The Power of the Stock: Accumulations in the Colombian Accusatory System Reform[†]

Camilo Olaya [*, 1] **, Gloria Díaz** [1] **, Ana María Ramos** [2]

[*]Corresponding author. *E-mail address*: colaya@uniandes.edu.co

[1] *Universidad de los Andes, Departamento de Ingeniería Industrial*
Calle 19A No. 1 – 37 Este, Bogotá, Colombia. Tel. +57 1 339 49 49

[2] *Corporación Excelencia en la Justicia*
Calle 94A No. 13 – 59, Bogotá, Colombia. Tel. +57 1 623 7365

Abstract

System dynamics models help decision makers to organize knowledge through assisting system design; the goal is to cope with complex systems. The criminal justice apparatus is an example of a complex system; it can be conceived as a purposeful arrangement of heterogeneous and loosely interrelated actors whose missions are to enforce the law and to prosecute and rehabilitate offenders. This article explores the problem of the growing congestion of criminal cases in a recently implemented reform introduced by the Colombian government; in particular the focus is the criminal process as stipulated by the new accusatory system formally implemented in 2005. A simulation model of this new criminal process system was developed for the Corporation for Excellence in Justice (*Corporación Excelencia en la Justicia*). The paper depicts the main aspects of the model. It also offers selected analyses focused on the way that criminal cases are accumulated and evacuated through the main stages of the new accusatory system. The model shows that this system is largely driven by accumulations which provide a pervasive inertia. But what is more important is that such distinctive characteristic seems to be unnoticed for decision making processes. The paper underlines the importance of understanding the significant dynamics associated with accumulations and how this learning can be promoted through simulation.

Key words: criminal justice, Colombia, case study, simulation, stocks, flows, inertia.

1. Introduction

A central claim of system dynamics (abbreviated, SD) is that it helps to cope with complex systems by considering the interrelations of flows of information, people, material, labor and money, placing a special emphasis on accumulations, non-linear relationships, feedback structures and time delays. The understanding of stock-and-flow dynamics has been recognized as a fundamental – though scarce - systems thinking skill in order to cope effectively with complex systems (Richmond, 1994; Sterman, 2002). Here "complex systems" refers to sets of

[†] This article compiles the main results of a project that was sponsored by the Corporation for Excellence in Justice.

interconnected parts whose relationships can lead to counterintuitive behavior (Sterman, 2000). Criminal justice systems are good examples of such systems. In particular, the way criminal processes are handled can be represented with a very basic stock-and-flow model: cases enter the system, they are accumulated through various stages and they are evacuated through diverse judicial mechanisms according to resources and productivity rates. However, regardless of such apparently simple depiction, the dynamics associated with these structures seem to elude policy makers. The task become harder if we bear in mind the intricacy of the apparatus of justice given the participation of different actors across different levels and with different interests. Nevertheless, the whole system pursues the demanding goal of administering justice. This challenge is usually placed in the middle of heated political debates around subjects like judicial congestion, impunity and corruption, among other matters. Colombia is no exception. Considering this framework then the task of designing policies for such complex systems turns out to be a significant enterprise; at this point, then, the mentioned claim of system dynamics becomes crucial; in this article, a SD model will be conceived as an aid to organize knowledge so as to enhance learning and systems design processes.

This paper presents the first results of a SD model developed for assessing the recent reform on the regulations of the criminal process made by the Colombian government. The article is organized as follows. The next section presents the Citizen Observatory for Justice a program developed by the Corporation for Excellence in Justice (*Corporación Excelencia en la Justicia*)[1] which assesses the results of the mentioned reform; in particular, the project presented here constitutes the third phase of the implementation of this Observatory. The third section presents a synopsis of the model building process which led to the articulation of a specific problem: the growing congestion levels in the criminal process system in spite of the new reform. The next section describes the main characteristics of the model, sectors, policy points and tests. The fifth section analyzes different policies and compiles the main results which stress the pervasive and enormous inertia of the system; it also shows how simulation promotes learning about such dynamics. The sixth part discusses the main lessons of these findings and it connects them with the significance of understanding stock-and-flow structures; this is shown through the examination of various indicators designed to observe and evaluate the criminal process system in Colombia. The last section outlines limitations and future steps to take.

2. A Citizen Observatory for the Criminal Process Reform

This short section describes the origin of this project which belongs to a larger effort for monitoring the effectiveness of the new accusatory system in Colombia.

The Colombian government introduced the new accusatory system through the Law 906 in 2005. The establishment of this regulation represents an important novelty in Colombia. After decades of having an inquisitive system the country is making its first steps into the philosophy of the new system. The main change has to do with the criminal process, that is, the formal manner in which criminal proceedings are conducted. The reform established the new criminal process through three main stages: inquiry, accusation, and trial. During the first stage — inquiry — the Judicial Police (JP, in Spanish *Policía Judicial*) collects evidence with the aim of finding if there has been an offence which should be prosecuted. The JP is informed through informations or because of police duty and the work of security corps in general. This first phase is driven by prosecutors helped by investigators of the JP. Once there is sufficient evidence then the leading prosecutor makes a formal accusation starting then the second stage in which is defined the penal

[1] This corporation is a non-governmental organization that studies the justice sector in Colombia on behalf of the Civil Society.

responsibility of the accused persons. With these elements a judge approves the case and establishes a date for trial. This last phase finishes with a sentence — either absolutory or condemnatory.

It must be mentioned that one of the main goals of this reform was to introduce celerity and efficiency to the criminal process. Several judicial instruments were introduced in order to make the process faster than the old one. For instance, a central improvement is that the new system promotes to end cases before trial; novel mechanisms were introduced for accomplishing this objective, e.g. the accused can agree to plea guilty and he can also bargain in order to obtain a lesser punishment. These types of innovations represent a challenge for the institutions of the judicial branch and this is the reason why various organizations are concerned with the observation and the evaluation of the reform.

The Corporation for Excellence in Justice (CEJ) has been one of the main promoters of the new accusatory system. The Corporation devoted several projects and conferences to analyze advantages and limitations that this reform could bring to Colombian justice. Once the constitutional reform was approved by the Congress, CEJ became the technical body for the commission that designed the bill that became the Law 906 which established the new accusatory system. Currently, CEJ plays the same role in the constitutional commission that monitors the implementation of the reform.

Consequently, CEJ has as a permanent line of work the promotion and the assessment of the new accusatory system. This task has been developed with the creation of an inter-institutional group, financially supported by USAID, which had to design a mechanism for collecting and analyzing information about the performance and the effectiveness of the new system. Within these lines, CEJ created the "CEJOSPA", the Citizen Observatory for the Accusatory System (in Spanish, *Observatorio Ciudadano del Sistema Penal Acusatorio*). This is an institutional space that acts on behalf of the Civil Society so as to create spaces for discussion and interaction among institutions, universities and citizens. Its main goal is to monitor the effectiveness and the impact of the reform. The observatory was designed for focusing on four specific domains: institutional expectations, citizens' expectations, behavior of the system, and identification of best practices. These domains have been addressed through four corresponding implementation phases. The observatory is supported by an expert committee that advises the technical team. Currently a full set of indicators has been developed for the first two phases.

The case presented here corresponds to the third phase of the observatory. The goal was established in the following way: "to develop a model that explains the dynamics of the accusatory system, that is, it should allow to comprehend the way in which relevant variables affect each other. The construction of this model should be based on the assumptions that represent the point of reference from which the effectiveness of strategies and policies will be analyzed". In particular, this phase should deliver a tool for helping to understand the justice apparatus as a *system*; moreover, the use of simulation was stressed in order to show how the modification of any variable affects the whole system. Specifically, this project was defined with the next goals in mind:

- To represent and to make explicit a set of assumptions about the reform.
- To examine the consequences of interactions of relevant variables.
- To estimate behavior patterns of the system in the long run.
- To promote fine-tuning intuition and learning processes about the system**.**

The rest of the paper depicts the first results of this project. The next section presents the manner in which the process was designed for fulfilling the goals above.

3. Problem Articulation : Congestion

The first step was a literature review; this section summarizes pertinent points for this project from previous SD models of criminal justice systems. It also depicts the model building process in order to present the way the problem was articulated, explaining in particular both the context and the focus on the criminal process system and its growing accumulation of cases.

3.1. Antecedents on Simulation Models of Criminal Justice Systems

It seems natural to recognize the necessity of a *systems* approach for supporting decision making and policy design for justice administration given its multilevel character, i.e. various heterogeneous actors, agencies and institutions — such as police, security corps, judicial agents, courts, penitentiary and correctional institutions, citizens — form together a complex and usually loosely connected apparatus that as a whole is expected to protect the community, combat crime, enforce the law and rehabilitate criminals.

Perhaps the first application of *systems* engineering to criminal justice was developed in 1965 by the Space-General Corporation (1965) which developed a cost-effectiveness study of the California criminal justice system[2]. Likewise, initial SD models applied to justice systems were developed showing the far-reaching scope of this approach for evaluating and formulating coherent and integrated policies, e.g. (Fey, Wadsworth, & Young, 1974; Riccio, 1971). Bard (1977) built a SD model which is particularly relevant for the case we present here; his work underlined the necessity of overcoming what he called a lack of perspective that "overlooks the effects that specifically targeted programs have on the other sectors of the system" (p.259); this model addressed three main sectors — police, court, corrections — and it analyzes their interactions through information-feedback loops and the role of nonlinear relationships and delays; the main point of this work is appropriate for the case we are presenting here: the resistance to change provided by multi-level structures of accumulations which were recognized by Bard as central attributes that should be taken into account in order to develop coherent policies for controlling crime.

A special interest for using SD techniques to address criminal justice seems to be recently upsurging. Sterman (2000) shows the role of feedback analysis for generating hypotheses in order to explain the growth of prison population in the US. MacDonald and Mojtahedzadeh (2007) introduce a simulation model of the New York criminal justice system tracking criminals flowing through the system; in particular they study the scenario of doubling the number of new criminals entering the system and doubling the productivity of police officers so as to examine the impact of technological improvement on the performance of the system; their analysis underlines the effect of negative feedback loops for compensating growth processes. Rouwette et al. (2007) present a group model building project aimed at examining the effects of an increment in case load and investments in different phases of criminal justice administration; the authors discuss in detail the process of model construction. Newsome (2008) focuses on how SD helped West Yorkshire Police (UK) to gain a better understanding of the high-level interrelationships between policing activity and performance outcomes. Hybrid models of criminal justice systems

[2] Bohigian (1977) presents an inventory of the first simulation models applied to criminal justice systems. Blumstein (1967) develops a first link between systems science and criminal justice; an updated overview of simulation models can be found in Blumstein (2002).

that combine SD with other simulation techniques have been also developed, e.g. an application with agent-based modeling is introduced by Boyle et al.(2003).

Regarding SD applications in Colombia, the direct antecedents are, on the one hand the work of Ariza and Sotaquira (2003) which presents various small models for studying different aspects of the 1991 reform, and on the other hand two works of Dyner based on economic theory of crime — as conceived by Becker (1968) — focused on the connections between social capital and criminality (Jaén & Dyner, 2005) and on prison overcrowding as explained by short-sighted policies that exclude long-run perspectives (Hernández & Dyner, 2001)

As a summary it can be stated that most of previous SD work on criminal justice share at least three important characteristics: (i) the presence of important feedback loops that explain both typical reinforcing processes, e.g. justice systems that reinforce crime, and typical balancing processes, e.g. crime control via rehabilitation, (ii) the recognition of heterogeneity, i.e. multiple actors and agencies interact in one single apparatus, and (iii) the necessity of comprehending the significant roles and distinctive characters of accumulations (prisoners, cases, criminals, police force, citizens, etc.) and flows (crime rates, arrest rates, annual sentences, etc.) for understanding the dynamics of these systems. The project that we present here builds on these latter two points and specifically it focuses on the dynamics generated by stock-and-flows structures.

3.2. Model Building Process

The preceding short review shows the potential of using SD as a way to enhance policy design processes for acting in complex and heterogeneous domains such as criminal justice systems. This project is no exception. Given the problems related to judicial cases congestion (see below), and the necessity of having a more efficient criminal justice system, the task of the phase 3 of the citizen observatory was conceived as the development of a model that should help to explain the dynamics of the new Colombian accusatory system; the original description of the goals of this phase underlines that this model should support the understanding of different relevant variables so as to design macro- policies able to match the complexity faced by the reform. Furthermore, this model should help to make explicit basic assumptions that should represent a point of reference in order to evaluate further strategies and policies.

In order to accomplish these goals a special team was formed for constructing the model. This team was supervised by the general manager of the Corporation for Excellence in Justice and it was audited by one of the authors of this paper. In addition, this team had both a supporting group of experts and a consulting committee formed by specialists in specific areas related to the Colombian criminal justice system, e.g. criminal process, human rights and defense attorneys, crime theory, constitutional law, prosecution, etc. The general methodology for building the model was proposed, discussed and agreed among the team members; it was conceived as an iterative process around three main activities: conceptualization, model formulation, and simulator development, fairly following the suggestions of Sterman (2000) — see Fig. 1. Periodic working sessions were scheduled in which discussions were held so as to build various versions of a first qualitative model; various meetings were exclusively devoted to specific topics related to the accusatory system while others were focused on addressing the “big picture”. These initial stages were the base for producing a quantitative model that was further discussed and updated; these subsequent versions of the model helped to revise purposes and expectations in subsequent meetings which in turn affected the definition of the final version.

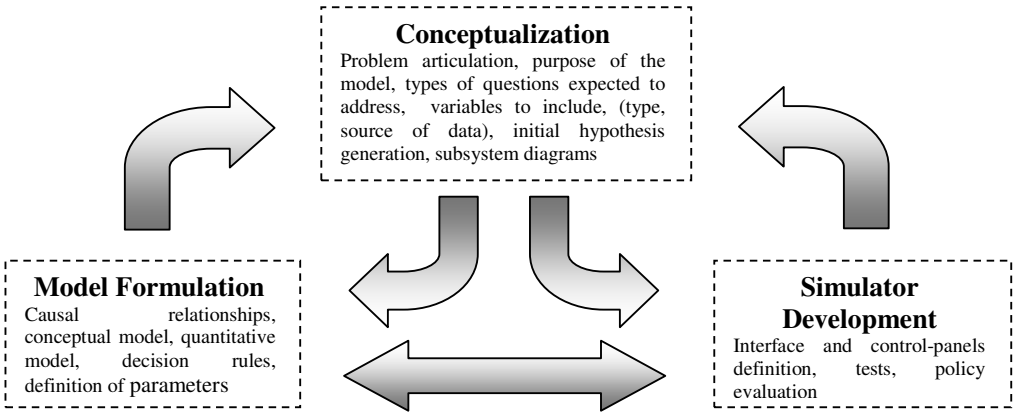


Figure 1
Modeling Process

A detailed description of the model building process is out of the scope of this paper. However, one particular aspect is worth mentioning to highlight the value of the modeling process itself and also to portray and justify the final boundary of the model. This issue is explained next.

3.3. Iterating: Definition of Purpose, Boundaries and Theme Selection

3.3.1. The Big Picture vs. The Criminal Process

A first general model was developed after various sessions (Fig. 2). This first version shared with previous works — e.g. the model of Bard (1977) mentioned above — the recognition of three main sectors: (i) community (citizens and criminality), (ii) the criminal process, i.e. resolution of cases through investigation and judicial procedures following the new accusatory system, and (iii) prisons. The initial concern with this initial macro-view pointed at the study of the effectiveness of the whole criminal justice system, e.g. to examine the impact of criminal rehabilitation on decreasing crime rates which in turn feed-back the system — this general balancing feedback loop explains one of the central missions regarding a justice system. In this first stage the economic theory of Becker (1968) was the cornerstone for explaining criminality. Another aspect identified as relevant was the role of the perception of citizens regarding the effectiveness of the system and how this perception feeds back its performance.

Yet, several questions arose in the project team regarding the scope of the reform and the central aspect that should guide the development of the model; in short: should the project focus on the efficient and correct resolution of cases through the new accusatory system or should it take the bigger look to the whole justice system? This was a highly debated issue because of the risk of misplacing or ignoring the overall goals of the justice system which were depicted in the first qualitative model, e.g. consider in Fig. 2 the loops “justice and rehabilitation” and “trust power” which face threats as the “school of crime” and “light punishment - crime incentive” loops. The final decision was to concentrate on the formal definition of the new *criminal process* established by the reform, i.e. the accusatory system. In other words, the team would initially concentrate on looking the reform as a plain “production machine” or “sentence factory” — indeed these expressions were coined by some members of the team suggesting a potential short-sighted model for the system as if it were a simple manufacturing line. As a result it was agreed that on subsequent phases this project will be expanded to the other two sectors, i.e. community and prisons. Fig. 2 highlights the accusatory system as the main sector to be studied.

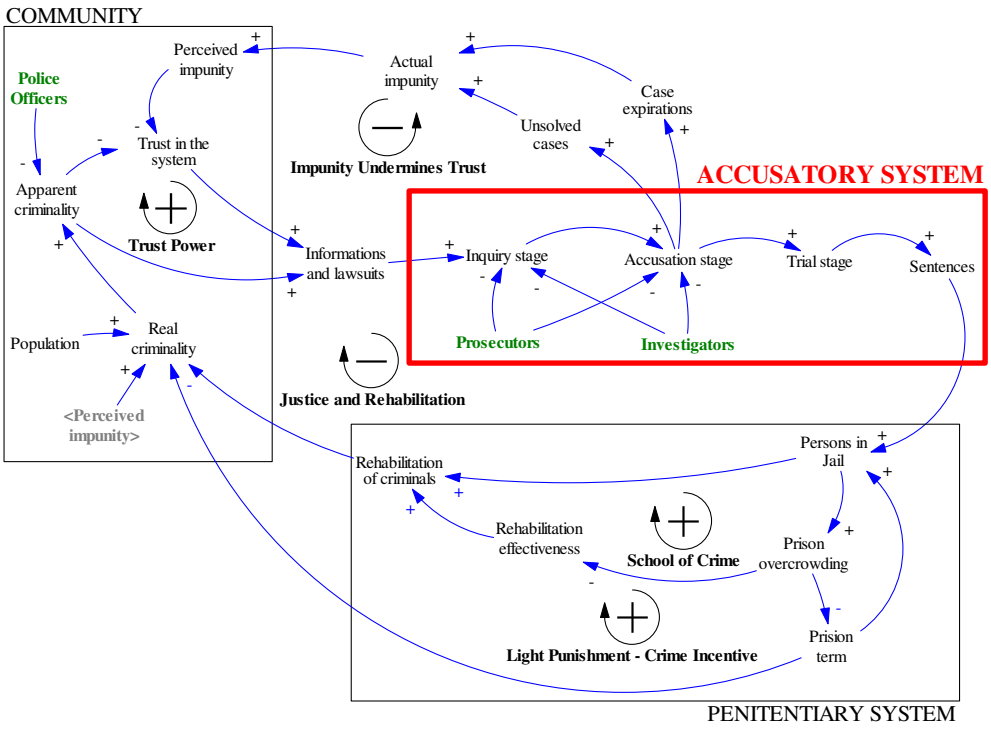


Figure 2 - First Macro-Model: the Context and Focus on the Accusatory System

3.3.2. The central Problem: Growing Accumulations

Consistent with the emphasis on the criminal process then the main problem to be addressed was identified as the growing accumulation of judicial cases that have to be evacuated and resolved by prosecutors and judges. Indeed, although the reform is only few years old, the stock of cases are already piling up across the three main stages of the accusatory system since its implementation in 2005 (Fig. 3).

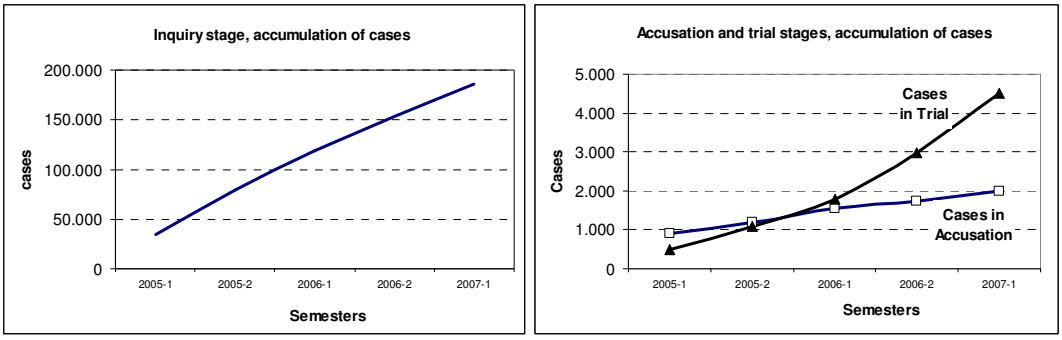


Figure 3
Accumulation of cases through the three stages of the accusatory system
Source: *Corporación Excelencia en la Justicia* and own calculations

In particular, 95% of the total accumulation corresponds to cases in the inquiry stage (approx. 189.000 in 2007) which is easy to understand if we consider that this is the only stage in which there is no limited amount of time for prosecutors for resolving cases (unless the case expires according to the particular characteristic of the offence).

A further aspect that was defined as relevant for addressing the problem concerns the pattern of the inputs to the system. These accumulations are fueled by the entry of cases ("lawsuits", in Spanish *noticias criminales*); given the initial focus on the criminal process as such, i.e. "manufacturing line approach", then these inputs were taken as exogenous variables for the project; in other words, to explain crime was not going to be considered as part of this first model. Therefore the inputs were projected according to historical data. Fig. 4 shows the aggregate inputs and a trend line.

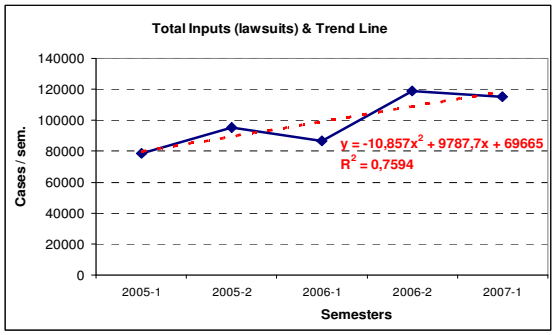


Figure 4
Total Inputs to the Accusatory System (cases per semester)
Source: *Corporación Excelencia en la Justicia* and own calculations

The stocks of cases in turn decrease with the volume and the productivity of resources associated to the resolution of cases; these resources can be defined specifically in terms of the volume of investigation teams of the General Attorney Office (investigators and prosecutors) and the number of judges. In fact, policies designed to reduce congestion usually target the modification of outflows, e.g. creating new outflows or increasing the rate of the existing ones. The team emphasized the importance of analyzing customary policies that seek to clear accumulations by increasing resources and productivity; it became clear that the model should be able to handle policies conceived to increment outflow rates, e.g. to augment productivity (or volume, or both). Time horizon was defined as no less than 10 years in the future.

4. Formulation of the Simulation Model

This section depicts the final simulation model, the types of policies and decisions that can be examined, tests, and a brief comment on the flight simulator that was developed.

The reform has been implemented through four phases; each phase corresponds to a set of judicial districts or regions. This study analyzes the initial phase which started the first semester of 2005 in four Colombian cities: Bogotá, Manizales, Armenia and Pereira. Currently, the Colombian authorities have implemented all of the four phases which cover all Colombian territory. The structure of the model maps the way the new criminal process system was defined by Law 906 (the reform); it should be noticed that the formal definition of the reform is identical for all of the phases; the only differences across them are parameters and initial conditions. The model was developed with the software *iThink*-v9.0.2. It is attached as supporting material to this paper. Data and parameters were estimated based on information and reports from the Corporation for Excellence in Justice and from statistics provided by governmental institutions of the judicial branch. In addition, the working sessions with experts supplied very important outlines for defining the logic of the model and its equations.

4.1. Main sectors

The model has four main sectors: settlement, criminal process, human resources and indicators. Figure 5 depicts this macro-structure. The settlement sector feeds the criminal process sector with cases that are not settled; the criminal process sector consists of three stages: inquiry, accusation and trial; through these stages there are basically two types of outputs: unresolved cases and sentences. The human resources sector characterizes the volume and the productivity of the workforce in order to model its impact on evacuation processes; this sector includes also the effect of learning processes which are boosted by experience. The indicators sector define relevant measures to evaluate system performance. These sectors are explained next.

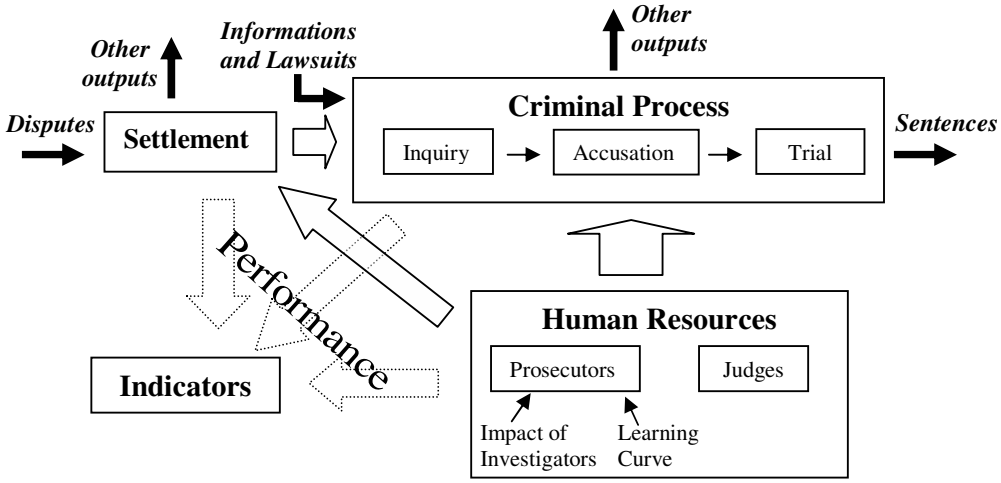


Figure 5 - Macro model. Main sectors

4.1.1. The settlement sector

This sector models those cases that can be settled according to law. In general, these cases consist of disputes which the system intends to resolve through conciliation and settlement processes provided by SAUs[3]. Whenever these disputes are not settled they are sent to the inquiry stage, whereby they become judicial cases. In the settlement sector there are different types of outcomes that represent different mechanisms or conditions that characterize the decision of sending out a case from this stage; these outflows are: cases that are extinguished, cases that cannot be attended by SAUs, cases that are temporarily filed, cases that are settled, and those other cases that are sent to the inquiry stage. Figure 6 shows the main flows associated with the stock of disputes cases.

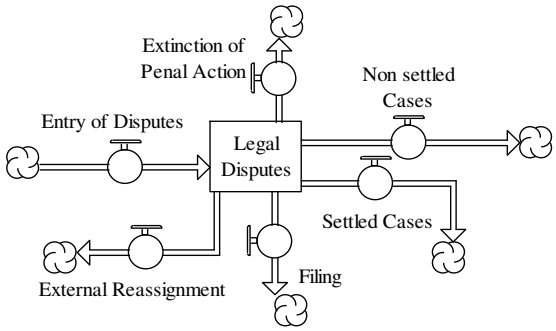


Figure 6
Settlement sector

[3] SAU is the Spanish abbreviation for *Salas de Atención al Usuario*. These are special attention units that receive disputes which are sent to settlement processes.

The number of cases that leave this sector depends on the volume and productivity of prosecutors who are dedicated to resolve them. The distribution of cases through the different outcomes depends on the probability associated with each one; these probabilities indicate the proportion of cases which leave the settlement sector through a particular outcome; or in other words, the total outflow of this sector is distributed across the different outflows according to the probabilities associated to each one; these probabilities were calculated with historical data.

4.1.2. The criminal process sector

This sector constitutes the backbone of the system. The model captures the three stages of the new accusatory system: inquiry, accusation and trial. These stages were modeled as consecutive accumulations with inflows and outflows associated with each stage. Appendix 1 shows a simplified version of this sector.

The first accumulation represents the inquiry stage which is fed by: lawsuits from URIs[4], any other lawsuits and informations that arrive directly to the inquiry stage, and cases that are not settled from SAUs. The main outflows represent the cases that are extinguished, cases that cannot be processed under Law 906, cases that are temporarily filed, other cases to which prosecutors apply the principle of opportunity — these cases go to suspension (maximum 3 years), and finally those other cases that are sent to the accusation stage. In addition, it is important to highlight that some cases can return to the inquiry stage from both the filed cases stock and the suspended cases stock. Similar to the settlement sector, the total outflow depends on the number of prosecutors dedicated to the inquiry stage and the productivity of these prosecutors. However, the prosecutors divide up their time between inquiry, accusation and trial activities; therefore the number of full-time prosecutors in the inquiry stage is equal to the number of prosecutors multiplied by the full-time equivalent of this stage. The distribution of the total outflow among the different outcomes of the inquiry stage is calculated according to the probabilities which in turn were estimated with historical data.

The second important accumulation in this sector represents the number of cases in the accusation stage. Inputs come from formal accusations which can be made either in the inquiry stage or directly from URIs. On the other hand, the outflows represent the cases that are extinguished, other cases to which prosecutors apply the principle of opportunity — these cases also go to suspension, cases in which plea of guilty or bargaining are reached, and finally, the cases that are sent to the trial stage. Furthermore, some suspended cases can return to the accusation stage if the prosecutor considers that the conditions imposed upon the accused were not fulfilled. Similar to previous stages, the total outflow depends on the number of full-time prosecutors in the accusation stage and the productivity of these prosecutors. The distribution of the total outflow among the different outcomes of the inquiry stage is again in function of probabilities calculated with historical data.

The third important accumulation represents the trial stage in which the judge pronounce sentence — either absolutory of condemnatory sentence. Some of the cases awaiting sentence are those involving a guilty plea, a bargaining plea, or the ones that have already been found guilty on trial. The trial stage is fed only by cases from the accusation stage. In contrast, the outflows of the trial stage represent one of two situations: the judge can either find the accused not-guilty or guilty; if the accused is found guilty the case goes forward to await sentence. And

[4] URI is the Spanish abbreviation for *Unidad de Reacción Inmediata*. These units handle any lawsuit that needs urgent action such as the collecting of evidence.

finally, the number of audiences generated by these dynamics are calculated — for instance to examine the number of required rooms in function of average duration per audience.

4.1.3. The human resources sector

This sector models the main resources associated with the criminal process. These resources are: prosecutors, investigators and judges.

The prosecutors divide up their activities through the three stages of the accusatory system. It was calculated that, according to the volume of cases, 94% of their time is related to the inquiry stage whereas only 2% of their time is involved in the accusation stage; the rest is needed for the trial stage. Additionally, productivity of prosecutors in the inquiry stage depends on two factors: the number of investigators per prosecutor and the average experience per prosecutor. One prosecutor increases his productivity in 0.6 cases per year with one additional investigator (Corporation for Excellence in Justice, 2007).

The productivity of the accusation stage also depends on the experience of each prosecutor. This experience can increase through the experience acquired by working and the new experience associated to the new prosecutors. The impact of the experience in the productivity has been modeled according to one traditional model of the learning curve theory which posits that productivity rises by a given percentage with each doubling of relevant experience:

$$\text{Productivity} = \left(\frac{\text{Average Experience}}{\text{Reference Experience}} \right)^{c} \qquad \text{[Eq.1]}$$

Reference Productivity is the productivity attained at the Reference Experience level. The exponent c determines the strength of the curve, see (Sterman, 2000; Zangwill & Kantor, 1998).

The model includes the increment of workforce since prosecutors dedicated to work in the previous inquisitive system (Law 600) progressively start to work in the Accusatory System. It is also possible to simulate hiring of new workforce; nevertheless in the base-case scenario this flow is zero since currently the General Attorney Office is not hiring new prosecutors. Finally, a first-order negative feedback structure models the dynamics of judges and their hiring process according to a goal and an adjustment time.

4.1.4. Indicators

This sector models measures of performance. The Citizen Observatory established specific indicators of effectiveness that are relevant for this case. These indicators, along with other ones defined by the project team, were integrated to the simulation model. Three types of indicators were modeled according to three criteria: accumulations, celerity and selectivity. The indicators of accumulations show the stage in which can be found the majority of cases, the celerity indicators give information about the speed of evacuation, and the selectivity indicators show the distribution of the total exit of cases among the different outcomes of the system, for instance the percentage of cases to which the principle of opportunity has been applied or the percentage of cases that have been filed.

Accumulations: as expected, the most important indicators of accumulation are the values of stocks, e.g. cases accumulated in each stage of the process. A further important indicator was defined as the ration of cases in each cases to the total of cases accumulated in the system:

$$Percentage\ of\ cases\ in\ Stage_{i,t} = \frac{Cases\ in\ Stage_{i,t}}{Total\ Cases_t}$$

$$i = inquiry,\ accusation,\ trial;\ t = time = 1,2,3... \qquad \text{[Eq. 2]}$$

Celerity: this indicator gives the ratio of cases still accumulated in the system to the total historical cases that have entered to the system, that is, it represents the congestion as a proportion of all cases that should have been evacuated at time *t*:

$$Congestion_t = \frac{Total\ Accumulated\ Cases\ in\ the\ System_t}{Total\ Accumulated\ Inputs_t} \qquad t = time = 1,2,3... \qquad \text{[Eq. 3]}$$

Selectivity: there are various indicators of this type and all of them follow the same logic: to measure the strength of the application of a particular instrument or decision as related to all possibilities of evacuations (outflows); for instance, the application of the principle of opportunity per semester is the ratio of the cases in which this mechanism was applied per semester to the total number of outflows in the same semester:

$$App.\ Principle\ of\ Opportunity_t = \frac{Total\ Outflows\ P.\ of\ Op._t}{Total\ Outflows\ per\ Sem._t} \quad t = time = 1,\ 2,\ 3... \qquad \text{[Eq. 4]}$$

4.2. Policy Points, Outflows and Resources

The previous section described the general sectors of the model. In general, the model focuses on the main accumulations and the evacuation of cases. These accumulations can only change through the different inflows and outflows whose changes represent decisions and policies.

For instance, the main inflows (disputes, lawsuits, informations) have been modeled as exogenous variables; several scenarios can be examined changing the patterns of these inflows which can represent also different policies, e.g. a law that re-define particular conducts as crimes which possibly affects the volume of inputs to the system; the flight-simulator that was developed with the model allows to explore these cases. Regarding decisions and other policies directly related to clearing congestion, the diverse outflows across the stages of the criminal process sector are related to different resources (prosecutors, investigators, judges) which can accelerate the evacuation process; changes in the volume or the productivity associated with them, e.g. hiring policies or training programs, can be explored. Additionally, it is possible to explore the intensification (or lessening as well) of specific judicial mechanisms, e.g. to increase case settling or to increase filing; these changes represent different policies which in practical form are materialized with such changes. Likewise it is possible to change the probability distribution of the outflows associated to one particular accumulation, e.g. representing new priorities for judicial evacuation mechanisms. Naturally, combinations of these policies are also possible, e.g. to give priority to specific instruments which in turn might imply changes in workforce productivities too. The fifth section shows examples of policies that can be explored.

4.3. Testing and Assessment of the Model

Various tests were done: dimensional consistency, behavior reproduction, boundary adequacy, structure assessment, and extreme conditions (Sterman, 2000).

Dimensional consistency and behavior reproduction analyses were the first tests that were done with reasonable results. Dimensional consistency was always verified as the model was built. The behavior reproduction was tested for four semesters according to available data since the implementation of the reform (2005). Appendix 2 shows graphs associated to this test.

Regarding boundary adequacy, the model concentrates on the way the criminal process system was formally defined in the reform, excluding other aspects of the larger justice system, e.g. community and prisons. This decision (sec. 3.3) produced important results since it directed the analysis to accumulations and the inertia of the system, permitting to find critical insights about these aspects (see below); as it was mentioned, this purpose was transformed through the different model building sessions with the group of experts until the current boundary of the model was agreed and established; nevertheless, further new phases of this project plan to expand these boundaries (see final section).

Two further boundary limitations should be noticed: the evacuation of cases in the trial stage does not depend on the number of audience rooms, assuming that audience rooms are an infinite resource; this assumption is significant because the purpose is to examine evacuation processes which could be affected in the case that this resource becomes scarce; but for now audience rooms are sufficient according to current practices. Similarly, the model does not take into account advocates availability which could affect various outflows since it is mandatory the presence of the advocate in many parts of the process; however, observed productivity of prosecutors and judges are result of the interaction of all the necessary resources and when the model considers these productivities, these resources which are not modeled are considered in an indirect way, through reference productivities. Moreover, these restrictions did not represent major obstacles; in the working sessions it was agreed that these elements could be excluded given that on a normal basis neither advocates nor audience rooms are currently slowing down the flow of processes as such; yet, it is important to highlight that this assumption is not necessarily correct in all scenarios, for example in extreme conditions. A final boundary limitation has to do with the workforce; for instance, if accumulated cases in the system increase dramatically, it should be possible that some resources such as prosecutors or judges should also increase as a reaction to this sudden increment; but these decisions are not considered in the model, the hiring of prosecutors and judges are considered as exogenous processes; neither investigators nor judges dynamics are affected by the number of cases in the system; this balancing loop would help to deepen the study of dynamics related to workforce policies designed to control accumulations; nevertheless the flight simulator allows to change hiring and outflows parameters for human resources. On the whole the boundary adequacy test suggests further steps to investigate provided that such restrictions in resources or that those decision processes are to be examined.

The level of aggregation and time units of the simulation were guided by the purpose of the model and by the specific questions which should be answered. However, the data to which this project had access was consolidated by semesters which could be troublesome; for example, it could be desirable to examine the dynamics of audiences as compared with capacity of audience rooms; the model estimates the number of audiences per semester which might be a too much

aggregated output; with more disaggregation it could be possible to examine particular types of audiences so as to analyze the utilization of related resources, e.g. rooms.

Additionally, extreme conditions tests were done. Basic tests were made through both direct inspection of equations and simulations; these examinations explored different variations in parameters (e.g. no inputs, no workforce) and diverse extreme scenarios showing realistic responses.

In summary, this model shows important robustness according to the purpose and the problem to be examined, i.e. the increment of accumulations since the implementation of the reform, which led to important insights associated mainly with stock-and-flows dynamics — see the fifth section. Nevertheless, the model also shows various limitations; the modelers are aware of these restrictions which turn out to be suggestions for further improvement.

4.4. Simulator

Since the main goal is learning through simulation then a flight-simulator was developed for decision makers, policy designers, and in general analysts interested in the criminal process system as it was defined by the new reform. Different policies can be tested with the simulator. It has various modules for simulating several policies; every module displays relevant variables and comparative graphs as well so as to analyze different scenarios or to develop sensitivity analyses.

The main panel displays the core variables of the model such as the number of cases accumulated across the stages of the criminal process i.e. inquiry, accusation, trial. For each stage there is a further module in which is possible to simulate different policies associated with the distribution of probabilities related to associated outflows. The indicators module displays the evolution of performance measures as estimated by the model. The human resources module permits to simulate different policies associated with workforce, e.g. hiring. The inputs module allows to conduct different experiments with disputes and lawsuits. There is a special module devoted to the principle of opportunity in which diverse variables associated with this mechanism can be changed, e.g. the percentage of suspended cases that return to inquiry and accusation stages. There is also a module for audiences, for instance in order to track requirements of rooms. Finally, there is a learning curve module in which is possible to explore diverse policies related to the strength of the curve, the average experience of new prosecutors, etc. Appendix 3 shows screenshots of the main panel and the workforce panel .

5. Policy Analysis and Learning

This section shows example of policies that led to simulations which produced valuable lessons regarding the inertia of the system. Moreover, these initial insights draw attention to a further discussion regarding the limitations and risks when the dynamics of stock-and-flows structures are not properly included in the design of policies aimed at clearing accumulations.

5.1. Examination of Policies

As it was mentioned, the policies to be examined seek to clear accumulations. One way of doing this is by increasing resources and productivity; another way is to stimulate particular outflow rates, for example by intensifying the application of specific judicial mechanisms that help to clear the stocks of cases. The next policies and experiments were studied with the simulator; this

overview will show the possibilities for enhancing the understanding of this system by using simulation.

5.1.1. Attacking Just One Policy Point: The Principle of Opportunity

A first natural approach has to do with decisions on outflows. It should be noticed that any of the outflow rates in the system is susceptible of being considered. As a prominent example we will briefly consider one of the possibilities: the principle of opportunity.

One of the key mechanisms of the Colombian reform is what is called the principle of opportunity — also called *prosecutorial discretion*. This instrument establishes discretion for prosecuting offences, i.e. to make a discretionary decision as to whether to prosecute or not, according to a broad criminal policy. The principle of opportunity stands in contrast to the principle of legality — or the principle of compulsory prosecution — typical in other systems in which the prosecution has no discretion: if the facts are sufficient to believe that a an offence has been committed, a prosecution must be instituted (Jehle, 2005; Ploom, 2000). The reform establishes that this mechanism can be applied discretionarily by prosecutors during both the inquiry stage and the accusation stage. Figure 7 shows the main inflows and outflows that affect the inquiry stage and the principle of opportunity as one of the mechanisms of the reform for decreasing congestion. An analogous outflow is present for the stage of accusation.

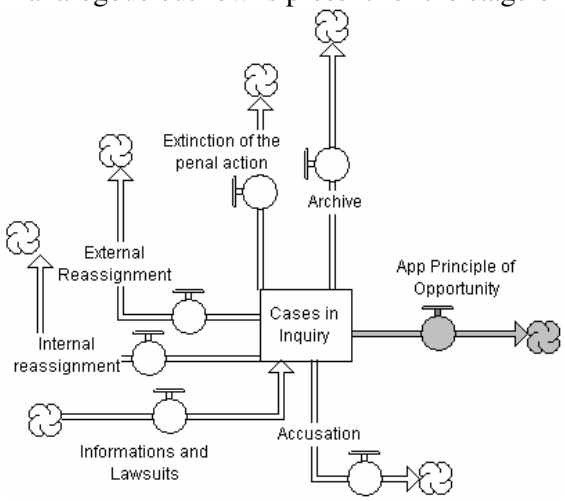


Figure 7
Macro-model of major outflows in the Inquiry Stage. The principle of opportunity is highlighted as one instrument for clearing congestion

Despite the fact that this instrument is one of the main innovations in the reform, currently it is almost not used in practice; indeed the principle of opportunity is currently applied only in the 1 percent of the cases in the inquiry stage and in the 8 percent of the cases in accusation stage (own calculations). These figures explain why a current goal of Colombian policy makers is to widen its application in order to reduce congestion. However, the use of the simulation model brings a first straightforward insight that poses an initial question for the application and the intended goals established for this instrument. Figure 8 displays different congestion values (Eq. 3) given strong increments in the application of the principle of opportunity from t = 0; application rises from 10% to 50%.

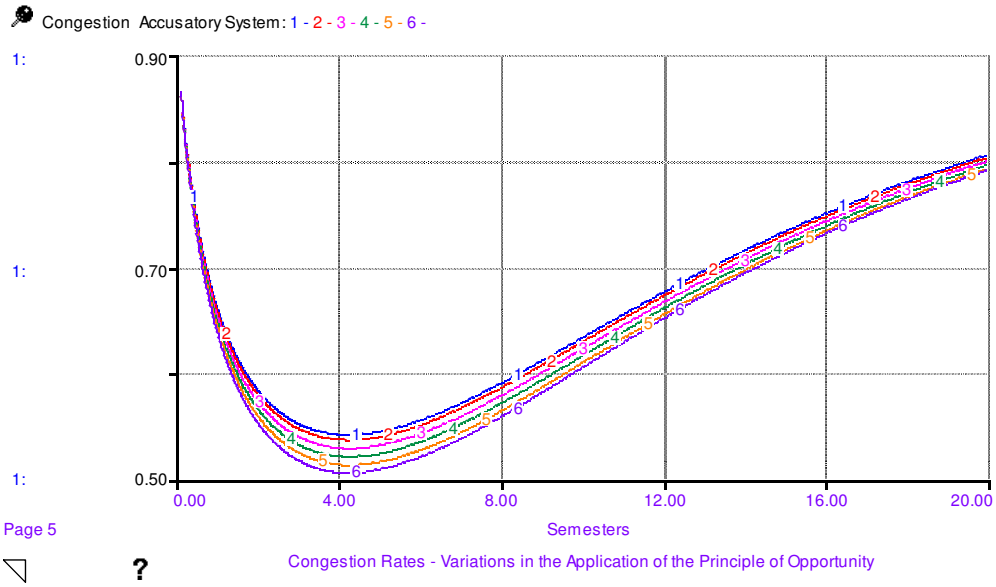


Figure 8

Values for *Congestion*$_t$ (Eq. 3) increasing the application of the Princp. of Opportunity in both Inquiry & Accusation Stages. 1: Base case. 2: increment to 10%. 3: 20%. 4: 30%. 5: 40%. 6: 50%

At first sight the results seemed counterintuitive: congestion rises fast reaching 60% near the 9th semester and above 70% after the 13th semester; furthermore, these increments seemed insensitive to the growing application of the principle of opportunity; even in the long run the difference is almost zero (run 1 vs. run 6). A closer look reveals that it is easy to see why: figure 9 displays the level of cases in inquiry stage for the same six simulations. The policy of increasing the application of the principle of opportunity in the first stage (inquiry) only re-directs outflows but the net evacuation rates are still the same; what really happens is that the cases in inquiry stage keep on growing at a fast rate and even though accusation and trial levels may slightly decline (i.e. inquiry stage send less cases to these stocks) this is not enough given the enormous volume represented by the inquiry stage as compared with the other phases of the process. This simple exercise shows that the mere increment of this mechanism without affecting the current levels of productivity and resources does not alleviate congestion at all — in spite of being one of its declared goals.

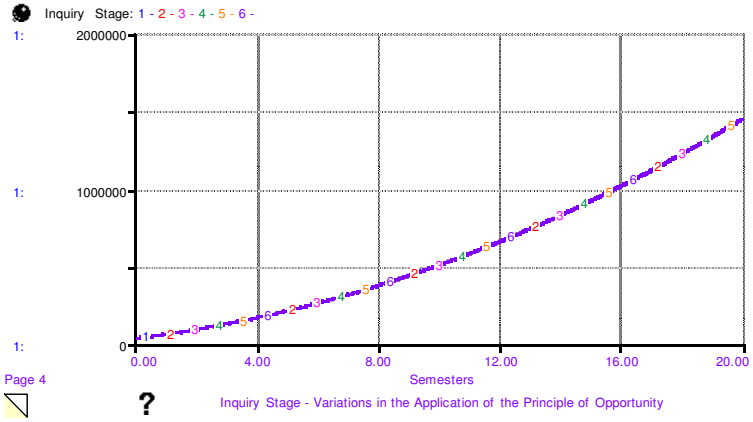


Figure 9

Cases in *Inquiry Stage* increasing the app. of the Pr. of Opportunity at both Inquiry & Accusation Stages
1: Base case. 2: 10% increment. 3: +20%. 4: +30%. 5: +40% 6: +50%

However, even though this first analysis sounds obvious the fact is that a first expectation of the reform is to ease congestion levels via the application of the principle of opportunity; this expectation does not necessarily imply to include an analysis on resources and productivity rates. This first result was already valuable and it proved to be obvious…from hindsight. It is almost counter-intuitive for policy makers to appreciate the fact that to impulse the application of a particular mechanism (or a combination of several of them) does not necessarily alleviate the congestion of the system as a whole.

In addition, the simulation of this kind of policies permitted to discuss and make some assumptions explicit. Indeed one team member argued that the application of the principle opportunity should imply to increase productivity because this process should take much less time than other evacuation processes and therefore each prosecutor could invest the remaining time in evacuating more cases; the simulations generated a discussion about this argument without reaching an agreement — one of the team members particularly claimed that the application of this mechanism does not necessarily require less time than other instruments. This kind of discussion permits to make explicit some assumptions which increase understanding for improving policy design. The illustration with the principle of opportunity is only one example; this reasoning is equally applicable to every other mechanism aimed at only stimulating the application of specific policies (or combinations of them), i.e. to favor some outflows instead of others. The first lesson: policies should not consider only to increase the application of one instrument at expenses of others but to include an increment of the very rates that can speed up the outflow processes. This point takes us to the next analysis: resources and productivity.

5.1.2. Resources vs. Productivity

A next exercise is to affect resources and productivity. To ask which of these two options causes the highest impact with lower costs is a normal inquiry. To increase workforce means at least to search, select, contract and train new employees. To boost productivity can be related to a number of options: to train personnel, to re-organize work units, to acquire technology, etc.

Frequently, to increase personnel implies higher costs under the premise of better results. Several experiments were done in order to combine increments of workforce, productivity and both. Table 1 shows the result of various simulations comparing increments in workforce vs. increments in productivity; although it is not the same to increase *X* % of workforce. vs. to increase a similar *X* % of productivity, this first analysis will be useful to notice the magnitude of the momentum given by accumulations. Here the increments in productivity affect only prosecutors so as to explore policies that are normally considered by the Office of the Attorney General which strongly drives the first two stages with its workforce, i.e. prosecutors. In the table these policies (and their combinations) are compared through a 15 year period at times t = 10, 20 & 30. The table shows the values for accumulated cases at the three stages.

Productivity incr. from $t = 0$ to $t = 30$
Workforce rises in 3 years; it is equally increased each semester from $t = 0$ to $t = 6$

		Accumulated Cases								
		t = 10th semester			t = 20th semester			t = 30th semester		
	Values: Productivity - Workforce	Inq.	Accu.	Trial	Inq.	Accu.	Trial	Inq.	Accu.	Trial
Base Case		557.891	15.650	12.422	1.549.495	28.762	30.670	3.016.207	41.058	51.623
Incr. 10% Productivity	*119*	538.207	15.064	14.121	1.505.325	27.448	34.697	2.945.146	38.944	58.252
Incr. 10% Workforce	*575*	521.218	11.154	23.257	1.460.394	17.839	59.564	2.874.678	23.709	98.581
Both policies		488.281	10.345	26.011	1.384.023	15.962	66.413	2.752.470	20.704	109.795
Incr. 20% Productivity	*130*	520.711	13.866	15.820	1.466.062	24.760	38.724	2.881.980	34.619	64.882
Incr. 20% Workforce	*627*	517.743	10.474	24.353	1.451.953	16.214	62.475	2.861.270	21.120	103.314
Both policies		453.157	7.983	29.997	1.301.935	10.428	76.546	2.621.280	11.865	126.348
Incr. 40% Productivity	*151*	481.343	13.203	19.077	1.377.720	23.271	46.441	2.739.856	32.225	77.588
Incr. 40% Workforce	*732*	510.795	9.604	26.411	1.435.070	14.099	67.963	2.834.454	17.760	112.233
Both policies		368.058	5.125	38.585	1.102.484	3.663	98.456	2.302.670	1.075	162.123
Incr. Productivity x 2	*216*	367.614	8.973	29.129	1.122.511	13.779	70.266	2.329.278	16.953	116.811
Incr. Workforce x 2	*1046*	490.721	7.089	32.356	1.386.299	7.989	83.817	2.756.985	8.055	137.999
Both policies		65.793	329	63.611	389.209	355	155.313	1.164.494	355	250.823
Incr. Productivity x 3	*324*	177.337	2.296	45.836	695.527	240	109.081	1.642.348	240	176.249
Incr. Workforce x 3	*1569*	457.136	2.882	42.303	1.304.701	235	108.739	2.627.373	235	174.490
Both policies		1.330	598	48.959	1.845	662	208.203	3.733	662	379.420
Incr. Productivity x 5	*540*	971	320	69.419	62.423	385	162.839	489.353	385	265.760
Incr. Workforce x 5	*2615*	389.965	316	58.581	1.141.504	329	145.437	2.368.151	329	234.172
Both policies		1.569	21	1.904	2.084	183	0	2.566	664	0
Productivity Base Case	108	*Cases evacuated per prosecutor per year*								
Workforce Base Case	523	*persons (prosecutors)*								

Table 1
Accumulations in inquiry stage, accusation stage and trial for increments in productivity and workforce at t = 10, 20 & 30.

The first thing to notice is the huge growth of cases in inquiry. Only the combination of two very strong (unrealistic indeed) policies seem to have impact on the long rung: to increase 3 times both productivity and workforce which leads to approximately 3.700 accumulated cases in inquiry (t = 30). It should be reminded that the current level of cases in inquiry is near 200.000 cases, already a very high number; nevertheless the figures of Table 1 considerably surpass these current levels (and ideally the goal should be close to zero!). In the long run any realistic policy regarding productivity and workforce seems totally ineffective. A further characteristic to comment is the collapse of the trial stage as long as the accumulations in the first two stages are effectively evacuated; this result underlines the inevitability of addressing the whole criminal process as a system; however, it is uncommon to have an integrated policy since the first two stages are strongly driven by the Office of the Attorney General[5], e.g. prosecutors and investigators, while the third stage concerns more to the Superior Council of Judicature[6], e.g. judges.

[5] The Office of the Attorney General of Colombia is part of the judicial branch of government; its mission is to prosecute offenders, investigate crimes, review judicial processes and bring offenders before courts of justice. It has administrative autonomy.

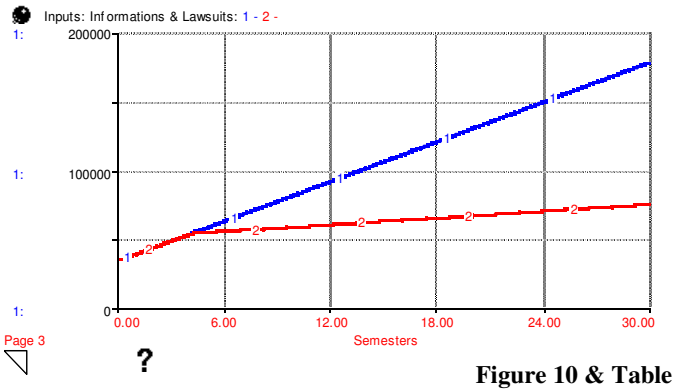
[6] The Superior Council of the Judicature (*Consejo Superior de la Judicatura*) serves for administrating the judicial branch in Colombia, which includes the proposal and implementation of rules for an efficient administration of justice.

In fact, using the model for simulating simple cases and scenarios the conclusion could be different: the question on "resources vs. productivity" seems wrong. The inertia of accumulations appears so high that another point seems to be implicitly suggested: to consider the rate of inputs: policies should include continuous increments of outflow capacities so as to be able to maintain the pace of the growing inputs — at least to sustain the current levels of accumulation! This insight leads to the next policy analysis.

5.1.3. Exploring Inflows

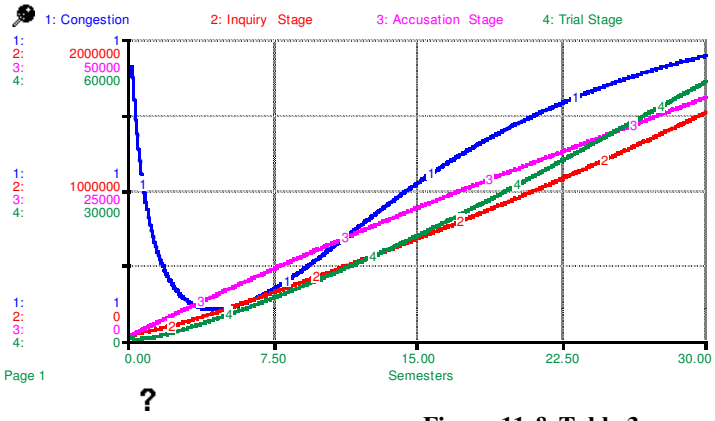
Various characteristics of the way inflows were modeled have been commented above. For instance, although in this first model the inputs are taken as exogenous variables, in reality they surely depend on the effectiveness of the system (see Fig. 2) and on other social factors as for instance population growth, i.e. more people, more crimes. Furthermore, informations and lawsuits were estimated as linear growths when in fact, because of population dynamics, patterns associated with exponential growth might be more precise, that is, even a stronger growth input pattern would be more accurate for the model — and thus congestion inertia is perhaps stronger as well. In any case, increasing rates for inputs seems safe to use with the model to explore the overall aspect of accumulations. Nevertheless, the actual input trend is a matter of debate. During the working sessions there were no consensus about this point. In the next experiments the input growth rate will be lowered. How can these inputs be affected? An actual example illustrates this option. It was mentioned earlier that one option to examine is the redefinition of which conducts constitute potential lawsuits; this change would affect the inflow rates of the system. A new reform was recently introduced (2007) in Colombia which is known as the reform of "small causes" (in Spanish, *Ley de Pequeñas Causas*); one of its most important aims is to reduce the input of minor offences to the accusatory system so as to reduce congestion (Ministry of Interior and Justice, 2008). These scenarios will be explored next.

The following experiment reduces input growth, specifically informations and lawsuits, see Fig. 10 and Table 2.



Decreasing Input starting t = 4			
Time	**1: Base Case**	**2: Reduction**	
0	39,343	39,343	
5	63,431	55,431	12.61%
10	87,519	59,519	31.99%
15	111,607	63,607	43.01%
20	135,695	67,695	50.11%
25	159,783	71,783	55.07%
30	183,871	75,871	58.74%

Figure 10 & Table 2
Reduction in growth of Informations and Lawsuits (*cases per semester*)
1: Base Case. 2: Slower Input Growth



Results				
Time	**Congest.**	**Inq. Stg**	**Acc.Stg**	**Trial**
Initial	0.84	34.756	909	318
5	0.57	229.805	8.414	5.043
10	0.65	460.031	15.504	12.463
15	0.74	707.375	22.176	21.041
20	0.81	971.835	28.433	30.767
25	0.86	1.255.071	34.481	41.252
30	0.89	1.558.747	40.529	51.782

Figure 11 & Table 3
Slower growth in input (informations and lawsuits): Results for Congestion (Eq. 3) and Stocks

As it can be seen in Fig. 11 and Table 3 a severe reduction in input growth does not ease congestion, let alone the accumulations at the main stages of the criminal process; indeed the pattern of congestion is unaltered (compare to Fig. 8 - curve 1, *Base Case*). It should be noticed that the input is still increasing — but at a much slower rate; indeed these are very high reduction rates (Table 2) and this very reflection points to the fact of the difficulty for achieving effective diminution of congestions and accumulations; these efforts would seem almost totally ineffective.

Up to this point another input variable has not been touched: legal disputes. The proportion of disputes which do not settle constitute an input for the accusatory system; currently 31 percent of disputes are not settled and enter to the inquiry stage (own calculations). Previous reforms have stimulated settlement processes in order to resolve minor offences so as to remove these cases from the prosecution system as a way to reduce congestion (Ministry of Interior and Justice, 2007). The next simulations assume additional reductions in legal disputes (Fig. 12)

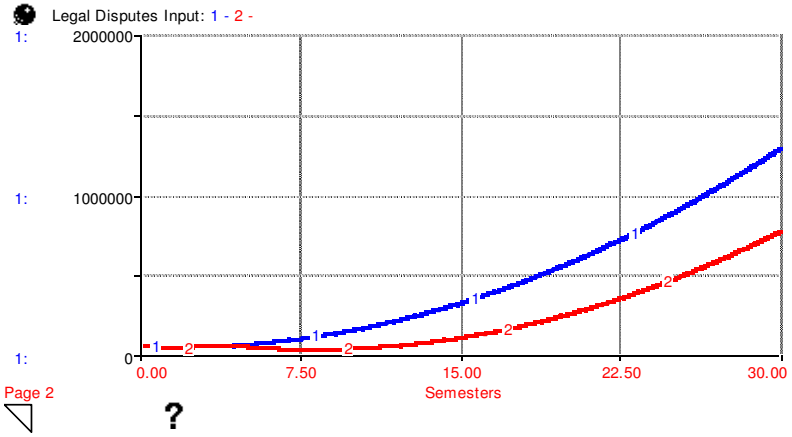
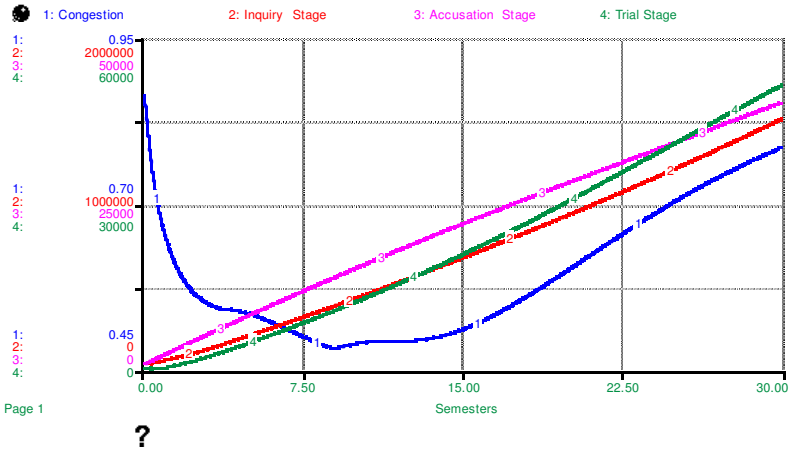


Figure 12
Strong reduction in legal disputes on t = 4. 1: Base Case. 2: Slower Growth

The results of the combination of both input reductions (informations, lawsuits & disputes) are shown in Figure 13 and Table 4.

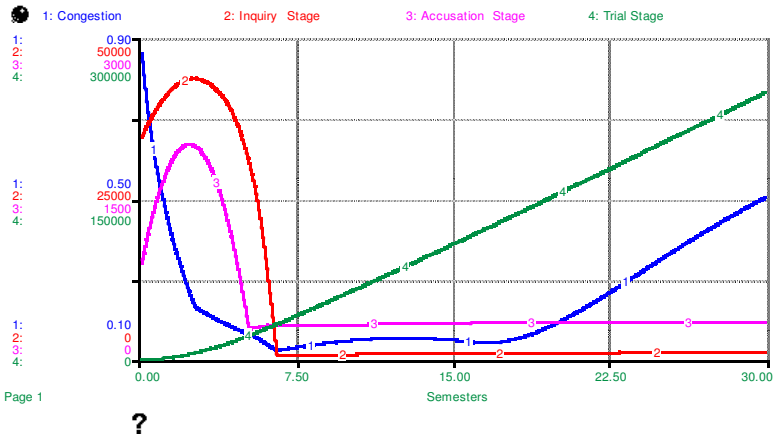


Results				
Time	Congest.	Inq. Stg	Acc.Stg	Trial
nitial	0.84	34.756	909	318
5	0.54	229.805	8.414	5.043
10	0.50	460.031	15.504	12.463
15	0.54	707.375	22.176	21.041
20	0.63	971.835	28.433	30.767
25	0.73	1.255.071	34.481	41.252
30	0.80	1.558.747	40.529	51.782

Figure 13 & Table 4
Slower growth in input (informations, lawsuits, and disputes):
Results for Congestion (Eq. 3) and Stocks

Regarding congestion the best results are achieved around t = 10 (5th year) with a reduction close to 45% (compare to Fig. 8); furthermore, this level is maintained through various semesters. However, it later starts to grow again though slower than in earlier simulations; it reaches a value close to 75% at t = 30 which is far better as compared with previous runs; it should be noticed that it is the first time that the pattern of congestion changes. But still, its values are high; moreover, the cases accumulated through the three stages do not change at all (compare Tables 3 & 4). This last point indicates a further lesson: reduction in inputs are ineffective unless the capacity of evacuation is modified as well; otherwise, the congestion index decreases but this does not imply that accumulations do so. Thus, this combination of affecting both input and output rates will be explored next.

The final option is to combine these reductions in input rates with increments in resources and productivity — as it was explored in the previous section. The next extreme case combines the previous strong reductions with a duplication of workforce and productivity. The results are shown in Figure 14 and Table 5.



Results				
Time	Congest.	Inq. Stg	Acc.Stg	Trial
Initial	0.84	34.756	909	318
5	0.19	50.178	453	21.355
10	0.14	891	298	60.934
15	0.13	948	311	102.135
20	0.23	1.005	324	144.936
25	0.41	1.046	324	188.804
30	0.55	1.498	324	232.737

Figure 14 & Table 5
Slower growth in input (informations, lawsuits, and disputes) plus duplication of resources and productivity:
Results for Congestion (Eq. 3) and Stocks

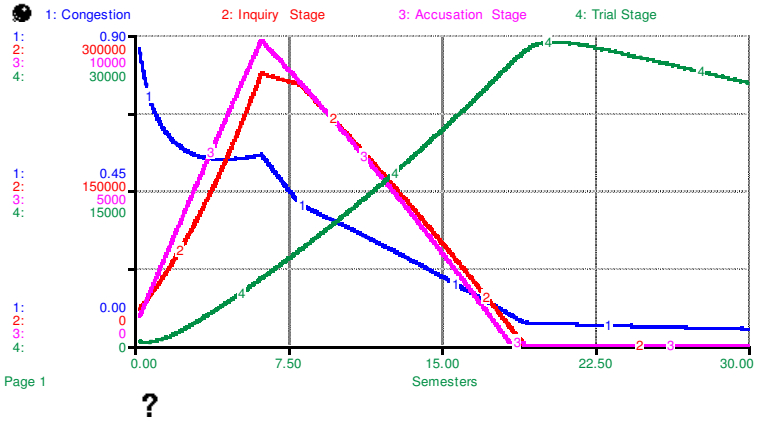
Finally this scenario seems to work. Although the congestion index starts to grow after 20 semesters the stocks of cases at the first two stages are almost depleted. The downside is the outbreak of cases on trial which grows faster than ever.

This last simulation is unrealistic and yet it shows three points. First, this extreme scenario indicates that the trial sector is not prepared for such over-efficiency of prosecutors; this characteristic, which was also stressed earlier, was commented by one expert that joined the last working session: "indeed judges actually seem to count on the inefficiency of prosecution". A coherent systemic policy for enhancing the effectiveness and efficiency of the accusatory reform should not be addressed only to boost prosecution or only to reduce inputs. Second, it is important to highlight that the system would possibly react hiring more judges when the number of cases increases, perhaps with an important delay; in order to focus on this particular issue it would be necessary to study alternative scenarios in which judges increase with the number of cases in the trial stage. Finally, the experiment in any case underlines the persistence of inertia: even with such strong reductions in inputs, combined with almost unfeasible boosting of workforce and productivity, congestion starts to grow after a minimum record rate of 13% reaching a 55% after 15 years, specially because of the remaining cases on trial.

5.1.4. Is Inertia Defeatable?

Stocks constitute the building blocks of SD simulation models. Furthermore, these accumulations provide momentum to systems. The precedent simulations showed the enormous difficulty for clearing (or at least decreasing) the main stocks of the accusatory system. The following last experiment shows the magnitude of this inertia.

Let us imagine that the inputs were shut down at t = 6, that means, if at the present time (actually June, 2008) the lawsuits, legal disputes and informations fall to absolute zero. Figure 15 and Table 6 display the results.



Results				
Time	Congest.	Inq. Stg	Acc.Stg	Trial
Initial	0.84	34.756	909	318
5	0.58	237.765	8.414	5.043
10	0.32	165.225	4.816	12.463
15	0.14	52.675	780	21.041
20	0.06	0	0	25.137
25	0.05	0	0	22.925
30	0.04	0	0	20.606

Figure 15 & Table 6
Total inflows = 0 starting at t = 6: Results for Congestion (Eq. 3) and Stocks

That is, given the current capacities and outflow rates of the present accusatory system, if the total input would fall to zero we would have to wait until the 19th semester of the reform for having zero cases in inquiry and accusation stages, that is, seven years from today! And still there would be cases in trial 15 years after the reform (year 2020); actually, given the current conditions of the system, it would took 42 years to empty the cases in trial.

6. Discussion: The Elusiveness of Stocks and Flows

The previous experiments were focused on the central problem defined for this project: the growing accumulation of cases in the recently implemented reform to the criminal process system. Various lessons were underlined with each new experiment, most of them related to the necessity of developing systemic thinking, in particular the skills to recognize *systems* and the dynamics associated with accumulations. This section connects these lessons with the pervasive struggle for recognizing the difference between stocks and flows and the further difficulty in understanding the dynamics associated with such concepts.

6.1. The Importance of Systemic Thinking

The first simulations in the precedent section tackled punctual points of the criminal process. After all, the usual way to deal with problems is in terms of cause and effect; for instance "How can we tackle congestion?" Answer: "increasing mechanisms for evacuating cases", that is, opening more outflow rates. These actions can include to strengthen the application of specific judicial mechanisms, e.g. the principle of opportunity, or to increase workforce, or to boost productivity, etc. But even the typical debate productivity vs. resources seemed almost irrelevant if we examine the results of such policies, e.g. the duplication of workforce showed to be almost totally ineffective. A parallel approach addresses the inputs, e.g. "How can we tackle congestion?" Answer: "decreasing the cases that enter the system", for instance the mentioned small-causes reform.

But perhaps the way to have a viable criminal process system does not have to do with punctual analyzes on resources, productivity, or any analysis focused on attacking punctual points, i.e. "causes". This approach uncovers our pervasive cause-effect mentality which is deceiving. These way of designing solutions is connected with our seemingly natural inability for recognizing connections, interrelations, systems. A more complete question should consider broader boundaries, e.g. Fig. 2, and should recognize that accumulations are related to *both* inputs and outputs. This point seems straightforward: to fight congestion simply let us guarantee that outflow rates are greater than inflow rates. This is correct. But, is it easy to grasp?

6.2. Dealing with Bathtub Dynamics: Observing and Measuring

The approach to ease congestion in the Colombian justice system seems to show a lack of awareness of "bathtub dynamics". This deficiency includes the unawareness of the distinction stocks vs. flows and a lack of understanding of the dynamics associated with these structures. The following examples illustrate these shortcomings by examining the way congestion and accumulations are observed, measured and evaluated.

The National Evacuation Index

The Superior Council of the Judicature in its 2003-2004 annual report to the Congress states as part of an improvement in the global efficacy of the judicial branch that "during 2003 outputs were 2% higher than inputs" (Superior Council of Judicature, 2004). This seems good news. And in fact this is a first correct approach. However, the report omits the initial value of the stock which in this case means to leave out 6 million cases accumulated in the whole judicial branch at that time (Fedesarrollo & CIPE, 2006); the same way of reasoning is present in its most recent report which states that during 2006 inputs were 5% higher than outputs (Superior

Council of Judicature, 2007). This ratio is defined by the Council as the "national evacuation index", i.e. the proportion of processes that are evacuated in one year in relation with the input of that year. And the interpretation on the behavior of the index is accurate in the report: below 100% means that cases are accumulating. However, these reports leave out the initial conditions of inventories and the respective projections or estimations regarding the effort required to clear such accumulations.

Congestion Indicators

The Office of the Attorney General establishes indicators for tracking congestion. But some of the defined measures overlook the distinction stocks vs. flows. A summary of this logic is shown in the latest statistical bulletin (Office of the Attorney General, 2007) which states: "if pending cases decrease then this means that inputs decrease" (p. 12), which in other words is equivalent to affirm that "a reduction in the stock indicates a reduction in the inflow". This approach is reflected in performance indicators. Three of them are selected here. The mentioned report (Office of the Attorney General, 2007) presents indicators for tracking congestion; the first two of them are defined exclusively in function of outflow rates: "*total congestion index*" and "*effective congestion index*", Eq. 5 & 6. The aim behind is to establish percentages of non-evacuated cases as proportions of the cases that should be evacuated according to a reference productivity.

$$\begin{aligned} \textit{Total congestion index} &= 100 - \textit{total evacuation} \\ &= 100 - \frac{\textit{total output during period } t}{K} \\ &\text{with } K = \text{constant for } \textit{average productivity} \end{aligned} \qquad \text{[Eq. 5]}$$

$$\begin{aligned} \textit{Effective congestion index} &= 100 - \textit{effective evacuation} \\ &= 100 - \frac{\textit{effective outputs during period } t}{K} \\ &\text{with } K = \text{constant for } \textit{average productivity} \end{aligned} \qquad \text{[Eq. 6]}$$

The problem is the misleading way in which these indicators might be interpreted. According to the report "a decrease in these indexes indicates higher decongestion" (p. 13) which is not accurate; such statement suggests that "more outflow (or more evacuation) implies less congestion". Inflows are skipped. And initial values of stocks as well.

A third indicator (Eq. 7) is based on the ratio outputs/inputs ("*marginal congestion index*"):

$$\begin{aligned} \textit{Marginal effective congestion index} &= 100 - \textit{marginal effective evacuation} \\ &= 100 - 100 * \left(\frac{\textit{effective outputs during period } t}{\textit{total inputs during period } t} \right) \\ &= 100 * \left(1 - \frac{\textit{effective outputs during period } t}{\textit{total inputs during period } t} \right) \end{aligned} \qquad \text{[Eq. 7]}$$

Similarly, the interpretation for this indicator is presented in the report as: "if the index falls then decongestion is greater" which is not accurate either; congestion falls only when the *net flow* is negative:

$$\textit{effective outputs during quarter } t > \textit{total inputs during quarter } t$$

which is true only when

$$Marginal\ congestion\ index < 0$$

That is, to be precise, congestion falls only when the index is negative. The definition of the index is useful but the precise interpretation is not utilized; the report only tracks the values of the index through time which are always greater than zero. Naturally it is better if the index falls; but as long as it remains positive, congestion will keep on rising

The Punctual Congestion Indicator

The next example follows a similar logic. It will be used to summarize the problem. The "punctual congestion indicator" is presented by the Office of the Attorney General as in Fig. 16.

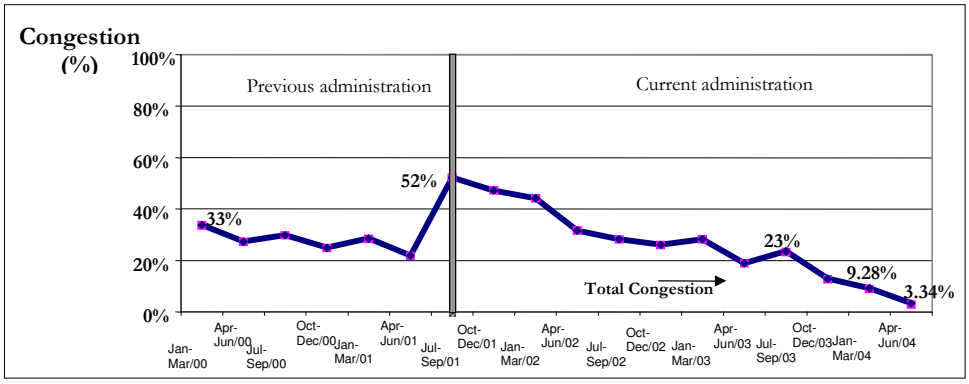


Figure 16

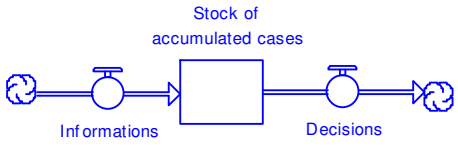
"Punctual Congestion Indicator". The source (Office of the Attorney General, 2004) entitles the graph: "Reduction of quarterly congestion in Office of the Attorney General"

The indicator shown in Fig. 16 is defined in the following way:

$$Punctual\ Congestion\ Indicator_t = \frac{Informations_t - Decisions_t}{Informations_t} \qquad [Eq.\ 8]$$

$Informations_t$: informations that entered during quarter t
$Decisions_t$: decisions and resolutions during quarter t

That is, using a stock-and-flow model this indicator refers to the following process:



Consider the net flow:

$$Net\ flow = Informations - Decisions$$

Thus, the indicator of congestion presented above (Eq. 8) actually is:

$$Punctual\ Congestion\ Indicator\ =\ \frac{Net\ Flow}{Informations}$$

An exercise of graphical integration shows that the stock falls only when the net flow is negative, i.e. *Decisions* > *Informations*, (*Informations* > *0*), otherwise the stock grows. That is, the accumulation of cases keeps on growing as long as this indicator is greater than zero which is what is shown in Fig. 16; this means that congestion is actually growing; nevertheless the corresponding section of the report (Office of the Attorney General, 2004) entitles the graph "reduction of quarterly congestion" which is used to support the final conclusion: "The data shows that the Office of the Attorney General efficiently manages the national budget resources … standing out the reduction of congestion..." (p. 11). The fact is that, according to such indicator, accumulated cases are still going up.

Annual Accumulation Factor

The last example is the *Annual Accumulation Factor – AAF*, (in Spanish *Factor Anual de Represamiento*) which is the central measure designed by the Superior Council of Judicature (2000) for allocating resources related to decongestions mechanisms; these resources are distributed among judicial offices depending on this indicator[7]. It is defined in function of two measures. The first one is intended to capture the amount of work that should be evacuated in one year; its definition is intuitive: the sum of initial inventory plus the amount of cases that enter during that year; this could be expressed as:

$$EW_t = Effective\ Workload_t = Stock_{t\text{-}1} + \sum Inflows_t \qquad \text{[Eq. 9]}$$

A second measure tracks the *effective output* during period t (cases decided and resolved):

$$EO_t = Effective\ output_t = \sum Outflows_t$$

These previous two indicators form the *Annual Accumulation Factor*:

$$AAF_t = 1 - \frac{EO_t}{EW_t}$$

Which becomes:

$$AAF_t = \frac{EW_t - EO_t}{EW_t} \qquad \text{[Eq.10]}$$

It seems logical. The accumulation factor measures the amount of cases that were not evacuated during one year as a proportion of the cases that should have been evacuated during that year. But, again, the distinction of stocks and flows is elusive. It should be noticed that the numerator adds stock with flows; this inconsistency comes from the definition of *effective workload* (Eq. 9). This is problematic. For instance, the stock distorts the whole operation, e.g. consider a high stock initial value as related to inflows and outflows. Appendix 4 shows a short simulation in which an office whose outflow doubles its inflow (a congestion fighter indeed) obtains values for AAF above 90% during 20 years.

[7] The indicator for allocating these resources is actually a composed index in which the *Annual Accumulation Factor* is the most important measure (out of 2).

A further new reform for the judicial branch has as a main goal to tackle congestion (Fedesarrollo & CIPE, 2006). The examples above show that observing and measuring accumulations can be tricky. However, those indicators are the base for observing, evaluating and taking significant decisions about the justice system. A clear understanding and a straight communication of the consequences of the dynamics of accumulations should be integrated in policy making. The use of simulation helps to clarify, understand, and communicate these stock-and-flow dynamics.

6.3. Static Models vs. Dynamic Inertia

Inertia provided by accumulations is one of the key aspects of system dynamics modeling. It is appealing to recall the "note" that Jay Forrester (2003) wrote in 1956 to the Faculty Research Seminar — the first ever M.I.T. "D-memo". In this communication he sketched the foundations of system dynamics. The very first point of his criticism to traditional models is that they overlook the resistance to change (momentum, inertia) and accumulations. These appreciations are still relevant and they constitute the key to understand basic dynamics related to the new Colombian criminal process system, that is, the fact that stocks absorb the differences between inflows and outflows creating disequilibrium dynamics, see e.g. (Mass, 1980; Sterman, 2000). Two central aspects related to the way we approach these problems are underlined next.

The Method of Searching Causes

The momentum generated by accumulations, delays, and the permanent disequilibrium of the system mean that a typical analysis correlating what we see as causes and effects can be deceiving, e.g. to define congestion in function of productivity (and possibly many other factors) does not help to understand *why* congestion is produced. "What *causes* congestion?" The answer to this question probably takes the form of a "laundry list", as in the analogy of Richmond (1993); and it seems commonsense but a further reflection shows that the dynamics of accumulations are complex enough in order to obscure such apparent direct causal connections that we impose upon the world that we observe. Therefore, a first broader lesson suggests to replace "laundry list thinking" for *operational* thinking, i.e. to understand *how things really work* (Richmond, 1993, 2000). This approach implies, among other things, to recognize stocks and flows.

What is a Stock? What is a Flow?

Decision makers have incorrect beliefs about the relationships between stocks and flows (Sweeney & Sterman, 2000); furthermore, such problem seems to be unrelated to particular domains of action, disciplines or background; indeed Cronin and Gonzalez (2007) suggest that this is a pervasive problem in human reasoning.

It has been extensively reported that decision makers hardly understand basic concepts related to dynamic complexity, even if they are highly educated people with training in mathematics and calculus (Sterman, 1989, 2002; Sweeney & Sterman, 2000). In decision making processes, this aspect of stock and flow dynamics is related to what Moxnes (1998) refers as "static mental models". A stock can grow even if the inflow is decreasing. And nevertheless if lawsuits and informations decline the fact that accumulations can still grow is a typical counterintuitive appreciation if decision makers do not include outflow rates in the analysis. And vice versa, to speed up outflows does not imply that stocks will drop; but still, resources and criteria for evaluating the effectiveness of expenditure tends to associate the effects on the stock that are

expected to be caused by changes in the outflow. In the specific case presented here it takes the form: “more prosecutors, less congestion” or “more investigators, less stock of cases”, a natural reasoning in policies aimed at tackling the inefficiency of the justice system; following this static way of thinking a decision maker would expect the stocks of cases to fall after the increment of resources or after the increment of productivity. These simplified static mental models are the rule rather than the exception. The use of simulation becomes an important aid to design policies as long as its systematic use leads to enhance understanding. For the case in hands this means to understand the inertial character of stocks and to recognize that any policy addressed to ease congestion should consider the concept of *rate* as distinct from the notion of *stock*.

7. Outlook

It was stressed that inputs were modeled as exogenous variables. However, this decision limits the feedback analysis. A broader way to approach the problem of congestion is related to the first model discussed in the initial stages of the project. The accusatory system should be conceived as more than a “production machine”. The next steps of this project will expand the boundaries of the model since the next questions are related to the overall purpose of the apparatus of criminal justice. For instance, important negative loops that balance the system could be the way not only to have a more efficient and responsive system but also to have a real integrated and coherent judicial apparatus anchored in a cohesive criminal policy; with such perspective then an examination of the interactions of decisive feedback loops (fig. 2) provide valuable directions.

Another observation is connected to the model building process itself. This project assumes that SD models enhance learning; in this sense, such a goal can be pursued not only through the use of the simulator but during the very process of developing the model. The case reflects the importance of following an iteratively process involving the clients of the model. This process delivered important lessons; the working sessions were conceived not as mere “advancement reports” or as “data demand and requirements by the modellers” but rather as participatory discussions which led to significant decisions; this way of working promoted inquiry, exposed hidden assumptions, motivated several tests and empowered clients, aspects that have been identified as essential for having successful modelling processes (Sterman, 2000).

This article showed the strength of simulation for understanding stock and flow dynamics. Accumulations can be deceiving. Policies addressed to tackle accumulations should take into account the inertia generated by them. This inertia can be explained by the dissimilar characteristics of inflow and outflow rates and the role played by stocks which absorb these differences. We should add the pervasive limitations of static mental models that lead to conceive flawed measures for observing and evaluating system performance. Consistent with previous research (Sterman, 1989; Sweeney & Sterman, 2000), the distinction of accumulations vs. rates, and the plain idea that outflows should work faster than inflows so as to evacuate cases and prevent accumulation, seem to be elusive. The distinction of stock and flows and the inertia associated with accumulations are seemingly simple, and yet valuable lessons are derived from this case.

One expert invited to the last working session, after discussing the main aspects of the model, realized and stressed that *a law is nothing more than a model*. Are legislators aware of the implications of such sentence? This project was conceived as an aid that should help to explain the dynamics of the new Colombian accusatory system; this aim included to understand the

interrelations of different variables that are assumed as relevant as defined by the law (model). To realize such assumptions, and to understand the consequences of those interactions are the central messages of this article.

It is worthy of note to close this paper with an old quotation that pictures the case presented here:

> *We stress the importance of being explicit about assumptions and interrelating them in a computer model… The most important difference between the properly conceived computer model and the mental model is in the ability to determine the dynamic consequences when the assumptions within the model interact with one another. The human mind is not adapted to sensing correctly the consequences of a mental model… The computer model…is a statement of system structure. It contains the assumptions being made about the system...Generally, the consequences are unexpected (Forrester, 1975) - pp. 213-215.*

Acknowledgement

We thank Alfonso Reyes for inviting us to contribute to the citizen observatory. We also thank the project team, specially Juan Carlos Arias, Isaac Beltrán and Oscar Flórez. Financial support for this project was provided by *Corporación Excelencia en la Justicia*; special thanks go to its general manager Gloria María Borrero.

References

Ariza, G., & Sotaquira, R. (2003). *Análisis Dinámico-Sistémico De La Política De Modernizacion Del Sistema Judicial Colombiano.* Paper presented at the I Congreso Latinoamericano de Dinámica de Sistemas, Instituto Tecnológico y de Estudios Superiores de Monterrey. Monterrey, México.

Bard, J. F. (1977). *Criminal Justice System: A Planning Model.* Paper presented at the 9th Winter Simulation Conference, National Bureau of Standards, Gaithersburg, MD.

Becker, G. (1968). Crime and Punishment: An Economic Approach. *The Journal of Political Economy, 76*(2), 169-217.

Blumstein, A. (1967). Systems Analysis and the Criminal Justice System. *Annals of the American Academy of Political and Social Science, 374*, 92-100.

Blumstein, A. (2002). Crime Modeling. *Operations Research, 50*(1), 16-24.

Bohigian, H. E. (1977). *Simulation Modeling of the Criminal Justice System and Process.* Paper presented at the 9th Winter Simulation Conference, National Bureau of Standards, Gaithersburg, MD.

Boyle, S., Guerin, S., Pratt, J., & Kunkle, D. (2003). *Application of agent-based simulation to policy appraisal in the criminal justice system in England and Wales.* Paper presented at the Agent 2003:Challenges in Social Simulation, University of Chicago, Chicago, IL.

Corporation for Excellence in Justice. (2007). Propuesta Organizacional de Planta y de Competencias. Bogotá, Colombia.

Cronin, M. A., & Gonzalez, C. (2007). Understanding the Building Blocks of Dynamic Systems. *System Dynamics Review, 23*(1), 1-17.

Fedesarrollo, & CIPE. (2006). The Justice Reform: An Adequate Redefinition of Priorities (La Reforma a la Justicia: Una Adecuada Redefinición de las Prioridades). *Economía y Política*(20), 1-7.

Fey, W. R., Wadsworth, H. M., & Young, D. B. (1974). Criminal Justice System Training Model: Final Report. Georgia Institute of Technology, School of Industrial and Systems Engineering (LEAA).

Forrester, J. W. (1975). Counterintuitive behavior of social systems. In *Collected Papers of Jay W. Forrester* (pp. 211-244). Cambridge: Wright-Allen Press.

Forrester, J. W. (2003). Dynamic Models of Economic Systems and Industrial Organizations. *System Dynamics Review, 19*(4), 331-345.

Hernández, J., & Dyner, I. (2001). *Crisis in Colombian prisons: Cause or Consequence of a flawed judicial system?* Paper presented at the The19th International Conference of The System Dynamics Society, Atlanta, GA.

Jaén, S., & Dyner, I. (2005). Espirales de la violencia. *Revista de Dinámica de Sistemas, 1*(1).

Jehle, J.-M. (2005). *The Function of Public Prosecution from a European Comparative Perspective. How International Research Can Contribute to the Development of Criminal Justice.* Paper presented at the UNDP-POGAR Conference, Cairo.

MacDonald, R., & Mojtahedzadeh, M. (2007). *Criminal Justice Simulation Model (CJSIM): Technology and the Flow of Criminals in the Criminal Justice System.* Paper presented at the 25th International Conference of the System Dynamics Society, Boston, MA.

Mass, N. J. (1980). Stock and Flow Variables and the Dynamics of Supply and Demand. In J. Randers (Ed.), *Elements of the System Dynamics Method* (pp. 95-114). Cambridge, MA: Productivity Press.

Ministry of Interior and Justice. (2007). Guía Institucional de Conciliación en Penal. Bogotá, Colombia.

Ministry of Interior and Justice. (2008). ABC de la Ley de Pequeñas Causas. Bogotá, Colombia.

Moxnes, E. (1998). Overexploitation of Renewable Resources: The Role of Misperceptions. *Journal of Economic Behavior & Organization, 37*, 107-127.

Newsome, I. M. (2008). Using system dynamics to model the impact of policing activity on performance. *Journal of the Operational Research Society, 59*(2), 164-170.

Office of the Attorney General. (2004). Indicators. Bogota, Colombia.

Office of the Attorney General. (2007). Statistical Bulletin 22. Bogota, Colombia.

Ploom, T. (2000). Principle of Opportunity. *Juridica Abstract, 6*, 369-376.

Riccio, K. J. (1971). A Model for the Study of Crime Control Administration. Lehigh University, Department of Industrial Engineering: (unpublished).

Richmond, B. (1993). Systems Thinking: Critical Thinking Skills for the1990s and Beyond. *System Dynamics Review, 9*(2), 113-133.

Richmond, B. (1994). Systems Thinking / System Dynamics: Let's Just Get on With it. *System Dynamics Review, 10*(2-3), 135-157.

Richmond, B. (2000). An introduction to systems thinking. Hanover: High Performance Systems.

Rouwette, E., van Hooff, P., Vennix, J., & Jongebreur, W. (2007). *Modeling crime control in the Netherlands: insights on process.* Paper presented at the 25th International Conference of the System Dynamics Society, Boston, MA.

Space-General Corporation. (1965). Prevention and Control of Crime and Deliquency. Report PCCD ((Prepared for the Youth and Adult Corrections Agency of the State of California). El Monte, CA.

Sterman, J. (1989). Misperceptions of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes, 43*(3), 301-335.

Sterman, J. (2000). *Business Dynamics. Systems Thinking and Modeling for a Complex World*. Boston, MA, USA: McGraw-Hill.

Sterman, J. (2002). All Models are Wrong: Reflections on Becoming a Systems Scientist. *System Dynamics Review, 18*(4), 501-531.

Superior Council of Judicature. (2000). Acuerdo No. 738 de 2000. Bogotá, Colombia.

Superior Council of Judicature. (2004). Report to the Congress of Colombia. Bogota.

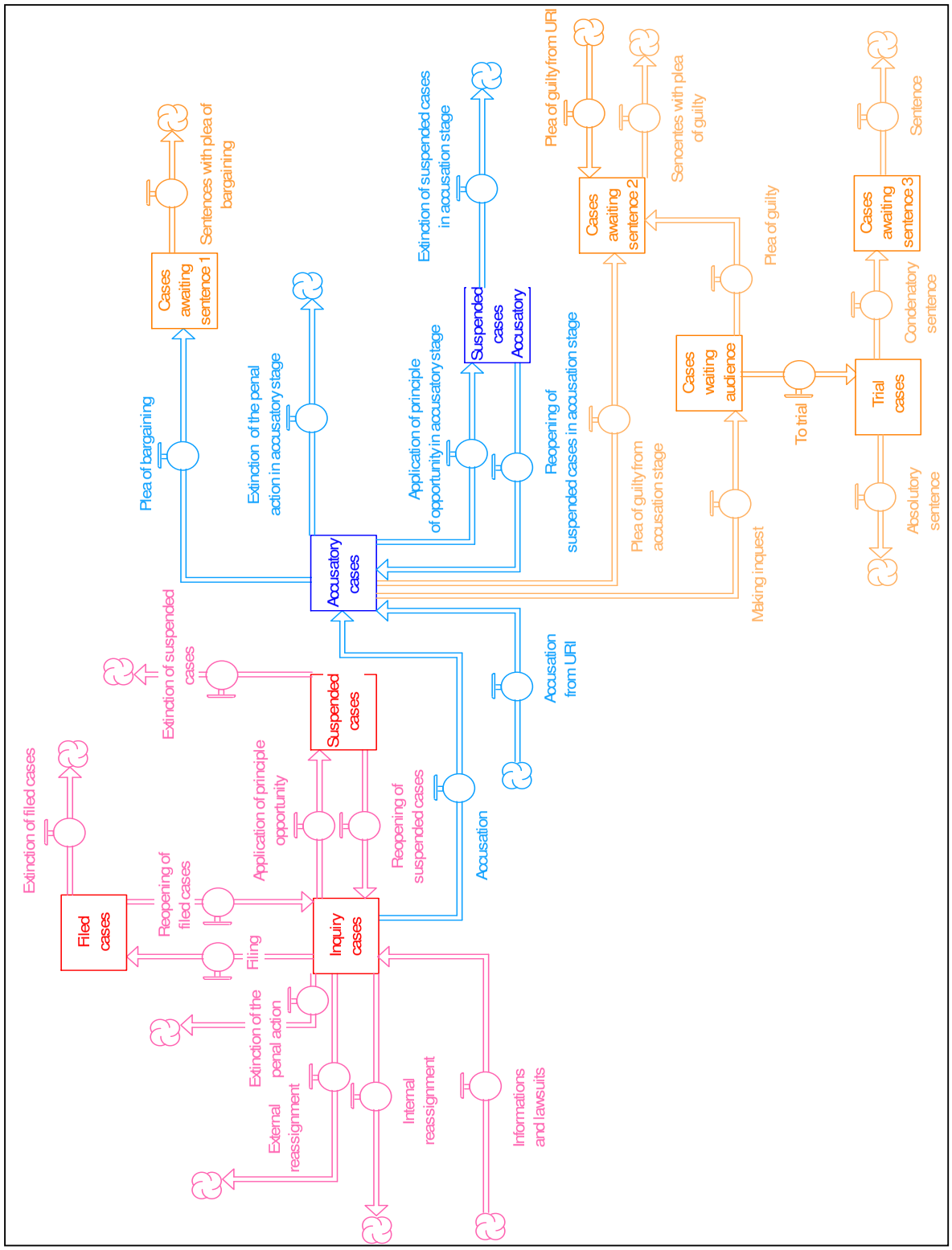
Superior Council of Judicature. (2007). Report to the Congress of Colombia. Bogota.

Sweeney, L. B., & Sterman, J. (2000). Bathtub Dynamics: Initial Results of a Systems Thinking Inventory. *System Dynamics Review, 16*(4), 249-286.

Zangwill, W. I., & Kantor, P. B. (1998). Toward a Theory of Continuous Improvement and the Learning Curve. *Management Science, 44*(7), 910-920.

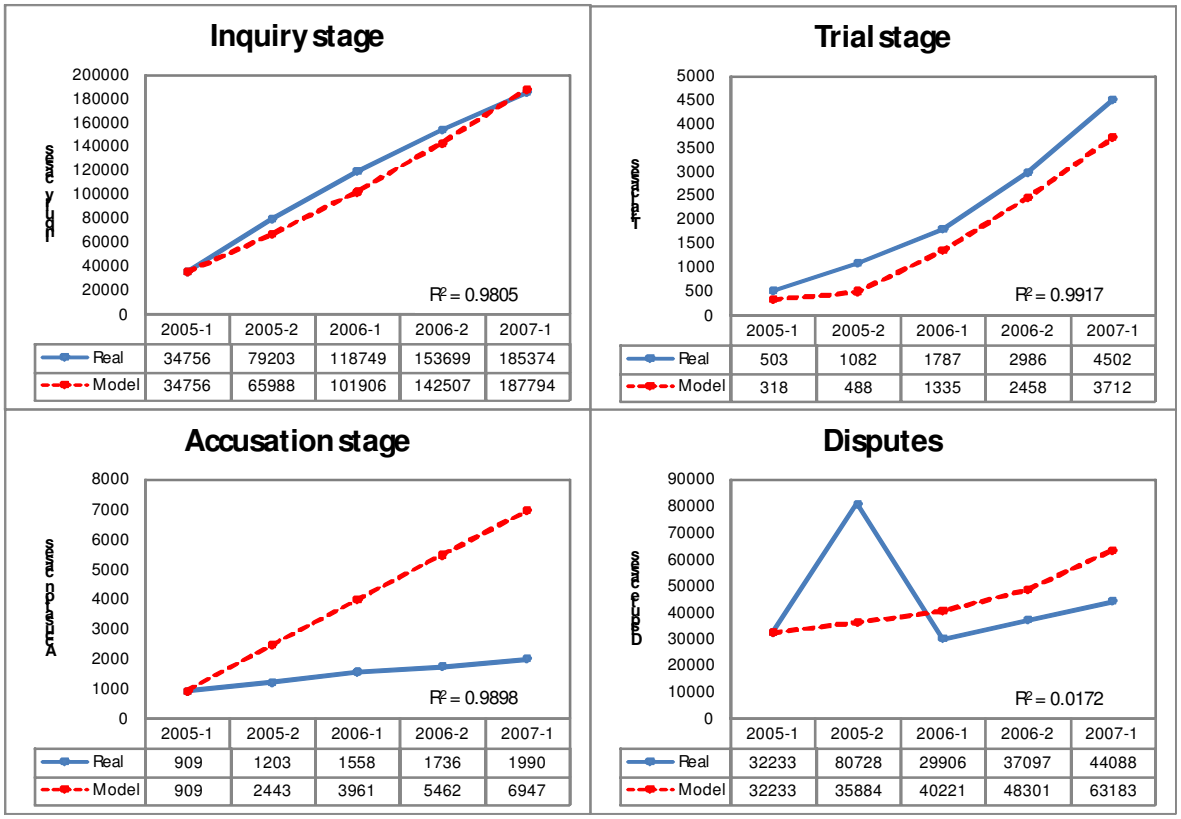
Appendix 1

The Criminal Process Sector



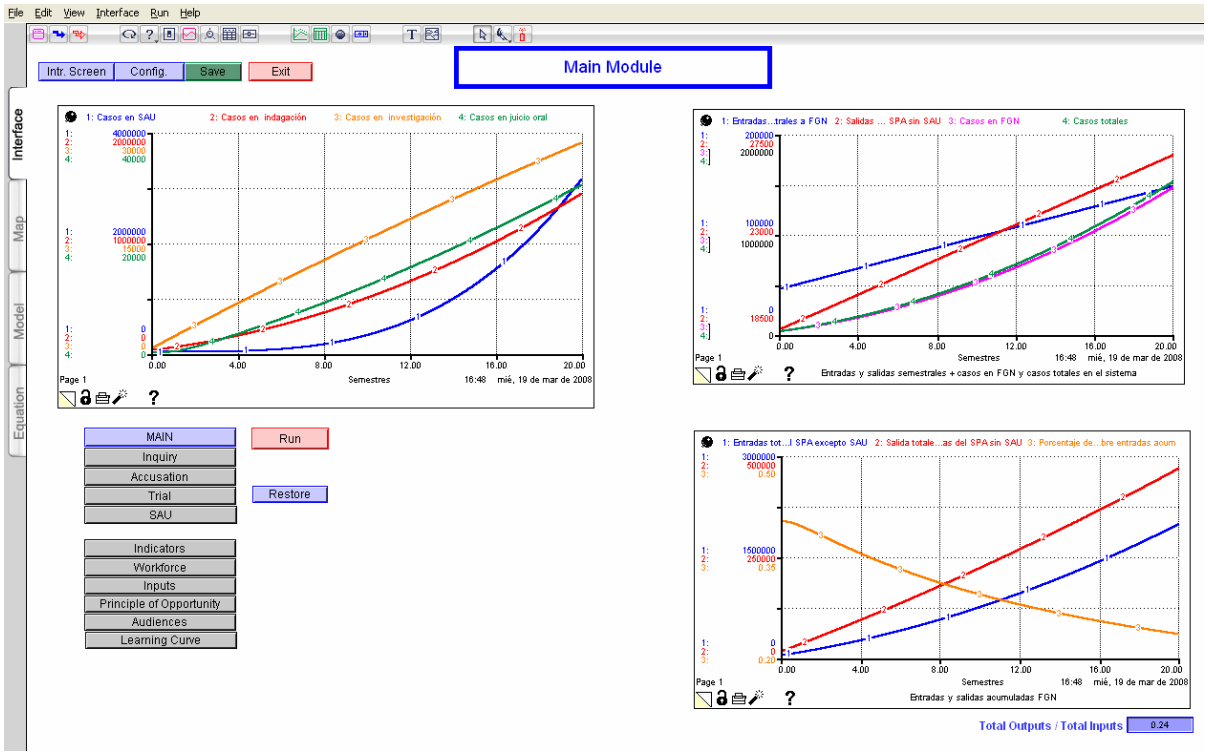
Appendix 2

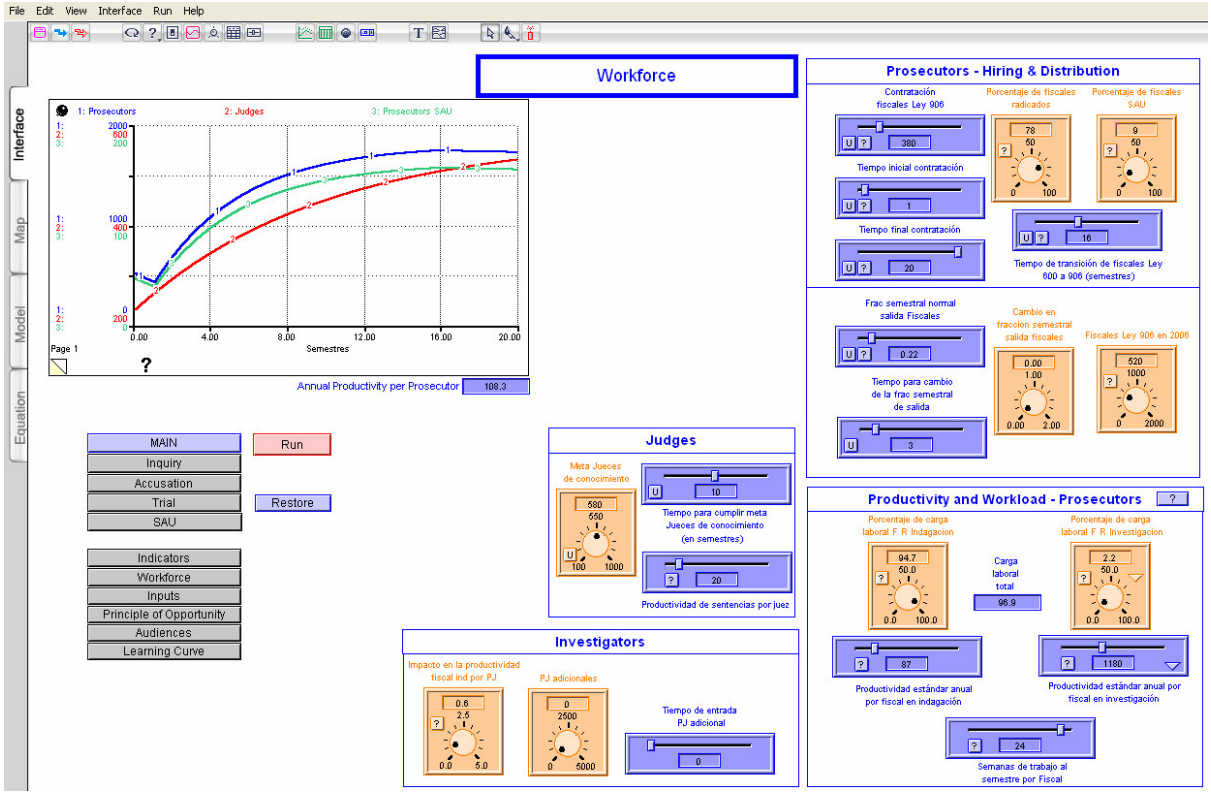
Behavior Reproduction Test



Appendix 3

Flight Simulator - Main Module and Workforce Module Screenshots





Appendix 4

Example for the *Annual Accumulation Factor*

YEAR	STOCK	ANNUAL INFLOW	ANNUAL OUTFLOW	AAF
Initial	2.000			0.955990
0	1.955	45	90	0.955990
1	1.910	45	90	0.955000
2	1.865	45	90	0.953964
3	1.820	45	90	0.952880
4	1.775	45	90	0.951743
5	1.730	45	90	0.950549
6	1.685	45	90	0.949296
7	1.640	45	90	0.947977
8	1.595	45	90	0.946588
9	1.550	45	90	0.945122
10	1.505	45	90	0.943574
11	1.460	45	90	0.941935
12	1.415	45	90	0.940199
13	1.370	45	90	0.938356
14	1.325	45	90	0.936396
15	1.280	45	90	0.934307
16	1.235	45	90	0.932075
17	1.190	45	90	0.929687
18	1.145	45	90	0.927126
19	1.100	45	90	0.924370