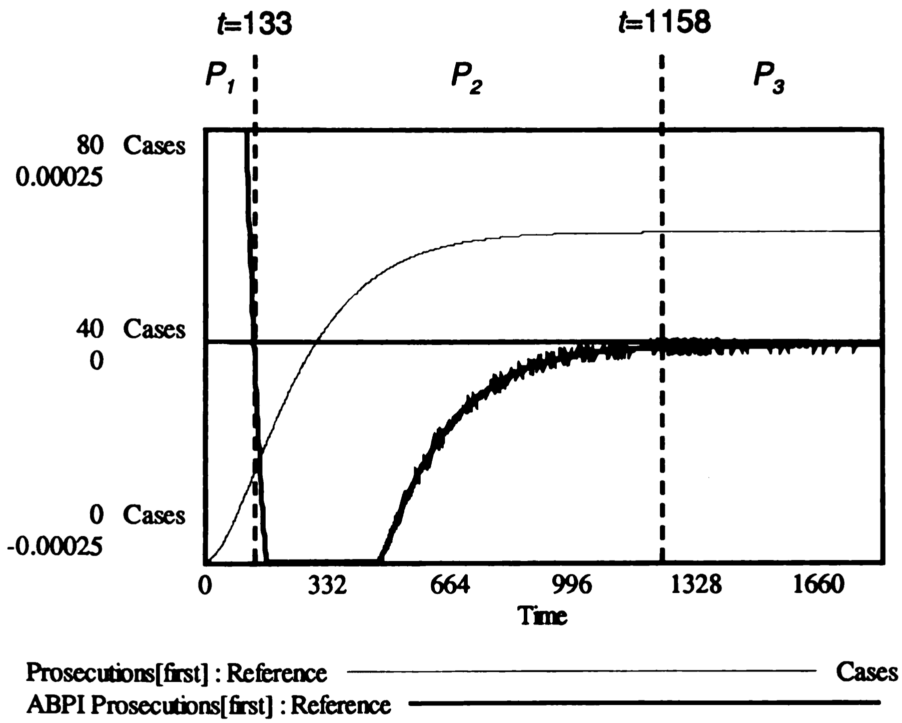
$$\frac{\partial x(t)}{\partial t} \approx \frac{\Delta x(t)}{\Delta t} = \frac{x_t - x_{t-\Delta t}}{\Delta t}.$$

One consequence of approximating the derivatives in this way is that two time points are required to estimate the second order derivative, which means the first value after a transition should be ignored. A second consequence is that the approximation will be sensitive to numerical integration errors. If the numerical integration errors are random or oscillate, then the approximated ABPI will also tend to oscillate. Generally, this is not a problem when the ABPI indicates an exponential or logarithmic behavior pattern. However, when the ABPI approaches zero, these oscillations can make it difficult to identify and locate the start of a linear behavior patterns. A linear behavior pattern was assumed in cases where (a) the oscillations of the ABPI were centered around zero, and (b) the amplitude of oscillations decreased over time. The start of such linear behavior patterns was taken as the first crossing of the horizontal axis.

4.7.7.1 Dynamics of managing prosecution caseloads

Figure 4.34 shows the reference atomic behavior pattern index for *Prosecutions* of first arrests (thick line) along with the plot of *Prosecutions* of first arrests (thin line). The atomic behavior pattern changes whenever the ABPI reaches or crosses the y-axis. To aid the identification of the specific time intervals in which the atomic behavior pattern is uniform, the horizontal axis is now in units of days, not years. To facilitate better examination of potential equilibrium conditions, the time horizon or horizontal axis has also been extended by one year, from 1461 days to 1826 days. The vertical dashed lines at $t = 133$ and $t = 1158$ indicate times when the ABPI crosses the y-axis. The corresponding intervals between these times of uniform atomic behavior pattern are labeled P_1, P_2, and P_3. P_1 corresponds to the situation where there is an increase in arrests when the caseload is empty. P_2 corresponds to the situation where the increase in arrests continues and prosecutor's office caseloads increases in response, along with an increase in cases being disposed through dismissals and guilty pleas. And P_3 corresponds to the case where prosecutor's office caseload has reached an equilibrium, that is the rate that arrests lead to new cases on the prosecutor's offices caseload matches the rate that cases are disposed through dismissals and plea agreements. The question is then to identify the immediate feedback loops that dominate the behavior of *Prosecutions* of first arrests during each of these intervals. *Prosecutions* has five candidate feedback loops to consider: B4 through B8 (see Table 4.28 on page 138).

Figure 4.34: Reference Atomic Behavior Pattern Index (ABPI) for First Arrest Prosecutions



During P_1 (from 0 to 133 days), feedback loop B7 dominates the behavior of *Prosecutions* of first arrests with a positive ABPI, indicating an exponential atomic behavior pattern. That is, the prosecutor's caseload for first arrests is growing exponentially for first arrests and this is driven by B7, the allocation of prosecutors to authorizing warrants. This makes sense since initially there are no cases to prosecute and the prosecutor's office domestic violence caseload is primarily focused on authorizing warrants.

The behavior of *Prosecutions* of first arrests changes from growing exponentially to growing logarithmically in P_2 (from 134 to 1158 days). No single feedback loop dominates *Prosecutions* during this time period. Instead, three pairs of shadow feedback loops dominate the behavior pattern of *Prosecutions*: B4 and B5; B4 and B6; and, B4 and B8. The B4 feedback loop is dismissals. The feedback loop B5 describes the prosecution of cases when there are enough resources to meet all the demands of the current prosecution caseload. B6 describes the prosecution of cases when there are not enough prosecutors to meet the demands of the current caseload. And, B8 describes the allocation of prosecutors to the prosecution of cases. This means that no single feedback loop dominates *Prosecutions* during P_2. Absent from these pairs is B7, which allocates prosecutors to the authorizing of warrants. This can be interpreted as a shift in focus of activity from reviewing warrants to prosecuting cases.

That each pair of shadow feedback loops dominate the behavior of *Prosecutions* of first arrests during this period means that a change in any one pair of shadow feedback loops can change the behavior pattern of *Prosecutions.* That is, one has to modify both B4 and one of the other pairs in order to see a fundamentally different behavior mode. For example, if one wanted to control prosecutor's office caseload during this period by modifying the structure of prosecutions in terms of feedback loops, one would have to modify both the dismissal mechanism (B4) and one additional feedback loop (B5, B6, or B8). Modifying only one feedback loop will have no effect on the qualitative behavior of the *Prosecutions.*

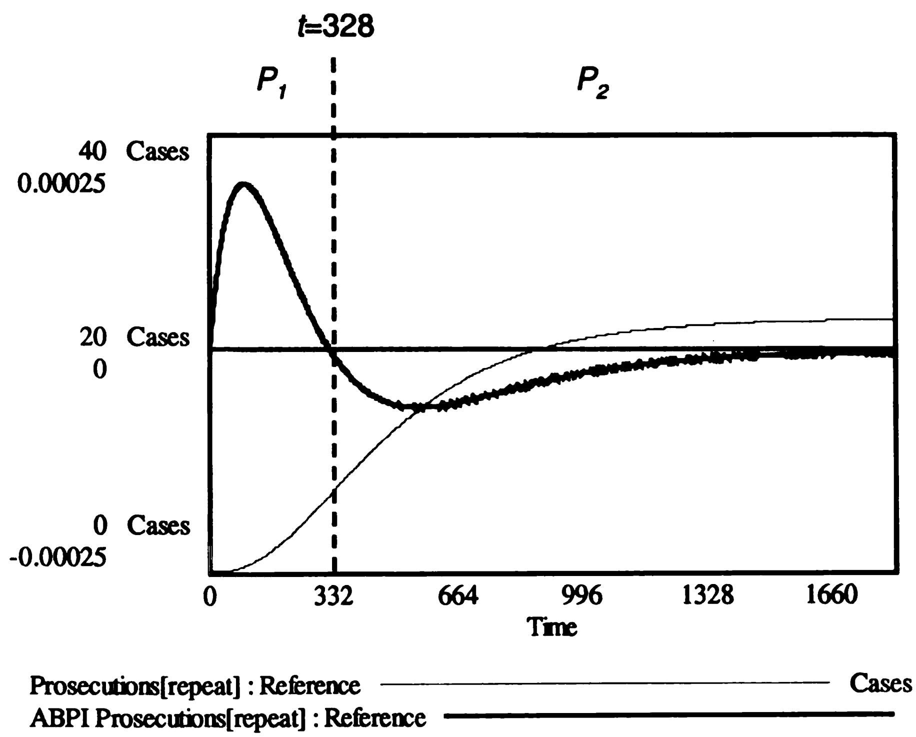
At $t = 1158$ days, the behavior pattern of *Prosecutions* of first arrests changes from logarithmic to linear in P_3 (from 1159 to 1826 days) as *Prosecutions* of first arrests reaches equilibrium. During P_3, all five feedback loops (B4, B5, B6, B7, and B8) dominate *Prosecutions* of first arrest, and a change in any one mechanism will result in either an exponential or logarithmic behavior pattern.

There are only two phases for *Prosecutions* of repeat arrests (see Figure 4.35 on page 144): P_1 (from 0 to 328) where *Prosecutions* displays an exponential behavior pattern, and P_2 (from 329 to 1826 days) where *Prosecutions* displays a logarithmic behavior pattern. Like the prosecution of first arrests, B7 dominates *Prosecutions* of repeat arrests during P_1. That is, the prosecutor's office caseload of repeat arrests grows exponentially and is primarily controlled by the allocation of prosecutor resources to authorizing warrants. Similarly, during P_2, there are four pairs of shadow feedback loops dominating *Prosecutions* of repeat arrests: B4 and B5; B4 and B6; and, B4 and B8. That *Prosecutions* for repeat arrests does not have a third phase is due to the limited time horizon.

There is essentially nothing surprising about the similarity between the dynamics of the *Prosecutions* with respect to first and repeat arrest since they are essentially the same in terms of the structure of their feedback loops. Somewhat surprising is the presence of shadow feedback loops during P_2. One might expect other feedback loops to directly control the behavior pattern of *Prosecutions* for first and repeat arrests. To consider this possibility, remote feedback loops in the arrest submodel were tested for their impact on *Prosecutions*.

During P_1 of first arrests, the feedback B3 also dominated the exponential growth behavior pattern of *Prosecutions*. B3 is the first arrest feedback loop. However, during P_2 of first arrests, the feedback loops in the arrest submodel (B1, B2, and B3) did not dominate *Prosecutions* of first arrest. Thus, during P_2, the behavior of *Prosecutions* of first arrests is entirely driven by feedback loops directly related to *Prosecutions*. During P_3, B3 (first arrests) again dominates the behavior of *Prosecutions*. In contrast, during P_1 of repeat arrests, *Prosecutions* of repeat arrests is dominated by both B2 (repeat arrests) and B3 (first arrests). And during P_2,

Figure 4.35: Reference Atomic Behavior Pattern Index (ABPI) for Repeat Arrest Prosecutions



Prosecutions of repeat arrests is dominated by B1 (cases exiting the stock of cases with prior DV arrest) and B3 (first arrests). The results are summarized in Table 4.29 on page 145.

Table 4.29: Dominance of Arrests Feedback Loops on Prosecutions Across Phases of Loop Dominance

	P_1	P_2	P_3
first arrests	B3	none	B3
repeat arrests	B2 and B3	B1 and B3	-

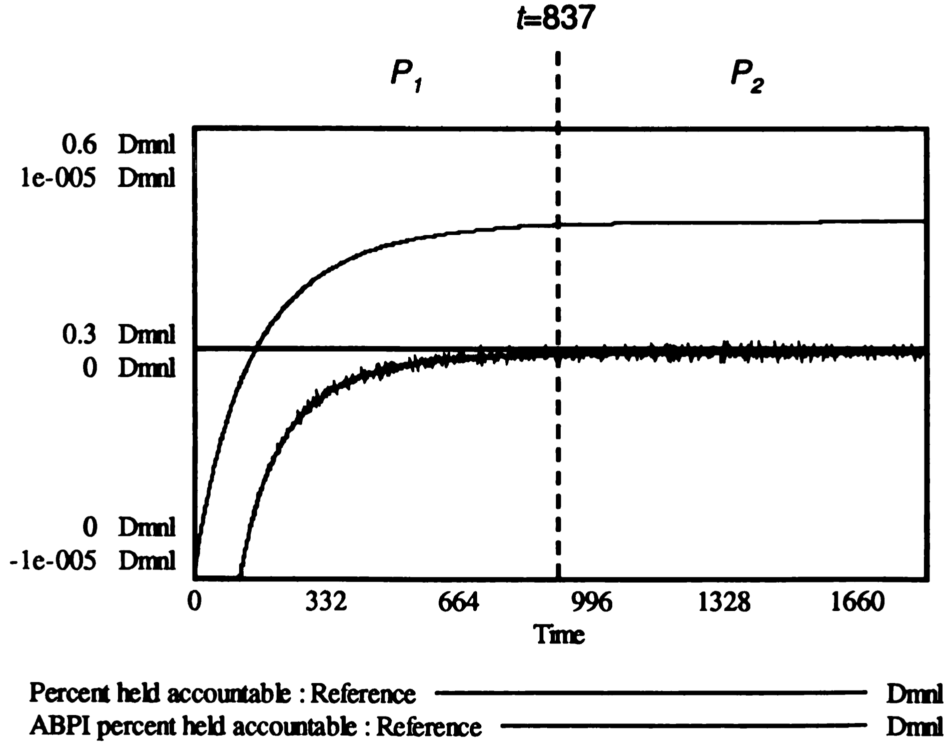
The main implication of this is that first arrests dominate the qualitative behavior of the prosecutor's office domestic violence caseload for both the prosecution of first arrests and repeat arrests in most phases of loop dominance. Interestingly, during P_2 where first arrests are being actively prosecuted, first arrests have no qualitative impact on the behavior of the prosecutor's office caseload with respect to first time arrests being prosecuted. In contrast, first arrests have a qualitative impact on the behavior of the prosecutor's office caseload of repeat arrests in both P_1 and P_2. Thus, even with this simple baseline model, the feedback dynamics of prosecutor's office caseload can involve some subtle complexities about the relationships between various mechanisms directly connected to the prosecutor's office and the relationship between the dynamics of arrest and the dynamics of the prosecutor's caseload.

4.7.7.2 Dynamics of accountability

The main question in this study is the dynamic relationship between assailant accountability and domestic violence caseloads. Accountability, as measured in (4.4), is dynamically related to the prosecutor's office caseload if one of the feedback loops regulating caseloads during a particular phase also regulates accountability. Note that the *Percent held accountable* includes subscripted terms from both first arrests and repeat arrests. Thus, *Percent held accountable* can be directly influenced by both first and repeat arrests, and the feedback loops B4 through B8 need to be tested separately for both first and repeat arrests.

Figure 4.36 shows that there are two phases of atomic behavior patterns to be considered with *Percent held accountable:* P_1 (from 0 to 837 days) and P_2 (from 838 to 1826 days). P_1 represents a period where the *Percent held accountable* grows logarithmically while P_2 corresponds to a period where *Percent held accountable* has reached an equilibrium at approximately 48%. That is, only 48% of the warrant requests result in a plea agreement. The remaining 52% either have their warrants denied or result in a dismissal.

Figure 4.36: Reference Atomic Behavior Pattern Index (ABPI) of Percent Held Accountable



During P_1, feedback loops B3 (first arrests), B5 (prosecutions with enough staff for first arrests), B6 (prosecutions when short staffed for first arrests), and B8 (allocation of prosecutors to prosecuting cases) each dominate the logarithmic growth of *Percent held accountable.* More specifically, each of these has the effect of limiting *Percent held accountable* from an otherwise exponential growth rate to a logarithmic growth rate.

During the subinterval between 0 and 133 days, B3 (first arrests) is the only feedback loop that is controlling both the growth of *Percent held accountable* and *Prosecutions* (for both first and repeat arrests). Thus caseload management mechanisms are not affecting accountability during this initial phase. That makes intuitive sense since the prosecutor's office caseload is fairly small.

During the subinterval from 134 to 837 (328 to 837 for repeat arrest), *Percent held accountable* and *Prosecutions* have feedback loops B5, B6, and B8 in common, although the nature of this dynamic relationship between *Percent held accountable* and *Prosecutions* is more complicated since B5, B6, and B8 influence the behavior pattern of *Prosecutions* only as members of a shadow feedback loop pair with B4. Thus, under ordinary conditions, *Percent held accountable* and *Prosecutions* are not related dynamically in terms feedback loop dominance. However, a change in the behavior pattern of *Prosecutions* would have an impact on the behavior pattern of *Percent held accountable,* since in order to change the behavior pattern of *Prosecutions* one would have to simultaneously modify B4 and either B5, B6, or B8. That is, during this subinterval, one would not be able to change the dynamics of how the prosecutor's office manages its domestic violence caseload without also impacting the feedback loops controlling accountability. The converse is not true. That is, one could change the behavior pattern of *Percent held accountable* by modifying B5, B6, or B8 without impacting the behavior pattern of *Prosecutions.*

During P_2 from 838 to 1826 days, feedback loops B1, B2, B4 (for both first arrests and repeat arrests), and B7 (for both first arrests and repeat arrests) dominate *Percent held accountable.* In the subinterval from 838 to 1158 days, *Percent held accountable* and *Prosecutions* of first arrests are both dominated by feedback loop B4 (dismissals), although B4 is only one member in a shadow feedback loop pair with B5, B6, and B8. So the analysis here is similar to the second subinterval of P_1, where changes in the behavior pattern of *Prosecutions* of first arrests will mean a change in the behavior pattern of *Percent held accountable.* And again, the converse is not true. Changing the behavior pattern of *Percent held accountable* by modifying B4 will not affect *Prosecutions.*

However, during the subinterval from 1158 to 1826 days of P_2, *Percent held accountable* and *Prosecutions* of first arrests both have B7 (allocation of prosecutors to authorizing warrants) as a dominant feedback loop. This corresponds to the period where the baseline model is in steady state or equilibrium. Note that during this period, the entire model has, in fact, reached equilibrium. Thus, the feedback mechanism associated with allocating prosecutors to review warrants is controlling both the prosecutor's office caseload of first arrest domestic violence cases and the proportion of warrant requests that result in the first arrest being held accountable. But, this only happens when the model has reached equilibrium. Thus, the baseline model indicates that there is a caseload management mechanism impacting accountability, but only during a specific phase of feedback loop dominance, namely, equilibrium. During the other phases, accountability and caseload are only conditionally controlled through the same mechanisms.

4.7.8 Summary of baseline model

The baseline model, which simply modeled flows of cases through the criminal justice system, did a reasonable job given its simplicity of reproducing the observed time series, although it was unable to reproduce what seem like important oscillation in the dismissals of repeat arrests (see Figure 4.20 on page 117). The baseline model also appears to be sensitive to higher frequency oscillations. Analysis of the baseline model in terms of dominance feedback loops suggests that accountability and prosecutor's office caseload have one controlling feedback loop in common when the system is in equilibrium, namely, the allocation of prosecutor resources to warrant reviews. Differences between prosecutions of first arrests and repeat arrests suggest a need to further analyze and disaggregate arrests.

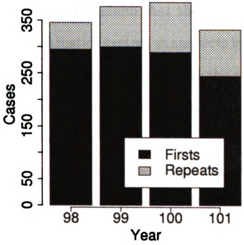
4.8 First versus repeat offenses

Figure 4.37 shows the general annual trends of warrant requests by first offenses, repeat offenses, and the type of relationship between victim and offender. Recall that first and repeat offenses refer not to previous convictions, but simply whether or not there was a prior warrant request on or after January 1, 1998. Repeat offenses include individuals who were arrested on or after January 1, 1998 where the warrant request was denied or the case dismissed. Individuals who had a domestic assault related arrest prior to January 1, 1998 would still be considered first offenses in this analysis. Furthermore, first and repeat offense would also include cases where victims were arrested and the warrant denied or the case dismissed.

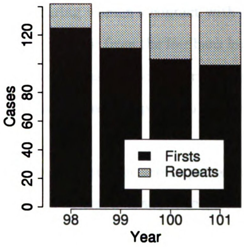
While Figure 4.37a shows a relatively stable number of first offense warrant requests per year with a slight decline starting in 2000. Repeat offenses for all cases, however, are increasing until 2001. Breaking down the warrant requests by the type of relationship reveals that four-year prevalence rate of domestic violence relationships is remaining constant, with a transfer of cases from first offenses to repeat offenses (see Figure 4.37b). The patterns for cases involving cohabiting and other types of relationships are more complicated (Figures 4.37c and 4.37d). Figure 4.37c shows an increasing trend from 1998 through 2000 that would be consistent with an "amplifying" mechanism, i.e. a positive feedback loop where the number first time arrests feed back via some mechanism to increase the likelihood of more first time arrests–the opposite of a general deterrence effect. Figure 4.37d suggests that from 1998 to 2000, the number of arrests equals the number of new cases becoming at risk of an arrest, that is, some type of replacement mechanism. The decline in 2001 could be the delayed effect of a general deterrence mechanism. Thus the general trends in Figures 4.37 suggest fundamentally different types of mechanisms with respect to the effects of arrest and prosecution that seem to depend on the type of relationship between the suspect and victim.

Figure 4.38 shows the results of breaking down the warrants, first by suspect/victim sex (male-to-female and female-to-male) and then by suspect/victim sex for a subset of cases where (a) both parties were 18 years of age or older, and (b) the type of relationship was domestic violence. The patterns depicted in both Figures 4.38a and 4.38b are similar to each other as well as the patten in Figure 4.37b, which is consistent with a closed system of assailants with no replacement of new cases. And more specifically, the overall decline in Figures 4.38a and 4.38b starting in 1999 and continuing through 2001 would be consistent with (a) general deterrent effects, (b) specific deterrent effects after the first contact with the criminal justice system, (c) treatment effects from intervention, or (d) displacement effects whereby assailants continue to abuse, but shift their tactics from criminal to non-criminal behaviors. All four of these have been proposed as explanations for declines in assailants' arrest

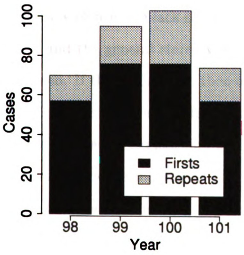
Figure 4.37: Warrant Requests by Type of Relationship and Year



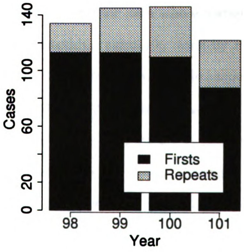
(a) All warrant requests



(b) DV cases



(c) Cohabitting cases



(d) Other cases

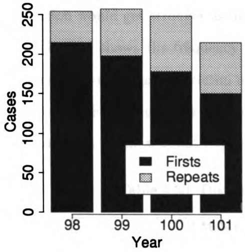
rates after contact with the criminal justice system.

Figures 4.38c and 4.38d are also consistent with each other, as well as with 4.37c but are in stark contrast to the patterns in Figures 4.38a and 4.38b. This lends support to the claim that the feedback loops structuring women's contact with the criminal justice system as suspects is fundamentally different from the feedback loops structuring men's contact with the criminal justice system as suspects. Furthermore, women's contact with the criminal justice system as suspects appears to involve a positive feedback loop.

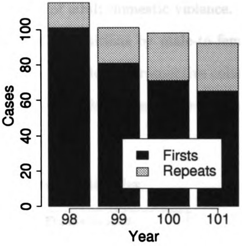
There are two possible explanations for differences between arrest patterns of male and female suspects in domestic violence cases. The first is that police generally correctly identify the assailant, but (a) the deterrence and intervention effects on suspects depend the suspect's sex, and (b) interventions with batterers have been geared toward male abusers. So while arrests, prosecution, and sanctions might deter male abusers, those same interventions have no impact or are detrimental to female abusers. Advocates of this position point to the gender "symmetry" of domestic violence, the lack of victim and advocacy services for men who experience abuse, and the general stereotype that abuse is something primarily done by men to women. If this is the case, then one would expect that female and male suspects would have similar trajectories after arrest within the criminal justice system.

The competing explanation is that the presence of female suspects is generally the result of police incorrectly identifying the assailant as the victim during a response to a domestic assault. This might happen, for instance, as assailants become more skilled at battering and manipulating law enforcement officers on the scene. Victims might then be at increased risk of being arrested in the course of defending themselves. If this is the situation, then one would expect the proportion of warrants being denied to be higher for female suspects than male suspects. And

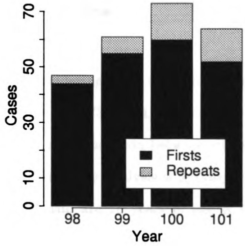
Figure 4.38: Warrant Requests by Suspect/Victim Sex and Year



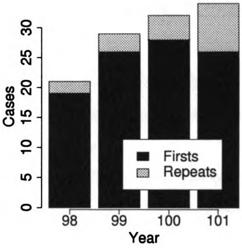
(a) Male to female



(b) Male to female adult DV



(c) Female to male



(d) Female to male adult DV

this association should get stronger as one narrows the selection of cases to adult domestic violence relationships. One would also expect a higher proportion of female suspects to have first appeared as victims than male suspects, and that this association would get stronger as one selected cases of adult domestic violence.

Table 4.30 shows the frequency distributions of dispositions by male-to-female and female-to-male suspect-victim relationships, along with their respective column percentages. Female-to-male cases are about 50% more likely than male-to-female cases to result in the warrant request being denied.

Table 4.30: Dispositions by Suspect/Victim Sex

	Male-to-female		Female-to-male	
Disposition	Freq.	%	Freq	%
Denied	354	37.8	137	57.3
Dismissed	86	9.2	22	9.2
Guilty at trial	4	0.4	1	0.4
Nolled	5	0.5	1	0.4
Not guilty at trial	10	1.1	0	0
Plead guilty	404	43.1	70	29.3
Plead no contest	74	7.9	8	3.3
Total	937	100.0	239	100.0

Considering only the subtable with cells of five or more cases in Table 4.31, $\chi^2(2, N = 1073) = 24.85$, $p = 4.0e-6$, and $\phi = 0.15$.[4] This suggests that there is indeed an association between suspect/victim sex and disposition.

If the same analysis is applied to the subset of cases where (a) both the suspect and victim are 18 years of age or older, and (b) there is clear evidence that the relationship is between intimate partners, cases involving female-to-male suspect-victim relationships are 75% more likely to have their cases dismissed (as opposed

[4] ϕ is a measure of association with values that range from 0 to 1, with 0 indicating no association between the variables and 1 indicating perfect association between two variables.

Table 4.31: Test of Association Between Disposition Suspect/Victim Sex

Disposition	Male-to-female Freq.	Female-to-male Freq
Denied	354	137
Dismissed	86	22
Plead guilty	404	70

$\chi^2(2, N = 1073) = 24.85$, $p = 4.0e-6$, $\phi = 0.15$

to 50% for all types of relationships) (see Table 4.32).

Table 4.32: Dispositions by Suspect/Victim Sex for Adult Domestic Violence Cases

Disposition	Male-to-female Freq.	%	Female-to-male Freq	%
Denied	127	32.1	63	56.3
Dismissed	37	9.3	10	8.9
Guilty at trial	2	0.5	1	0.9
Nolled	3	0.8	1	0.9
Not guilty at trial	1	0.2	0	0
Plead guilty	192	48.5	36	32.1
Plead no contest	34	8.5	1	0.89
Total	396	100.0	112	100.0

Selecting the same subtable, $\chi^2(2, N = 465) = 17.56$, $p = 1.5e-5$, and $\phi = 0.19$, indicating that the association gets stronger ($\phi = 0.19$ for adult domestic violence cases versus $\phi = 0.15$ for all types of relationships) (see Table 4.33).

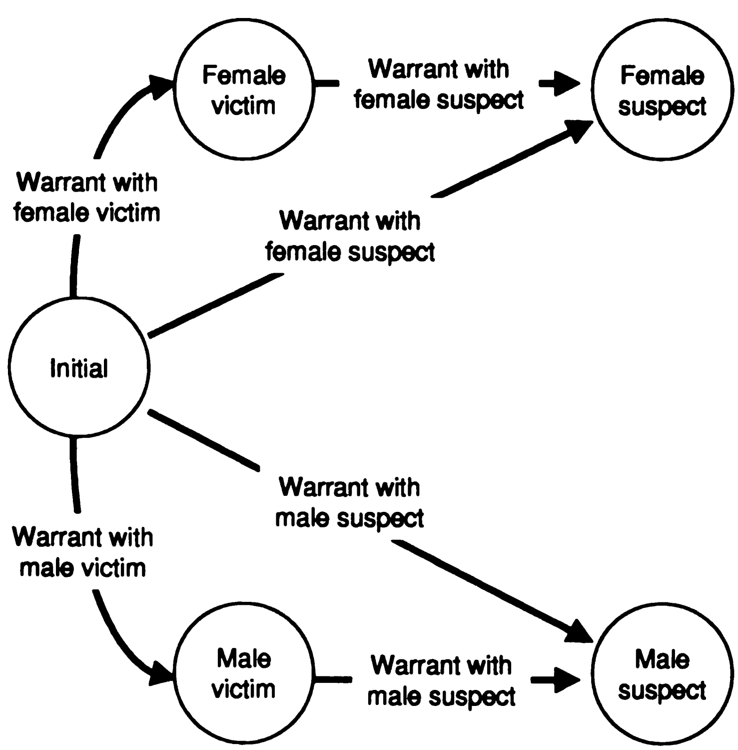
To test whether female suspects were more likely than male suspects to have first appeared as a victim on a warrant, the finite state machine shown in Figure 4.39 was used to generate the frequencies for each the transitions depicted. The test is whether the number of transitions from female victim to female suspects relative to the number of transitions from the initial state to female suspects is greater than the corresponding relation for male suspects.

Table 4.33: Test of Association Between Disposition Suspect/Victim Sex for Adult Domestic Violence Cases

Disposition	Male-to-female Freq.	Female-to-male Freq
Denied	127	63
Dismissed	37	10
Plead guilty	192	36
Plead no contest	34	1

$$\chi^2(2, N = 465) = 17.56,\ p = 1.5e - 5,\ \phi = 0.19$$

Figure 4.39: Finite State Machine to Test Paths of Suspects by Sex



The results of running this analysis for all cases are listed in Table 4.34. Indeed, female suspects are much more likely than male suspects to have first been recorded as victims of domestic violence. Running the same analysis on cases involving adult domestic violence showed a similar result (see Table 4.35), with the strength of association increasing from $\phi = 0.26$ to $\phi = 0.35$.

Table 4.34: Suspects by Previous State and Sex of Individual

Previous state	Male Freq.		Female Freq	
Initial	784	92.9%	208	73.2%
Victim	60	7.1%	76	36.5%
Total	844	100.0%	284	100.0%

$\chi^2(1, N = 1128) = 77.4$, $p = 1.4e - 18$, $\phi = 0.26$

Table 4.35: Suspects by Previous State and Sex of Individual for Adult DV

Previous state	Male Freq.		Female Freq	
Initial	320	94.4%	72	69.9%
Victim	19	5.6%	34	32.1%
Total	339	100.0%	106	100.0%

$\chi^2(1, N = 445) = 53.9$, $p = 2.1e - 13$, $\phi = 0.35$

4.9 Key informant interviews

The formulation and analysis of the baseline model generated questions for key informant interviews about the model's formulation, validity of various trends and structure, and additional feedback loops. A total of seven interviews were conducted with five key informants. Three of the key informants had victim advocacy positions, while the other two key informants worked for the prosecutor's office. All five key informants had attended a variety of domestic violence trainings. All five key informants had experience working with victims of domestic violence, with a range of between 4 to approximately 10 years experience.

4.9.1 Victim arrests

The three key informants with victim advocacy positions reported a trend of victims of domestic violence being arrested by police. These were not dual arrests, but single arrests where the police arrested the victim instead of the assailant. One described the trend as increasing from 1998 through 2000, and getting better in 2001. Several reasons were given for the decline starting in 2001. First, victim advocates would talk with individual police officers on a case-by-case basis after a victim was arrested. The nature of these conversations focused on helping officers identify victims of domestic violence and helping officers understand the dynamics of domestic violence. Second, the domestic violence shelter and police departments had increasingly disagreed on issues of sharing information about victims between victim advocates and law enforcement. This conflict deteriorated victim advocates' access to law enforcement officers from 1998 through 2000. In 2000, however, changes in the leadership of both the police department and domestic violence shelter resulted in a renewed sense of cooperation and access to police officers in cases where victims were arrested. Third, the prosecutor's office started in 2001 to have only one assistant prosecuting attorney handle all of the domestic violence cases, with exception of sexual assaults, as opposed to dividing the work between two assistant prosecuting attorneys. This simplified communication with the prosecutor's office for victims and victim advocates, making it much easier to respond quickly to cases where the victim had been arrested.

4.9.2 Batterer intervention programs

All of the key informants with knowledge of batterer intervention programs and state standards (four out of the five) reported that the batterer intervention programs in their community were ineffective and did not hold assailants accountable. None of the programs followed the minimum standards of a Duluth-like accountability model.

4.9.3 Trends

Key informants reported no change in the severity or intensity of domestic violence over the four year period. Nor did key informants report any general trends in the demographics of victims or assailants. Two key informants did, however, report that the increased outreach efforts from 1998 to 2001 had resulted in an increased awareness in the community of domestic violence and an increased demand for services. One key informant reported that a number of repeat assailants that had been consuming a significant amount of time and energy moved out of the jurisdiction.

4.9.4 Victim advocacy

Three of the five key informants emphasized the importance of having a skilled domestic violence victim advocate working with the prosecutor's office. This position was funded under the grant supporting the coordinated community response and filled by a person with substantial victim advocacy experience. The victim advocate was given a list of all meetings scheduled with victims for the week. The victim advocate would meet with the victim before the initial meeting with the assistant prosecuting attorney, be with the victim during the meeting, and with time permitting, follow up with the victim after the meeting. Unlike victims seeking domestic violence shelter or calling the crisis hotlines, who were more likely to identify their experience as domestic violence, victims in the criminal justice system were more likely to see the assailant's behavior as an isolated incident. The victim advocate would assess for patterns of domestic violence, provide the victim with some basic information about domestic violence (e.g., power and control wheel, cycles of abuse), facilitate referrals to local resources, and help the victim with safety planing.

4.9.5 Prosecution

Staring in 1995, the prosecutor's office implemented an evidence based no-drop prosecution policy, which meant that cases would be prosecuted without requiring victim testimony and the prosecutor would not dismiss the case if the victim wanted the charges dropped. Starting in 2001, the prosecutor's office moved to vertical prosecution, which meant that the same prosecutor handled the case from the warrant review to the final disposition.

While most of the cases that were prosecuted resulted in plea agreements or dismissals, some did go to trial. The assistant prosecuting attorney selected the cases that would go to trial on the basis of wanting to set a precedent for future plea negotiations. For example, a case where the suspect had only one criminal charge but a pattern of abusive behaviors might be pushed to trial as a way to raise the "going rate" for that type of case, where the "going rate" would be the likely sanction for a case with similar attributes. Meanwhile, the defense bar was closely monitoring the types of cases that were pushed to trial and their outcomes.

Key informants did report some effects of caseload pressure on decisions affecting case disposition, but this was primarily limited to cases receiving more generous plea offers. These were described as times of "bad judgment" where the assistant prosecuting attorney might offer a delay of sentence for an assailant with a pattern of abuser where ordinarily no such offer would be made. These periods were reported as being infrequent and limited in time, on the order of one or two weeks.

4.9.6 Resource issues

Key informants reported three major resource issues. First, raising awareness of domestic violence in the community through increased outreach and victim services increased demand for victim services. A federal grant made the increased outreach and victim services possible. However, the demand for services remained steady even after the grant ended. Thus the domestic violence shelter and victim advocates were faced with the problem of having to meet the same demand for services, but with fewer resources.

Second, funding a domestic violence victim advocate for the prosecutor's office allowed the victim advocate to be available for most meetings with the prosecutor. When the position was no longer funded, the victim advocate was forced to divide time between the prosecutor's office and domestic violence shelter, with shelter crises taking priority over meetings in the prosecutor's office. The result was less victim advocacy in the prosecutor's office.

Third, between 10 and 25% of the assistant prosecuting attorney's time assigned to domestic violence cases was devoted to seeking and justifying funding for the domestic violence unit. This time was not distributed evenly, but concentrated in brief periods when proposals and grants were being prepared. Meetings with victims would often be rescheduled during these periods to the following weeks, which made it less likely that victims would show up for the meetings with the assistant prosecuting attorney. The assistant prosecuting attorney was also much less likely to visit crime scenes and interview other witnesses during such periods, delegating these tasks to police officers. The net effect was less assessment of the pattern of violence, less safety planning, and weaker criminal cases.

4.9.7 Rural issues

Closely related to resource issues were what might be called rural issues. Several key informants identified specifics of providing services to domestic violence victims in rural communities. For example, jurisdictions were geographically larger, which resulted in longer travel time. Longer travel times meant longer delays in terms of meeting with victims to provide advocacy and more resources for investigating and interviewing crime scenes. Rural communities are also small communities, and several key informants pointed out that individuals fighting against domestic violence were quickly identified and often became targets for assailants. For example, assailants were more likely to effectively access and use local media to slander individuals. Although not true for all rural communities, this one was outside the circuit of speakers and various professional resources that one might tap into for support and continuing education. The result was that professionals felt more isolated, and used more financial resources for travel to conferences and other types of professional development opportunities.

4.10 Revised model

Analysis of the baseline model and key informant interviews identified five new feedback structures for the revised model: safety planning, community awareness, resource allocation, "going rate", and prosecutor meetings with victims. Each of these structures represents an additional submodel, with a corresponding need for numerical time series data. The numerical data that might be used to generate such time series, especially the shelter outreach and contact data, was unavailable at the time of this writing because the original grant funding the data collection had been extended another six months, ending in March 2003. Thus, in contrast to the descriptive and baseline models, the revised model is conceptual in nature.

In order to distinguish the existing structure of the baseline model and revised model, italics will be used to denote the new variables in the revised model. Variables in the revised model represent general concepts, which would typically be disaggregated in a numerical model.

Safety planning represents a critical balancing feedback loop (B1) to positive feedback loop effect of arrests of assailants resulting in assailants learning how to get victims arrested (R1) in Figure 4.40. *Safety planning* is the outcome of *Meetings with victims* and counteracts the effect that some assailants will learn how to manipulate the system. Specifically, victim advocates and the assistant prosecuting attorney can, during meetings with victims, assess for patterns of domestic violence and help the victim develop safety plans, which might help the victim identify more options and anticipate some of the tactics that the assailant might use as a consequence of his contact with the criminal justice system.

First arrests increases the level of *Community awareness,* and this increases the likelihood of more incidents of domestic violence being reported, resulting in a reinforcing feedback loop (R2) in Figure 4.41. For example, a coworker or neighbor who one did not suspect of being a batterer is arrested for domestic violence. It is possible that *Repeat arrests* also contributes to *Community awareness*, but it is less clear whether the effect is to increase or decrease *Community awareness* as something that increases reporting. For example, *Repeat arrests* might discourage individuals from reporting as they feel increasingly frustrated with arrests having no desirable impact on the abuser.

The resource allocation structure shown in Figure 4.42 introduces four feedback

Figure 4.40: Safety Planning in Revised Model

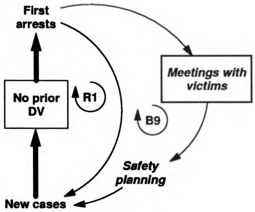
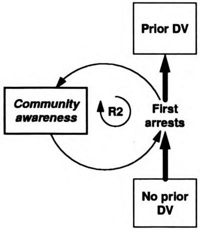
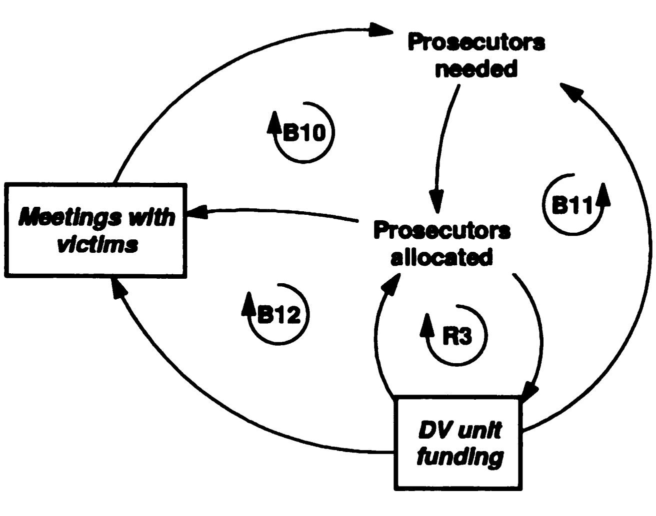


Figure 4.41: Community Awareness in Revised Model



loops. First, there is the effect that as more prosecutors are allocated to seeking funding for the domestic violence unit, there are more prosecutors to allocate, and more prosecutors to seek funding (R3). *DV unit funding* also creates a demand on *Prosecutors needed*, which increases the *Prosecutors allocated* to seeking funding (B3). This might happen, for example, when funding levels go below the desired goal and there is an increased pressure to find new grants. When more prosecutors can be allocated, *Prosecutors allocated,* there can be more *Meetings with victims*, which decreases the demand for prosecutors, *Prosecutors needed*, forming a balancing feedback loop (B2). *Meeting with victims* is also within another feedback loop. Specifically, as there are more *Meetings with victims,* there is more demand for prosecutors to meet with victims relative to demand to for funding the DV unit, thus *Prosecutors allocated* to finding funding for the DV unit goes down, the *DV unit funding* goes down, and the funding for the victim advocate position decreases, which decreases number of or delays *Meetings with victims* (B4 in Figure 4.42).

Figure 4.42: Resource Allocation in Revised Model



The *Going rate* is a stock variable that is perceived, imperfectly, by the assistant prosecuting attorney and defense attorneys (see Figure 4.43). When the *Going rate* is high, more cases can be disposed through *Plea agreements* faster, which lowers the *Prosecution caseload*, and decreases the cases that are *Dismissed.* Fewer dismissals has the effect of increasing the *Going rate.* Thus, a higher *Going rate* has the effect of increasing the *Going rate*, forming a reinforcing feedback loop (R5).

If one expects there to be some deterrence effect from arresting and prosecuting assailants, then a higher *Going rate* should lower *Repeat arrests,* which would decrease the *Warrant requests* and the number of *Warrants denied.* And fewer warrants being denied will have the effect of further increasing the *Going rate,* forming another reinforcing feedback loop (R4).

The assistant prosecuting attorney develops a better understanding of the case by meeting with the victims (see Figure 4.44). So more *Meetings with victims* means more *Plea agreements,* which lowers the *Prosecution caseload,* and this lowers the number of *Prosecutors needed,* for meeting with victims, and thus fewer prosecutors are allocated to meet with the victim, *Prosecutors allocated*, which reduces the number of *Meetings with victims.* This forms the balancing feedback loop B5. The resulting revised model is shown in Figure 4.45.

4.11 Conclusion

This chapter described the results, including descriptions of the violence against women (VAW) database, comparisons with census statistics, and dispositions. Three descriptive models were presented. Analysis of the fit statistics suggested that a descriptive model based on dispositions being modeled as a function of caseload performed better than models where dispositions were modeled as a function of case

Figure 4.43: "Going Rate" in Revised Model

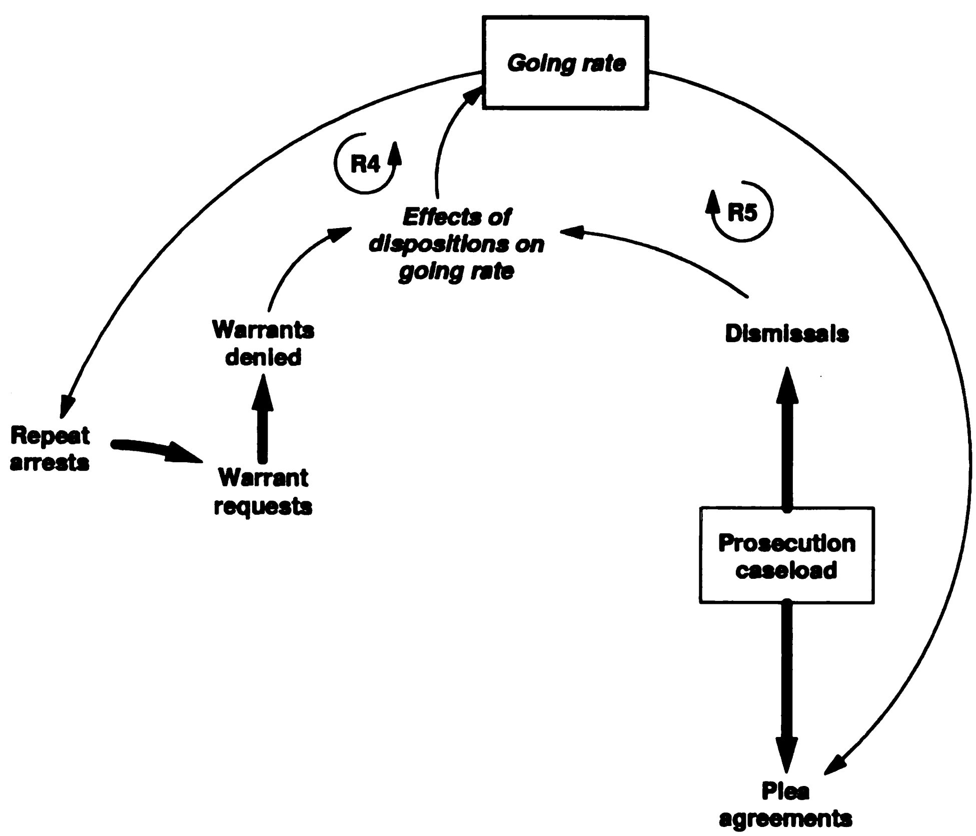


Figure 4.44: Prosecutor Meetings with Victims in Revised Model

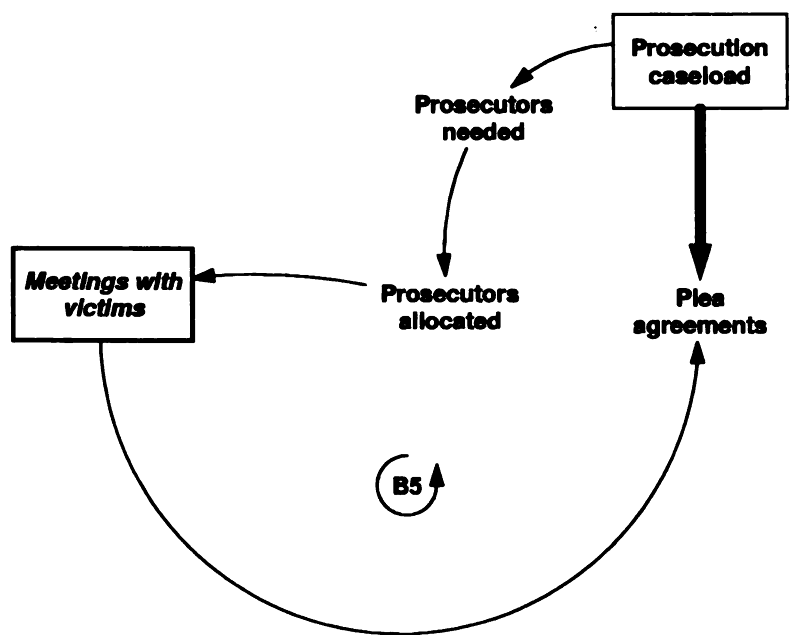
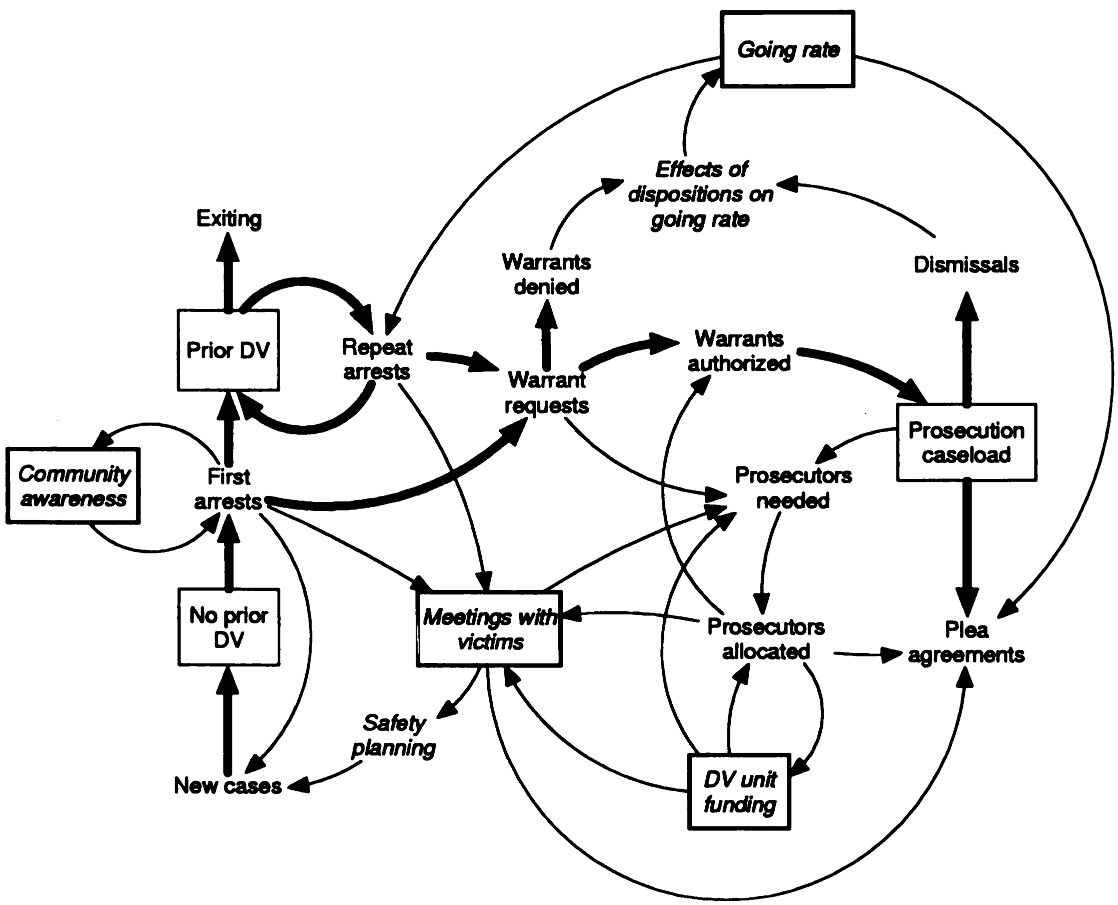


Figure 4.45: Overview of Revised Model



attributes. The chapter then presented the baseline model. Analysis of the baseline model showed that it did a reasonable job of reproducing the behavior of the real system given its simplicity.

The main result of the baseline model in terms of the study was that accountability and the prosecutor's office caseload of first arrests were both controlled by the feedback loop representing the mechanism of allocating prosecutor resources to warrant reviews, but only when the system was in equilibrium.

One problem in the baseline model concerned the aggregation of first and repeat arrests. Disaggregating arrests by other variables led to an important insight into a potential positive feedback loop between the first time arrests of male assailants and the risk of arrest of female victims.

Finally, the baseline model was used to motivate questions for key informant interviews. The key informant interviews provided further support for the claim that there exists a positive feedback loop linking first arrests of male assailants and the arrest of female victims. Victim advocacy emerged as a critical component of prosecuting domestic violence cases, and key informants suggested several feedback mechanisms that related to the prosecution of criminal domestic violence cases. These were summarized in the presentation of feedback loops that would be added to the baseline model to form a revised model.

Chapter 5

Conclusion

5.1 Overview

This chapter provides a summary of the major findings, along with their policy implications, strengths and limitations of the approach taken in this research, and directions for future research.

5.2 Major findings and discussion

There were nine major findings in this study:

Finding 1: Descriptive models with dispositions written in terms of caseloads outperformed similar models with dispositions written in terms of case attributes. The major implication of this finding is that, at some level, dispositions are affected by the size of the prosecutor's caseload. That is, there is evidence in supporting the caseload pressure hypothesis.

Finding 2: The baseline model, formulated as a simple flow of cases from arrest to dispositions, did a reasonable job of reproducing the observed time series.

The model was structurally unable to reproduce what might be some important oscillations in the dismissals of repeat arrests. The major implication of this is that caseflow dynamics alone can explain the general four-year trends, but not some oscillations like those seen in the dismissals of repeat arrests. That is, the baseline model is a good robust model for studying the general trends of caseflows, but additional structures are needed for studying some key oscillations like dismissals of repeat arrests.

Finding 3: Sensitivity analyses point to problems in the baseline model's response to higher frequency oscillation. Higher frequency oscillations can be indications of information mechanisms. Information feedback mechanisms include such things as perception of caseloads or the "going rate", awareness of domestic violence resources, court rulings, and actual paper work. The implication is that additional structures are needed for studying oscillations like dismissals of repeat arrests, which are likely to be information mechanisms. That is, things like perception of caseload, the "going rate", and the awareness of domestic violence resources need to be added to the baseline model in order to reasonably represent some variables like dismissals of repeat arrests.

Finding 4: The baseline model was sensitive to reductions in prosecutor resources. Specifically, cutting prosecutor resources changed the dynamics of warrant reviews from authorizing warrants on the basis of case attributes to denying warrants because of limited resources. The implication is that the primary impact of cutting prosecutor resources on the dynamics of prosecuting domestic violence cases is on the denial of warrants, not the dismissal of cases or plea offers being accepted. Specifically, cutting prosecutor resources essentially puts a cap on the number of cases that can be authorized, and this makes the level of prosecutor resources the main determinate of warrants being authorized, as opposed to whether or not there is probable cause. That is, cutting prosecutor resources causes warrant reviews to be dominated by caseload pressures as opposed to professional norms.

Finding 5: The baseline model was insensitive to increases in new cases. Increasing the rate that individuals flowed into the criminal justice system generally did not, surprisingly, affect the dynamics prosecutor caseloads or case dispositions. The implication is that the structure of prosecutor's office caseflows is robust with respect to increases in new cases. What this means is that, according to the baseline model, policy changes that increase the number of new cases and thus arrests will generally not affect the prosecution dynamics. That is, when there are adequate resources for the prosecuting domestic violence cases, the prosecutor's office can absorb relatively large increases in new cases without affecting the caseflow dynamics of prosecution outcomes or having to increase prosecutor resources.

Finding 6: Accountability and prosecution caseload of first arrests shared a common dominating feedback loop, allocation of prosecutors to authorizing warrants, but only when the model was in equilibrium. The implication is that accountability and prosecution caseload are controlled by a common feedback loop when caseloads are in a steady state, that is, when caseloads are neither increasing nor decreasing. Moreover, structural changes during steady state to increase accountability would also have the unintended consequence of also changing the dynamics of caseloads. That is, if prosecution caseloads were stable, one would have to counteract the effects on prosecution caseload of policy changes to increase accountability in order to avoid changes in the dynamics of case dispositions.

Finding 7: There were strong indications of a positive feedback loop linking the first arrest of male assailants and the subsequent risk of arrest for female victims. Victim advocacy and safety planning with female victims at the prosecutor's office are potentially vital interventions that might be counteracting some of the effects of the positive feedback loop linking first arrest of male assailants and the subsequent risk of arrest for female victims. The main implication is that female domestic violence arrests represent female victims being arrested, not female perpetrators of domestic violence.

Finding 8: An underlying assumption in this study was that the prosecutor's office operated under fixed funding constraints. Thus, resource allocation was achieved by managing caseloads as opposed to increasing revenues. However, key informants pointed out that a significant portion of their time (up to 25%) went into grant writing and reporting (i.e., maintaining or increasing their resources). The main implication is that resource allocation is dynamic. That is, trying to increase prosecutor resources can consume a significant portion of resources, which ultimately affects the outcomes of prosecutions. This can lead to even more demand for prosecution resources. Thus, the resource allocation problem for the prosecutor's office is dynamic.

Finding 9: An assumption in the baseline model had been that the infrequent nature of trials meant that they had little impact on the prosecutor's office caseload and holding assailants accountability. However, key informants reported that trials were critical for establishing the "going rate" of a criminal offense, and that this "going rate" determined subsequent negotiations of plea agreements between the assistant prosecuting attorney and defense attorneys.

That is, the feedback loops between trials and the "going rate" need to be included in future system dynamics models.

5.3 Limitations of study

There are five major limitations of this study: community racial/ethnic demographics, lack of replication, use of sparse time series, uncertainty in validity of identifying first time arrests, and peripheral role of stakeholders. The last four stem from an original decision to focus on understanding and developing methods for subsequent studies. This design decision was partly motivated by the possible availability of additional data from an ongoing evaluation study of the county's coordinated community response effort to domestic violence. However, the funding for the evaluation study was extended by six months, which effectively postponed the availability of additional data sets until April, 2003. This limited the numerical data available for this study to the existing prosecutor's office VAW database, which increased the importance on being able to extract and evaluate time series data from an otherwise messy database, and this eventually led to the postponement of the initial key informant interviews.

5.3.1 Community racial/ethnic demographics

The population was mostly white non-Hispanic (see Table 4.2). Race/ethnicity, as a case attribute of victims and offenders, was thus less likely to have a measurable effect on the disposition of cases. In turn, this would increase the effects of caseloads on case dispositions relative to case attributes. Case attributes could well play an equal if not larger role than caseload on case dispositions in more diverse communities.

5.3.2 Lack of replication

The second major limitation of this study and its findings was the single case study design. Single case study designs are inherently limited in their lack of replication and the results cannot be generalized beyond the single case. The decision to focus on a single case of a prosecutor's office was originally motivated by (a) the desire to understand and develop methods that might be subsequently applied in a multiple case study design, and (b) the potential availability of numerous data sets from other researchers working on an evaluation study of the county's coordinated community response to domestic violence. The main implication of being limited to the Violence Against Women (VAW) database from the prosecutor's office was that there were no independent means of validating its contents, and there were obvious indications of inconsistency in data entry. This resulted in more emphasis being placed on understanding the contents of VAW database, data cleaning, construction of variables for analysis, and checking variables.

5.3.3 Sparse time series

Deciding that it was important to handle sparse time series introduced major problems in terms of both representing the real data and evaluating the final results of the model. First, using a smoothing algorithm had the effect of introducing transients that were not present in the original raw time series and did not, at a conceptual level, make sense. This was specifically true for rates such as *First arrests* and *Repeat arrests*, which appeared to gradually increase in 1998 when the raw data indicated that they started out in 1998 near their mean. While this did not affect the baseline model's performance, it made it difficult to interpret the results in real terms. This gradual increase or transient response could have been offset by adjusting the initial values of the smoothing algorithm, but this would have complicated the initialization of the baseline model. Ordinarily, this would not be a problem, but since the model was required to satisfy a range of inputs and outputs from different smoothing algorithms (a consequence of the approach to handling sparse time series), each smoothing algorithm would have implied a different procedure to initialize the model. Similar problems appeared with the smoothed time series being delayed. Both approaches might be solved by selecting a class of smoothing filters that minimizes some of these distortions or providing ways of compensating their output in terms analytically related to the parameters used in the smoothing algorithm.

Second, analyzing the sensitivity of the simulated model to variations in smoothing parameters (smoothing delays) turned out to be more elusive than expected. While summary statistics provided a means of comparing model and real data in a systematic way, it was difficult to assess what different patterns of comparisons actually meant. More theoretical work on this approach is needed in terms of analytical results, simulation studies, and demonstrations with other data sets.

Third, generating the summary statistics manually drastically limited the amount of variation that could realistically be studied. For example, it would have been good to not only vary the length of the delay in the smoothing algorithm, but the order of the delay and even the type of smoothing algorithm. But, doing so would have generated hundreds of data sets that needed to be compared. This is more of an implementation issue. Automating some of these procedures clearly helped, but more sophisticated routines could generate a family of smoothed data and run the comparisons with the real data.

5.3.4 Validity of identifying first time arrests

The fourth major limitation of this study concerned the procedure used to identify first time arrests. Specifically, first time arrests were identified according to their first appearance in the Violence Against Women (VAW) database starting January 1, 1998. Anyone with an arrest prior to January 1, 1998 would have the first arrest after that date count as a first arrest as opposed to an actual repeat arrest. This would have the effect of elevating the number of cases appearing as first time arrests. The distortion would be the highest immediately after January 1, 1998, and then gradually decline as one considered more recent offenses. This introduced a question about the validity of the identification of first time arrests in this study. Additional prosecution, police data, and shelter data might have helped answer this question.

5.3.5 Stakeholders

The fifth major limitation of this study was that stakeholders had a peripheral role at the end of the model building. The problem of improving the capacity of a prosecutor's office to hold assailants accountable by understanding how domestic violence caseloads were managed was not defined in terms of the prosecutor's office. The prosecutor's office was, in fact, much more concerned about victim safety and supporting the critical role of the victim advocate. Ideally, stakeholders should be included at the beginning of a project in defining the problem, the purpose of the model, and the criteria for evaluating the model. Deciding to involve stakeholders late in the modeling effort was part of the trade-off between focusing on methods and problem solving. But future system dynamics studies should, unless purely concerned with method, begin by meeting with potential stakeholders and identifying the problem to be solved.

5.4 Policy implications

As a single case study design, the policy implications are only provisional. The main policy implication of this study concerns the positive feedback loop from first arrest of male assailants to female victims being arrested, which represents an unintended consequence of mandatory arrest policies. That is, one unintended consequence of mandatory arrest policies might be that some assailants learn, through their first contact with the criminal justice system the criteria for an arrest, and begin using that knowledge in their manipulation of the criminal justice system to get victims arrested as a battering tactic. Key informants pointed out that this also happened with child protective services (CPS), where the assailant would learn the CPS worker's criteria, and use that knowledge to manipulate CPS to gain custody over children, and thereby use the children as a battering tactic to manipulate their mothers.

Experienced victim advocates understood these risks and would work aggressively with victims in terms of safety planning. This was especially critical in criminal cases, where victims might not identify themselves in a domestic violence relationship and minimize the severity of the pattern of abuse. Information about patterns of domestic violence, helping the victim assess her own safety, and providing resources such as shelter and legal advocacy appeared to counteract the effects of some assailants using the criminal justice system and CPS against the victim. But this required victim advocates with significant experience working with victims in the area of domestic violence. Thus, the main policy recommendation would be for more stable funding of domestic violence victim advocates.

5.5 Directions for future research

The immediate next step is developing the revised model and formulating the feedback relationships described by key informants. The key informant interviews provided a basis for identifying several dynamic problems. Thus, the revised model can be better oriented toward solving a specific problem. It should also be possible to calibrate and test the revised model with the availability of additional numerical data from police departments and domestic violence shelters.

A major limitation of this study was the lack of replication. So an obvious next step is to consider several diverse prosecutor's offices, and compare the results of specifically calibrating the revised model to each county's prosecutor's office. Qualitative differences between the fit should indicate structural differences between the counties in terms of feedback loops. This would lead to further refinement and understanding of how accountability and prosecutor's office caseloads are related.

A general model should also be able to, with proper parameterization, reproduce related phenomena. One possibility would be to extend the model to the arrest and prosecution of driving under the influence (DUI) cases, which also result in a high volume of referrals to diversion or counseling type interventions programs. This would be a good test for both the baseline and revised model, although it might ultimately be more difficult to successfully adapt the revised model to the prosecution of DUI cases. Differences in performance between the two domains would indicate either similar structures that had been excluded from the model or structural differences between domestic violence and DUI cases.

A fourth direction for future research concerns the theory of analyzing feedback loop dominance. In system dynamics, the relationship between structure and behavior of complex dynamic systems is thought of in terms of feedback loop dominance.

Despite recent efforts to formalize and automate the analysis of feedback loop dominance, the results have been disappointing. The alternative has been to rely on the existing ad hoc methods that can lead to misperceptions of model behavior. If the goal of system dynamics is to make a contribution to the understanding of complex systems, then it seems essential that both (a) the underlying theory of feedback loop dominance be extended, and (b) that methods be developed for more efficiently and systematically identifying patterns of feedback loop dominance. For example, it seems that one should with at least simple models be able to analytically relate various features of feedback loop dominance. It seems equally important to develop or extend existing procedures for automatically identifying patterns of feedback loop dominance, especially for larger models.

5.6 Conclusion

One of the challenges for this study was to develop methods for studying community responses to domestic violence that paid attention to problem of modeling small populations. The emphasis on developing skill and a method for applying system dynamics to the general problem of increasing assailant accountability led to a disproportionate effort being spent on a variety of technical issues as opposed to problem formulation, model building, and involvement of stakeholders. A quicker and more productive route to actually developing a model that might have some immediate benefits to stakeholders in terms of implementation could have been produced with earlier involvement of stakeholders at the stage of defining the problem and purpose of the model, but not without sacrificing a certain degree of understanding the differences between a system dynamics approach and other methods. Specifically, paying attention to issues such as working with sparse time series and differences between descriptive and mechanistic accounts of the data led to theoretical insights that had real consequences in terms of how one might evaluate claims involving complex systems.

For example, it is common in statistical model building to focus on estimating various parameters as if these parameters are inherently meaningful. But there are, in fact, many possible formulations that can describe the same set of behaviors. To develop a sense of what a parameter means, one must have a sense of what it does not mean and why. This is not something that can simply be done through an operational definition. It requires a familiarity with the field of possible meanings. And to achieve that, one must not look narrowly at fitting one model, but develop and consider a variety of models that all claim to account for a similar concept. This was, abstractly speaking, the approach that emerged in Chapter 3 when considering the problem of sparse time series, and in Chapter 4 when comparing and contrasting the descriptive and baseline models.

Accounting for feedback loops and analyzing the relationship between their structure and resulting behavior is critical if one is going to try and understand the interactions and behaviors of complex systems. System dynamics is a promising method for doing this, especially when one can use both numerical time series and involve stakeholders in defining the purpose of modeling and as key informants. However, more work needs to be done in relating the structure of feedback loops to understanding the system's behavior in terms of patterns of feedback loop dominance. Otherwise, fully analyzing system dynamics models of complex and emerging social problems will have limited application in social work and the social sciences.

References

Albonetti, C. A. (1987). Prosecutorial discretion: The effects of uncertainty. *Law and Society Review, 21*(2), 291–313.

Albonetti, C. A. (1992). Charge reduction: An analysis of prosecutorial discretion in burglary and robbery cases. journal of quantitative. *Criminology, 8*(3), 317–333.

Albonetti, C. A., & Hepburn, J. R. (1996). Prosecutorial discretion to defer criminalization: The effects of defendant's ascribed and achieved status characteristics. *Journal of Quantitative Criminology, 12*(1), 63–81.

Anderson, B. D. O., & Moore, J. B. (1979). *Optimal filtering.* Englewood Cliffs, NJ: Prentice-Hall, Inc.

Becker, R. A., Chambers, J. M., & Wilks, A. R. (1988). *The new s language: A programming environment for data analysis and graphics.* Pacific Grove, CA: Wadsworth and Brooks/Cole Advanced Books.

Bennett, L., & Piet, M. (1999). Standards for batterer intervention programs: In whose interest? *Violence Against Women, 5*(1), 6–24.

Cahn, N. R. (1992). Innovative approaches to the prosecution of domestic violence crimes: An overview. In E. S. Buzawa & C. G. Buzawa (Eds.), *Domestic* *violence: The changing criminal justice response* (pp. 161–180). Westport, CT: Auburn House.

Coulter, M. L. (1999). Police-reporting behavior and victim-police interactions as described by women in a domestic violence shelter. *Journal of Interpersonal Violence, 14*(12), 1290–1298.

Davis, R. C., Smith, B. E., & Nickles, L. B. (1998). The deterrent effect of prosecuting domestic violence misdemeanors. *Crime and Delinquency, 44*(3), 434–442.

Davis, R. C., & Taylor, B. G. (1999). Does batterer treatment reduce violence? A synthesis of the literature. *Women and Criminal Justice, 10*(2), 69–93.

Dawson, J. M., Smith, S. K., & DeFrances, C. J. (1993). *Prosecutors in state courts, 1992.* (Government document). Washington, DC: Bureau of Justice Statistics, Office of Justice Programs, U. S. Department of Justice.

Dawson, M., & Dinovitizer, R. (2001). Victim cooperation and the prosecution of domestic violence in a specialized court. *Justice Quarterly, 18*(3), 593–622.

DeHart, D. D., Kennerly, R. J., Burke, L. K., & Follingstad, D. R. (1999). Predictors of attrition in a treatment program for battering men. *Journal of Family Violence, 14*(1), 19–34.

Ebbesen, E. B., & Konečni, V. J. (1980). On the external validity of decision-making research: What do we know about decisions in the real world? In T. S. Wallsten (Ed.), *Cognitive processes in choice and decision behavior* (pp. 21–45). Hillsdale, NJ: Lawrence Erlbaum Associates.

Eisikovits, Z. C., & Edleson, J. L. (1989). Intervening with men who batter: A critical review of the literature. *Social Service Review, 63*(3), 384–414.

Fagen, J. (1996). *The criminalization of domestic violence: Promises and limits.* (Tech. Rep.). Washington, DC: National Institute of Justice, U.S. Department of Justice.

Ford, D. A., & Regoli, M. M. (1992). The prevalence impacts of policies for prosecuting wife batterers. In E. S. Buzawa & C. G. Buzawa (Eds.), *Domestic violence: The changing criminal justice response* (pp. 182–208). Westport, CT: Auburn House.

Ford, D. N. (1999). A behavioral approach to feedback loop dominance analysis. *System Dynamics Review, 15*(1), 3–36.

Forrester, J. W. (1971). *Principles of systems.* Waltham: Pegasus Communications, Inc.

Forrester, J. W. (1980). Information sources for modeling the national economy. *Journal of the American Statistical Association, 75*(371), 555–566.

Forrester, J. W. (1999). *Industrial dynamics.* Waltham: Pegasus Communications, Inc. (Original work published 1961)

Forrester, J. W., & Senge, P. M. (1980). Tests for building model confidence in system dynamics models. In G. P. Richardson (Ed.), *Modelling for management: Volume ii: Simulation in support of systems thinking.* (pp. 413–432). Brookfield, VT: Dartmouth Publishing Company, Ltd.

Garner, J., Fagan, J., & Maxwell, C. (1995). Published findings from the spouse assault replication program: A critical review. *Journal of Quantitative Criminology, 11*(1), 3–18.

Gondolf, E. W. (1990). An exploratory survey of court-mandated batterer programs. *Response*, *13*(3), 7–11.

Gondolf, E. W. (1995). *Batterer programs: What we know and need to know* (Tech. Rep.). Washington, DC.

Gondolf, E. W. (1999). A comparison of four batterer intervention systems: Do court referral, program length, and services matter? *Journal of Interpersonal Violence*, *14*(1), 41–61.

Gondolf, E. W. (2002). *Batterer intervention systems: Issues, outcomes, and recommendations.* Thousand Oaks, CA: SAGE.

Härdle, W. (1990). *Smoothing techniques: with implementation in s.* New York: Springer-Verlag.

Healey, K., Smith, C., & O'Sullivan, C. (1998). *Batterer intervention: Program approaches and criminal justice strategies* (Tech. Rep.). Washington, DC: National Institute of Justice, Office of Justice Programs, U. S. Department of Justice.

Hedström, P., & Swedberg, R. (Eds.). (1998). *Social mechanisms: An analytical approach to social theory.* New York: Cambridge University Press.

Heuman, M. (1975). A note on plea bargaining and case pressure. *Law and Society Review*, *9*, 515–528.

Hirschel, D., & Hutchison, I. W. (2001). The relative effects of offense, offender, and victim variables on the decision to prosecute domestic violence cases. *Violence Against Women*, *7*(1), 46–59.

Hollander-Blumoff, R. (1997). Getting to guilty: Plea bargaining as negotiation. *Harvard Neogtiation Law Review, 2*, 115–148.

Homer, J. B. (1983). Partial-model testing as a validation tool for system dynamics. *Proceedings of the 1983 Intl. System Dynamics Conference*, 260–263.

Hovmand, P. S. (2000). *Domestic violence and court mandated batterer intervention: A community model.* Paper presented at the international system dynamics conference, Bergen, Norway.

Kaufman, G. (1992). The mysterious disappearance of battered women in family therapists' offices: Male privilege colluding with male violence. *Journal of Marital and Family Therapy, 18*(3), 233–243.

Lakoff, G. (1987). *Women, fire, and dangerous things.* Chicago, IL: University of Chicago Press.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: the embodied mind and its challenge to western thought.* New York: Basic Books.

Lane, D. C. (2001). Rerum cognoscere causas: Part 1 - how do the ideas of system dynamics relate to traditional social theories and the voluntarism/determinism debate? *System Dynamics Review, 17*(2), 97–118.

Largo, T. W., Smith, P. K., Thrush, J. C., McCrohan, N., & Rafferty, J., A.and Christensen. (1999). *Violence in the lives of michigan women: Results of a statewide survey* (Tech. Rep.). Michigan Department of Community Health.

Lewis, H. R., & Papdimitriou, C. H. (1981). *Elements of the theory of computation.* Englewood Cliffs, NJ: Prentice-Hall, Inc.

Martin, M. E. (1994). Mandatory arrest for domestic violence: The courts' response. *Criminal Justice Review, 19*(2), 212–227.

Maxwell, S. R. (2000). Sanction threats in court-ordered programs:examining their effects on offenders mandated into drug treatment. *Crime and Delinquency, 46*(4), 542–563.

McAllister, H. A., & Bregman, N. J. (1986). Plea bargaining by prosecutors and defense attorneys: A decision theory approach. *Journal of Applied Psychology, 71*(4), 686–690.

McCoy, C. (1984). Determinate sentencing, plea bargaining bans, and hydraulic discretion in california. *The Justice System Journal, 9*(3), 256-275.

McCoy, C. (1993). *Police and plea bargaining: Victim's rights in california.* Philadelphia: University of Pennsylvania Press.

McDonald, W. F. (1985). *Plea bargaining: Critical issues and common practices* (Government document). Washington, DC: National Institute of Justice, U.S. Department of Justice.

Moxnes, E. (2000). Not only the tragedy of the commons: misperceptions of feedback and policies for sustainable development. *System Dynamics Review, 16*(4), 325–348.

Moyer, S. (1982). *Pretrial diversion: An overview* (Government document). Ottawa, Canada: Solicitor General.

Mullender, A. (1996). *Rethinking domestic violence: The social work and probation response.* New York: Routledge.

Özveren, C. M., & Willsky, A. S. (1990). Observability of discrete event dynamic systems. *IEEE Transactions on Automatic Control, 35*(7), 797–806.

Pence, E. (1999). Some thoughts on philosophy. In M. F. Shepard & E. L. Pence (Eds.), *Coordinating community responses to domestic violence: Lessons from duluth and beyond* (pp. 25–40). Thousand Oaks, CA: Sage Publications.

Pence, E., & Paymar, M. (1993). *Education groups for men who batter: The Duluth model.* New York: Springer Publishing.

Pennsylvania Coalition Against Domestic Assault. (1992). *Accountability: Program standards for batterer intervention services.* Retrieved November 23, 2002 from http://coment.org/bisc/standards/pcadv.html.

Petrik, N. D., Gildersleeve-High, L. G., McEllistrem, J. E., & Subotnik, L. S. (1994). The reduction of male abusiveness as a result of treatment: Reality or myth? *Journal of Family Violence, 9*(4), 307–316.

Rhodes, W. M. (1976). The economics of criminal courts: A theoretical and empirical investigation. *Journal of Legal Studies, 5*, 311–340.

Richardson, G. P. (1991). *Feedback thought in the social science and systems theory.* University of Pennsylvania Press.

Richardson, G. P., & Pugh, A. L. (1986). *Introduction to system dynamics modeling with dynamo.* Cambridge, MA: MIT Press.

Rigakos, G. S. (1997). Situational determinants of police responses to civil and criminal injunctions for battered women. *Violence Against Women, 3*(2), 204–216.

Robinson, A. L. (2000). The effect of a domestic violence policy change on police officers' schemata. *Criminal Justice and Behavior, 27*(5), 600–624.

Rozé, L., & Cordier, M. (2002). Diagnosing discrete-event systems: Extending the "diagnoser approach" to deal with telecommunication networks. *Discrete Event Systems: Theory and Applications, 12*, 43–81.

Sampath, M., Raja, S., Lafortune, S., Sinnamohideen, K., & Teneketzis, D. C. (1996). Failure diagnosis using discrete-event models. *IEEE Transactionss on Control Systems Technology, 4*(2), 105–124.

Sampath, M., Sengupta, R., Lafortune, L., Sinnamohideen, K., & Teneketzis, D. (1995). Diagnosibility of discrete-event systems. *IEEE Transactions on Automatic Control, 40*(9), 1555–1575.

Schechter, S. (1982). *Women and male violence: The visions and struggles of the battered women's movement.* Boston, MA: South End Press.

Schmidt, J., & Steury, E. H. (1989). Prosecutorial discretion in filing charges in domestic violence cases. *Criminology, 27*(3), 487–510.

Scott, D. W. (1992). *Multivariate density estimation: theory, practice, and visualization.* New York: John Wiley and Sons, Inc.

Shepared, M. F. (1999). Evaluating a coordinated community response. In M. F. Shepard & E. L. Pence (Eds.), *Coordinating community responses to domestic violence: Lessons from duluth and beyond* (pp. 169–191). Thousand Oaks, CA: Sage Publications.

Sherman, L. W., & Berk, R. A. (1984). The specific deterrent effects of arrest for domestic assault. *American Sociological Review, 49*, 261–272.

Sherman, L. W., Smith, D. A., Schmidt, J. D., & Rogan, D. P. (1992). Crime, punishment, and stake in conformity: Legal and informal control of domestic violence. *American Sociological Review*, *57*(5), 680–690.

Smith, A. (2000). It's my decision, isn't it? a research note on battered women's perceptions of mandatory intervention laws. *Violence Against Women*, *6*(12), 1384–1402.

Sonkin, J. D. (1988). The male batterer: Clinical and research issues. *Violence and Victims*, *3*(1), 65–79.

Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world.* Irwin McGraw-Hill.

Suppes, P., & Zinnes, J. L. (1963). Basic measurement theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 1–76). New York: John Wiley and Sons, Inc.

Taha, A. (2001). The equilibrium effect of legal rule changes: Are the federal sentencing guidelines being circumvented? *International Review of Law and Economics*, *21*, 251-269.

Taylor, B. G., Davis, R. C., & Maxwell, C. D. (2001). The effects of a group batterer treatment program: A randomized experiment in brooklyn. *Justice Quarterly*, *18*(1), 171–201.

Theil, H. (1966). *Applied economic forcasting.* Chicago: Rand McNally and Company.

Tjaden, P., & Thoennes, N. (1998). *Prevalence, incidence, and consequences of violence against women: Findings from the national violence against women* *survey* (Government document). National Institute of Justice and Centers for Disease Control and Prevention.

Tolman, R., & Edleson, J. (1989). Cognitive-behavioral intervention with men who batter. In B. A. Thyer (Ed.), *Behavioral family therapy* (pp. 169–190). Springfield: Charles C Thomas Publisher.

Venables, W. N., & Ripley, B. D. (2000). *S programming.* New York: Springer-Gerlag.

Wittgenstein, L. (1974). *Tractatus logico-philosophicus* (D. F. Pears & B. F. McGuinness, Trans.). New York: Routledge. (Original work published 1921)

Yin, R. K. (1994). *Case study research: Design and methods* (second ed.). Thousand Oaks, CA: Sage Publications.

Young, I. M. (1990). *Justice and the politics of difference.* Princeton, NJ: Princeton University Press.

Zorza, J., & Woods, L. (1994). *Analysis and policy implications of the new domestic violence studies* (Tech. Rep.). National Center on Women and Family Law.