 Supporting Material is available for this work. For more information, follow the link from the Table of Contents to "Accessing Supporting Material".

USING GROUP MODEL BUILDING TO INFORM WELFARE REFORM POLICY-MAKING IN NEW YORK STATE: A CRITICAL LOOK[†]

Aldo A. Zagonel

Nelson A. Rockefeller College of Public Affairs and Policy
University at Albany, State University of New York
E-mail address: Zagonel@aol.com

Proceedings of the 2003 International Conference of the System Dynamics Society.
New York City, U.S.A., July 20-24. (Draft dated May 2003)

Building system dynamics models directly with groups requires skillful combination of problem analysis and group facilitation. This paper provides an overview of an applied research project – using this growing and innovative approach – to model and analyze a complex social issue in the public sector. It provides a synopsis of the design and products of this intervention involving three counties in New York State. It reports how GMB was used to inform welfare reform policy-making, drawing upon the perspectives and knowledge of key welfare service providers. This paper is focused upon the final product of this intervention – a system dynamics model used to experiment with management strategies and to explore scenarios. It addresses the policy aspect of the research, and is focused upon documenting the model elicited, built, simulated, evaluated, tested and extensively used throughout this intervention. The analysis is organized around three policy questions. The conclusion is a summary of the findings, and a discussion of the limitations of the analysis in terms of policy evaluation for the purpose of system's redesign.

Key words: Group model building, system dynamics, decision conferencing, welfare reform, TANF, Safety Net, boundary-object, micro-world, management flight simulator.

Intellectual heritage

Approaches to systems thinking, strategic planning, decision analysis, decision support, and decision conferencing are increasingly coming to rely upon the practice of building models directly with management teams and decision-making groups. The objectives of these researchers and practitioners are manifold, ranging from improving group decision-making processes to enhancing group, team and organizational effectiveness and productivity (Andersen *et al.*, 1997).

[†] This research was conducted in the context of a team effort, with John Rohrbaugh, David Andersen, George Richardson, Robert Johnson and Tsuey-Ping Lee. Although I played an important role in this project, I cannot report on this work detached of their contributions. In my work, and in writing this paper, I am deeply indebted to David, John and George for their crucial mentorship. I also would like to acknowledge the support in this project of Irene Lurie and Naiyi Hsiao. While sharing with my research colleagues in the joy and ownership of this research, I solely take responsibility for any mistakes contained in the model and in this analytic report. This work was successful in large part due to the level of interest presented by the Commissioners of Social Services in the participating counties: Jane Rogers, Robert Allers and Irene Lapidez, and the genuine engagement of their management teams. I am very grateful for having played a part in this collaborative study, and to now contribute with this partial testimony.

The importance of involving the clients in the process of model building has been recognized early on in the field of system dynamics (Forrester, 1961; Roberts, 1978). Stenberg (1980) first introduced the practice of drawing upon a reference group, as a “mini-universe of the part of the real world under study” (p. 303). A recent development involves more active client-group engagement especially but not exclusively in the conceptual phase of model building, in the form of group meetings or conferences. This line of research and practice has been termed group model building (Richardson *et al.*, 1992; Vennix *et al.*, 1997). Vennix (1996) characterizes it as a kind of group decision support for helping teams tackle strategic problems (p. xi). Gradually, a unified body of knowledge containing methodological guidelines to develop group model building procedures is flourishing (Richardson and Andersen, 1995; Vennix, 1996; Andersen and Richardson, 1997). As they are experimented with, these procedures are also being evaluated (Rouwette *et al.*, 2002).

Group model building is based essentially in the system dynamics model building method. However, deeply involving a client group in the process of model construction has required theoretical and applied input from other fields, such as sociology, social psychology, and small-group research (Vennix, 1999, p. 379). In the applied research on the subject of group model building being conducted in Albany, this influence has been refined and expanded in a framework called decision conferencing (Rohrbaugh, 2000).

While one could probably apply group model building to modeling any sort of dynamic problem, it seems that particular kinds of problems or situations “attract” the application of this approach. Group model building interventions will often address problems that involve multiple stakeholders that contribute with partial views of the system, but who are affected by the system as a whole (Huz *et al.*, 1997). They are also particularly useful in situations in which there is strong inter-personal disagreement in the client group, regarding the problem and/or regarding the policies that govern system behavior. Vennix (1999) refers to the latter as messy problems, *i.e.*, “a situation in which opinions in a management team differ considerably” (p. 379).

Zagonel (2002a) traced a genealogy of group model building, from the point of view of the approach used by the research group working at the University at Albany. Two schools of thought were identified as contributing to this approach, as illustrated in Figure 1.

Figure 1. Tracing a genealogy of group model building

(See p. 44)

System dynamics is at the root of the *policy* thread. The system dynamics model building method can be described in phases that begin with a clear definition of the problem of interest, and end with a conclusive statement about this problem, containing policy recommendations aimed at its solution or mitigation (Richardson and Pugh, 1981, pp. 15-17). This method is based upon an endogenous feedback view of system causes and effects. Solutions to the perceived problem are revealed through feedback thinking, the key expertise offered by system dynamicists (Serman, 1994).

The second thread, labeled as the *decision* thread, is formed by a confluence of schools that gave shape to the decision conferencing framework. Those are group dynamics, decision analysis and decision support (Rohrbaugh, 2000). People who conduct decision conferences consider themselves technique/process experts, and they focus upon the appropriate techniques and the best processes used to arrive at decisions (Reagan *et al.*, 1991). They help structure problem solving while focusing upon facilitation and elicitation strategies and techniques.

An important characteristic of group model building is its diversity in objectives and expectations, maybe resulting from the confluence of the varied influences giving shape to it (system dynamics, small-group dynamics, decision support, etc.). A superficial examination of its genealogy alone reveals a tension between policy *versus* decision, and between content *versus* process. To some extent these tensions overlap.

The decision conferencing influence emphasizes a decision to be made, and focuses upon the processes that lead up to this decision. *Decision or process oriented objectives* in group model building may be stated as accelerating a management team's work, problem structuring and classification schemes, generating commitment to a decision, creating a shared vision and promoting alignment, and creating agreement or building consensus about a policy or decision. Alternatively, *policy or content oriented objectives* may be stated as improving shared understanding regarding the system or problem at hand, system improvement, and system process and outcome change. These involve changing the mental models of individuals in the group or organization, guided by reliable insights produced using the modeling tools and methods.

Ideally, as Eden (1990) appropriately pointed out, *astute analysis (content) and skillful facilitation (process) should be combined*: “within the context of group decision support it may be suggested that the two skills can become integrally tied together so that they are fully interdependent” (p. 49). In the context of group model building, this may be stated as promoting organizational learning and organizational change, or promoting insightful collaboration and cooperation amongst interdependent stakeholders.

While this may be an ideal, Zagonel (2002a) suggests that there are tensions related to bringing these two skills together. The objectives and procedures for using the model building process as a tool for creating a shared understanding of an interpersonal or inter-organizational problem –in the form of a “boundary-object” model, and as a tool for exploring a “micro-world” representation of reality –to address this particular problem, are not necessarily aligned. While GMB interventions are designed to achieve this ideal goal by intertwining these two threads, in reality, one or both may be sacrificed.

This paper takes a critical look at an applied research project –using this innovative approach– to model and analyze a complex social issue in the public sector. It reports how GMB was used to inform welfare reform policy-making, drawing upon the perspectives and knowledge of the client teams. It addresses the policy aspects of the research, by answering three policy questions using the model elicited, built, simulated, evaluated, tested and extensively used throughout the intervention. The conclusion

contains a summary of the findings, and a discussion of the limitations of the analysis *in terms of policy evaluation for the purpose of system's redesign*.

Synopsis of the process and products of the intervention

This work goes under the title of *Group Model Building for Welfare Reform (GMB/WR)* project, or simply *Welfare Reform (WR)* project. It took place in 1997 and 1998, and it lasted approximately 22 months. It evolved gradually from qualitative reflection to quantitative inquiry,¹ from small to large county, and from modeling sectors to a fully combined model. However, the apparently sequential stream of activities was actually formed by *parallel* streams that were linked (back and forth) at certain points in time. Figure 2 serves as an illustration. It is the timeline of this project. It specifies vertically the distinct but interconnected streams of activities, and reveals horizontally how each activity unfolded over time. Most of the actual meetings with the client-teams were documented in the form of meeting reports, referenced in Appendix 1. A detailed view of the “road-map” to this project is available at the end of this appendix.²

Figure 2. Streams of activities of the Welfare Reform project

(See p. 45)

The first half of this paper contains a synopsis of the design and of the products of this project, as background information to the overview, in the second half, of the final welfare reform model, produced as an overall result of this group model building effort. I begin this description with some comments regarding project negotiation and design. Then I cover each of the streams of activity, as they unfolded, chronologically. First, I describe the initial development of the TANF model in Cortland County; second, the development of the safety-net model in Dutchess; third, the joining of these two sectors to create a single model; fourth, the work involved in preparing the full model for roll out, buy-in and use in Nassau. Also, I outline general efforts (not county-based) to reach out to policy makers and managers in the welfare system.

¹ Several authors have found useful to describe the system dynamics method in terms of its phases (Andersen and Richardson, 1980, p. 93; Richardson and Pugh, 1981, p. 16; Roberts *et al.*, 1983, p. 8; Sterman, 2000, p. 87). Zagonel (2002a) depicted these phases as a sequence of iterative *steps*, as in climbing up and down a ladder (p. 12, Figure 4). The first two steps are comprised of the phases associated with *qualitative reflection*: problem identification and definition, and model conceptualization. Next come three steps associated with *quantitative inquiry*: model formulation and simulation, model testing and evaluation, and model based problem analysis and policy experiments. Iteration takes place both within each cluster of steps, and across clusters. At any point in the process, there exists some degree of understanding and discernment regarding the problem and the system under study. But, as one climbs toward the higher steps, the level of understanding and discernment improves and gains accuracy (p. 13).

² Appendix 1 also contains a complete list of references of papers, presentations, workshops, briefings and manuals related to this project.

In discussing each stream of activity, I highlight the main products achieved and briefly record how these products served as inputs to the other streams. It should become clear how we gradually built and refined the final model that will be discussed in some detail in the second half of this paper. But first, as a prelude, a brief statement of the issues involved in this policy domain is in order. The next section concisely summarizes the evolution of social welfare services in the United States, with emphasis placed upon the enactment and implementation of the 1996 Welfare Reform Act, and the envisioned consequences of this bill in terms of its impact in the welfare system.

Welfare reform

The history of social welfare in the United States has come full cycle. In 1935, amidst the Great Depression, the federal government assumed greater responsibility for the needy through the enactment of the Social Security Act, and the implementation of a federally funded program, called *Aid to Families with Dependent Children (AFDC)*. Recently, there was devolution of this task to state and local governments (Thompson, 1996). This devolution occurred in the form of the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA), enacted in 1996, and gradually implemented in five years, between 1997 and 2002.

In brief, the act ended a federal entitlement, and it limited benefits to a cumulative five-year period per family --after which the family loses eligibility to the federally funded *Temporary Assistance to Needy Families (TANF)* program, which replaced AFDC. The law also switched the funding mechanism from matching to block grants, thereby capping federal spending while providing states and localities with more managerial flexibility.

The shift in focus, from monetary assistance (through AFDC) to services designed to promote work and avoid dependency (through TANF), was actually initiated earlier, with the Family Support Act, of 1988, designed to promote welfare-to-work initiatives (Lurie, 1997). While this transitional experience contributed to understanding the new mode of operation, and to initial restructuring of services and programs, the full impact of the newest legislation was to unfold only as states and municipalities strove to fully implement it. Thus, enactment of PRWORA was accompanied by criticism that its measures were too harsh on the poor, or simply not thought out carefully enough (Edelman, 1997).

The new Act, in fact, represented a political compromise involving the President, Congress and state governors, in which a *greater degree of flexibility* in designing and operating the programs was traded for a *federal spending cap* represented by the shift in the funding mechanism from matching to block grants. This Act resulted from frustration in the political realm with the incapacity of welfare program administrators to contain increases in program costs; coupled with a sense that the welfare culture, focused on assistance, was stubborn to change. Thus, the new law gave state governments authority to devise their own solutions to end welfare dependency, or bear the costs of sustaining

those people unable to attain self-sufficiency, after five cumulative years of receiving federally funded benefits.

In this compromise, it is possible that short-term political (and perhaps electoral) gains overshadowed long-term uncertainties and dangers. Federal welfare spending was actually higher than it would have been without the law, due to the fact that states could choose, among the years between 1992 and 1995, the level of spending at which their block grant was capped. Since caseloads, in general, had been falling, “the states got a sweet deal in the short term” (Lurie, p. 77). On the other hand, inflation, future economic downturns, and loss of eligibility after five years, are some of the threats to financial viability that the states and locals face in the future.

Despite skepticism from economists, who argued that redistributive programs are best handled at the federal level (Lurie, p. 74), even critics of welfare reform cheered the fact that the new Act promoted a new vision for welfare programs, based upon the concept of individual responsibility; and that the reform also shifted decision-making power closer to where the action is. As previously stated, there is no doubt that initiatives redirecting welfare programs from mere monetary assistance to services that promote self-sufficiency were already in place. But, PRWORA put an entirely new emphasis on them. Whereas the old system was said to have contributed to chronic dependency (Edelman, p. 43), the end of the entitlement, the absence of matching funds, and the time limitation for drawing upon federal dollars, forced the states and recipients to “address this crisis or suffer the consequences” (Lurie, p. 76). Under this kind of pressure, it was quite likely that cultural changes in welfare agencies would occur sooner, rather than later.³

The new act increased state and local flexibility and discretion in designing and running programs. States could now choose to transfer a portion of federal dollars to child-care or other social services deemed important to accomplish the legislative intent of the Act. They could reduce monetary assistance in favor of other means of support. They could establish governmental and non-governmental partnerships to carryout programs. And, they could promote innovative means to reduce caseloads, such as an increased emphasis on child support enforcement services. With enriched resources in the short term, and with added flexibility, states *with foresight and discipline* could pursue the goals of the legislation without immediate pressure to impose restrictions on eligibility or benefit levels (Lurie, p. 77).

Why is foresight and discipline necessary? There is a lot of uncertainty regarding the effects of this new policy structure in the medium to long term. Can it be realistically expected that recipients will be able to move into permanent, well-paying jobs? Exactly

³ Lurie and Riccucci (2001) examined more closely this issue of cultural change in welfare offices post welfare reform. While they observed changes in rules, structures and processes (within the control of agency executives), it was not clear that the caseworkers’ basic assumptions had actually changed.

how big will the damage be in five years and beyond, when hard-core welfare recipients begin to lose eligibility to federally funded TANF (beginning in 2002)?

WELFARE-TO-WORK, JOBS AND POVERTY. The first question has to do with the availability of employment opportunities for TANF recipients. This issue may be broken down into several existing barriers. First, there are natural limits to the recipients' abilities to move into permanent work. Studies have indicated that attaining 50 percent of the caseload meeting work requirements is near maximum one can achieve (Edelman, p. 53), and that one in every three persons coming on to welfare remains on the rolls for over five years (Pavetti, 1995). The fact is that, often, local jobs for welfare recipients are simply not there. Second, monetary policy impedes unemployment to fall below certain levels. Third, there are other reasons –regardless of the availability of jobs– preventing recipients from being able to work (e.g., drug dependency). Fourth, the jobs that are available present low earning levels (*i.e.*, low wages), and do not provide the means for recipients to rise above the poverty level. These and other barriers may stand in the way of meeting PRWORA's work targets.⁴ If so, the modest 20 percent exemption allowed in the caseload –with respect to the five-year limit requirement (Edelman, p. 50)– may not be enough to bring managerial relief to program administrators.

WHAT HAPPENS AFTER 2002? In 1997, when this intervention was taking place, it seemed clear that even the best experts on the subject had their eyes glazed at the size and complexity of this problem. Welfare as we knew it had ended, but what would take its place? Welfare services providers asked themselves and experts: "How big will the damage be in five years and beyond?" This question combined the above uncertainties plus some. How was this going to impact the community? Would it increase homelessness and violence? By how much will it increase? Would private charities and the alike provide more shelters and food pantries for the needy? And so on. In addition to the uncertainties raised by PRWORA, critics argued there was not enough in terms of consolidation to promote flexibility; historically low benefit states were capped at an disadvantage; and it was unclear just how the federal government would enforce time limits, given that it relied on the states themselves for information (Lurie, 1997).

Welfare reform implementation in New York State (NYS)

That same year, Governor Pataki proposed a package of reforms at the state level that would provide for the implementation of PRWORA within the specific NYS context. The NYS reform maintained the entitlement by creating the *Safety Net (SN)* program, thus, respecting Article 17 of the State Constitution, which mandates that the State provide for its poor. As Governor Pataki's reform package was scheduled for debate in the Legislature, researchers within the NYS Department of Social Services approached the

⁴ In PRWORA, the minimum target for work participation rose to 50 percent by the year 2002 for single-parent families, and was set at 75 percent in 1997, raising to 90 percent in two years, for two-parent families (Edelman, p. 50).

Center for Policy Research at the University at Albany with a proposal to undertake a series of group model building sessions designed to focus on how *local* providers might respond to both federal and state reforms. (In NYS, it is the counties that actually provide social welfare services, and those services must conform to state and federal guidelines and funding policies.)

These group model building sessions were scheduled to take place in one small rural county, one medium-sized county, and one large county. *The purpose of these efforts was three-fold: (1) to assist the participating counties to think through welfare reform strategies using a system dynamics modeling framework, (2) to provide state policy makers with an opportunity to observe the group model building process and, hence, to learn how local communities were likely to respond to state and federal initiatives, and (3) to create a management flight simulator for welfare reform viewed from the county level.*⁵ The three counties directly involved in this project were Cortland, Dutchess, and Nassau.

Drawing upon the collective knowledge of management teams,⁶ the modeling team⁷ at the University at Albany produced a series of system dynamics models. In Cortland, the focus of our discussions was on TANF; in Dutchess, it was on the safety net. In preparation for Nassau, we developed a joined model. In building these models, we followed the group model building method devised in Albany (Richardson and Andersen, 1995; Andersen and Richardson, 1997; Rohrbaugh, 2000; Zagonel, 2002a), while incorporating into this approach specific frameworks developed by others (Checkland, 1981; Eden, 1989; Hodgson, 1994; Wolstenholme, 1994; Bryson, 1995; Vennix, 1996; Eden and Ackermann, 1998).

⁵ Project proposal (March 3, 1997). "System Dynamics Modeling for Welfare Reform." Center for Policy Research, University at Albany, State University of New York.

⁶ The *management teams* are the practitioners engaged in the group model building effort, also referred to as the client groups or client teams. In this project, the participants were primarily, but not exclusively, high-level local managers of social welfare agencies. Other agencies involved were health, mental health, labor, and education. State personnel often observed the sessions and sometimes participated more actively (in particular, when the TANF and safety-net models were joined). On occasion, a local legislator was present. Some non-governmental actors also participated (in particular, representatives of private charities). At each site, the results of the modeling work were presented to a broader group of community leaders.

⁷ Richardson and Andersen (1995) referred to the *modeling team* in a GMB session as ideally composed of individuals performing five roles: (1) facilitator, (2) modeler/reflector, (3) process coach, (4) recorder, and (5) gatekeeper. Except for the gatekeeper, a role filled by "a person within, or related to, the client group" (p. 115), all of the other roles were performed by Rockefeller College faculty and students. In addition to these roles, this particular intervention had a model-builder (6), was supported by a data and parameter recorder (7), and an interface developer (8). Dr. Irene Lurie—an economist who is also a specialist in social welfare policy—assisted the modeling team, taking the role of substance expert (8).

PROJECT NEGOTIATION AND DESIGN.⁸ The basic plan for this project was conceived in discussions between John Rohrbaugh –a faculty member at the Rockefeller College, University at Albany, and Robert Johnson –a researcher at the NYS Office of Temporary and Disability Assistance – OTDA (then Department of Social Services). Johnson made contacts within OTDA and promoted the idea of three county-level models culminating in a management flight simulator, and Rohrbaugh lined up a university-based modeling team to complete the technical modeling portions of the project.

Rohrbaugh suggested to Johnson a group model building approach, based upon prior experience using system dynamics modeling in decision conferencing (detailed in Rohrbaugh, 2000). Johnson frequently used a team approach in his work in OTDA, and he knew that Cortland County had an active management team, led by an energetic commissioner. He approached the Cortland team and arranged for the first full-day modeling conference. They had had prior exposure to systems approaches and were enthusiastic about working on the project. The other two counties involved were chosen later, along the same lines. Thus, in all three counties, we tried to work with commissioners who were open to a system's view, and *who were willing to involve their management teams* in the processes of problem definition, decision-making and change.

A first TANF model (February-April, 1997)

This intervention began with a sequence of four full-day meetings based upon prior experience in group model building documented in Richardson and Andersen (1995) and Andersen and Richardson (1997). The first meeting was an *issue elicitation* (i.e., problem definition) meeting, on February 11, 1997 (C-1)⁹. The pair of second and third meetings occurred five weeks later, for the purpose of *model elicitation* (i.e., model conceptualization), on March 17 and 18 (C-2/3). The fourth meeting was held after another five weeks, for the purpose of *model presentation*, on April 29 (C-4). In addition to the meetings themselves, documented in detail in meeting reports (Zagonel *et al.*, 1997a-d), there was extensive in-between-meetings work, both to prepare elicitation (and presentation) roles and scripts,¹⁰ to plan the meetings,¹¹ and to develop the model.

⁸ The proposal to this project contains a figure with the proposed timeline. There are significant differences between this proposed timeline and the actual timeline depicted in Figure 2. Perhaps the most significant is that the proposal laid out the project in 24 weeks, while actual implementation took nearly four times longer (22 months).

⁹ The meetings are identified in Appendix 1 with a letter and number. The letters stand for: C-Cortland; D-Dutchess; J-joined model (State); N-Nassau; B-briefings to OTDA; W-workshops; P-presentations; A-articles and papers; and T-dissertations. The numbers are sequential chronologically.

¹⁰ The different *roles* of the members of the modeling team are discussed in Richardson and Andersen (1995). *Scripts* are “sophisticated pieces of small group process” (Andersen and Richardson 1997, p. 107), “planned and rehearsed for accomplishing sub goals in the course of a group model building workshop” (Richardson and Andersen 1995, p. 130). For detailed

Preliminary contacts between the modeling team in Albany, and the client team in Cortland, involved only Rohrbaugh, Johnson, and Jane Rogers –the Commissioner in Cortland. The exact purpose of the modeling effort was not defined ahead of time, and the decision to model the TANF program resulted from focused discussions carried out during the first formal day of work with the local management team (C-1; Zagonel *et al.*, 1997a). It is not possible to expand here the description of what actually took place in these meetings and between meetings. For more information, the reader is invited to see the reports and papers available (see Appendix 1). A separate paper focused upon the *processes* of group model building used in this intervention is being written. It will describe in detail the roles played and the scripts used for the Cortland stream of this intervention (per Zagonel, 2002b).¹²

The model assembled for the Cortland management team was built in layers, with additional complexity being added one stage at a time. We began with concept models¹³ (Welfare 1-3)¹⁴ to spur their interest and engagement. The model elicitation meetings strove for the *clients' view* of the system (C-2/3). Each day ended on a high note, with the expert modeler reflecting back to the participants the key feedback-rich insights the group generated during the day.¹⁵ This moment of the session has been termed as “modeler’s

illustrations of scripts designed to achieve different purposes, see Andersen and Richardson (1997).

¹¹ For more information on planning the meetings (*scheduling the day*), see Andersen and Richardson (1997, pp. 112-117).

¹² Upcoming as the third essay of my dissertation (in progress, expected Fall/03): “Modeling TANF in Cortland County: Chronology of a group model building intervention” in *Reflecting on Group Model Building Used to Support Welfare Reform in New York State*.

¹³ *Concept models* are used to introduce system dynamics icons, to demonstrate the connection between structure and behavior, and to initiate discussion with the client team about the real system (Richardson and Andersen, 1995, p. 130). These models are very simple and usually purposively “wrong,” and they serve the additional objective of motivating the client team to engage, take control, and begin re-formulating the model to their liking. Some guidelines in building and using concept models can be found in Richardson and Andersen (pp. 130-131), and Andersen and Richardson (1997, pp. 116-117).

¹⁴ Appendix 1 refers to the particular model and interface versions used in each of the meetings of the project.

¹⁵ In “Teamwork in Group Model Building,” Richardson and Andersen (1995) polished the role of the system dynamics modeler:

[The modeler acts] as a reflector on the group’s discussion, a “contemplator” whose job [is] to refine and crystallize the thinking of the group ... there is great value to having a person reflecting on the group’s thinking and reflecting it back to them. The modeler/reflector can perceive subtleties the facilitator might miss,

feedback insights” (Zagonel *et al.*, 1997b). The model that was actually built captured many of these insights, refined through this process of reflection based upon system dynamics expertise. Version “Phase_6e” of the TANF model was used both at the model presentation meeting in Cortland (C-4; Zagonel *et al.*, 1997c), and in reporting the first part of this work (A-1).¹⁶

The Commissioner of Cortland County summed up this early part of the intervention with the following statement: (Rogers *et al.*, 1997, pp. 610-611)

The model provokes the participants to examine the impact of changes in the administration of welfare programs as having community-wide consequences. Traditionally, changes in this department’s budgeting have been viewed as being an isolated problem, which allowed a “business as usual” response for providers of services... The development process involved promotes the creation of a grass roots plan... The model provides a structure for discussions for community-wide strategy sessions by identifying validated high impact points in the system. Welfare reform partners are able to create community-acceptable goals... Most importantly, the simulations vividly expose the system of welfare by portraying the community-wide opportunities for investment and risk.

There were four other full days of group meetings in Cortland County, associated with model improvement, community presentations and model use (C-5/6/7/8). These meetings will be discussed ahead in the section entitled “responding to client needs.” At this point in the project, however, the focus of our attention turned to the second county.

The safety-net model (May-July, 1997)

The first four meetings in Dutchess (D-1/2/3/4) followed the same format described above. The Commissioner Robert Allers was familiar with the TANF work done in Cortland, and had decided before-hand that the focus of the Dutchess meetings should be in developing a model for the safety net. During the first half of the first day (D-1), on May 30, 1997, we explored the nature of this sector, and the resources, policies and actions available in it. The second half of the day was used to present to the client team the existing TANF model,¹⁷ and to show to the group the base-run, and two policy simulations. A few days later, on June 3 and 4, the structure and parameters of the safety-net model were elicited (D-2/3). The fourth meeting, for model presentation (D-4), was

can identify linkages and systems insights that emerge only from reflection, and can punctuate the discussion with points of important emphasis. (p. 124)

¹⁶ Model diagrams and equations for TANF, Phase_6e, are contained in Zagonel *et al.* (1997d) and Johnson *et al.* (1997). These reports also include the content of meetings C-1 through C-4.

¹⁷ At this point, already version Phase_8a, containing improvements following the model presentation meeting in Cortland (C-4).

held seven weeks later, on July 29. These meetings were also documented in detail (Rohrbaugh *et al.*, 1997).

The safety-net model assembled for the Dutchess management team also was built in layers. At the same time, Cortland's TANF model was adjusted to reflect Dutchess' numbers and parameters. The model presentation meeting drew upon both models: Safety-net (SN_2), and TANF (Phase_8e). There were two other full-day group meetings in Dutchess County, one for model improvement (D-5), and the other for a community presentation of the modeling work and insights (D-6). These meetings are discussed in the next sections.

RESPONDING TO CLIENT NEEDS (JUNE-OCTOBER, 1997). In this project we had one paying client, and at least three other non-paying clients. As previously stated, we had a contract with OTDA, in which we promised to deliver a stand-alone management flight simulator for welfare reform, viewed from the county level. In order to build the WR model, we were drawing upon the county-based management teams for model elicitation and validation. However, the development of the models (in meetings 1 through 4) spurred interest on the part of the local client teams to actually *use* the models for some purpose at the local level.

The Commissioners in both Cortland and Dutchess were interested in using the models to share their views of the welfare system, and of welfare reform, with their communities at large. The larger boundary depicted in the models made vividly clear the community-wide implications of delivering social services, revealing inter-organizational effects and dependencies. But, they felt the models need to be validated before they were willing to open them to the community. While they were very pleased with the structure of the models, and the behaviors of the simulations, they were concerned with making sure that the models depicted "their numbers" correctly.¹⁸ Also, there were plenty suggestions to add detail complexity to the model. The modeling team strove to strike a balance between these individual client needs, and the overall purpose of the modeling effort.

In terms of model improvement and group model building practice, the most fruitful outcomes of these requests were the development and use of new scripts designed for group processes related to *model parameterization and calibration*. Andersen and

¹⁸ The closing remarks of the model presentation meetings in Cortland and Dutchess (C-4 and D-4) contain the following comments and suggestions made by the participants: "Huge caution: the numbers everywhere in the model need to be checked" (Zagonel *et al.*, 1997c, p. 45), and "fine tune the numbers in the model to be completely consistent with Dutchess County numbers" (Rohrbaugh *et al.*, 1997, p. 44). We found this to be a recurring theme in this project. In fact, none of the counties involved were willing to engage in a generic model, and were only willing to use the model after it was parameterized and calibrated to their numbers. The last version of the management flight simulator also reflects this problem, giving the user the ability to select as a scenario the size of the county.

Richardson (1997) already described scripts for supporting equation writing and parameterization (pp. 122-124), and these were used in the course of meetings 1 through 4 (see Zagonel *et al.*, 1997d, and Rohrbaugh *et al.*, 1997). However, the concerns raised and the modeling team's attention devoted to resolving them, led to some creative ideas and some innovations in these procedures. An initial attempt to whiz through these issues with the management team in Cortland County (C-5), on June 6 (only two days after the intensive model elicitation meetings in Dutchess, D-2/3), led to more careful preparation and planning for a second meeting in Dutchess (D-5) several months later, on October 16.¹⁹

While the parameterization and calibration meeting in Cortland was not documented (C-5), the meeting in Dutchess was (D-5). The report (Zagonel *et al.*, 1997e) identifies which parameters were elicited using expert judgments, how resources were computed and aggregated, and which table functions were elicited directly from the participants (only five). Two new protocols were used in this meeting. Because of the large number of table functions used in the models (and the time constraint to elicit each and every one of them directly from the expert team), a column-by-column and row-by-row procedure was used to identify the most significant table functions in the safety-net model (SN-2), to guide the elicitation, and to triangulate logically to the other table functions, in both safety-net and TANF (Phase_8f) models (pp. 11-14). Finally, cross-sectional calibration (in equilibrium) was done directly with the client group, using a spreadsheet equipped with stock-and-flow visualization (pp. 18-22).²⁰ These parameterization and calibration activities were pursued both to accommodate the concerns of the county teams, and to the advance the goal of developing a generic welfare model at the state level.

While this work was being done, development of a prototype for the management flight simulator interface was also underway. To advance this main purpose of the project, David Andersen and George Richardson decided to incorporate the WR work into a sequence of systems thinking *workshops* scheduled with the NYS Governor's Office of Employee Relations (GOER). Since many of the participants of this workshop were social welfare services managers and staff, it seemed to be an ideal circumstance to develop and experiment with a training module based upon hands-on use of the model. The workshop began on June 9 and 10 (W-1), when Systems Thinking was introduced, and soft systems analysis based upon a specially developed case was conducted.²¹ This

¹⁹ The Cortland meeting (C-5) used model version Phase_8a, while the Dutchess meeting (D-5) used both models, versions SN_2 and Phase_8f.

²⁰ Lee *et al.* (1998c) provides a review of estimation techniques and describes the procedures used in the Dutchess parameterization and calibration meeting (D-5). The paper also contains a discussion of their strengths and weaknesses, and suggestions for future work (A-5). Zagonel *et al.* (1998) document a similar meeting in Nassau County (N-2).

²¹ "Systems Thinking: A case-study on welfare reform in *Chugwa* County."

meeting ended with a demonstration of the TANF model (Phase_8a), in anticipation of the participants' hands-on use of the model in the follow-up meetings.

In the mean time, the Cortland management team was more comfortable with the modeling effort and eager to show the work to their partners in the community. A facilitated *community-wide presentation meeting* was held on July 8 (C-6). This presentation included a detailed overview of the structure of the TANF model (Phase_8d), and several simulations of both policies and scenarios, tailored by subgroups and contrasted by all. The community-wide implications of welfare reform were explored linking model behavior to system structure. Model based insights contrasting high and low leverage strategies were explained in terms of the stock-and-flow structure, and reinforcing and balancing loop-structures in the system (the essence of which is documented in Rogers *et al.*, 1997). During breaks, the participants experimented, hands-on, with a preliminary version of the management flight simulator interface (based on model version Phase_8d).

The follow-up to the GOER promoted workshop was held on September 8 and 9 (W-2). The structure of the TANF model was reviewed (Phase_8e). Then, the participants engaged in hands-on work with the second version of the simulator (based on model version Phase_8e), in guided simulation exercises (in pairs) using a tailored user's manual (Richardson *et al.*, 1997a). Whole-group debriefing followed the exercises.

These workshops at the state level were followed by a *resource allocation conference* in Cortland County, on September 29 and 30 (C-7/8). In this case it was Rohrbaugh who decided to volunteer his time and, upon request, promote model use for the purpose of decision-making. The conference counted with 45 participants from 30 agencies and organizations in the county. (Appendix 2 contains the list with their names and affiliated organizations. The ten members of the original management team are highlighted with an asterisk (*).) The conference began with the presentation of the TANF model (Phase_8e), and policy and scenario simulations. Based upon insights derived from this model and group discussion, the participants agreed on a set of planning assumptions, which led to the creation of three new countywide task forces (Rohrbaugh, 2000, p. 530):

- Teen life skills – to target resources to prevention services;
- Day care – to target resources to improved employment services and self-sufficiency promotion; and
- Jobs center – to target resources to a one-stop service center for job skills assessment and training, job finding, and job mentoring.

Also, they developed consensus behind five proposals, costing \$675 thousand dollars, which were widely agreed as the most important, immediate initiatives that could be undertaken in Cortland County (Johnson *et al.*, 1997):

- Job center (\$150K);
- Resource center (150K) – to coordinate community efforts toward diversion;

- A program to support employed self-sufficiency (\$200K) – involving case managers and employers in job counseling to keep individuals on the job;
- Computer-based comprehensive assistance (150K) – to link all providers and to develop a common resource information and referral database; and
- Expansion of child-care services (\$25K).

A second resource allocation conference was conducted a year later, on September 28, 1998 (C-9). On this occasion we used the final WR model described in the second-half of the paper (Jnd_6n). The Dutchess team did not have a resource allocation conference, but they did bring together their own facilitated community-wide presentation meeting (D-6), similar to the one done in Cortland (C-6). But they waited until much later, on April 28, of 1998, when the models were already joined (Jnd_6) and parameterized.

Joining the two models (November, 1997 – January, 1998)

After the model parameterization and calibration meeting in Dutchess in October (D-5), we began preparing to join the two models. However, a host of issues needed clarification, including important questions related to model purpose and boundary. To clarify these issues, two meetings were assembled. The first meeting involved a small group of financial experts at the state level. The purpose of this meeting was to review the finances sector of the models, and to propose and discuss corrections and improvements to be incorporated in the joined TANF/safety-net model. Some of the specific issues of concern were the separation of program and administrative costs, and how to model the administrative caps imposed through federal and state regulations. This meeting took place on December 1, 1997 (J-1; Zagonel, 1997).²²

The other, more important meeting had the participation of all of the major clients of this modeling effort, including the Commissioners of Cortland and Dutchess Counties. Based upon a preliminary discussion of model purpose and audience, the group made decisions regarding the client-flow structure of the model, and the form of aggregation of services to form resource clusters. This second meeting took place on December 16 (J-2). These meetings drew upon the TANF model and its numbers for Dutchess and Cortland (Phase_8e and Phase_8f, respectively), the Dutchess' safety-net model (SN_2), and the working version of the management flight simulator.²³ Very specific questions were listed, and several decisions were made in these two meetings, carrying significant implications in terms of model boundary and structure. From these decisions, we rebuilt the entire model, this time with both TANF and safety-net sectors.

²² Tsuey-Ping Lee wrote her dissertation on the topic of welfare finance at the county level (Lee, 1999). In her study, she explored the unexpected effects that offset the intended impacts of several welfare fiscal policies from a system dynamics point of view, using a highly aggregated version of the welfare reform model (T-1). Lee (2001) contains a summary of her findings (A-7).

²³ A revised version of the user's manual was prepared for this occasion (Richardson *et al.*, 1997b).

REACHING OUT TO POLICY MAKERS. Throughout this project we reached out to the people that were actually implementing welfare reform. We drew upon them as the primary source of expertise to design a simplified yet comprehensive view of the welfare system and of welfare reform. Actually, these welfare managers and executives were not just implementing reform. Rather, they were making the reform happen. This was recognized openly by Brian J. Wing, the Commissioner of the Office of Temporary and Disability Assistance:

With welfare reform, everyone's role is changing. For us at the state level, we have to stay out of the business of dictating how things get done at the local level. The modeling project is a good example of this kind of positive change in state/local relations and cross-agency teamwork. In the past, we probably would have told counties –in great detail and with incredible specificity– how we wanted them to implement welfare reform. Now we are making every effort to provide sophisticated, yet practical tools such as the welfare reform simulator, that local communities can use to think through policy implementation and arrive at their own solutions. I'm convinced that flexibility will be a key to successful welfare reform. (Extracted from quotation in Rohrbaugh and Johnson, 1998)

These efforts to work closely with state and local policy-makers are evident not only from the meetings previously discussed, but also from a series of briefings (B-1 through 3) and presentations (P-1 through 4) carried out between April of 1997 and November of 1998.

Of these, one of the most important meetings was the *Management and Leadership for the 21st Century Training Forum*, on September 22-25, of 1997 (P-2). This meeting was organized around the theme of welfare reform, and it counted with the participation of the state commissioner, and of 29 local commissioners. This three-and-a-half-days forum included presentations and discussions conducted by experts on the topic, as well as discussions led by the commissioners themselves. One afternoon was devoted to presenting and experimenting with the welfare model (Phase_8e). The work was presented in shared ownership by the commissioners of Cortland and Dutchess Counties, and by a liaison at the state level. The modeling team described the group model building process and facilitated the use of the model. Interested commissioners were invited to take the project into their counties. The conversations evolving from this meeting led to the decision to extend the modeling work into Nassau County, as the third project site.

A return to Dutchess County for a community-wide presentation using the joined model (Jnd_6) and a tailored version of the simulator occurred on April 28, 1998 (D-6). This meeting was very similar in format to the meeting in Cortland (C-6). Allers *et al.* (1998) contains an update of the project up to this point in time, and it depicts the essential simulations showed and discussed in this meeting (A-4). A new product was created especially for this occasion: a parameter booklet carefully documenting the type,

values, units, sources, and history of each model parameter and table function (Lee *et al.*, 1998a).²⁴ With a firm handle on the numbers, the leadership of Dutchess County stated (Allers *et al.*, 1998, p. 10):

The group modeling sessions provided an opportunity for representatives from several agencies within the community to come together in one setting to plan for welfare reform. As discussions regarding the complexity of social services programs took place, the value of computer support and modeling was amply recognized. The model and the simulations permitted detail exploration of the interactions between different scenarios and alternative welfare reform policies. Finally, the experience to work with the University at Albany modeling team –in the process of model conceptualization, formulation and calibration, for the purpose of examining the impact of changes in the administration of welfare programs –was both challenging and rewarding.

These experiences in Cortland and Dutchess Counties were communicated in vehicles such as *Empire State Report* (A-2, *ESR*, 1998) and *Government Technology* (A-3, Rohrbaugh and Johnson, 1998):

“We were hoping that, with the development of the model, communities could identify the high-leverage points in their system and shift their strategies accordingly,” said David Avenius, OTDA deputy commissioner. (*GT*)

The flight simulator makes the relationships between government and non-government agencies such as charities visible, thereby helping to make the entire welfare system easier to understand. (*ESR*)

The simulator makes it easy to pose a large number of “what if” questions about alternative strategies, changes in the economic and social environment and alternative approaches to funding services. (*GT*)

Rolling out the Welfare Reform model (February-September, 1998)

With these two success stories behind, we moved on to work with the third and final county. It was not clear how this work was to be developed further in Nassau. The Commissioner Irene Lapidez requested that the model (Jnd_4) be presented to the local management team on February 19, 1998 (N-1).

When presented the model, the first thing the local client team wanted was to have it parameterized and calibrated to their county. This provided us with yet another

²⁴ This booklet also contained a record of the cross-sectional contrast between model-generated and actual numbers for key stocks and flows (in equilibrium); and, spreadsheets documenting the formulas and sources for aggregation of actual spending items into model-based resource clusters.

opportunity to build upon the design and use of group processes related to model parameterization and calibration. First, however, we finished building up to the final version of the model, and preliminarily parameterized and calibrated the model to their numbers. For this we drew upon the experiences in Cortland and Dutchess Counties (meetings C-5 and D-5), identifying the key information needed to parameterize the model, and requesting in writing those data that would be easily obtainable (without the need for expert judgments). We also prepared the scripts and materials for the several elicitation exercises using expert judgments (elicitation of unknown parameters, relative strength of the table functions, and key table functions), and for reconciling cross-sectional model-generated stock-and-flow values, versus actual data.

Everything prepared, the parameterization and calibration meeting was held on June 16 (N-2), almost four months later. These increasingly sophisticated parameterization and calibration procedures were apparently requiring much more time in terms of preparation. (In Cortland it happened within five weeks, and in Dutchess within two-and-half months.) Alternatively, it could be that there was less pressure to move forward on the project, or less interest on the part of the local client team.²⁵

Out of this meeting surfaced the desire to replicate the historical behavior of the TANF caseload.²⁶ The assumption was that if the simulated caseload corresponded to the actual caseload, then this would build confidence in the model. The calibration of the model resulted in a reasonably good fit for the TANF caseload. When we returned to Nassau three months later to present the final results, on September 22 (N-3), there were no challenges to the model structure or to its simulated behaviors.²⁷ The historical run, the base run, and the strategies and scenarios simulated were received with neither enthusiasm nor criticism. Unfortunately, there was no follow-up of this work in Nassau County, at least not facilitated by the modeling team.

At this point in the project, the final version of the management flight simulator was delivered to the state, and to the participating counties. The project had run its course. The budget had been used up. The deliverables had been met. Without additional funding or the specific request for expanding this work into new domains, the project simply closed quietly.

²⁵ Nassau County's model parameterization and calibration meeting was documented in Zagonel *et al.* (1998).

²⁶ I do not remember and could not find any evidence of who had this initiative.

²⁷ We also prepared and delivered a parameter booklet for this version of the model (see Lee *et al.*, 1998b).

An overview of the welfare reform model

The final welfare reform model (Jnd_6n) contained nearly 700 equations. It was built in 33 views, as illustrated in Figure 3. The first seven views of the model describe the patterns of client-flow through the welfare system, as conceptualized by the management teams involved in the group model building sessions.

Figure 3. A sample view of the WR model: Job maintenance services
(See p. 46)

Figure 4 is a summary view of this client-flow structure. The model contains a TANF sector and a safety-net sector. In the former, families are differentiated as high-need or low-need families on TANF. In the latter, there are both families and individuals: post-TANF families on the SN, individuals on the SN, and other families (non-TANF) on the SN.²⁸

Figure 4. Summary view of the stock-and-flow structure of the WR model
(See p. 47)

The main stock-and-flow dynamics of both sectors are:

- An inflow from the mainstream economy into the TANF and safety-net systems, representing families and individuals whom become at risk (e.g.: fall into poverty, loss of employment, teen-age pregnancy, etc.);
- In the safety-net sector, this inflow represents direct enrollment into the SN program;
- In the TANF sector, before families are enrolled, they are accounted for in the stock of families at risk eligible for TANF;
- An outflow of clients from welfare programs into employment; this is a “pump” that moves clients from monetary assistance into paid jobs;
- An outflow back into the mainstream economy, representing the attainment of self-sufficiency; these families and individuals are no longer at risk;
- Those families and individuals served and placed into employment that are *not* able to attain self-sufficiency normally return directly into the programs in a process called recidivism, or eventually fall back at risk (in the TANF sector).

The consequence of this client-flow structure is that, at any one point in time, only a fraction of the clients served in the welfare system actually makes it back into the mainstream economy. Most families and individuals in this system are *cycling* (as

²⁸ Post-TANF families on the SN are those families that exhausted their federally funded TANF eligibility (after five cumulative years on the TANF rolls) and, therefore, fell into the state funded Safety Net program. The individuals on the SN constitute the clientele of the existing Home Relief program, now incorporated onto the Safety Net program. Other (non-TANF) families are those families that do not qualify for TANF for particular reasons –such as legal aliens– but who qualify for certain services provided by the Safety Net program.

opposed to passing through), coming in and out of programs, finding only temporary employment or being able to maintain only low-paying jobs, and eventually returning to the welfare rolls. This is why, with welfare reform, and with the implementation of time limits in the TANF program, it was expected that many initially eligible recipients would eventually run out of time, and lose TANF eligibility. Thus, the connection between the two sectors is represented by the rate of loss of TANF eligibility –a one-way flow from the TANF sector to the safety-net sector.

The management teams perceived as important and incorporated into the model two other stock-and-flow processes: the possibility of diverting many of the potential TANF enrollees into diversion programs, for the purpose of avoiding usage of time on TANF; also, they wished to experiment with sanctioning, as an important tool for monitoring recipients (e.g., for non-compliance with drug-free regulations). The remaining views of the model contain the formulations of the resources and scenarios that control the rates of flow in the stock-and-flow structure, and compute the costs of the welfare system.

Figure 5 illustrates, in a high-level conceptual view, how resources and scenarios affect the rates of flow in the TANF sector. In the “front-end” of the system, *prevention* resources help to reduce the inflow of families falling from the mainstream economy, into the welfare system. In the “middle,” TANF *assessment and monitoring* promotes diversion programs, helps clients get jobs and other services needed to move into employment, and sanctions clients who are not complying with drug-free regulations or mandated participation in work preparation and workfare. *Employment services*, in turn, help to “pump” clients into paid jobs. In the “back-end,” *self-sufficiency promotion* helps families maintain the jobs they obtained, move-on to better jobs and, eventually, transition back into the mainstream economy. Finally, *child-support enforcement* has widespread effects in the welfare system: it prevents families from becoming at risk; it helps them to depart the rolls; it leads to more sanctioning; and it helps some single-parent employed families rise above poverty level.

Figure 5. Stock-and-flow structure of the TANF sector with resource impacts
(See p. 48)

As illustrated in Figure 6, child-support enforcement has similar effects for families in the safety-net sector. And, except for prevention, all other resources have counterparts in the safety-net sector. These are *monitoring*, *employment services*, and *job maintenance services*. (Figure 3 provides a glimpse at the typical formulation for the resource sectors.)

Figure 6. Stock-and-flow structure of the safety-net sector with resource impacts
(See p. 49)

In addition to resources, there are also scenarios controlling the rates of flow. For example, *unemployment* has wide-ranging effects on the rates. Also, “step” and “ramp” changes can be made to the inflow from the mainstream economy.

When simulating the model:

1. Funds can be reallocated across resources;
2. Scenarios can be changed;
3. Different budgeting mechanisms can be chosen (from rigid to flexible); and/or
4. Assumptions regarding client behavior can be changed.

Figure 7 illustrates the control panel developed for the stand-alone version of the model, in the form of the management flight simulator.

Figure 7. Control panel of the management flight simulator of the WR model

(See p. 50)

The simulation of the model computes the dynamic behavior of the client-flow, as well as the costs of managing the system, which includes funding the above mentioned resources, as well as paying for “census-driven” costs --such as *diversion services and payments, TANF payments, basic services, and emergency services*. In all, there are 13 spending categories. (Appendix 3 contains a detailed list of the actual services provided under each resource cluster.) The rate of loss of TANF eligibility is controlled by a co-flow structure that mirrors the TANF sector of the model, and that keeps track of the time used-up by the clients on the TANF rolls.

In sum, the richness of the model permits investigation of system performance through testing a wide-range of strategies utilizing a mix of resource allocations (prevention, monitoring, employment, self-sufficiency, etc.), as well as experimenting with alternative scenarios (e.g., unemployment). Within its boundary, the model also helps to identify interdependencies among different service providers within the welfare system (education department, social welfare agency, labor department, charities, etc.).

Policy question # 1:

Will there be a burden to NYS and its counties?

I will use the base run of this model to address the concern that welfare reform may cause a long-term burden in states and counties, as welfare recipients expend their TANF eligibility. The question is: *Will loss of eligibility shift caseloads and costs from the federally funded TANF to the state (and locally) funded safety net? Will there be a burden to New York State and its counties?* I recall from the discussion of the Welfare Reform Act that the law gave state governments authority to devise their own solutions to end welfare dependency, or bear the costs of sustaining those people unable to attain self-sufficiency, after five cumulative years of receiving federally funded benefits (p. 5). Thus, can we address the question of *exactly how big will the damage be in five years and beyond, when hard-core welfare recipients begin to lose eligibility (beginning in 2002)?*

The base run

The base run of the model is described in terms of the resulting effects of the *time limits* imposed in the welfare reform bill on the TANF program. It conveys a rather undramatic view of welfare reform, in which the total size of the welfare population does not really change. This is not surprising because the simulation portrayed in the graphs reflects what would be happening to a county, under a stable economy, and future client behavior—lengths-of-stay, recidivism, etc.—conforming with historical patterns of behavior, without implementation of any particular set of new managerial policies or strategies (*other than time limits on TANF*). Only, there is a shift in caseloads from TANF to the safety net, accompanied by a marginal decrease in total expenditures, and a marginal increase in costs at the state and local level.

THE EFFECTS OF TIME LIMITS. The base run of the model is essentially driven by the co-flow structure that keeps track of the time used-up by clients on the TANF rolls. In Figure 8, the average time used-up for both high and low-need clients (lines 1 and 2) begins to rise in 1997, when the clock starts ticking for TANF recipients. In 2002, high-need TANF recipients who have been on the rolls uninterruptedly begin to lose eligibility (line 3). The rate of loss of TANF eligibility for high-need clients spikes in that year, with approximately 700 families losing eligibility in 2002, but subsequently stabilizes at a lower level, with approximately 300 families losing eligibility per year in equilibrium. For low-need clients, loss of eligibility raises gradually to a much lower level (line 4), with approximately 50 families losing eligibility per year in equilibrium. The average time used-up for high-need clients peaks at approximately 2.7 years in 2002, but stabilizes at around 2.2 years per family in equilibrium (line 1). For low-need clients, the time used-up stabilizes around the average value of 1.1 years per family (line 2). These averages are kept low because the clients who use up their five years are expelled from the TANF program. Otherwise, these averages would rise to approximately 1.3 and 3.7 years per family for the low-need and the high-need client-groups, respectively.

Figure 8. Base run: loss of eligibility

(See p. 51)

As welfare recipients lose TANF eligibility, they move on to the safety-net sector. Figure 9 portrays this shift in caseloads after the year of 2002. The sizes of the TANF and SN *rolls* converge (lines 1 and 2). Families on TANF drop by approximately 19 percent, from 6,400 to 5,200 families, while the SN caseload rises by about 55 percent, from 2,200 to 3,400 families. However, overall, the total population on the welfare rolls does not change meaningfully (line 3), remaining stable at around 8,600 families.

Figure 9. Base run: shift in caseloads

(See p. 52)

In Figure 10, the same process can be observed for the sum of the populations in each sector (on rolls, employed, at risk, etc.). The sizes of the *sectors* also converge (lines 1 and 2), but this change is less accentuated because the size of the population in the

sector is considerably larger than the size of the population on the rolls. The TANF system shrinks by approximately 8 percent (from 16,700 to 15,300 families), while the safety-net system expands by approximately 28 percent (from 5,100 to 6,500 families). Therefore, those TANF recipients who lose eligibility and become SN recipients have a large impact on the size of the rolls, and a relatively smaller impact on the size of the sectors. The sum of all populations in the welfare system –including *both* TANF and safety-net sectors (line 3)– remains stable at around 21,800 families.

Figure 10. Base run: populations by sector

(See p. 53)

Because TANF is largely funded by the federal government and the safety net is mostly funded at the state and local levels of government, the shift in client populations results in a marginal increase of the state and local share of costs, as illustrated in Figure 11 (line 3). The sum of state and local expenditures rises by about 2.2 million dollars (line 2), an increase of approximately 3 percent. The base run also portrays a small decrease in overall expenditures (line 1), of approximately 2 percent or 2.4 million dollars.

Figure 11. Base run: expenditures and share of costs

(See p. 54)

The increase in state and local expenditures combined with the decrease in total expenditures results in a rise in the state and local share of costs of about three percentage points (from 56 to 59 percent). *Surprisingly, the rather dramatic shift in caseloads observed in Figure 9 causes a rather undramatic shift in the share of costs.* This is largely due to the facts that: direct payments to TANF recipients account for only 46 percent of all expenditures (59.1 million dollars in a welfare budget of nearly 130 million dollars); and most safety net programs are also partially funded with federal dollars (examples given: basic services at 23 cents on the dollar, and employment services at 66 cents on the dollar).²⁹

THE ANSWER TO QUESTION #1. The results of the base run simulation confirm that we should expect a shift in caseloads, from TANF to the safety net. The size of the caseloads should change noticeably, particularly for the safety net. However, surprisingly, this should not result in a significant financial burden at the state and local level. According to information obtained directly from the counties, the federal government contributes financially to programs run under the safety net umbrella. As a result, we should expect only a marginal increase in costs at the state and local level.

The base run assumed a stable economy. In fact, the economy has been fairly good in Nassau County between 1998 and 2002, with unemployment ranging between

²⁹ The federal share for each program expense was calculated based upon information obtained directly from the County during the parameterization of the model (see Lee *et al.*, 1998b, p. 4).

2.2 and 3.7 percent (which is somewhat below average for this county). As of March of this year, 466 TANF cases had been converted to the safety net,³⁰ and 161 families had less than one year remaining.³¹ This suggests that the model overestimated by two-fold the number of families losing eligibility (700 families in the year of 2002 and 300 in equilibrium). This could be a result of the prosperous economy during this period. In any case, it would be useful to revisit the model, now that there is new data, to empirically verify some of its assumptions (e.g., the table of distribution of loss of eligibility).

THE HISTORIC RUN. How can it be known if the behaviors observed in the base run simulation have anything to do with reality? This is an important but difficult question to concisely address here. The validity of this model as a simulation tool to project the consequences of welfare reform into the future is contingent upon a number of things.³² Such as: this group model building effort must have been successful in revealing and capturing the underlying structure of the welfare system, as well as accurately and reliably estimating behavioral and financial parameters embedded in this system. Also, the workings of the system, and the changes under implementation as a result of reform must have been modeled coherently. Finally, there must be correspondence between the historical behaviors of this system, and the simulated behaviors of the model.

Throughout this project, much effort was made to address the above considerations, and to improve the model and its simulated output, *in accordance with the issues and considerations raised by the management teams, and using available data and expert judgments.*³³ After many presentations and workshops, including participants knowledgeable in the workings of this system and familiar with the requirements of welfare reform –both within and outside of the client teams– the modeling team and the clients gained confidence that the model was a useful instrument for various

³⁰ New York State Office of Temporary and Disability Assistance - Sixty Month Time Limit Report - WINR-8116.

³¹ Database: New York State Office of Temporary and Disability Assistance - PA Caseload Tracking Report - WINR-8113.

³² Guidelines for model validation in the field of SD are contained in Forrester and Senge (1980), Graham (1980), Richardson and Pugh (1981), Sterman (1984, 2000), Barlas (1989 and 1994), Barlas and Carpenter (1990), Oliva (1995 and 1996), among others.

³³ Confidence building processes in this intervention relied on both structural and behavioral tests. The structure of the model emerged from the group process. The parameters were based on administrative data everywhere possible. The most important table functions were elicited from expert groups using processes that checked for consistency in judgments and often tested for convergence between separate elicitation approaches. In addition to the replication of historical time series for caseloads, the model was calibrated and evaluated for detailed cross-sectional comparisons for particular years. Also, every simulation done in the context of the group was followed by debriefing and story telling that aimed to diagnose and explain behavioral modes, shifts in behavior, tracing the behavior back to the structure of the model (and of the real system).

constituencies to examine the unfolding of welfare reform. The model endured and passed several tests of logical coherence, and structural and behavioral correspondence. Such as: welfare managers and experts found the model structure and the simulated behaviors depicted by the model reasonable and plausible.

In Nassau County, where model parameterization and calibration received fullest attention during the project, the actual historical caseloads were tracked by the simulated caseloads, for a thirteen-year period, between 1984 and 1997, as depicted in Figures 12-A and B.³⁴

Figures 12 (A-B). Historic run: simulated v. actual caseloads

(See p. 55)

The simulated dynamics (line 3) observed in the historic run depict how the model structure reacts to the stimulus imposed by the actual fluctuation in unemployment in the county (line 1).³⁵ The actual caseloads for AFDC/TANF and for HR/SN are provided for contrast (line 2).³⁶ Their summary statistics for the goodness of fit are provided in Table 1.

Table 1. Theil summary statistics: simulated v. actual caseloads (1984-97)

Population	N	R ²	U	Bias	Unequal variation	Unequal covariation
AFDC/TANF	14	.91	.05	.10	.01	.89
HR/SN	14	.74	.23	.47	.17	.36
TOTAL	14	.86	.09	.21	.14	.65

The TANF sector produced the best fit ($R^2=.91$).³⁷ It shows a very small bias (.10), slightly overestimating the minimum value in caseload, and similarly

³⁴ The calibration of the model was done by hand, within the ranges for the parameters, which, in turn, were estimated with data and judgments below the level of aggregation. Lyneis and Pugh (1996) present a critique of the hand-calibration process. Oliva (unpublished) suggests heuristics and techniques for taking advantage of automated calibration algorithms without bending the model to fit the data.

³⁵ In this model, unemployment is an exogenous variable that serves as an input to test scenarios.

³⁶ The actual unemployment and caseloads were recalculated using running medians of three adjacent values to smooth out high-frequency noise in the data (Mosteller and Tukey, 1977, pp. 52-58).

³⁷ The fit can also be discussed in terms of Theil's *Inequality Coefficient* (U), a measure with similar properties to the *Root Mean Square Percentage Error* (RMSPE). It's a standardized

overestimating its maximum value. It shows virtually none unequal variation (.01). High unequal covariation (.89) indicates that the error terms are imperfectly correlated, that is, they differ mostly point-by-point. For this economic cycle, the simulated caseload contains a lag of one to two years in a prosperous economy (when caseload is falling), but appears to track very well the peak in unemployment (when caseload is rising). High unequal covariation can be considered an unsystematic error (unless the purpose of the model is to study the cycles in the data, in which case phasing would be important).

The safety-net sector calls our attention for the need for improvement in the model, with a much poorer fit ($R^2=.74$). It shows a large bias (.47), with the simulated caseload having predominantly a lower value than the actual caseload. There is also some unequal variation (.17) due to the shift in bias before and after the year of 1986.

According to Sterman (2000, p. 879):

The proper use of the behavior reproduction test is to uncover flaws in the structure or parameters of the model and assess whether they matter relative to the purpose... These discrepancies mark the trails that can guide you to erroneous parameter estimates and inappropriate assumptions you should revise before using the model for policy analysis.

The primary focus of this intervention on the TANF component of welfare reform, and the objective of building a simulator for the purpose of comparing management strategies, and for experimenting with alternative scenarios (without ambition for accurate forecasting), served as an excuse to move forward, and not resolve these discrepancies with the management team.

As a result, the goodness of fit for the sum of the two populations (.86) was an average of the very good fit for the TANF sector and the relatively poorer fit for the safety-net sector, showing both some bias (.21) and some unequal variation (.14). While only a weak test of model validity (Sterman, 2000, pp. 328-330 and 879), this was still a reassuring preliminary result, considered sufficiently good to provide the basis for using the model to address “what if” type of questions of interest to the client teams.

Policy question # 2:
What’s the “best” strategy?

I will use two policy simulations to address the question: *What is the best strategy to improve system performance?* I recall that the 1996 Welfare Reform act shifted decision-making power closer to where the action is, by increasing state and local flexibility and discretion in designing and running programs: States could now choose to transfer a

measure of the size of the error and values closer to zero indicate a better fit. I chose to discuss it in terms of R^2 simply because the latter is more commonly used and understood in social science.

portion of federal dollars to child-care or other social services deemed important to accomplish the legislative intent of the Act; they could reduce monetary assistance in favor of other means of support; they could establish governmental and non-governmental partnerships to carryout programs; and, they could promote innovative means to reduce caseloads, such as an increased emphasis on child support enforcement services (p. 6). Thus, after extensive hand-on strategizing with this model, based upon the solutions envisioned by the state and local policy-makers themselves, *can we suggest ways of improving system performance?*

Improving system performance

This model and its earlier versions were used with the client groups in search for management strategies to improve the performance of the welfare system. Managers were allowed to experiment with their “pet” strategies, investing in alternative mixes of welfare services (prevention, assessment and monitoring, diversion, employment services, child support, and self-sufficiency promotion). This process of strategy selection, implementation and evaluation, using the model as a laboratory, proved to be an excellent means of surfacing and testing the assumptions of members of the management teams. It also provided an opportunity for exchange of ideas among team members, and learning. However, *it turns out that the “best” strategies are heavily dependent upon what one seeks to accomplish, and how one measures success in this system. There is no clear-cut optimal strategy to improve system performance, as I will illustrate in this section.*

I will underscore the difficulty involved in finding the best strategies by drawing upon two alternative approaches to managing this system. The first approach is the classic form of solution, where the department of social services invests in those services closely associated with its mission, and over which it has most control. I will label this as the “middle” investment strategy, but it can also be referred to as the “welfare-to-work” approach. It involves allocating more resources towards:³⁸

- Assessment and monitoring (and diversion)
- Employment services

The department of social services alone cannot accomplish the second approach, as it requires investing in services more closely associated with the missions of other actors in the community, both public and private (e.g., education, labor, private charities). I will

³⁸ The “middle” investment strategy, or “welfare-to-work” approach, is simulated with enrichment in intensity of resources in the order of 25 percent for the following services: TANF Assessment and Monitoring (TAM), SN Monitoring (SNM), Employment Services to High-need Families on TANF (HiS), Employment Services to Low-need Families on TANF (LoS), and Employment Services for SN Clients (ES).

label this as the “edges” investment strategy, but it can also be referred to as the “community-wide” approach. It involves allocating more resources towards:³⁹

- Prevention
- Self-sufficiency promotion (or job maintenance)

The “middle” investment strategy devotes resources towards services that have an effect upon the rates that pump clients from one stock to another in the middle of the system (e.g., employment services that help TANF recipients find jobs). The “edges” investment strategy devotes resources towards services that have an effect upon the rates that pump people into and out of the welfare system (e.g., job maintenance services that help those clients who found jobs maintain those jobs, find better jobs, and move back into the mainstream economy). These effects are illustrated in Figure 13. (Appendix 3 contains a detailed list of the actual services provided under each resource cluster.)

Figure 13. The effects of the “middle” and “edges” strategies on the TANF sector
(See p. 56)

The way the investment priorities work themselves in the model is linked to the idea of resource intensities. Initial intensities are a function of on-going expenditures divided by the number of clients currently enjoying services. A decision to prioritize a particular service is made by establishing a desired intensity that is greater than the initial intensity. The gap between desired and actual intensities determines investment and triggers additional spending in the resource, until desired and actual intensities match. In these simulations, the budget is allowed to increase (or decrease) to adjust to the requested resources. The only delay is the time needed to reallocate resources (budget cycle), and to build (or do away with) resources.⁴⁰ Figure 3 provides a glimpse at the typical formulation for the resource sectors. Changing the target for the service creates the gap between desired and actual intensities.

ILLUSTRATING AN INVESTMENT. Figure 14 illustrates the process of an investment on TANF Assessment and Monitoring (TAM). Observe how the actual resource intensity (line 2) “seeks” the desired intensity (line 1), and how the actual available resource (line 4) “seeks” the proposed (line 3). The dynamics in the behaviors after 1998 include not only

³⁹ The “edges” investment strategy, or “community-wide” approach is simulated with enrichment in intensity of resources in the order of 25 percent for the following services: Prevention (PR), Self-sufficiency Promotion on TANF (SSP), and Job Maintenance Services on the SN (JMS).

⁴⁰ The simulations can be performed under different budgeting assumptions, ranging from a fully flexible budget based upon desired intensities of resources, to a fully rigid budget that cannot respond to desired intensities. In fact, the flexibility in the budget can be set independently for each resource in question. Here, for the sake of illustration, I assume that resources can be adjusted freely to match desired intensities. In other words, managers have the ability to allocate resources freely, and to adjust spending as necessary according to priorities. (The base run was also run under this assumption.)

this process of resource adjustment to new priorities, but also changes in the client-flow. The investment strategy itself begins to have its effect (after 1998), causing clients to move around in the system, and changing needs (proposed resources). The shift in the client populations, from TANF to the safety net, due to loss of eligibility (after 2002), causes additional changes in the client-flow, and leads to further need to adjust resources. All the while, *actual* intensity and resource (lines 2 and 4) are “seeking” to reach *desired* intensity and resource levels (lines 1 and 3, respectively).

Figure 14. E.g.: investment on TANF Assessment and Monitoring (TAM)
(See p. 57)

Figure 15 illustrates what is accomplished by an increased investment on TAM. Although the scales do not help to visualize the changes, this investment is increasing enrollment in diversion programs (line 1), and helping clients enrolled on TANF to find jobs (line 2). Thus, it causes the TANF caseload (line 3) to fall slightly (between 1998 and 2002), and employed-families post-TANF to rise (line 4). After 2002, the most significant changes are related to the loss of TANF eligibility, and not to the investment on TAM. As a result, observe that the number of families employed (line 4) rises in the short term but falls towards the end of the period, due to the lower numbers in the TANF system.

Figure 15. E.g.: results of investment on TAM
(See p.58)

In the next sub-section I contrast the results of the “middle” versus “edges” strategies. This exercise serves two purposes. It illustrates the advantages and disadvantages of the “welfare-to-work” and of the “community-wide” strategies, and it underscores the problem of finding an optimal strategy.

MULTIPLE CRITERIA OF SUCCESS. In this section, I examine 19 criteria to evaluate the success of the “middle” versus “edges” strategies. Six criteria are related to the size of the caseloads and populations in the welfare system:

- Number of families on TANF
- Number of families and individuals on the SN
- Size of the TANF system (sum of all stocks in the TANF system)
- Size of the safety-net system (sum of all stocks in the safety-net system)
- Number of families in diversion^(*)
- Populations employed but at risk (in the welfare system, but not on the welfare rolls)^(*)

Eight criteria are related to the rates of flow in the system:

- In-flow per year into the system from the mainstream economy
- Families enrolled per year into diversion programs^(*)

- Families enrolled per year into TANF^(*)
- TANF families finding jobs per year
- TANF families losing eligibility per year
- Recidivism per year
- Flow back at risk per year
- Out-flow per year back into the mainstream economy

Finally, five criteria are related to costs to finance the system, and who pays for those costs:

- Actual expenditures
- Local expenditures
- State and local expenditures
- Local share
- State and local share

Two observations are in order. First, it is less clear what are best and worst results for four criteria (identified with an asterisk).⁴¹ Therefore, they will be left out of the discussion. Second, because this GMB effort was undertaken with *state and local* priorities in mind I assume to be “best” those results that reduce expenditures and shares for those levels of the federation. *The results of the “middle” versus “edges” strategies are summarized in Table 2. From this table it is clear that neither strategy is best with respect to all criteria, or with respect to time.*

Table 2. Condensed results of the “middle” versus “edges” strategies
(See p. 59)

The “welfare-to-work” approach, emphasizing the “middle” of the system, is best at pumping clients out of the rolls, and into jobs, as illustrated in Figure 16-A (line 2). It is the best in terms of lowering the size of the SN caseload. It is also best, but in the short term only, in reducing the size of the TANF caseload (see Figure 18-B), thus avoiding loss of eligibility, also in the short term. This strategy results in the lowest state and local share of costs for the short term. However, it performs the worst in terms of “trapping” the populations of the welfare system (both TANF and SN), as illustrated for TANF in Figure 16-B. It leads to the largest flows of recidivism (see Figure 17-A) and back at risk. It results in the highest local expenditures in the short term, and the highest overall expenditures in the long term (see Figure 18-A).

⁴¹ I assume that diverting families from TANF is “good” because it prevents them from using up their eligibility period. Thus higher numbers in diversion is “best.” For similar reason, the rate of flow enrolled into diversion is “best” when larger, and the rate of flow enrolled into TANF is “best” when smaller. A larger population employed but at risk is “good” because (even though they are still in the welfare system) they are not on the welfare rolls.

Figures 16 (A-B). E.g.: pros and cons of the “middle” strategy
(See p. 60)

The “community-wide” approach, emphasizing the “edges” of the system, is best at preventing families from coming from the mainstream economy into the welfare system, and helping families in the system transition back into the mainstream economy. Because people succeed in transitioning back into the mainstream economy, it results in the lowest flows of recidivism and back at risk, as illustrated for recidivism in Figure 17-A (line 3). For these reasons, it deflates the welfare system and produces the best results in terms of the reduction in size of the TANF and SN populations (see for example Figure 16-B). It also achieves the best results, in the long term, for the size of the TANF caseload (see Figure 18-B), thus avoiding loss of eligibility, also in the long term. It is the best strategy in the long term for the volume of local expenditures. Finally, it is a winner in terms of reduction of the local share of costs. However, it results in the worst performance for job finding (see Figure 16-A), for state and local expenditures, as illustrated in Figure 17-B, and for state and local share. It is also the worst strategy in the short term for overall expenditures (see Figure 18-A).

Figures 17 (A-B). E.g.: pros and cons of the “edges” strategy
(See p. 61)

For some criteria, the base run or “status quo” strategy is best. For obvious reasons, it results in the lowest overall expenditures, as illustrated in Figure 18-A (line 1), and the lowest state and local expenditures (see Figure 17-B). This is largely because no new investments are made. And, for the same reason, in the short term, it yields the lowest local expenditures. However, the favorable results of the “middle” and “edges” strategies make the “status quo” strategy pale by comparison, in terms of the size of the caseloads (both for TANF and SN), as illustrated in Figure 18-B, and in terms of the out-flow back into the mainstream economy. The local share of costs is always worst for this strategy, and the state and local share of costs is best only in the long term.

Figures 18 (A-B). E.g.: pros and cons of the “status quo” strategy (base run)
(See p. 62)

THE ANSWER TO QUESTION #2. This exercise has served to illustrate the difficulty involved in devising robust strategies to improve the performance of the welfare system. *It is unlikely that any one strategy will produce the best results with respect to all relevant criteria, both in the short and long term. The goals in the minds of the decision-makers, and the positions in the system that they occupy (with respect both to the agency and the level of government, and if public or private provider), are likely to play a big role in their views of their favorite strategies.* There are short and long term tradeoffs. There may even be differences in perception if one is looking at changes in the value of expenditures or, alternatively, relative changes in the federal-state-and-local shares. Individual providers are likely to ask themselves: Are we spending more than before? Are we spending more relative to other levels of government? Has the system improved with respect to the criteria affecting our workload? Has it improved with respect to the

performance measures that we are evaluated upon? Can we afford in the short term to wait for the long-term results expected?

However, the use of the model to examine “what if” questions related to management policies and strategies revealed a wealth of insights with respect to things that are absolutely necessary to improve system performance. Most importantly, it seems that *success or failure in welfare reform lies at large in the community*, and not strictly within the realm of the departments of social services (Rohrbaugh, 2000). Thus, *what is needed to succeed is a concerted countywide response involving multiple community organizations, both public and private, in a broad leadership coalition, who should make shared decisions about reallocation of county resources* (p. 530).

Also, key reinforcing and balancing feedback loops indicate that *self-sufficiency promotion is a powerful policy vis-à-vis employment services* (Zagonel *et al.*, 1997c, pp. 40-42; Rogers *et al.*, 1997, pp. 609-610; Allers *et al.*, 1998, pp. 8-9; Richardson *et al.*, 2002). While it does not win in every respect, it has the power of deflating the welfare system, and avoiding cycling within. In the long term, it is likely to produce the best outcomes for the welfare of the recipients themselves, and for the welfare of society at large.

Policy question # 3:
Can management really make a difference?

The last question I will address is related to the significance of the state of the economy on the welfare system vis-à-vis the role of management. The question is: *Can “endogenous” management really make a difference, or is this system run by the state of the economy –i.e., “exogenously”?*

The historic simulation run depicted how the model structure reacted to the exogenous stimulus imposed by the actual fluctuation in the economy (see Figures 12-A and B). A comparison with the actual caseloads denoted a reasonably good fit for the period between 1984 and 1997. Most of the behavioral patterns observed in the caseloads were explained by fluctuations in unemployment alone.

Fluctuations in unemployment between 3 and 6 percent produced fluctuations in caseload between 6,500 and 11,400 cases, denoting an impact in caseload of the order of approximately ± 30 percent. Policy experimentation using this model, on the other hand, produced simulations denoting a much smaller impact in caseload, of the order of approximately ± 5 percent. These results are illustrated in Figures 19-A and B, and in Table 3, for the historic run, and for the middle and edges strategies.

Figure 19 (A-B). Amplitude of management and unemployment effects upon total caseload
(See p. 63)

Observe that the scale for the amplitude of management effect (Figure 19-B) has been magnified for visibility purposes. The range of amplitude has been reduced four-fold, from 5,700-11,500 to 7,900-9,300. Still, the effect of management is noticeably smaller. In fact, simulations of management policies and unemployment scenarios during this intervention suggested that unemployment effects dominate system performance (± 20 to 40 percent), and that endogenous management makes a smaller difference (± 5 to 10 percent).

Table 3. Amplitude of management and unemployment effects upon total caseload

Policy/ Scenario	Percent change	Minimum value	Equilibrium	Maximum value	Percent change
Middle	-4%	8,280	8,600	—	+4%
Edges	-4%	8,280	8,600	—	+4%
Unemployment	-24%	6,530	8,600	11,430	+33%

This would indicate that the decline in caseload experienced during the intervention (1997-98) was largely driven by the decline in unemployment. Thus, an expectation that caseload would rise with unemployment in the future. In order to verify these results, now five years later, I gathered new data and ran it through the model. An update of the behavioral reproduction tests is illustrated in Figures 20-A and B, and Theil's summary statistics are provided in Table 4.

Figure 20 (A-B). An update of behavioral reproduction tests
(See p. 64)

Table 4. Theil summary statistics (1984-97 v. 1984-02)

AFDC/ TANF	N	R ²	U	Bias	Unequal variation	Unequal covariation
1984-1997	14	.91	.05	.10	.01	.89
1984-2002	19	.81	.16	.29	.34	.37

HR/SN	N	R ²	U	Bias	Unequal variation	Unequal covariation
1984-1997	14	.74	.23	.47	.17	.36
1984-2002	19	.80	.22	.36	.27	.37

TOTAL	N	R ²	U	Bias	Unequal variation	Unequal covariation
1984-1997	14	.86	.09	.21	.14	.65
1984-2002	19	.85	.14	.02	.50	.48

The time period added, 1998 to 2002, contains an upswing in unemployment (line 1). According to the assumptions embedded in the model, the simulated behaviors for both TANF and SN (line 3) contain a corresponding upswing in caseloads. However, the actual caseload for TANF (Figure 20-A, line 2) continued to decline throughout the period. Moreover, the caseload declined by a surprisingly large amount. The TANF sector now produced a poor fit ($R^2=.81$). It shows both bias (.29) and unequal variation (.34).

The same drop in caseload did not occur in the safety-net sector (Figure 20-B), where actual caseload (line 3) seems to have followed a predictable pattern (line 2). While the fit is not poorer than before, it is not possible to say that it is any better either ($R^2=.80$). There appears to have been an exchange between bias (an improvement, down to .36) and unequal variation (a deterioration, up to .27), but the lack of fit cannot be attributed to unsystematic error (*i.e.*, the error terms do not differ mostly point-by-point and are correlated).

THE ANSWER TO QUESTION #3. The answer to question number 2 involves addressing another question: *Could a management strategy implemented since 1997 be responsible for this sharp drop in caseload?* While this is possible, the model probably would not have predicted it. Figures 21-A and B contain a simulation of the middle strategy, coupled with the actual unemployment scenario (line 4). The caseloads in these figures are the sum of TANF and SN (total caseload). Figure 21-B is an amplified version of Figure 21-A, with the time period illustrated shortened to the last five years (1997-2002). *The model-simulated caseload with a strong emphasis on the welfare-to-work component shows a marginal change, reflecting the assumption embedded in the model that endogenous management makes a small difference, and unemployment dominates system performance. The real system is behaving much differently. Actual caseload has dropped (between 1997 and 2002) by an amount that suggests that endogenous management may have an effect of similar magnitude as exogenous fluctuations in the economy.*

Figures 21 (A-B). Actual behavior unexplained (exogenously nor endogenously)
(See p. 65)

These tests call for further research and the need for improvement in the model. As a first step, I am reporting these results to the client team in Nassau County, and to the members of the modeling team in Albany. I am interested in exploring whether the flaws observed are due to endogenous causes, such as:

- Inadequate formulation of the system's structure and/or estimation of parameters;
- Underestimation of the effects of investments, vis-à-vis unemployment effects;
- Disregard for management strategies under consideration and/or implementation;
- An actual change in the welfare culture, affecting client paths through the system and/or lengths of stay.

Or, if there is an exogenous explanation (other than the economy), for what is happening with the TANF program in Nassau County.⁴²

Conclusion

This model was built for the purpose of helping people go through a difficult conversation regarding a topic filled with uncertainty. They needed facilitation to have this conversation, including technology to help them “remember” their assumptions, “integrate” feedback-rich causal relationships, and deal with dynamic complexity. This intervention was apparently successful in terms of producing among the participants a consensual view of the welfare system and of welfare reform. The structure of the model and the behavioral patterns that it produced –base, policy and scenario runs– were found reasonable and plausible, even to audiences who had not engaged directly in the model building process.

The model proved useful also for the purpose of addressing “what if” questions. The model was built based upon causal relationships elicited from the participants, and it lent itself adequately for pattern analysis: “if we do so and so, then we expect so and so to happen...” The usefulness of this modeling work is not to predict how much something is rising or falling, or what will be the payoff of a particular investment. But to understand the multiple and interacting relationships in the model, and to be able to link behavior back to structure.

However, this case study is a good illustration of the multiple purposes usually involved in group model building efforts, and the difficulties related to the absence of clarity in problem definition and to having multiple clients/audiences. The overall purpose was broadly stated in terms of the creation of a management flight simulator for welfare reform viewed from the county level. But individual clients pursued different uses of this simulation environment.

Cortland and Dutchess Counties wanted to build a shared view of the welfare system and of welfare reform, more than to solve a particular problem. But, once the models were built, the clients sought to “test” their pet strategies for making welfare reform successful, and to discover high impact points in the system. As they built consensus on a shared view of this system, and confidence in some of the lessons emerging from this work, they aimed to communicate and involve their partners and

⁴² The county is recovering from a serious financial crisis, and the risk of bankruptcy (Toy, *The New York Times*, 2002):

The county has come a long way since the spring of 2000 when state officials created the Nassau County Interim Finance Authority, commonly known as NIFA, as part of a \$105 million state bailout to help pull the county away from the jaws of bankruptcy....

communities at large in the welfare reform effort. Thus, the model was used also as a focusing point for decision-making and resource reallocation efforts.

But, for this, the clients felt they needed to validate the model, which in their minds meant to have the model reproduce their numbers. It was only after the modeling team was able satisfy this client need, in terms of cross-sectional and longitudinal fit, that we were able to deviate the clients' attention from the numbers to the patterns of behavior. However, it was tempting to use the model as a forecasting tool, as revealed in this passage published in the *Empire State Report*:

The computer model is essentially a flow diagram, with arrows and boxes showing how families would move through the welfare system and formulas calculating how factors such as an increase in the unemployment rate or additional job training programs might affect the overall social services budget and welfare population. For example, in Cortland County, the welfare caseload is about 600 families, requiring about \$5.4 million in assistance per year. *The Flight Simulator indicated that a strategic emphasis on helping people keep the jobs they have or move to better ones might save the county \$500,000 a year.* [Emphasis added]

While the modeling team attempted to bring about the necessary focus to this model building effort, and provide clarity regarding the usefulness and limitations of this modeling work, the somewhat ambiguous purpose limited the confidence in the findings derived from this work.

In this paper, I have stretched the usefulness of this model to attempt to address three policy questions related to welfare reform. So, I have also redirected the use of this modeling work to policy analysis and evaluation for the purpose of system redesign. This modeling objective is at the heart of the system dynamics tradition, and I was interested in investigating the usefulness and limitations of this work for this particular purpose. Can this model be useful to guide welfare policy analysis and evaluation? While I believe the model is a useful tool to address these questions, I should do so with caution. Here I try to summarize my findings, and to further specify the limitations of this analysis.

POLICY QUESTION #1. *Will there be a burden to NYS and its counties?* In order to arrive at a model of welfare reform, we went through a lengthy process of model negotiation and design. The final assumptions contained in the model suggest that the potential crisis anticipated is not as harsh as one would expect. We can expect the TANF caseload to fall as a result of time limits, and we should be concerned with a rise (possibly steeper) in the safety net caseload. But, this shift from one program to the other will not imply a significant financial burden at the state and local level. This is because the federal government appears to contribute financially, one way or another, to programs run under the safety net umbrella.

The model has a broad boundary, and includes the elements involved to tackle a number of problems and questions. Given all of the other issues under consideration during this modeling effort, *limited attention was paid to providing this type of forecasting capability*. A more focused problem statement and purpose would probably have directed our efforts more efficiently and effectively toward gathering the necessary data, and using the best analytical methods to address this question. For example, we could have attempted to collect data (if such data existed) on the distribution of the clients' time on TANF, in order to improve the formulation of loss of eligibility. Also, we could have spent considerably more time in optimization and sensitivity analyses, to search for the "right" parameters, and to draw confidence intervals for the caseloads and financial variables.

Fundamentally, there is a counter-intuitive insight here that needs further investigation. We could have asked the question: Why is it, and how is it that the federal government contributes, directly or indirectly, to programs run under the safety net umbrella? Paradoxically, this finding contradicts one of the stated intents of the Welfare Reform Act, which was for states to devise their own solutions to end welfare dependency, *or bear the costs* of sustaining those people unable to attain self-sufficiency, after five cumulative years of receiving federally funded benefits (see p. 5).

POLICY QUESTION #2. *What's the "best" strategy?* Numerous simulations of alternative management strategies involving investments in different parts of the system highlighted that successful implementation of welfare reform depended upon a coordinated community-wide effort. Welfare reform goes beyond the reach and responsibilities of the departments of social services. The leverage appears to be at the edges of this system, because only at the edges can we effectively deflate the system. However, involving the community in this effort is no easy task. It is unlikely that any one strategy will produce the best results with respect to all criteria of interest, both in the short and long term. The goals in the minds of the decision-makers, and the positions they occupy in the system are likely to play a big role in their views of their favorite strategies. Ultimately, a concerted countywide response involving multiple community organizations, both public and private, depends on effective leadership and building a shared vision of what it means to be successful.

But, we are far from being finished with our analysis, and community leaders need our help to successfully engage in these important discussions. We need to look more critically at the list of criteria of success (pp. 29-30). Are some indicators more meaningful or important than others (e.g., recidivism)? Which indicators need to be looked at in the context of others (e.g., rates as a function of stocks, as in ratios, possibly)? We need to derive more robust guidelines for using the indicators, and highlight the dangers of looking at the indicators for their face value (e.g., job finding). Thus, a lot more can (and probably should) be done to help people derive better indicators, and modeling could help to guide the development of an information database and performance measurements to tackle this problem. An important question that needs to be addressed is whether the existing performance measures and incentive mechanisms

in place create the adequate setting to arrive at the most successful outcomes possible for the communities? And, subsequently, can we suggest what the “right” incentives might be?

No matter how far we take this analytical effort, it is important to keep in mind that it is still best when people understand the causal links in the system, and how behavior can be traced to causal structure. So, if we make a lot of progress developing indicators, and guidelines to use indicators, and even the “correct” incentive system, we still need to help people understand why a particular strategy works or not. Probably, no amount of directing people to the right answers will replace the value of having them understand what happens if we do so and so, and trace that explanation back to system structure.

POLICY QUESTION #3. *Can management really make a difference?* When eliciting the information to build this model, the people in the room told us that they made a small difference in this system, and that the state of the economy was the major player. This assumption was embedded in this model and reflected in the simulations. The historical time-series used to calibrate the model, and to assess goodness of fit was highly correlated with an exogenous factor: unemployment. Now, with new data, we observe an actual pattern of behavior that could not have been predicted using this model.

Can we attribute this new deviation to management effects? Should we fix the model to account for inadequate formulations of the system’s structure and parameters? Did we overlook management strategies or underestimate their effects? Has there been an actual change in welfare culture, affecting both clients and caseworkers, and consequently affecting the clients’ paths through the system and/or lengths of stay? Is there another reason for this discrepancy?

I draw two conclusions from this “weakness” or “limitation” of the model and of the analyses (then, in 1997, and now, in 2002), one from the point of view of system dynamics analysis, and the other from the point of view of using models to deal with social phenomena involving human beings. First, this serves to illustrate the point that “all models are wrong,” while they may be useful for certain purposes (Richardson and Pugh, 1981, p. 310).⁴³ Had we been primarily concerned with juxtaposing the relative power of explanation of the economic context, vis-à-vis the managerial context, in the

⁴³ In discussing validity, Richardson and Pugh (1981) quote Greeberger *et al.* (1976):

No model has ever been or ever will be thoroughly validated. “Useful,” “illuminating,” or “inspiring confidence” are more apt descriptors applying to models than “valid.”

They also stress: “it is meaningless to try to judge validity in the absence of a clear view of model purpose (p. 310).”

pattern of behavior of the caseloads, we probably should have been more precise in building, parameterizing and testing this model, *for this particular purpose*.

The way we designed the model, to simply address “what if” types of questions, led us to build and parameterize the model with simple formulations leading directly from input to output. The clients were then asked to assess the relative strengths of unemployment and investments in services (inputs), in terms of flows in the system (outputs). The table functions elicited forced the clients to make judgments that probably were outside the range of their expertise. (This is the equivalent of asking them to work out non-linear differential equations mentally.) Thus, we might have asked the clients to tell us what they were not able to know. Alternatively, we could have formulated the model in more careful ways to ensure that those judgments (or data) were obtainable more accurately and reliably. (However, this is not a trivial task.)

The other conclusion also illustrates a point widely recognized in the systems thinking community, but that we sometimes stubbornly neglect. If the clients state that the main causes of their problems, and the main factors that determine their performance are outside the reach of their control (*i.e.*, they are exogenous), it is probably crucial to challenge this assumption (Senge, 1990, pp. 19-21).⁴⁴ Why do they think they cannot make a difference? What makes them think so? Can we provide evidence to the contrary, to get the clients to work harder at things that they can do to *really* make a difference?

References

- Allers R, R Johnson, DF Andersen, TP Lee, GP Richardson, J Rohrbaugh and AA Zagonel. Group model building to support welfare reform, part II: Dutchess county. *Proceedings of the 1998 International Conference of the System Dynamics Society*. Québec, Canada (July 20-23, 1998).
- Andersen DF and GP Richardson. 1980. Toward a pedagogy of system dynamics. In AA Legasto Jr, JW Forrester and JM Lyneis (eds.). *System Dynamics. TIMS Studies in the Management Sciences* 14. New York: North-Holland: 91-106.
- Andersen DF and GP Richardson. 1997. Scripts for group model building. *System Dynamics Review* 13 (2): 107-129.
- Andersen DF, GP Richardson and JAM Vennix. 1997. Group model building: adding more science to the craft. *System Dynamics Review* 13 (2): 187-201.

⁴⁴ As argued by Senge (1990, pp. 19-21):

There is in each of us a propensity to find someone or something outside ourselves to blame when things go wrong... When we focus only on our position, we do not see how our actions extend beyond the boundary of that position. When those actions have consequences that come back to hurt us, we misperceive these new problems as externally caused... *True proactiveness comes from seeing how we contribute to our own problems.*

- Barlas Y. 1989. Multiple tests for validations of system dynamics type of simulation models. *European Journal of Operational Research* 42 (1): 59-87.
- Barlas Y. 1994. Model validation in system dynamics. *Proceedings of the 1994 International Conference of the System Dynamics Society*. Stirling, Scotland: 1-10.
- Barlas Y and S Carpenter. 1990. Philosophical roots of model validation: two paradigms. *System Dynamics Review* 6 (2): 148-166.
- Bryson JM. 1995. *Strategic planning for public and nonprofit organizations: A guide to strengthening and sustaining organizational achievement*. San Francisco: Jossey-Bass.
- Checkland P. 1981. *Systems Thinking, Systems Practice*. Chichester, England: John Wiley & Sons.
- Edelman P. 1997. The worst thing Bill Clinton has done. *The Atlantic Monthly* (March): 43-58.
- Eden C. 1989. Strategic options development and analysis. In J Rosenhead (ed.). *Rational Analysis in a Problematic World*. London: John Wiley & Sons.
- Eden C. 1990. The unfolding nature of group decision support: two dimensions of skill. In C Eden and J Radford (eds.). *Tackling Strategic Problems: The Role of Group Decision Support*. London: SAGE Publications, 48-52.
- Eden C and F Ackermann. 1998. *Making Strategy: The Journey of Strategic Management*. London: SAGE Publications.
- Empire State Report*. January 1998. Understanding Welfare: 17.
- Forrester JW. 1961. *Industrial Dynamics*. Portland, Oregon: Productivity Press.
- Forrester JW and PM Senge. 1980. Tests for building confidence in system dynamics models. In AA Legasto Jr., JW Forrester and JM Lyneis (eds.). *System Dynamics. TIMS Studies in the Management Sciences* 14. New York: North-Holland, 209-228.
- Graham, AK. 1980. Parameter estimation in system dynamics modeling. In J Randers (ed.). *Elements of the System Dynamics Method*. Cambridge, MA: Productivity Press, 143-161.
- Greenberger M, MA Crensen and BL Crissy. 1976. *Models in the Policy Process*. New York: Russell Sage Foundation.
- Hodgson AM. 1994. Hexagons for systems thinking. In JDW Morecroft and JD Sterman (eds.). *Modeling for Learning Organizations*. Portland, OR: Productivity Press: 359-374.
- Huz S, DF Andersen, GP Richardson and R Boothroyd. 1997. A framework for evaluating systems thinking interventions: an experimental approach to mental health system change. *System Dynamics Review* 13 (2): 149-169.
- Johnson R, DF Andersen, GP Richardson, J Rohrbaugh, T.P. Lee, M Mojtahedzadeh, D Rosenstreich and AA Zagonel. 1997. Report on the group model building conference in Cortland County: issue elicitation (February 11), model conceptualization (March 17 and 18), model presentation (April 29), resource allocation (September 29 and 30), model diagrams, model equations, and Conference paper (System Dynamics Society, Istanbul, Turkey). Welfare Reform Project. Center for Policy Research. University at Albany.
- Lee TP. 1999. The dynamics of New York State welfare finance at county level: A feedback view of system behavior. Doctoral dissertation. Nelson A. Rockefeller

College of Public Affairs and Policy. University at Albany, State University of New York.

- Lee TP. 2001. The dynamics of New York State social welfare reform finance at the county level: A feedback view of system behavior. *Proceedings of the 2001 International Conference of the System Dynamics Society*. Atlanta, Georgia (July 23-27)
- Lee TP, DF Andersen, J Rohrbaugh and AA Zagonel. 1998a. Parameter booklet for the Dutchess County joined TANF and safety net model (April). Welfare Reform Project. Center for Policy Research. University at Albany.
- Lee TP, DF Andersen, J Rohrbaugh and AA Zagonel. 1998b. Parameter booklet for the Nassau County joined TANF and safety net model (September). Welfare Reform Project. Center for Policy Research. University at Albany.
- Lee TP, AA Zagonel, DF Andersen, J Rohrbaugh, GP Richardson. 1998c. A judgment approach to estimating parameter in group model building: A case study of social welfare reform at Dutchess County. *Proceedings of the 1998 International Conference of the System Dynamics Society*. Québec, Canada (July 20-23).
- Lyneis JM and AL Pugh. 1996. Automated vs. "hand" calibration of system dynamics models: an experiment with a simple project model. *Proceedings of the 1996 International Conference of the System Dynamics Society*. Boston, MA: 317-320.
- Lurie I. 1997. Temporary Assistance for Needy Families: a green light for the states. *The Journal of Federalism* 27 (2): 73-87.
- Lurie I and NM Riccucci. 2001. Changing the "culture" of welfare offices: From Vision to the front lines. Department of Public Administration and Policy, University at Albany, State University of New York.
- Oliva R. 1995. A Vensim module to calculate summary statistics for historical fit (D-4584). System Dynamics Group, Massachusetts Institute of Technology, Cambridge, MA. (Available from: <http://www.people.hbs.edu/roliva/research/sd/>)
- Oliva R. 1996. Empirical validation of a dynamic hypothesis. *Proceedings of the 1996 International Conference of the System Dynamics Society*. Boston, MA: 405-408.
- Oliva R. (unpublished). Model calibration as a testing strategy for system dynamics models. (Available from: <http://www.people.hbs.edu/roliva/research/sd/calibration.html>)
- Pavetti, LD. 1995. Who is affected by time limits? *Welfare Reform Briefs* 7. Washington, DC: Urban Institute.
- Reagan-Cirincione P, S Schuman, GP Richardson and SA Dorf. 1991. Decision modeling: tools for strategic thinking. *Interfaces* 21 (6): 52-65.
- Richardson GP and DF Andersen. 1995. Teamwork in group model building. *System Dynamics Review* 11 (2): 113-137.
- Richardson GP, DF Andersen, N Hsiao, TP Lee, J Rohrbaugh and AA Zagonel. 1997a. User's manual to the welfare reform TANF flight simulator, Version 1.0 (September). Welfare Reform Project. Center for Policy Research. University at Albany.
- Richardson GP, DF Andersen, N Hsiao, TP Lee, J Rohrbaugh and AA Zagonel. 1997b. User's manual to the welfare reform TANF flight simulator, Version 2.0 (November). Welfare Reform Project. Center for Policy Research. University at Albany.

- Richardson GP, DF Andersen, J Rohrbaugh and W Steinhurst. 1992. Group model building. *Proceedings of the 1992 International Conference of the System Dynamics Society*. Utrecht, The Netherlands: 595-604.
- Richardson GP, DF Andersen and YJ Wu. 2002. Misattribution in welfare dynamics: The puzzling dynamics of recidivism. *Proceedings of the 2002 International Conference of the System Dynamics Society*. Palermo, Italy (July 28-August 1).
- Richardson GP and AL Pugh III. 1981. *Introduction to System Dynamics Modeling with DYNAMO*. Cambridge, MA: Productivity Press.
- Roberts EB (ed.). 1978. *Managerial Applications of System Dynamics*. Cambridge, MA: The MIT Press.
- Roberts N, D Andersen, R Deal, M Garet and W Shaffer. 1983. *Introduction to Computer Simulation: A System Dynamics Modeling Approach*. Portland, Oregon: Productivity Press.
- Rogers J, R Johnson, DF Andersen, J Rohrbaugh, GP Richardson, TP Lee, and AA Zagonel. 1997. Group model building to support welfare reform in Cortland County. In *Proceedings of the 1997 International Conference of the System Dynamics Society*. Istanbul, Turkey (August 19-22): 605-612.
- Rohrbaugh J. 2000. The use of system dynamics in decision conferencing: implementing welfare reform in New York State. In G. David Garson (ed.). *Handbook of Public Information Systems*. New York: Marcel Dekker, 521-533.
- Rohrbaugh J, DF Andersen, TP Lee, I Lurie, GP Richardson and AA Zagonel. 1997. Report on the group model building conferences in Dutchess County (August): issue elicitation (May 30), model conceptualization (June 3 and 4), model presentation (July 29), model diagrams and equations. Welfare Reform Project. Center for Policy Research. University at Albany.
- Rohrbaugh J and R Johnson. 1998. Welfare reform flies in New York. *Government Technology* (June): 58.
- Rouwette EAJA, JAM Vennix and T van Mullekom. 2002. Group model-building effectiveness: a review of assessment studies. *System Dynamics Review* 18 (1): 5-45.
- Senge, P.M. 1990. *The fifth discipline: the art and practice of the learning organization*. New York: Doubleday.
- Stenberg L. 1980. A modeling procedure for public policy. In J Randers (ed.). *Elements of the System Dynamics Method*. Cambridge, MA: Productivity Press, 292-312.
- Sterman JD. 1984. Appropriate summary statistics for evaluating the historical fit of system dynamics models. *Dynamica* 10 (Winter): 51-66.
- Sterman JD. 1994. Learning in and about complex systems. *System Dynamics Review* 10 (2-3): 291-330.
- Sterman JD. 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Boston, MA: The McGraw-Hill Companies, Inc.
- Thompson FJ. 1996. Devolution and the states: challenges for public management. Paper presented at the Annual Meeting of the Association for Public Policy Analysis and Management. Pittsburgh, Pennsylvania (October 31 - November 2).
- Toy VS. 2002. Tough love from angel on Nassau's shoulder. *The New York Times* (December 15, Sunday, Late Edition): s. 14LI, p. 1 (Long Island Weekly Desk).

- Mosteller F and JW Tukey. 1977. *Data Analysis and Regression: A Second Course in Statistics*. Reading, MA: Addison-Wesley Publishing Company.
- Vennix JAM. 1996. *Group Model Building: Facilitating Team Learning Using System Dynamics*. London: John Wiley & Sons.
- Vennix JMA. 1999. Group model-building: tackling messy problems. *System Dynamics Review* 15 (4): 379-401.
- Vennix JMA, DF Andersen and GP Richardson (eds.). 1997. *Group Model Building*. *System Dynamics Review* 13 (2).
- Wolstenholme EF. 1994. A Systematic approach to model creation. In JDW Morecroft and JD Sterman (eds.). *Modeling for Learning Organizations*. Portland, Oregon: Productivity Press: 175-194.
- Zagonel AA. 1997. Report on the group model building conference to join the TANF and SN models: Financial issues (December 1). Welfare Reform Project. Center for Policy Research. University at Albany.
- Zagonel AA. 2002a. Model conceptualization in group model building: a review of the literature exploring the tension between representing reality and negotiating a social order. *Proceedings of the 2002 International Conference of the System Dynamics Society*. Palermo, Italy (July 28-August 1).
- Zagonel AA. 2002b. Reflecting on group model building used to support welfare reform in New York State. Dissertation proposal. Nelson A. Rockefeller College of Public Affairs and Policy. University at Albany, State University of New York.
- Zagonel AA, DF Andersen and J Rohrbaugh. 1997a. Report on the group model building conference in Cortland County: issue elicitation meeting (February 11). Welfare Reform Project. Center for Policy Research. University at Albany.
- Zagonel AA, DF Andersen, GP Richardson, J Rohrbaugh, M Mojtahedzadeh and D Rosenstreich. 1997b. Report on the group model building conference in Cortland County: model conceptualization (March 17 and 18). Welfare Reform Project. Center for Policy Research. University at Albany.
- Zagonel AA, DF Andersen, GP Richardson, J Rohrbaugh, TP Lee. 1997c. Report on the group model building conference in Cortland County: model presentation (April 29). Welfare Reform Project. Center for Policy Research. University at Albany.
- Zagonel AA, DF Andersen, TP Lee, M Mojtahedzadeh, GP Richardson, J Rohrbaugh and D Rosenstreich. 1997d. Report on the group model building conferences in Cortland County (May): issue elicitation (February 11), model conceptualization (March 17 and 18), model presentation (April 29), model diagrams and equations. Welfare Reform Project. Center for Policy Research. University at Albany.
- Zagonel AA, DF Andersen, TP Lee and J Rohrbaugh. 1997e. Report on the group model building conference in Dutchess County: model calibration (October 16). Welfare Reform Project. Center for Policy Research. University at Albany.
- Zagonel AA, DF Andersen, TP Lee and J Rohrbaugh. 1998. Report on the group model building conference in Nassau County: model calibration (June 16). Welfare Reform Project. Center for Policy Research. University at Albany.

Figure 1. Tracing a genealogy of group model building

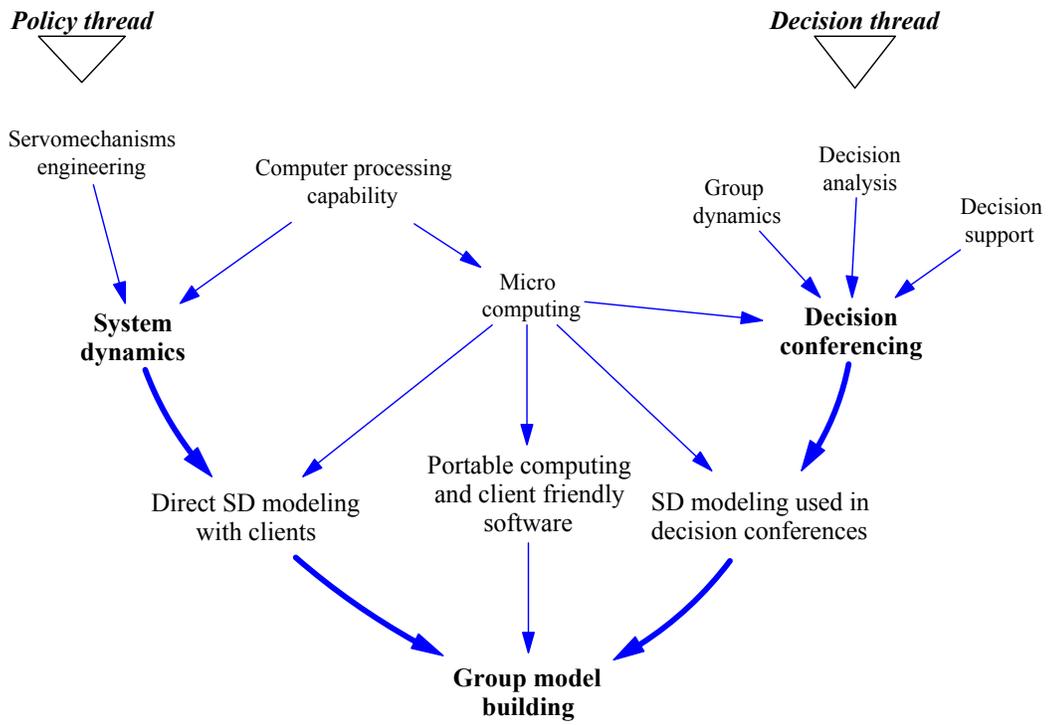


Figure 2. Streams of activities of the Welfare Reform project

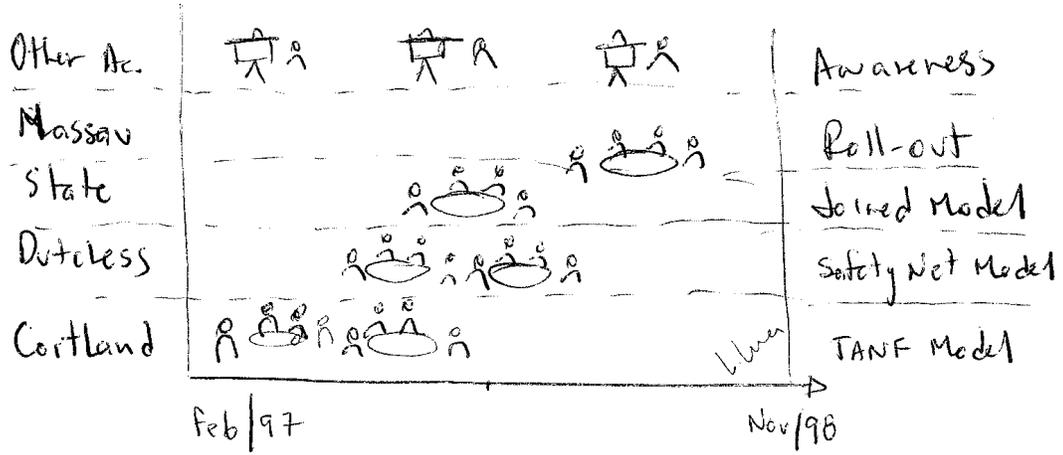


Figure 3. A sample view of the WR model: Job maintenance services (JMS)

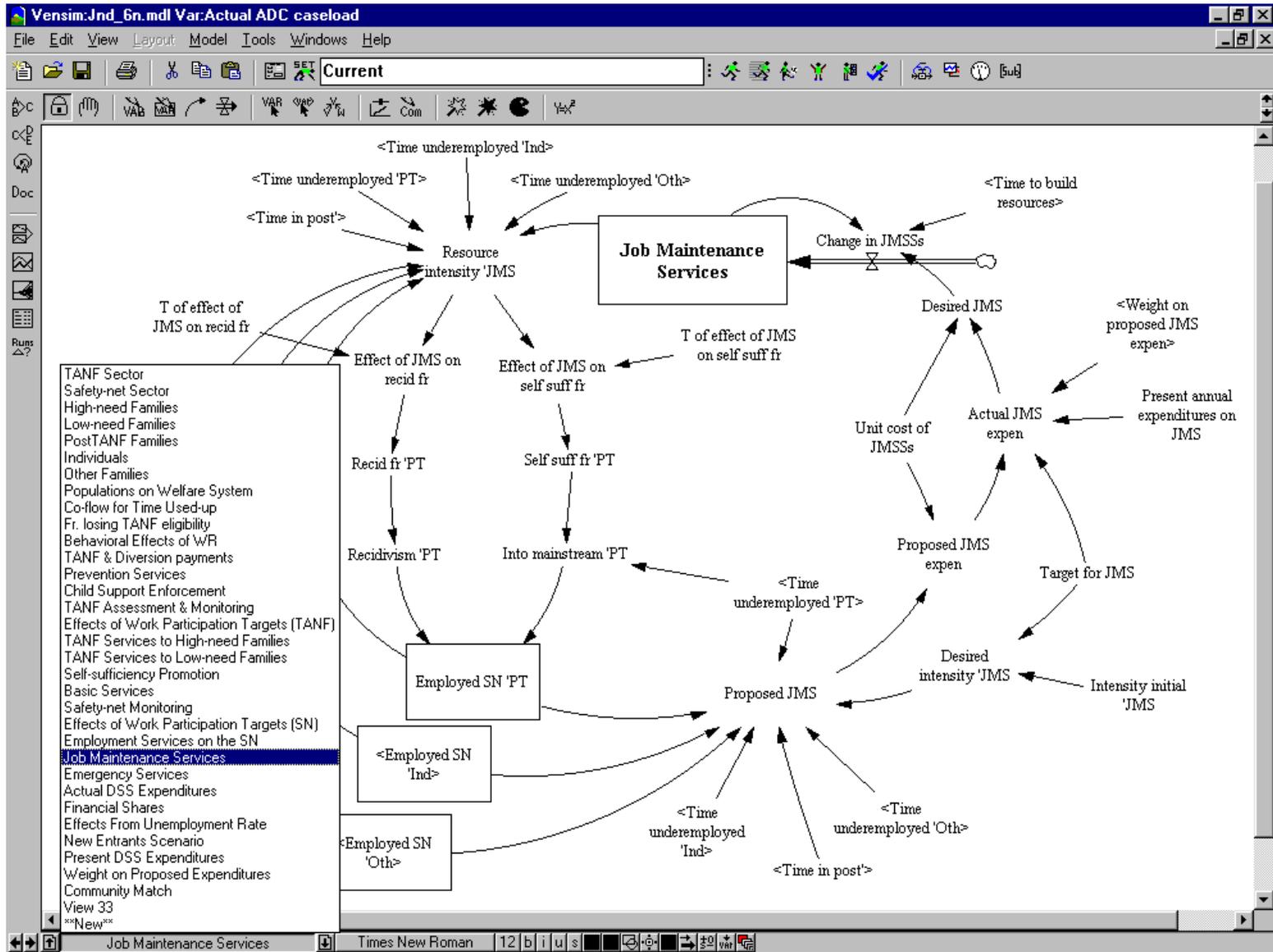


Figure 4. Summary view of the stock-and-flow structure of the WR model

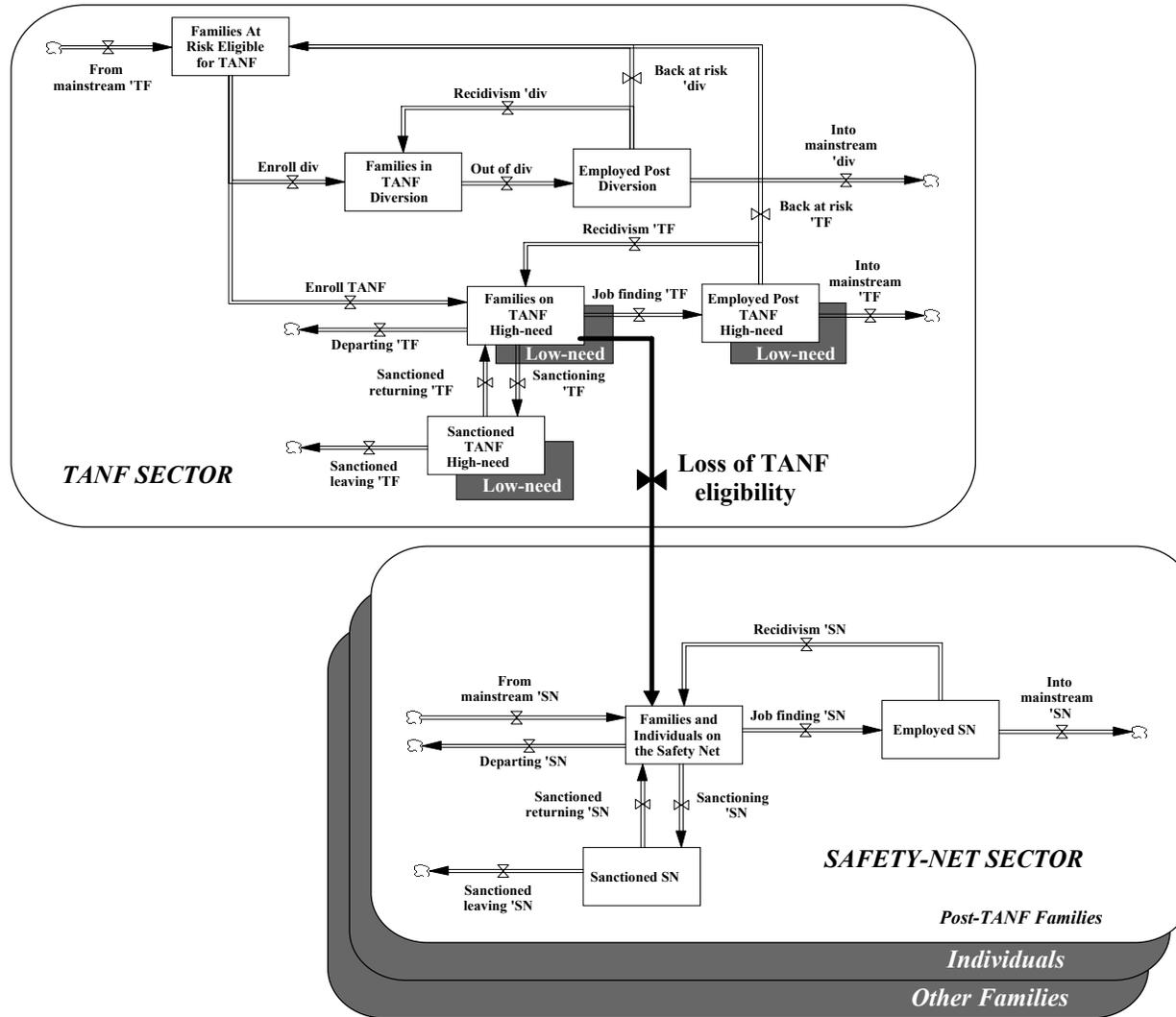


Figure 5. Stock-and-flow structure of the TANF sector with resource impacts

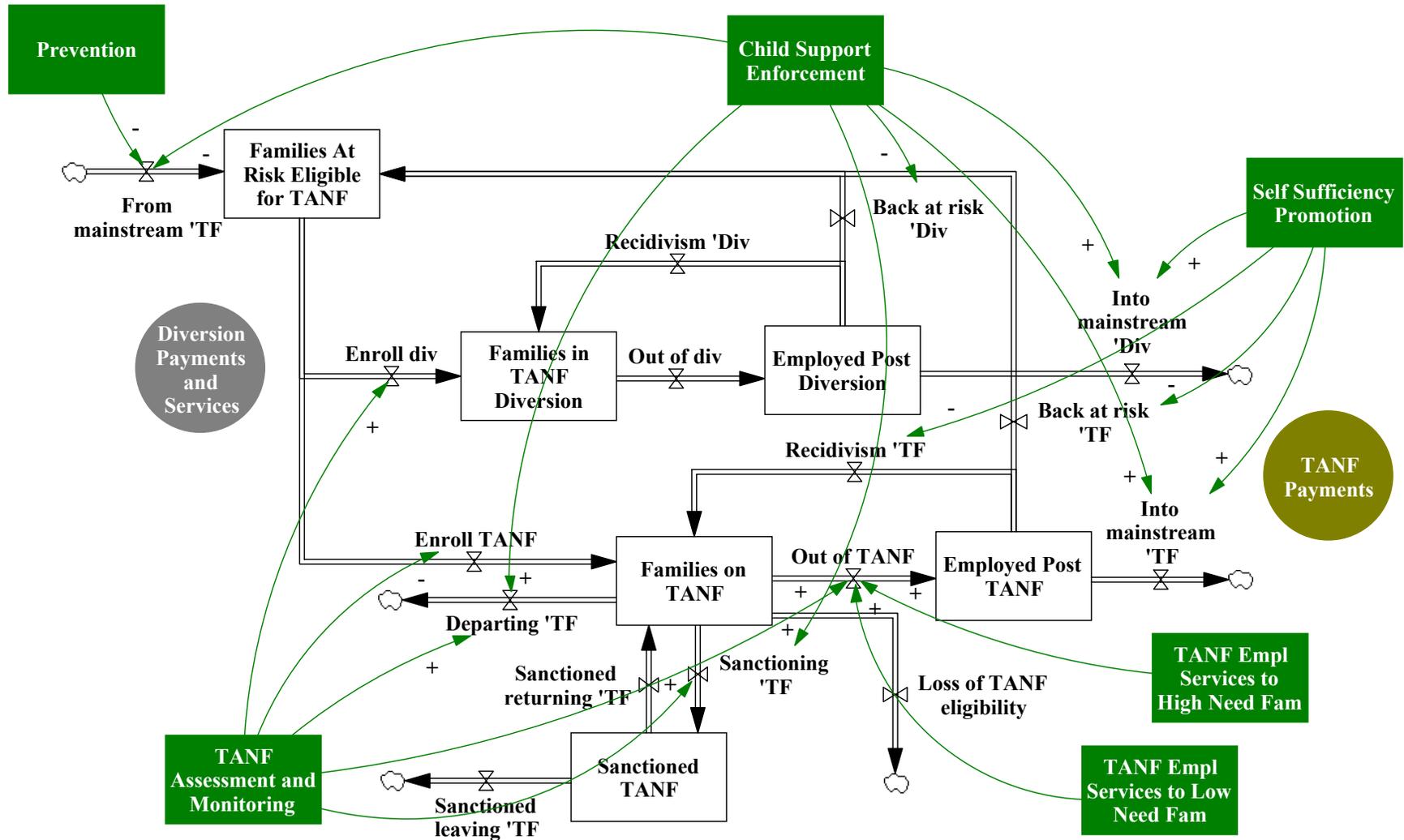


Figure 6. Stock-and-flow structure of the SN sector with resource impacts

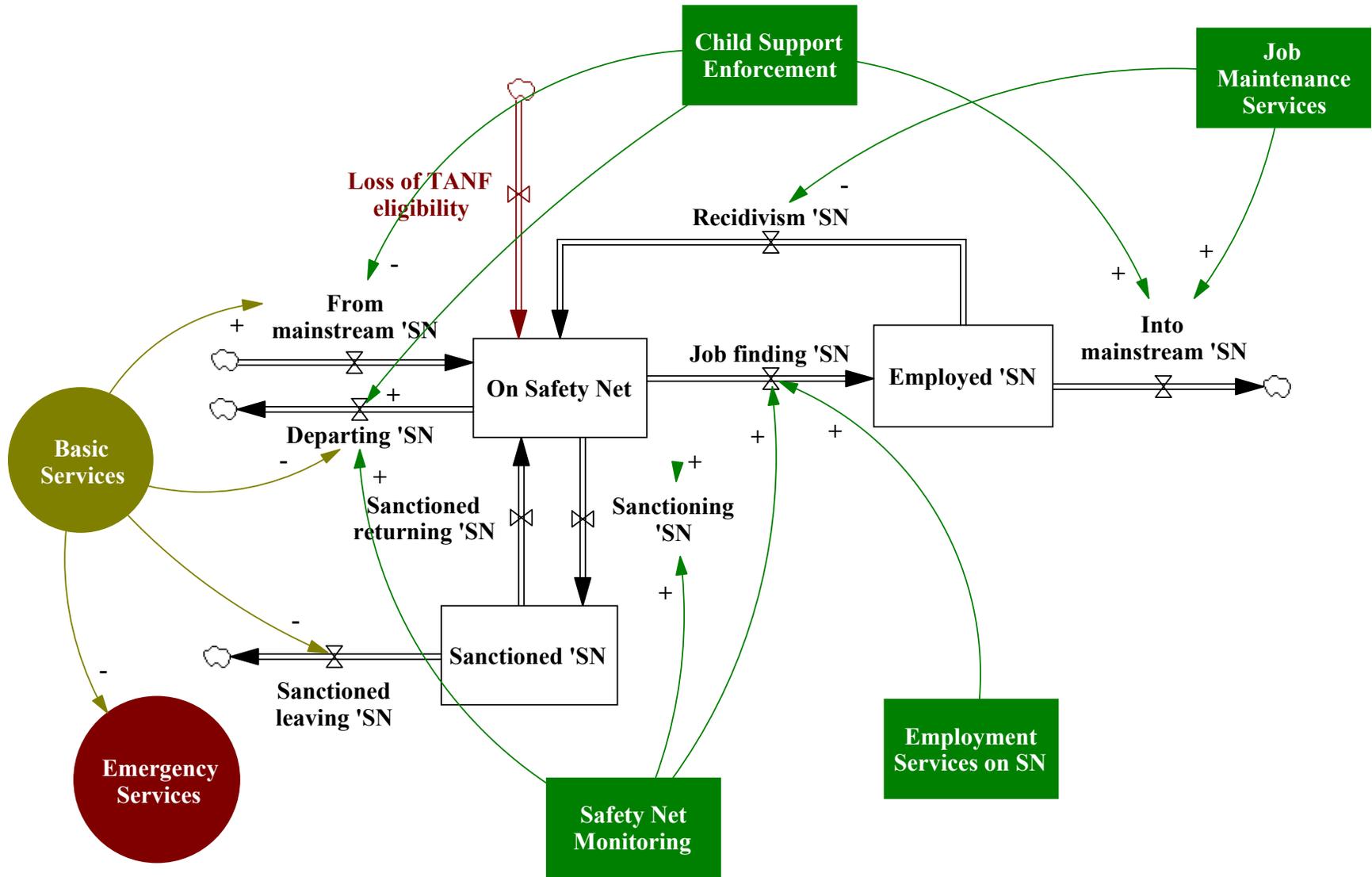


Figure 7. Control panel of the management flight simulator of the WR model

The figure displays two screenshots of the Welfare Reform Simulator control panel. The left screenshot, titled "Base World Characteristics", shows six assumptions with their respective settings and "Details" buttons. The right screenshot, titled "Tailoring Your Policy for Simulation", shows a list of adjustable services, a "Resource Target Multiplier" section with sliders, and a "Current Year = 1997" indicator.

Base World Characteristics

- Assumption 1: Assuming one of the two client behaviors**
 - Historic Optimistic
 - Details
- Assumption 2: Choosing one of the following scenarios**
 - Environment responds to Base Safety Net Package
 - Environment does not respond
 - Details
- Assumption 3: Community match**
 - Set community match ratios
 - Details
- Assumption 4: County size**
 - Small Medium Large
 - Details
- Assumption 5: Budget flexibility**
 - Set budget flexibilities
 - Details
- Assumption 6: Unemployment scenario**
 - Set unemployment over time
 - Details

Back to the Prior Screen

Tailoring Your Policy for Simulation

Adjustable Services

- Prevention
- TANF Assessment and Monitoring
- TANF High Need
- TANF Low Need
- Self-Sufficiency Promotion

- Safety Net Monitoring
- SN Employment Services
- SN Job Maintenance Services
- Child Support Enforcement

Resource Target Multiplier

Resource intensities are ratios of resources to need. Setting a slider to 1 targets a resource intensity equal to 1997 conditions.

Setting a slider to 0 sets the target for that resource to 0.

Setting a slider to 2 strives to increase that resource-percentage to double the 1997 level.

Setting a slider to any value between 0 and 2 sets a target resource intensity for that multiple of 1997 conditions.

After tailoring, you should ...

Current Year = 1997

Advance Simulation

Help for Services | Policy Diagrams | New Simulation

Figure 9. Base run: shift in caseloads

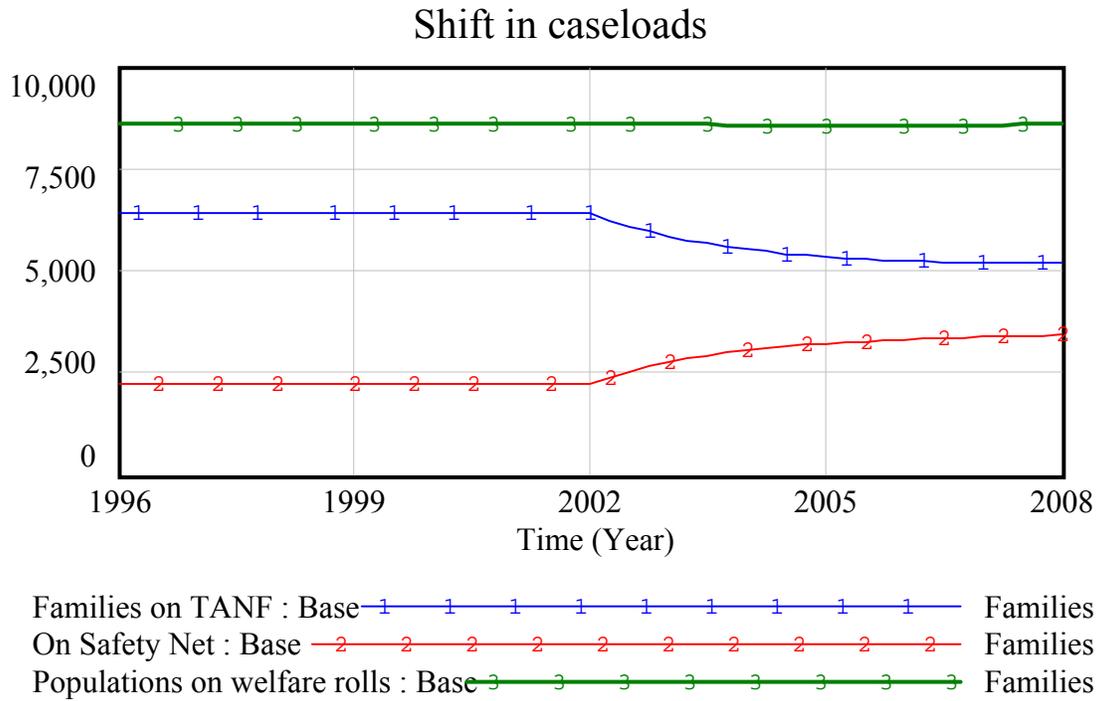
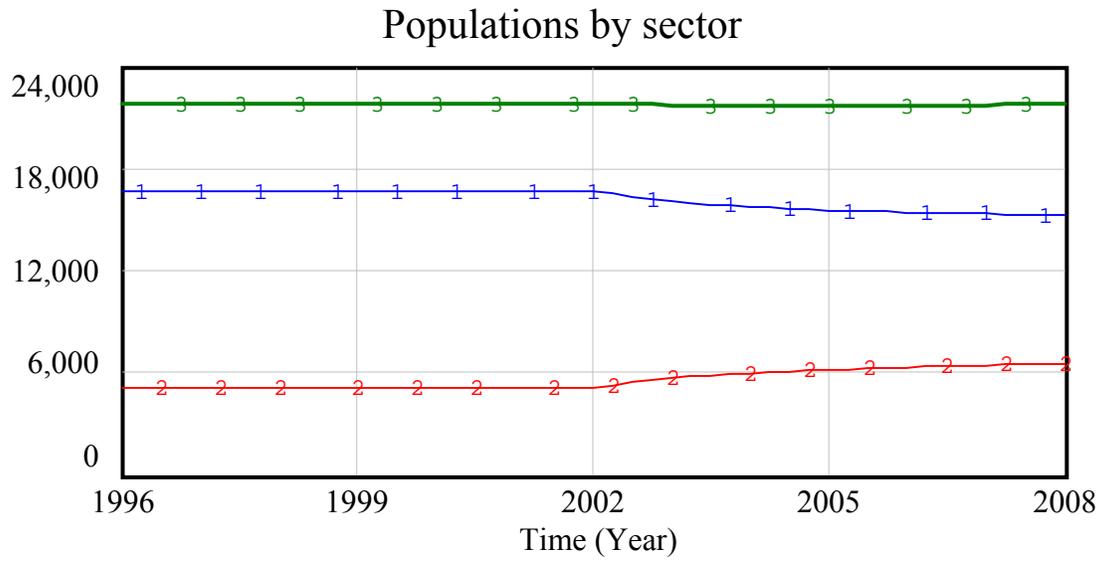
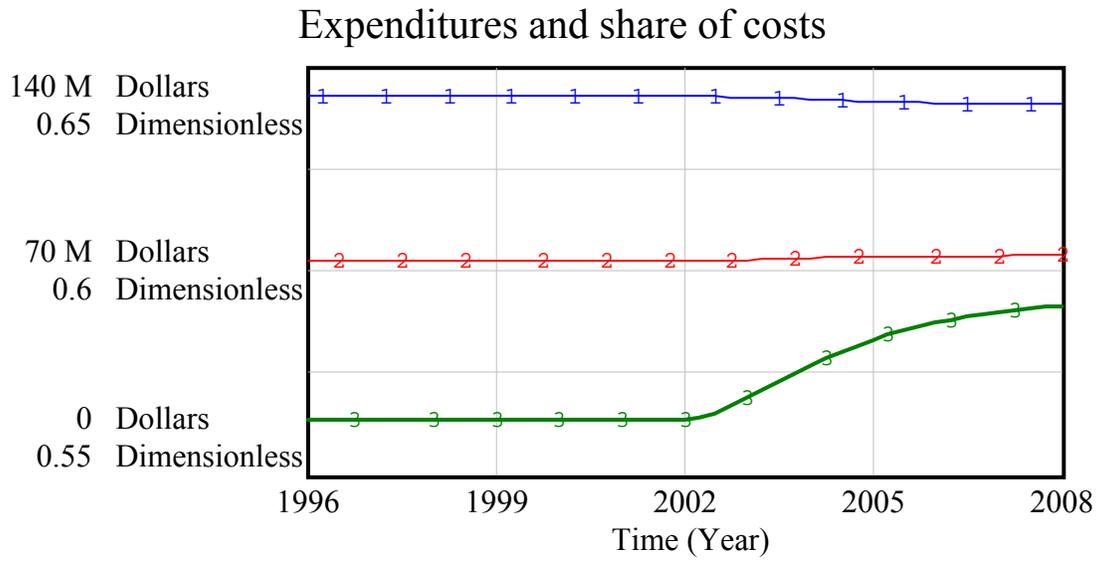


Figure 10. Base run: populations by sector



Populations on TANF system : Base 1 1 1 1 1 1 1 1 Families
Populations on SN system : Base 2 2 2 2 2 2 2 2 Families
Populations on welfare system : Base 3 3 3 3 3 3 3 3 Families

Figure 11. Base run: expenditures and share of costs



Actual DSS expen : Base 1 1 1 1 1 1 1 1 1 1 1 Dollars
 State and local DSS expen : Base 2 2 2 2 2 2 2 2 2 2 Dollars
 State and local share : Base 3 3 3 3 3 3 3 3 3 3 Dimensionless

Figures 12 (A-B). Historic run: simulated v. actual AFDC/TANF caseload

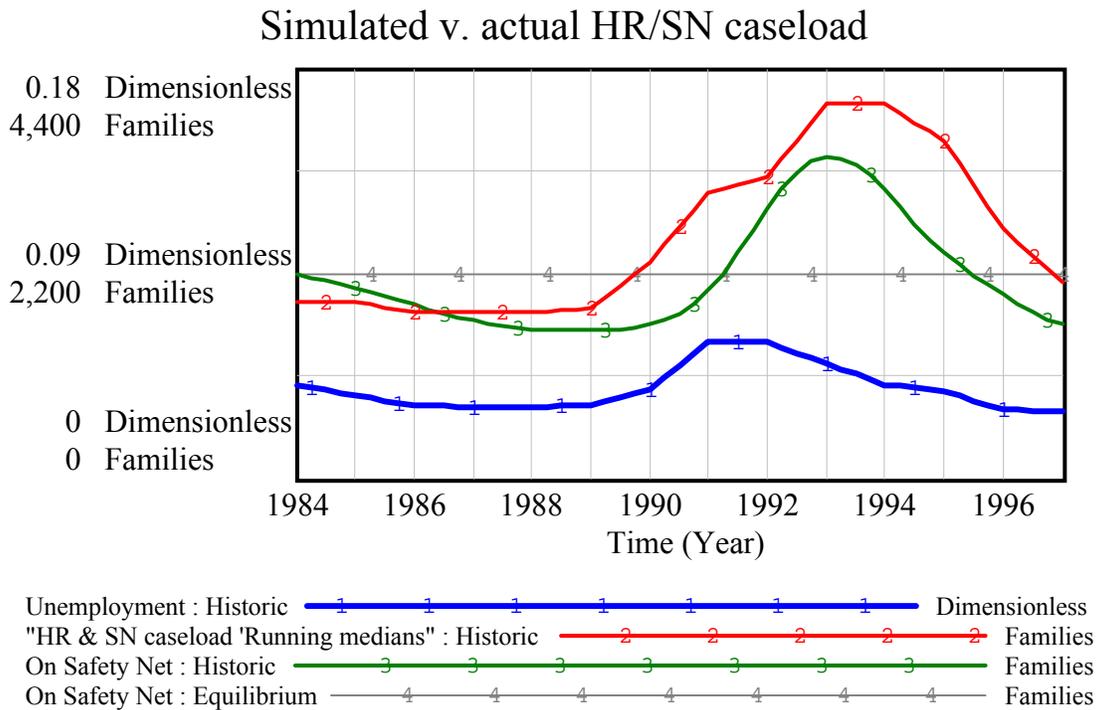
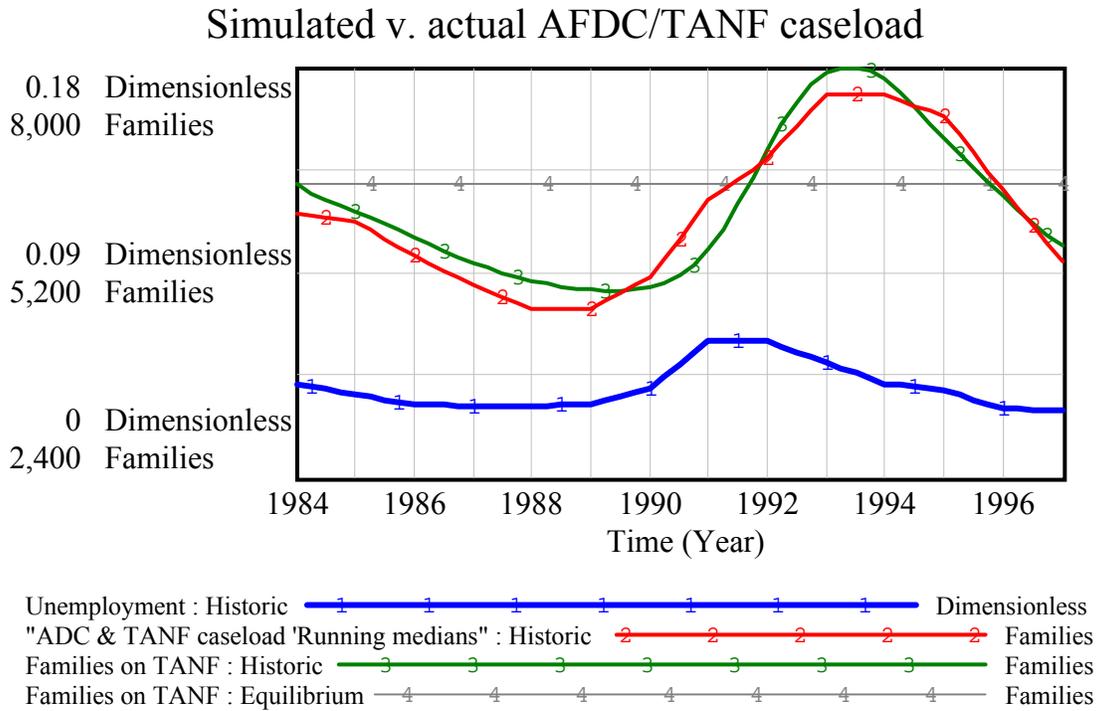


Figure 13. The effects of the “middle” and “edges” strategies on the TANF sector

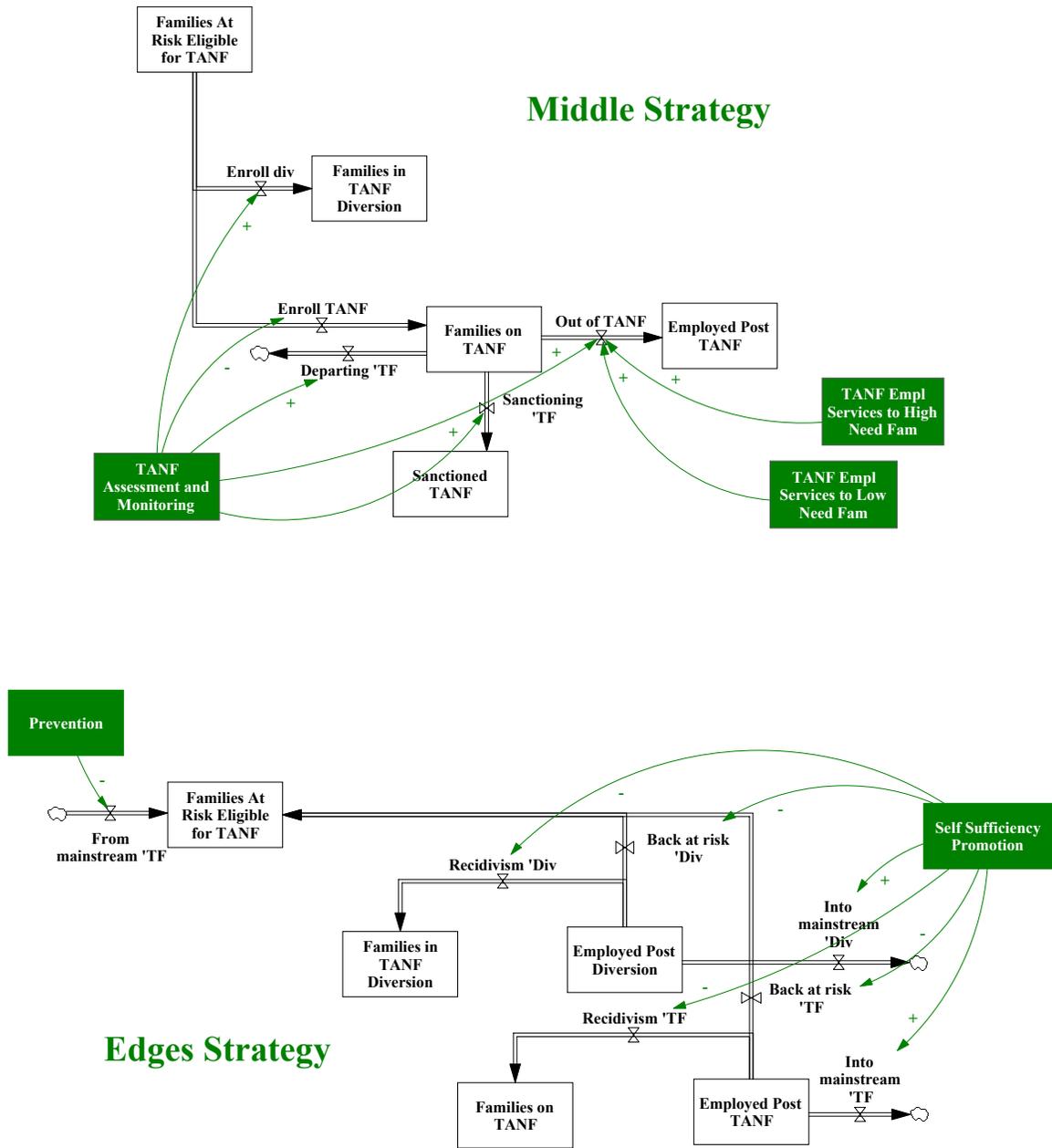


Figure 14. E.g.: investment on TANF Assessment and Monitoring (TAM)

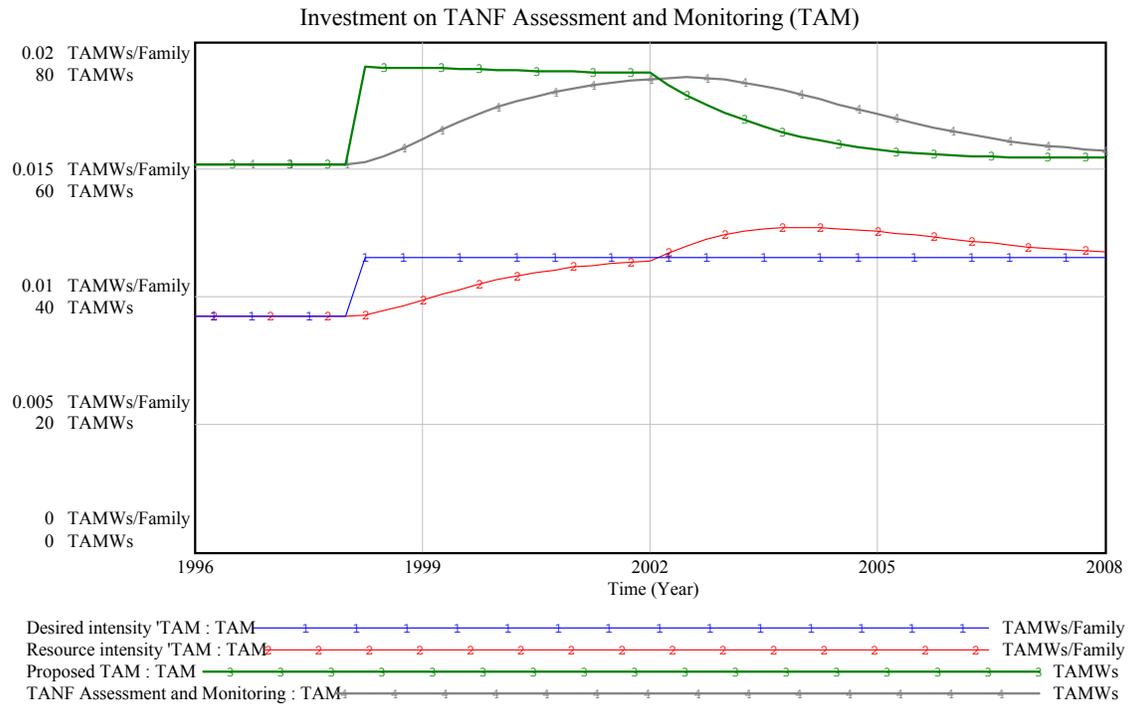


Figure 15. E.g.: results of investment on TAM

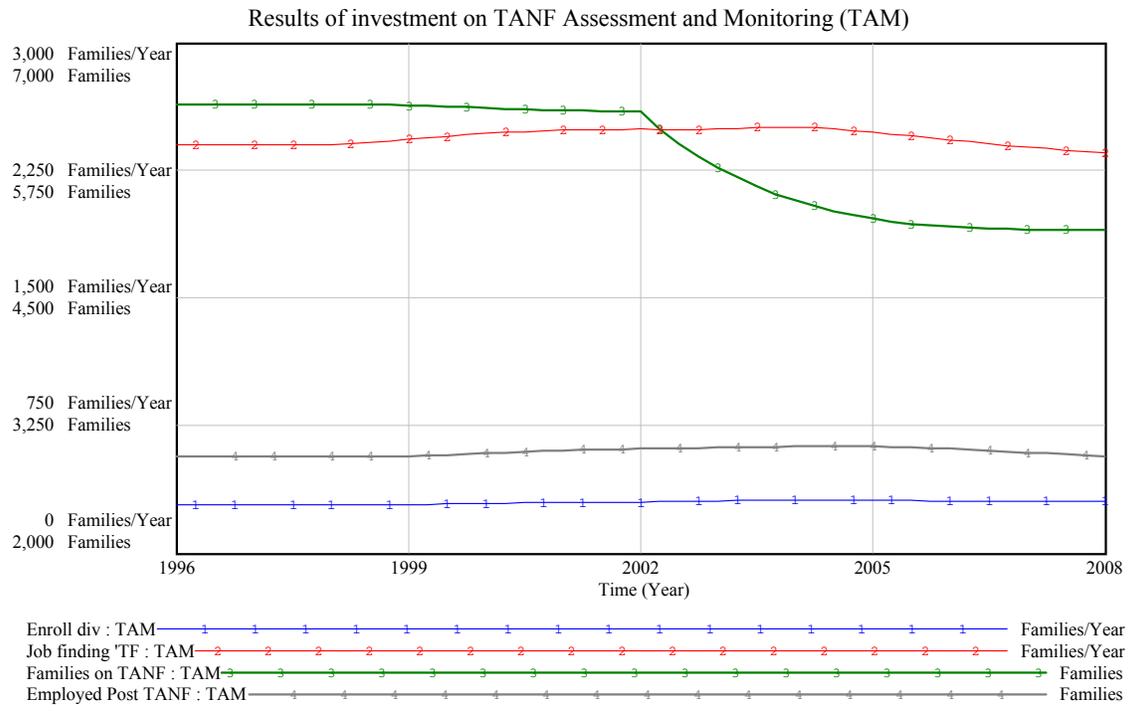


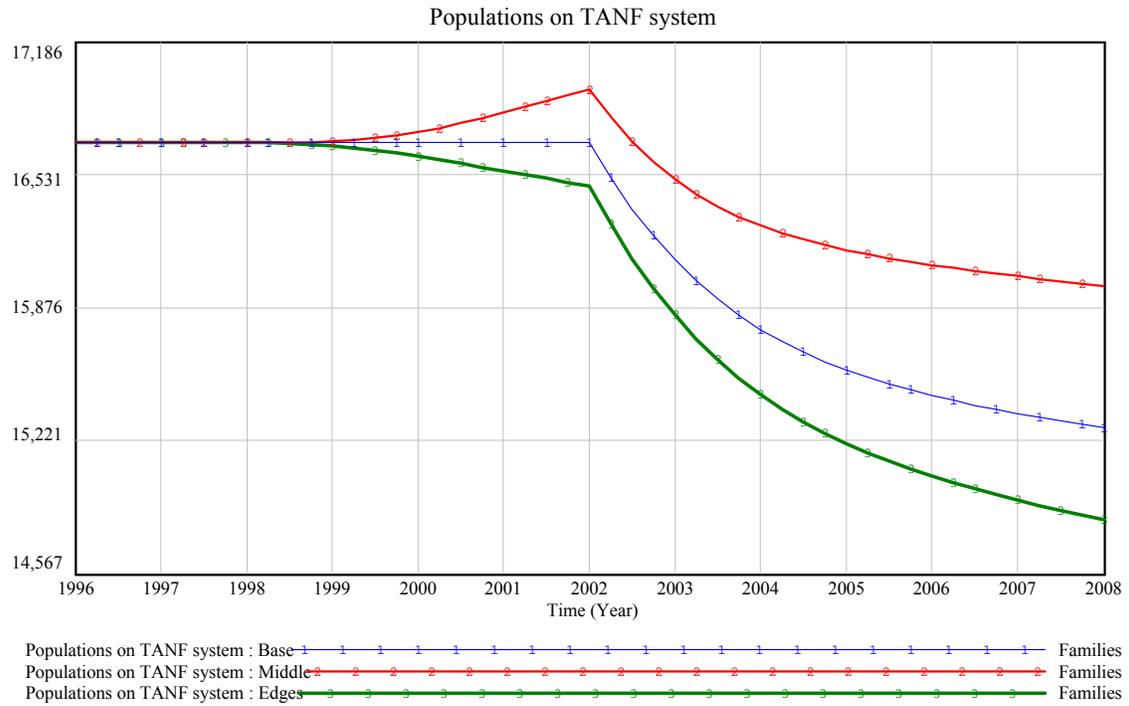
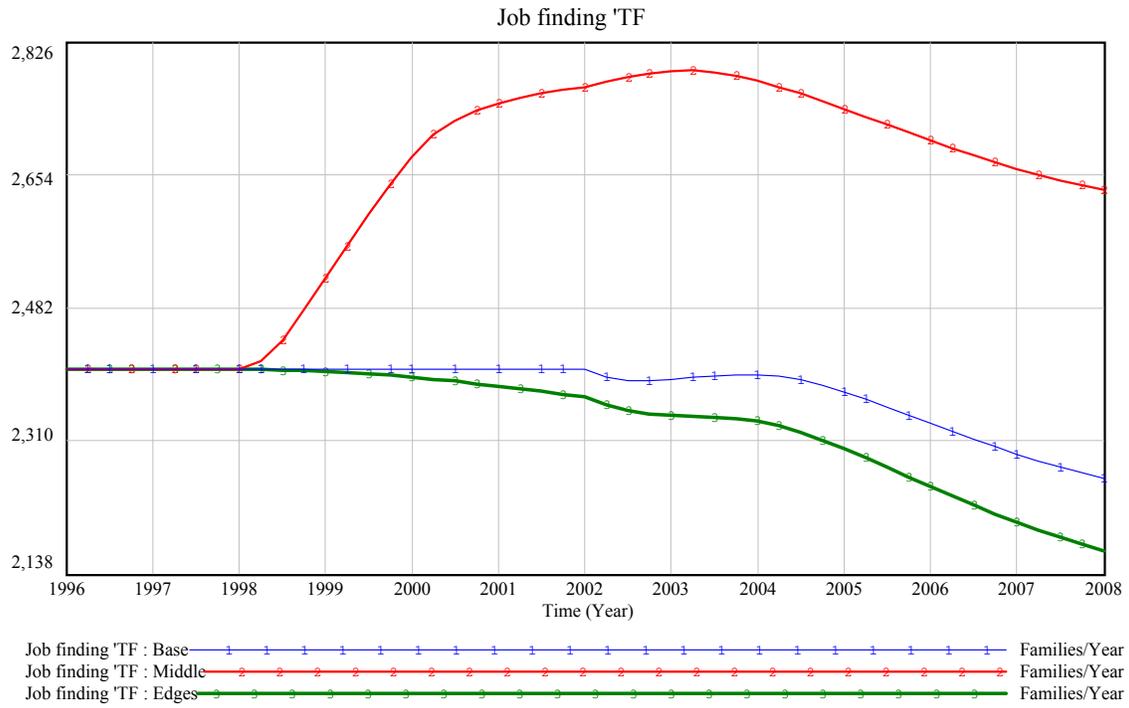
Table 2. Condensed results of the “middle” versus “edges” strategies

Populations:	Base run:	Middle strategy:	Edges strategy:
Families on TANF	Worst	Best in short term (until 2004)	Best in long term (after 2004)
Families and individuals on the SN	Worst	Best	---
Populations on TANF system	---	Worst	Best
Populations on the SN system	---	Worst	Best
Families in TANF diversion	---	Best?	Worst?
Populations employed but at risk	---	Best?	Worst?

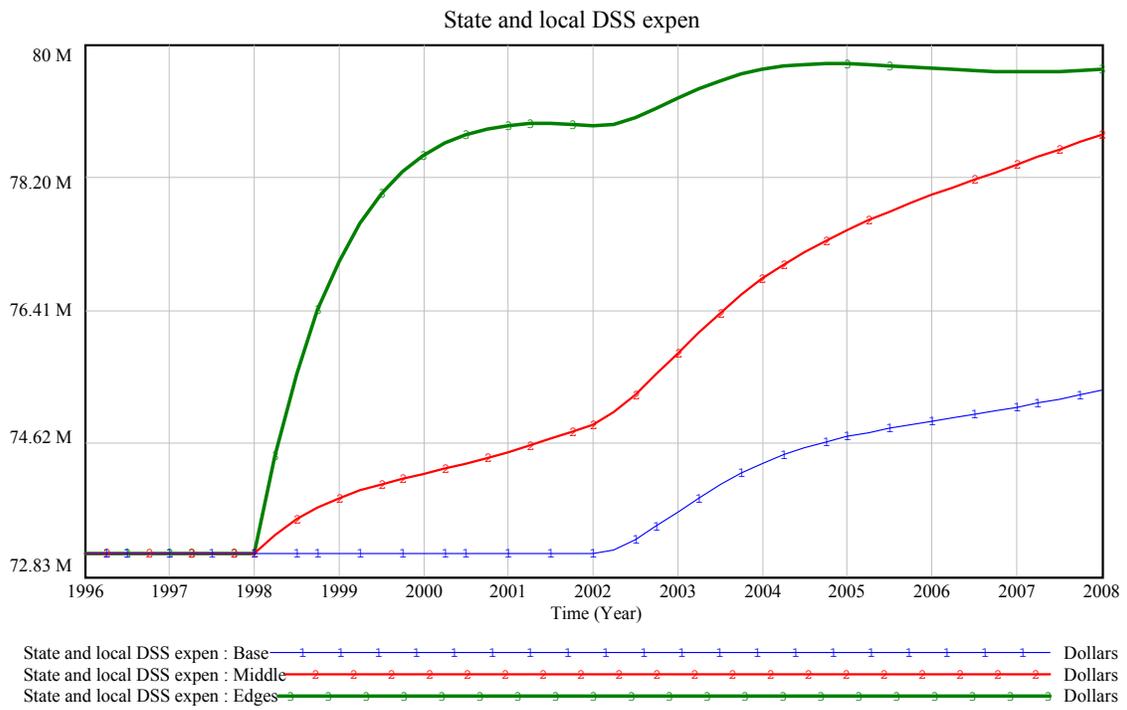
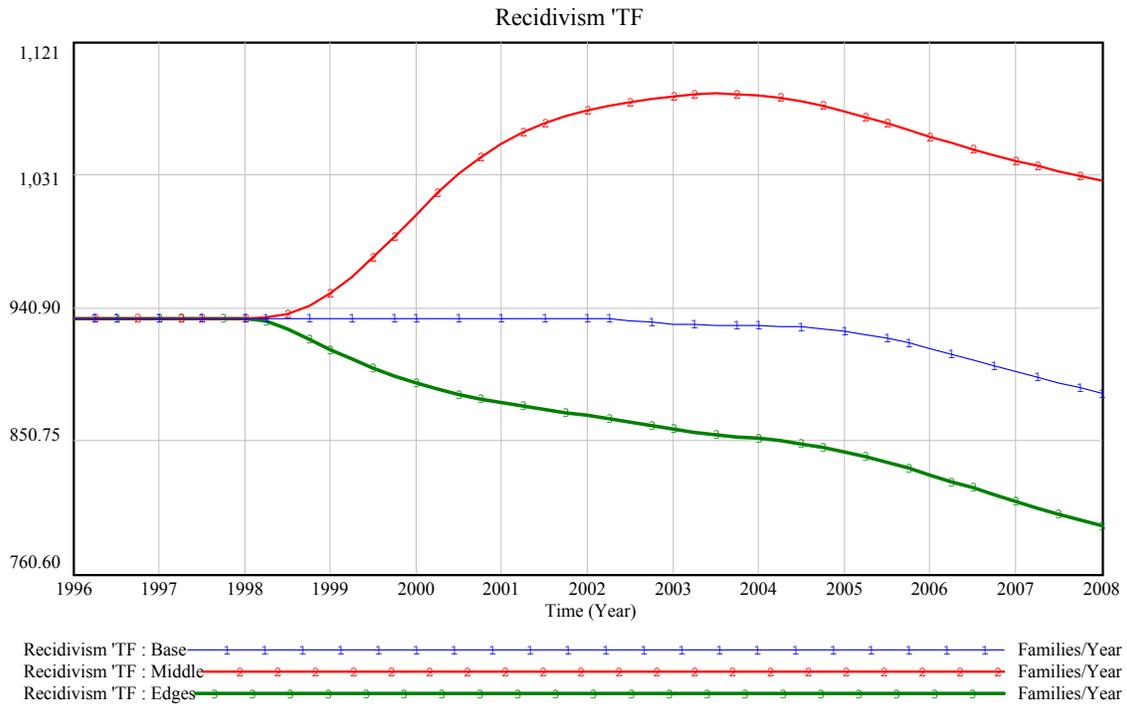
Rates of flow:	Base run:	Middle strategy:	Edges strategy:
From mainstream economy	---	---	Best
Enrolled in diversion	---	Best?	Worst?
Enrolled in TANF	---	Worst?	Best?
Job finding	---	Best	Worst
Loss of TANF eligibility	---	Best in short term (until 2006)	Best in long term (after 2006)
Recidivism	---	Worst	Best
Back at risk	---	Worst	Best
Into mainstream economy	Worst	---	Best

Costs:	Base run:	Middle strategy:	Edges strategy:
Actual expenditures	Best	Worst in long term (after 2006)	Worst in short term (until 2006)
Local expenditures	Best in short term (until 2002)	Worst in short term (until 2001)	Best in long term (after 2003)
State and local expenditures	Best	---	Worst
Local share	Worst	---	Best
State and local share	Best in long term (after 2002)	Best in short term (until 2002)	Worst

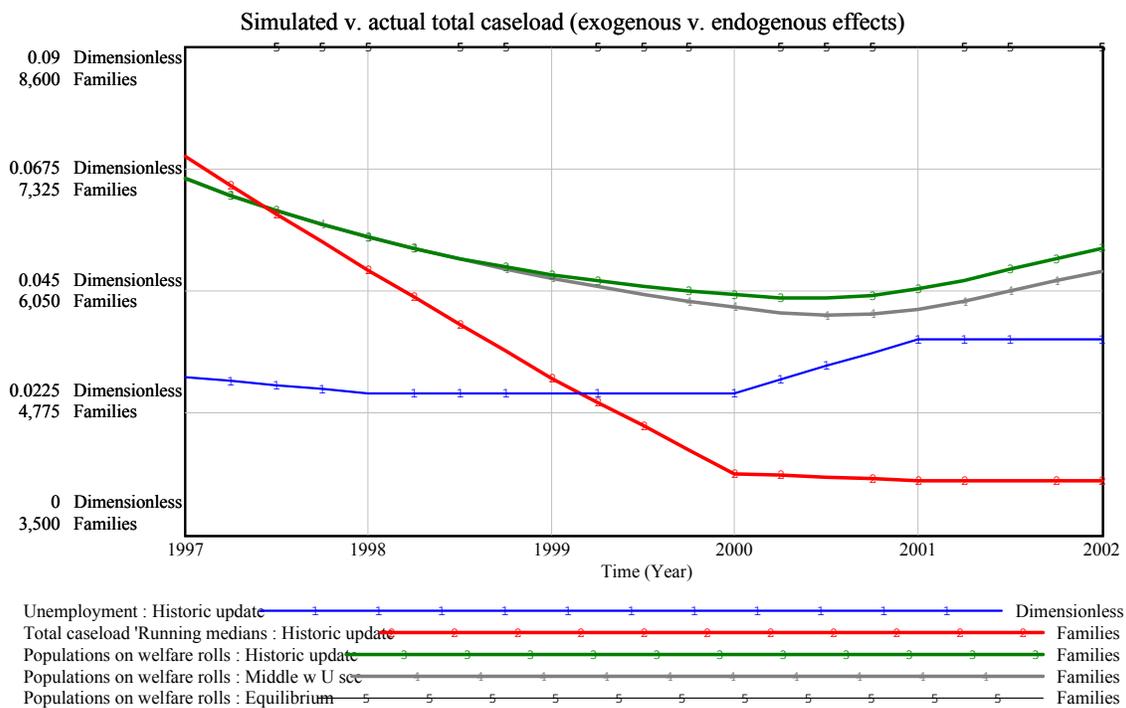
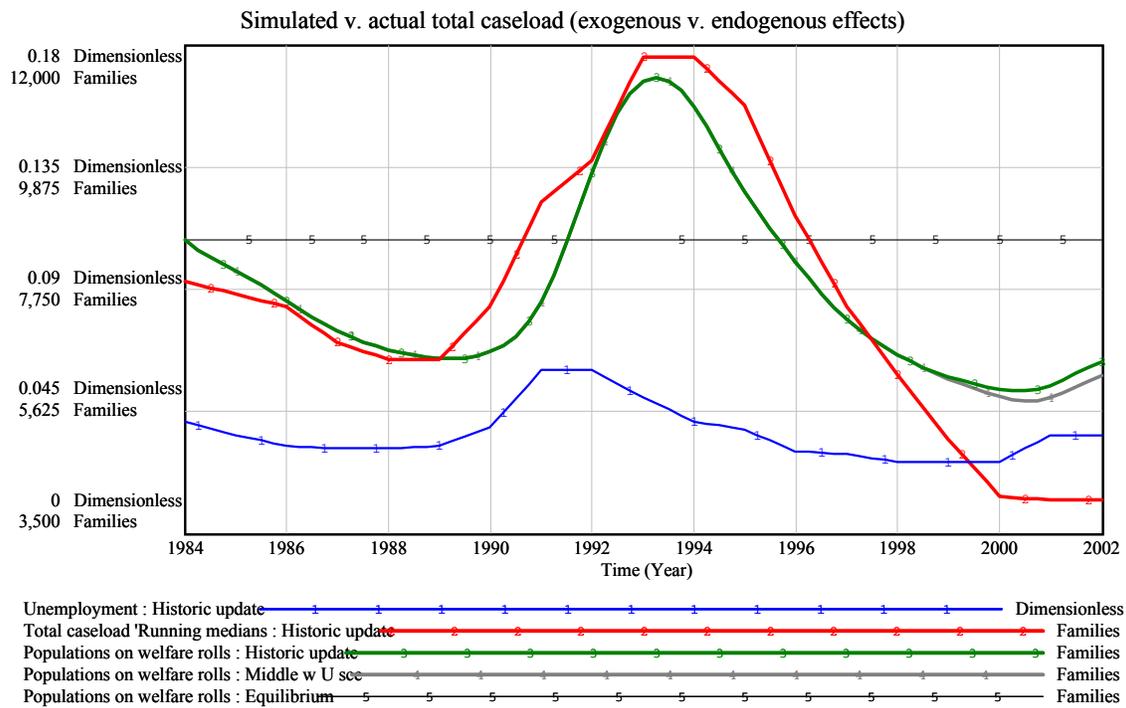
Figures 16 (A-B). E.g.: pros and cons of the “middle” strategy



Figures 17 (A-B). E.g.: pros and cons of the “edges” strategy



Figures 21 (A-B). Actual behavior unexplained (exogenously nor endogenously)



Appendix 1. Reports, presentations, papers, workshops, and dissertations

Group Model Building for Welfare Reform Project (1997-98)

List of references:

PROJECT REPORTS AND MANUALS

Cortland County:

C-1	Zagonel <i>et al.</i> 1997a	Zagonel AA, DF Andersen and J Rohrbaugh. Report on the group model building conference in Cortland County: <u>issue elicitation</u> (February 11). Welfare Reform Project. Center for Policy Research. University at Albany.
C-2/3	Zagonel <i>et al.</i> 1997b	Zagonel AA, DF Andersen, M Mojtahedzadeh, GP Richardson, J Rohrbaugh and D Rosenstreich. Report on the group model building conference in Cortland County: <u>model conceptualization</u> (March 17 and 18). Welfare Reform Project. Center for Policy Research. University at Albany. – WELFARE 1-3.MDL
C-4	Zagonel <i>et al.</i> 1997c	Zagonel AA, DF Andersen, TP Lee, GP Richardson and J Rohrbaugh. Report on the group model building conference in Cortland County: <u>model presentation</u> (April 29). Welfare Reform Project. Center for Policy Research. University at Albany. – PHASE 6E.MDL
C-1/2/3/4	Zagonel <i>et al.</i> 1997d	Zagonel AA, DF Andersen, TP Lee, M Mojtahedzadeh, GP Richardson, J Rohrbaugh and D Rosenstreich. Report on the group model building conferences in Cortland County (May): <u>issue elicitation</u> (February 11), <u>model conceptualization</u> (March 17 and 18), <u>model presentation</u> (April 29), <u>model diagrams and equations</u> . Welfare Reform Project. Center for Policy Research. University at Albany. – WELFARE 1-3.MDL and PHASE 6E.MDL
C-1/2/3/4 7/8	Johnson <i>et al.</i> 1997	Johnson R, DF Andersen, TP Lee, M Mojtahedzadeh, GP Richardson, J Rohrbaugh, D Rosenstreich and AA Zagonel. Report on the group model building conferences in Cortland County (December): <u>issue elicitation</u> (February 11), <u>model conceptualization</u> (March 17 and 18), <u>model presentation</u> (April 29), <u>resource allocation</u> (September 29 and 30), <u>model diagrams</u> , <u>model equations</u> and Conference paper (System Dynamics Society, Istanbul, Turkey). Welfare Reform Project. Center for Policy Research. University at Albany. - WELFARE 1-3.MDL , PHASE 6E.MDL and PHASE 8E.MDL

Dutchess County:

D-1/2/3/4	Rohrbaugh <i>et al.</i> 1997	Rohrbaugh J, DF Andersen, TP Lee, I Lurie, GP Richardson and AA Zagonel. Report on the group model building conferences in Dutchess County (August): <u>issue elicitation</u> (May 30), <u>model conceptualization</u> (June 3 and 4), <u>model presentation</u> (July 29), <u>model diagrams and equations</u> . Welfare Reform Project. Center for Policy Research. University at Albany. – PHASE 8A.MDL , SN 2.MDL and PHASE 8E.MDL
D-5	Zagonel <i>et al.</i> 1997e	Zagonel AA, DF Andersen, TP Lee and J Rohrbaugh. Report on the group model building conference in Dutchess County: <u>model calibration</u> (October 16). Welfare Reform Project. Center for Policy Research. University at Albany. – SN 2.MDL and PHASE 8F.MDL
	Lee <i>et al.</i> 1998a	Lee TP, DF Andersen, J Rohrbaugh and AA Zagonel. <u>Parameter booklet</u> for the Dutchess County joined TANF and safety net model (April). Welfare Reform Project. Center for Policy Research. University at Albany. – JND 6.MDL

Manuals:

W-2	Richardson <i>et al.</i> 1997a	Richardson GP, DF Andersen, N Hsiao, TP Lee, J Rohrbaugh and AA Zagonel. <u>User's manual to the welfare reform TANF flight simulator</u> , Version 1.0 (September). Welfare Reform Project. Center for Policy Research. University at Albany. – PHASE_8E.MDL
	Richardson <i>et al.</i> 1997b	Richardson GP, DF Andersen, N Hsiao, TP Lee, J Rohrbaugh and AA Zagonel. <u>User's manual to the welfare reform TANF flight simulator</u> , Version 2.0 (November). Welfare Reform Project. Center for Policy Research. University at Albany. – PHASE_8F.MDL

State:

J-1	Zagonel 1997	Zagonel AA. Report on the group model building conference to join the TANF and SN models: <u>Financial issues</u> (December 1). Welfare Reform Project. Center for Policy Research. University at Albany. – PHASE_8E.MDL, PHASE_8F.MDL and SN_2.MDL
-----	-----------------	--

Nassau County:

N-2	Zagonel <i>et al.</i> 1998	Zagonel AA, DF Andersen, TP Lee and J Rohrbaugh. Report on the group model building conference in Nassau County: <u>model calibration</u> (June 16). Welfare Reform Project. Center for Policy Research. University at Albany. – JND_6N.MDL
	Lee <i>et al.</i> 1998b	Lee TP, DF Andersen, J Rohrbaugh and AA Zagonel. <u>Parameter booklet</u> for the Nassau County joined TANF and safety net model (September). Welfare Reform Project. Center for Policy Research. University at Albany. – JND_6N.MDL

CLIENT-ORIENTED BRIEFINGS AND PRESENTATIONS

B-1	April 1997	OTDA briefing 1 – “Pilot” TANF model (April 18, 1997) – PHASE_4A.MDL
C-4	April 1997	Cortland, Day 4 – Model presentation (April 29, 1997) – PHASE_6E.MDL
D-1	May 1997	Dutchess, Day 1 – Problem definition and TANF model presentation (May 30, 1997) – PHASE_8A.MDL
B-2	June 1997	OTDA briefing 2 – Project update (June 5, 1997) – PHASE_8A.MDL
C-6	July 1997	Cortland, Day 6 – Community-wide presentation (July 8, 1997) – PHASE_8D.MDL
B-3	July 1997	OTDA briefing 3 – Project update to NYS Commissioner (July 1997)
D-4	July 1997	Dutchess, Day 4 – Model presentation (July 29, 1997) – SN_2.MDL and PHASE_8E.MDL
N-1	February 1998	Nassau, Day 1 – Project presentation (February 19, 1998) – JND_4.MDL
D-6	April 1998	Dutchess, Day 6 – Community-wide presentation (April 28, 1998) – JND_6.MDL
N-3	September 1998	Nassau, Day 3 – Model presentation (September 22, 1998) – JND_6N.MDL
P-4	November 1998	Onondaga County Presentation (November 30, 1998) – JND_6N.MDL

WORKSHOPS

W-1/2	June 1997	Andersen DF and GP Richardson. "Systems Thinking: A case study on welfare reform in <i>Chugwa County</i> ." – PHASE_8A.MDL
W-1	June 1997	Andersen DF and GP Richardson. A system's approach to welfare reform at the county level. <i>The Leadership and Collaboration Classroom: Challenges in Transforming Human Services Systems (GOER)</i> . Auburn, New York (June 8-12, 1997). – PHASE_8A.MDL
W-2	September 1997	Andersen DF and GP Richardson. Policy and implementation strategies for welfare reform. <i>The Leadership and Collaboration Classroom: Challenges in Transforming Human Services Systems (GOER)</i> . Hudson Valley Community College. Troy, New York (September 8-9, 1997). – PHASE_8E.MDL

Cortland County:

C-7/8	1997	Cortland, Days 7&8 – Resource allocation conference (September 29-30, 1997) – PHASE_8E.MDL
C-9	1998	Cortland, Day 9 – 2 nd Resource allocation conference (September 28, 1998) – JND_6N.MDL

CONFERENCE PAPERS

Cortland County:

A-1	Rogers <i>et al.</i> 1997	Rogers J, R Johnson, AA Zagonel, J Rohrbaugh, DF Andersen, GP Richardson and TP Lee. Group model building to support welfare reform in Cortland County. <i>Proceedings of the 1997 International Conference of the System Dynamics Society</i> . Istanbul, Turkey (August 19-22, 1997). – PHASE_6E.MDL
-----	---------------------------	---

Dutchess County:

A-4	Allers <i>et al.</i> 1998	Allers R, R Johnson, DF Andersen, TP Lee, GP Richardson, J Rohrbaugh and AA Zagonel. Group model building to support welfare reform, part II: Dutchess county. <i>Proceedings of the 1998 International Conference of the System Dynamics Society</i> . Québec, Canada (July 20-23, 1998). – JND_6.MDL
A-5	Lee <i>et al.</i> 1998c	Lee TP, AA Zagonel, DF Andersen, J Rohrbaugh, GP Richardson. A judgment approach to estimating parameter in group model building: A case study of social welfare reform at Dutchess County. <i>Proceedings of the 1998 International Conference of the System Dynamics Society</i> . Québec, Canada (July 20-23, 1998). – SN_2.MDL

Other:

A-7	Lee 2001	Lee TP. The dynamics of New York State social welfare reform finance at the county level: A feedback view of system behavior. <i>Proceedings of the 2001 International Conference of the System Dynamics Society</i> . Atlanta, Georgia (July 23-27, 2001)
-----	----------	--

OTHER CONFERENCE PRESENTATIONS

P-1	July 1997	Rohrbaugh J, DF Andersen, GP Richardson, R Johnson, AA Zagonel and TP Lee. Systems modeling for welfare reform. <i>NYPWA Conference</i> . Saratoga, New York (July 15, 1997). – PHASE 8D.MDL
P-2	September 1997	Andersen DF, D Childs, J Rogers, R Allers and AA Zagonel. Systems modeling for welfare reform: The steps and purposes of building and using management flight simulators. <i>Management and Leadership for the 21st Century: Training Forum I</i> . White Eagle, New York (September 22-25, 1997). – PHASE 8E.MDL
P-3	April 1998	Rohrbaugh J, R Johnson, GP Richardson and AA Zagonel. Welfare reform collaboration in New York State: Decision support tools for welfare reform. <i>LINKS 1998, 6th Annual Conference</i> . Boston, Massachusetts (April 15-18, 1998). – JND 4.MDL
P-5	July 2000	Andersen DF, GP Richardson, J Rohrbaugh, AA Zagonel and TP Lee. The impact of US welfare reform on states and counties: Group facilitated system dynamics modeling, policy analysis and implementation. <i>GDN 2000, INFORMS Section on Group Decision and Negotiation</i> . University of Strathclyde. Glasgow, Scotland (July 3-7, 2000). – JND 6N.MDL
P-6	July 2002	Richardson GP, DF Andersen and YJ Wu. Misattribution in welfare dynamics: The puzzling dynamics of recidivism. <i>Proceedings of the 2002 International Conference of the System Dynamics Society</i> . Palermo, Italy (July 28–August 1, 2002).
P-7	November 2002	Andersen DF and GP Richardson. Addressing “what if” questions about welfare reform. <i>APPAM Fall Research Conference, Asking “What If”- Assessing the Public Policy and Management Implications of Social Science Research</i> . Dallas, Texas (November 7-9, 2002). – JND 6N.MDL

PUBLISHED ARTICLES AND BOOK CHAPTER

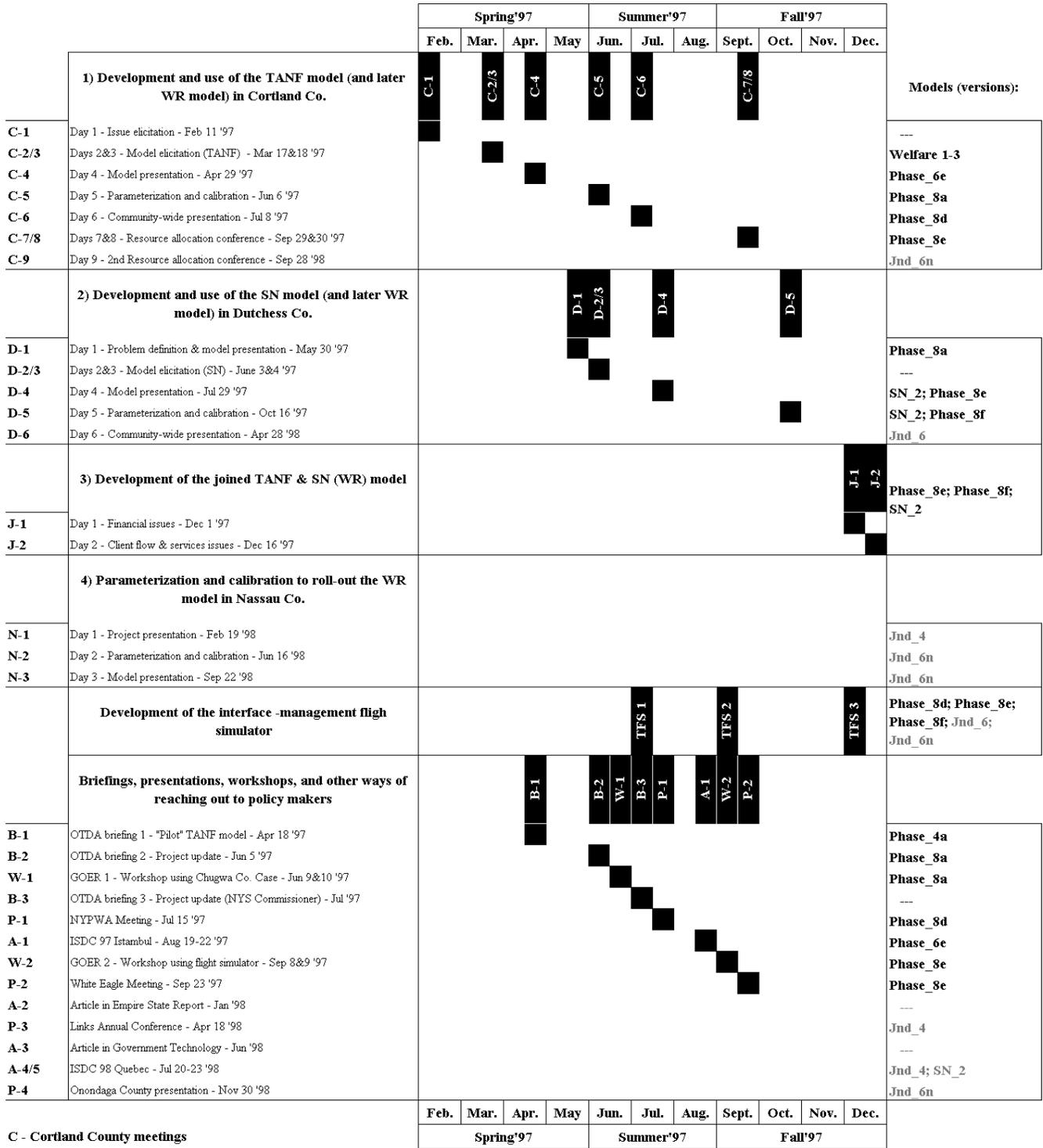
A-2	ESR 1998	Understanding Welfare. <i>Empire State Report</i> (January 1998): 17.
A-3	Rohrbaugh & Johnson 1998	Rohrbaugh J and R Johnson. Welfare reform flies in New York. <i>Government Technology</i> (June 1998): 58.
A-6	Rohrbaugh 2000	Rohrbaugh J. The use of system dynamics in decision conferencing. In GD Garson (ed.). <i>Handbook of Public Information Systems</i> . New York: Marcel Dekker: 521-533.

DISSERTATION WORK

T-1	Lee 1999	Lee TP. The dynamics of New York State welfare finance at county level: A feedback view of system behavior. Doctoral dissertation. Nelson A. Rockefeller College of Public Affairs and Policy. University at Albany, State University of New York.
T-2	Zagonel 2002b	Zagonel AA. Reflecting on group model building used to support welfare reform in New York State. Dissertation proposal. Nelson A. Rockefeller College of Public Affairs and Policy. University at Albany, State University of New York.

Figure 2 (detailed). Schedule and streams of activities in the Welfare Reform project

- 1997 -

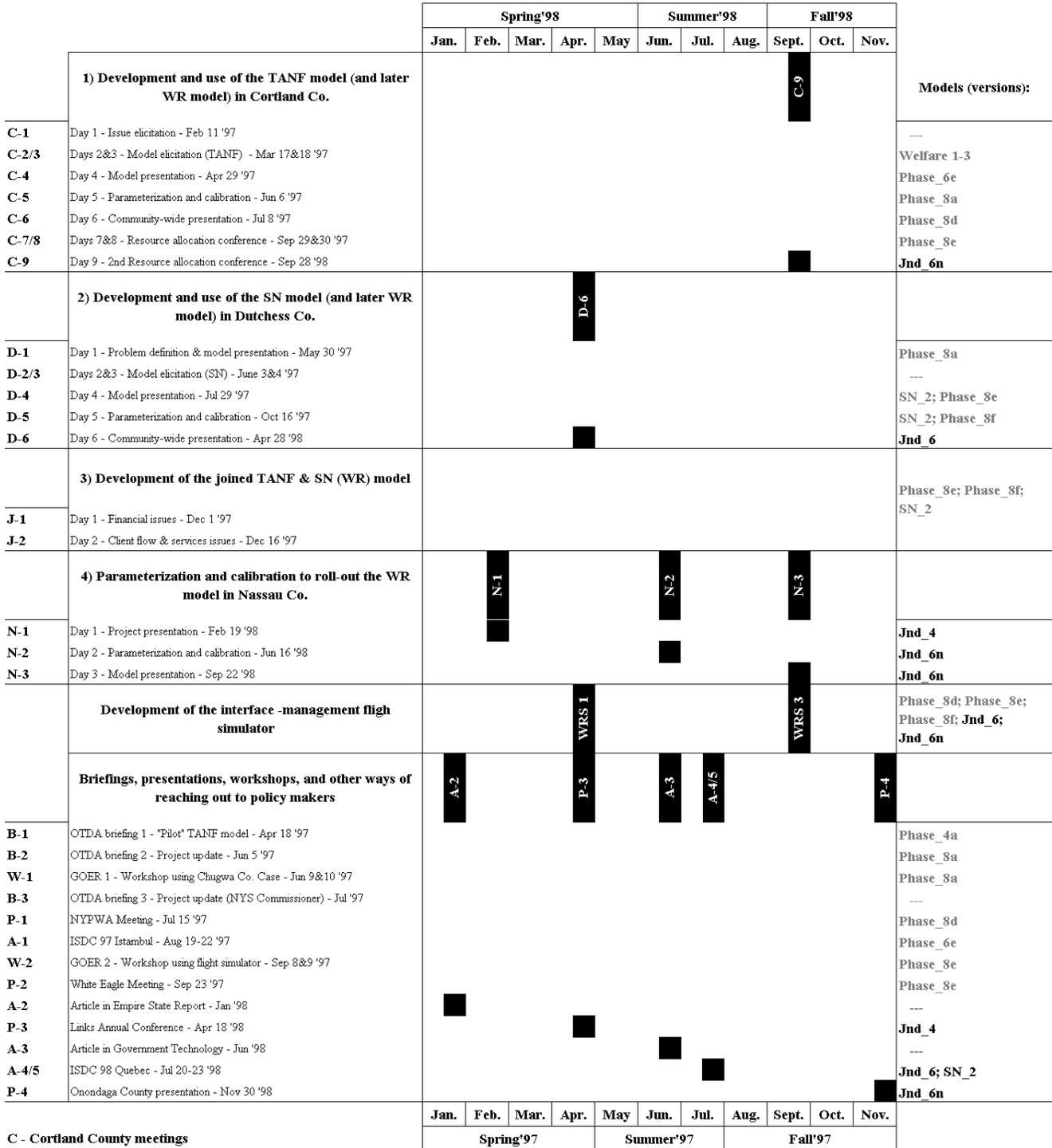


- C - Cortland County meetings
- D - Dutchess County meetings
- J - State meetings to join TANF and SN models
- N - Nassau County meetings
- B - Briefings to OTDA
- W - Workshops
- P - Presentations (other than Cortland, Dutchess, Nassau and OTDA)
- A - Articles and papers

Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sept.	Oct.	Nov.	Dec.
Spring'97				Summer'97			Fall'97			

Figure 2 (detailed). Schedule and streams of activities in the Welfare Reform project

- 1998 -



C - Cortland County meetings

D - Dutchess County meetings

J - State meetings to join TANF and SN models

N - Nassau County meetings

B - Briefings to OTDA

W - Workshops

P - Presentations (other than Cortland, Dutchess, Nassau and OTDA)

A - Articles and papers

Models (versions):

Welfare 1-3
Phase_6e
Phase_8a
Phase_8d
Phase_8e

Phase_8a

SN_2; Phase_8e
SN_2; Phase_8f
Jnd_6

Phase_8e; Phase_8f;
SN_2

Jnd_4
Jnd_6n
Jnd_6n

Phase_8d; Phase_8e;
Phase_8f; Jnd_6;
Jnd_6n

Phase_4a
Phase_8a
Phase_8a

Phase_8d

Phase_6e
Phase_8e
Phase_8e

Jnd_4

Jnd_6; SN_2
Jnd_6n

**Appendix 3. Resource definitions (clusters, services and spending items)
Welfare reform model (Jnd_6n)**

TANF sector:

	Community resources ⁴⁵	DSS budgeted resources ⁴⁶
Prevention	<i>Teen pregnancy reduction and out of wedlock programs</i>	DSS contributes little to community wide Prevention
TANF assessment and monitoring	Portions of Case management resources Assessment resources Examiner resources	Portions of administrative cost of TANF & Diversion Services
Diversion payments and services	Emergency services for job continuation and job getting Treatment services at assessment stage Diversion case management	At Risk Day Care Emergency Assistance to Families (EAF) Partial share of Non-Public Assistance (NPA) Food Stamps
TANF payments	TANF cash payment	Aid to Dependent Child (ADC), ADCU, PG-ADC Partial share of Food Stamps Public Assistance (PA) Total Total Food Stamps MIX
TANF employment services to: ... High-need families ... Low-need families	Education, training slots and referrals for jobs Substance abuse and mental health treatment VESID & JM Murray Center for Disabled Workfare and emergency services Job readiness programs DOL + JTPA + private Transportation Federal dollars for training (JTPA) Moneys for grant diversion Transitional Medicaid Registered & Licensed day-care centers and other types of child-care Supports for children, e.g. transportation	Employment services Day Care IV-A DC Partial share of Day Care Block Grant Partial share of Day Care training for IV-A, ADC/Emergency Assistance to Families (EAF)
Self-sufficiency promotion	Child-care Counseling and case management services Services that help maintain employment Grant diversion to support job creation Retaining (while on the job) Medicaid & Food-Stamps Unemployment benefits	Partial share NPA Food Stamps Partial share of Day Care Block Grant Low Income Day Care Day Care training for Low Income DC Day Care IV-A Transitional

⁴⁵ Community resources as defined in group model building sessions.

⁴⁶ Local DSS budgeted resources as defined and reported on claim forms.

Appendix 3. Resource definitions (continued)

Safety-net sector:

	Community resources	DSS budgeted resources
Child support enforcement	Child support workers Establish paternity support	Child support
Safety-net monitoring	Family counseling Adequate Staffing "case management" Social Services	Portions of administrative costs of other services for Safety Net
Basic services	Food Housing Basic medical services through Medicaid Shelter All basic support for everyone	Home Relief/Safety Net Partial share of Food Stamps Total PA
Employment services on safety-net	Employment services Training slots "JTPA and DOL and DSS" Adequate job training opportunities DOL and DSS employment workers Referral for jobs Transitional services Jobs Day care for T ANF ineligible	Partial share of Day Care Block Grant Partial share of Day Care training for IV-A, ADC/EAF
Job maintenance services	Portions of Case Management resources Transportation Transitional Services Day Care	Small DSS contribution. Most services are provided by DOL
Emergency services	Emergency Assistance to Adults (EAA)	EAA