5. Virtual Prostitution, Real Complexity

In the last chapter we considered complexity theory as an alternative to more conventional views of causality and scientific research. Here we will explore further the complexity perspective and its implications for HIV prevention by conducting a computer assisted *gedankenexperiment*. By using empirical data from the prostitution scene in a mid-sized German city as a starting point, we will examine step-by-step how complexity theory can be employed to conceptualize HIV transmission, to model its dynamics, and to consider what interventions may be useful in changing the course of the epidemic.

Social simulation allows for a sort of dynamic theorizing process in which ideas and concepts take on a virtual life of their own. We actively create this life and observe how it behaves under our influence. At this point, the life we are creating has an admittedly simple and thus less than human form. However, the potential for analogy is great, thus activating new and creative ways of approaching the interrelationships present in the "real" world.

5.1 Introduction

The spread of HIV among male street prostitutes (*Stricher*) and their clients is an issue for prevention programs in Germany and elsewhere. Male street prostitutes work in bars, train stations and in other public places, receiving money or favors in exchange for sex. This group is to be distinguished from professional male prostitutes (*Callboys*) who are generally self-employed and who are less likely to require basic social services. The starting point for our experiment is data collected by the author since 1999 on this population in Germany and on the organizations providing health promotion services to them (Wright 2000b and Wright in press). There are several advantages to using this group as an example: (1) Male street prostitution takes place within a describable "scene," having relatively clear boundaries; (2) Basic information is available on several key aspects of this scene; (3) Standards have been developed for providing HIV prevention to this population which specify the important factors believed to influence the spread of the epidemic in this group (AKSD, in press).

5.2 HIV and Male Street Prostitutes in Germany

The abovementioned studies conducted by the author provide information concerning the situation of male street prostitutes in Germany and the prevention of HIV for this population. The first study, commissioned by the Deutsche AIDS-Hilfe (Wright 2000b), is an evaluation of the work of the five projects in the country funded specifically to provide services to this group. The analysis presents several estimates concerning primary features of the population as well as information regarding how HIV may be spread in the prostitution scene. The second study (Wright in press) is a needs assessment commissioned by a coalition of public health authorities, AIDS service organizations, and drug treatment agencies in the Rhine-Ruhr area. The multi-site design enabled the gathering of quantitative and qualitative information on the needs of male street prostitutes in this region as well as on the service structures to address those needs. Both studies include a comprehensive review of the international and German literature on the topic. The above-named standards for HIV prevention in this population (AKSD, in press) provides a third source of information.

On the basis of this work and the findings presented in Chapter 2 on the causes of HIV transmission, we can apply complexity thinking in order to construct a model of HIV-risk for male street prostitutes (Figure 8). A more conventional view may focus on characteristics of the individual sex workers and postulate a primary cause for the spread of the disease in this population (e.g. unsafe sex). Our model instead incorporates the following main features: (1) a multi-level structure, (2) interactions between the various levels, (3) the lack of a main or primary cause for disease spread, with the focus being on the dynamic as a whole.

5.2.1 Level 1: The Individual

The primary individuals involved in the prostitution scene are the prostitutes and their clients. Unfortunately, little is known about the characteristics of the clients (e.g. Kleiber et al. 1995) as both research and practice have to date concentrated on the sex workers. At the individual level we identify the following characteristics as being primary, based on the findings in the aforementioned sources:

Level of psychological and social need (instability) – The degree to which the sex worker experiences an instable living situation based on the degree of psychological and social need (the more need, the more instability).

Level of disadvantage – There appear to be three groups of male street prostitutes who are particularly disadvantaged regarding HIV and other risks. These are younger sex workers (including all minors, but particularly those under 16); drug users (that is, those sex workers whose primary motivation for prostitution is raising money to buy drugs); and non-national sex workers. Each of these groups is faced with particularly precarious situations based on a higher degree of need and/or lack of access to existing services. Presumably, membership in more than one of these groups would increase the overall risk accordingly.

Safer sex norm – Each individual has a baseline readiness to engage in unprotected sex when he goes into a sexual encounter. This is influenced by such factors as level of information, past experience, beliefs concerning vulnerability, etc.

Infectiousness – At any moment in time each individual carries a certain probability of transmitting HIV. In the absence of the virus, the probability is zero. When the virus is present, the level of infectiousness can vary.

5.2.2 Level 2: The Dyad

Each sexual transmission of HIV occurs within the context of two people performing a sexual act. In the male street prostitution scene, sex takes place not only between client and prostitute, but also between the sex workers themselves. Here the following factors can be identified:

Contact – Not all members of the prostitution scene have contact with one another. Particularly in larger cities, one cannot assume that all sex workers and clients know each other or that they see each other regularly. Having contact with one another is the necessary pre-condition for all other interactions.

Sexual contact – Sex workers and clients may know each other from the scene for years, but never have sex. Sexual contact would presumably take place less often than other forms of contact in the scene.

Type of sexual contact – Sexual contact itself does not necessarily pose a risk for HIV infection; the type of contact is the deciding factor. What type of contact takes place is the result of myriad factors in the interaction between the two partners.

5.2.3 Level 3: The Prostitution Scene

The prostitution scene is the place in which sex workers and clients meet and it may differ considerably from city to city. Three factors will be named here:

Size – The number of clients and sex workers in the scene affects several dynamics, for example, how often people meet and where.

Mobility – An important characteristic of male street prostitution is the mobility of both clients and prostitutes. People come and go on the scene, thus regularly changing the composition of the population.

HIV prevalence – The degree to which HIV is present in a prostitution scene affects the likelihood of having contact with someone who is infectious and is thus an important factor in terms of HIV risk.

5.2.4 Level 4: The Societal Context

The prostitution scene operates within a larger society which is generally antagonistic to its existence. Thus, an important factor at this level is:

Discrimination – The degree to which the society discriminates against the prostitution scene affects the working conditions of the sex workers, influencing such factors as where prostitution takes place, the risk of legal prosecution, etc.

5.2.5 The Interactions of the Four Levels

When presented in this form, the factors affecting HIV risk within the male street prostitution scene are highly suggestive of interactions within and among the four levels. Based on the information gathered to date, it seems likely that the following dynamics are operative:

The **level of psychological and social need** experienced by individual sex workers is the result of baseline biographical influences (e.g., psychological trauma, family problems, etc.) compounded by the **level of disadvantage** (younger age, non-national status, or drug user) as well as by the **level of discrimination** exerted on the scene from the outside. This neediness influences the readiness to engage in unprotected sex, with a greater level of need associated with a higher level of readiness. The level of need thus increases the baseline probability that a sex worker will engage in unprotected sex (**safer sex norm**). The **size, mobility, and HIV prevalence** within the scene as a whole are important in determining how often people meet up with each other as well as the population-level probability that one's sex partner is carrying the virus. For a transmission to take place, two members of the scene need to have **contact**, this contact needs to be **sexual in nature**, **higher risk sex** needs to occur as a result of the dynamic in the dyad, and the probably of **viral infectivity** (based on the type of sex performed and the infectiousness of the person with HIV) needs to be sufficient.

5.3 Creating the Virtual Environment

Having created a framework for describing the phenomenon of HIV transmission within the male prostitution scene, we now need a way to think through what the ramifications of such a framework are in terms of transmissions dynamics and prevention alternatives. The relatively large array of interactive variables defy a straightforward analysis based on logic and a linear understanding of causality. We could further postulate at a theoretical level regarding potential scenarios based on the proposed relationships between the variables, but such an exercise is not able to look at changing patterns over time or at effects based on the level of the phenomenon. In recent years, computer software has been developed precisely for this purpose. In the computer simulated environment, many variables can be considered simultaneously and over time, thus allowing us to experiment with different scenarios in order to better understand the proposed model.

5.3.1 Agent-Based Models

The software NetLogo from Northwestern University (Wilensky 1999) was employed to write a program for an agent-based model of a larger male street prostitution scene based on data from Düsseldorf. In this way a virtual interactive social environment was produced in which the spread of HIV could be observed and prevention experiments conducted.

As mentioned in Chapter 4 under the discussion of social simulation as a source for complexity theory, agent-based models are only one of several approaches which can be employed. Gilbert & Troitzsch (1999) provide an interesting overview of current approaches in social simulation, describing a history of their development (Figure 9). As the figure shows, the agent-based approach (or "multi-agent models") is relatively recent, being strongly influenced by the work on artificial intelligence.

As discussed in Chapter 4, social simulation is particularly amenable to applying complexity theory because it allows several factors to be addressed simultaneously in an interactive environment. Gilbert & Troitzsch (1999) provide a concise description of the advantage of simulation over mathematical models for the social sciences (pp. 5-6):

Mathematics has sometimes been used as a means of formalization in the social sciences, but has never become widespread except, perhaps, in some parts of economics. There are several reasons why simulation is more appropriate for formalizing social science theories than mathematics. First, programming languages are more expressive and less abstract than most mathematical techniques, at least those accessible to non-specialists. Second, programs deal more easily with parallel processes and processes without a well-defined order of actions than systems of mathematical equations. Third, programs are (or can easily be made to be) modular, so that major changes can be made in one part without the need to change other parts of the program. Mathematical systems often lack this modularity. Finally, it is easy to build simulation systems which include heterogeneous agents–for example, to simulate people with different perspectives on their social worlds, different stocks of knowledge, different capabilities and so on– while this is usually relatively difficult using mathematics.

The latter quality mentioned in the above quotation—the ability to model a heterogeneous population—is maximized in the agent-based approach, and is the reason why this form of social simulation was chosen here. Agent-based programs create a virtual world in which the "people" are mobile, have different qualities, and interact with each other based on these differences. The resulting models are directly analogous to verbal accounts of social phenomena and therefore have a direct intuitive appeal for applications addressing social interventions, being understandable to both researchers and practitioners alike. This is not to say, however, that insights into prostitution and HIV may not also be gained from other methods of social simulation based on simpler patterns of interaction (e.g. cellular automata[17]).

The basic logic behind a social simulation as outlined by Gilbert & Troitzsch (1999) was applied (Figure 10). We see in the figure that the *target* (or population of interest, here male street prostitutes and their clients) can be explored in two ways: by

[17] This form of simulation, also mentioned in Chapter 4, is comprised of a grid of squares or cells. Each square is either "on" or "off," as symbolized by a specific color. Whether a square is turned "on" is determined by a simple rule, for example, if the majority of surrounding squares are "on." This approach is simplier than agent-based models, for example, in that the squares do not move and they are all homogenous (the same rules apply to all squares). Regardless of this relatively simple structure, extraordinarily complex phenomena can be generated. The most well-known application of cellular automata is called "The Game of Life" and can be accessed by the curious reader through various internet sites.

gathering data directly (as was done in the aforementioned studies) thus producing *collected data*, or by abstracting from the target to produce a *model* of what is believed to take place. So far, this resembles the usual inductive and deductive aspects of scientific reasoning. The aspect unique to social simulation is that the model itself can then be activated to produce *simulated data*. The comparison of this data with that collected in the "real" world forms a basis for further theory building and, in some cases, for prediction. Thus, simulation creates a microcosm of the problem *in silico* (to quote Epstein & Axtell 1996), generating data of a quantity and quality which is impossible in the social sciences, either *in vivo* or *in vitro*, because of the enormous difficulty in conducting experimental research, particularly at a macro level.

There are many unanswered questions and yet unresolved problems regarding the application of social simulation. For example, there is no standardized platform for writing simulation programs, which makes the reproduction of experiments practically impossible. The program NetLogo was chosen because it was the only program found which provided powerful capabilities without requiring the prior knowledge of a programming language. In addition, NetLogo comes with an online guide, several examples of models (including code), and a user-friendly interface. Other problems encountered in designing simulations can be summarized under the headings *verification*, *validation*, and *sensitivity analysis*.

Verification refers to the need to test whether or not the model is running as was intended. As will be seen below, the number of randomized variables and the interconnectedness of various aspects of the model make this process challenging. Each run can produce different results than the run before. In the development of the model applied here, verification was tested step-by-step, the number and complexity of variables being increased gradually. When in doubt, sections of code were removed and re-built in order to determine whether the intended relationships were actually being established. The frequent monitoring of the data stream produced by the model made this determination possible.

Validation needs to be applied to all scientific models. It is not readily clear, particularly in more complex explanations, to what degree the model resembles the "real" phenomenon it was designed to simulate. Ultimately, comparing the simulated data to collected data is the best form of validation. This will be applied here, to the extent possible. However, as Gilbert & Troitzsch (1999) explain, there are several potential pitfalls in such a comparison. For example, there is the problem of deciding the validity of how variables have been operationalized—but this is not unique to social simulation. In addition, one has to consider such questions as:

1. To what degree do the model runs reflect the range of possible behaviors which can be generated by all value combinations of the model parameters, including random variables?
2. What data in the "real" world correspond best to those being generated in the model?
3. What aspects of the "real" world can and cannot be reproduced?

Sensitivity analysis refers to how vulnerable the model is to specific assumptions, particularly to the given parameter values. In other words, the robustness of the model needs to be tested. Given the enormous range of possible variable combinations in complex models, one needs to explore how sensitive observed patterns are to specific changes. The number of random variables compounds this problem. Concretely, this results in deciding how often a model has to be run and under what conditions before the researcher can say that s/he has observed the full range (state space) of the model's behavior. Here, there are no hard and fast rules. NetLogo provides a useful application to help in this process. More details will be discussed below.

The actual model was written in the interface language found in NetLogo (Appendix 3). The model will be described here in words, clarifying how the program operates.

We begin with the graphical interface of the program (Figure 11). The two purple-colored buttons set up (*Set Up*) and run (*Start*) the model. Underneath the purple buttons and to the right of the screen are eleven longish green "sliders" which set certain parameters. Each represents a range of possible values. The sliders *stricher-no* and *freier-no* represent the number of sex workers and their clients, respectively, at the beginning of each simulation. *HIVstricher* and *HIVfreier* show the prevalence among the respective populations at the start of each run. The sliders *infectivity-max*, *instability-max*, and *unsafe-prob-max* represent the upper end of the range of possible values for the corresponding variable. *Infectivity-max*[18] is the maximum probability of an infection taking place in an encounter when one agent is positive (scale 0-3.2%).

[18] *Infectivity-max-s* is the probability for sex workers, and *infectivity-max-f* is the probability for clients.

positive (scale 0-3.2%). *Instability-max* is the maximum level of psychosocial need found among the sex workers (scale of 0-100). *Unsafe-prob-max* is the maximum base probability of an agent having unprotected sex (scale 0-100%)[19]. *Discrim* is the level of discrimination in society against the particular prostitution scene being modeled (scale 1-100). The "switch" labeled *migration* determines whether or not the population is closed or open. When turned off, all of the original agents remain in the model. When turned on, both sex workers and their clients come and go. The larger black area is the space in which the agents operate. Sex workers are red and yellow and their clients are blue and white. The yellow and white agents have HIV. The size of the black area affects how often the agents can have contact to one another; for example, a few agents in a large area would be less likely to meet up than many agents in a smaller area. All agents move about this area on an invisible grid. The area itself is not a plane but rather a flattened torus (or doughnut shape), meaning that agents wrap around to the other side when crossing over an edge. The green squares with white windows represent "monitors" which report certain values continually as the model runs. The graph keeps record of HIV prevalence over time.

5.3.2 Setting-Up the Model

For the initial exploratory runs (reported below), the program executed the following:

Upon pressing "Set Up" a batch of sex workers and clients was created. The respective values of 234 and 209[20] as well as the HIV prevalence of 15% are direct estimates based on collected data (Wright in press). The HIV prevalence of 0.05% for clients reflects that for the general population, given the lack of more exact information[21]. Since the number of client agents is well below 2000 (0.05% of 2000 is 1), this results in no clients being infected at the start of the model. The button "Set Up" thus creates the given total numbers of sex workers and clients as well as the given percentages of those who are HIV positive.

---

[19] *Unsafe-prob-max-s* is the probability for sex workers, and *unsafe-prob-max-f is the probability for clients.*

[20] The cited study did not estimate the population of clients directly. The ratio of 1.12 clients per 1 prostitute was, however, used. This is the average ratio of clients to prostitutes observed in Düsseldorf.

[21] Based on the RKI (2001) estimate of approximately 38,000 people living with HIV in Germany. This amounts to a prevalence of about 0.05%.

In addition to **HIV status**, each agent is assigned certain qualities, based on the assumptions of the model. In the case of the sex workers, each agent has values for:

- a level of **instability**
- a level of **disadvantage**
- a probability of engaging in **unprotected sex**
- the amount of **time** he will spend on the scene
- and in the case of those who are HIV positive, a level of **infectiousness**.

The HIV status is assigned randomly such that 15% of the sex workers are positive. Each of those who has HIV is randomly assigned a level of infectiousness from 0-3.2%. This level is based on estimates of the probability of sexual transmission as demonstrated in population-based studies (Vittinghoff et al. 1999)[22]. Although operationalized here as "infectiousness" it is not only influenced by the actual contagiousness of the person with HIV but also, for example, by the biological receptivity of the partner and the type of sex performed.

The level of disadvantage is based on a random distribution of membership in one or more of the three aforementioned groups (minors, non-nationals, drug users), such that the distribution reflects data collected in Düsseldorf.[23] The value of 0.25 is assigned to those with membership in one group, 0.50 to those with membership in two groups, and 0.75 with membership in all three. Those with membership in none of the groups receives a disadvantage score of 0. A baseline level of instability is set

---

[22] The study cited estimates the per-contact probability of HIV transmission in a sample of 2,189 gay men in a multi-site prospective cohort study in the US. The estimates are actually not of infectiousness per se but of the probability of sexual transmission which also depends on the type of sex engaged in, etc. For example, unprotected anal intercourse with an HIV positive partner carried a risk of 0.82% (95% CI: 0.24-2.76) and protected anal intercourse with a partner of unknown status a risk of 0.04% (0.01-0.11). The range of possible values for the actual population mean was taken here to represent a best estimate of a spectrum of infectivity which, in certain cases, is probably higher. The actual transmissibility at any given moment between two people cannot be reliably calculated, given the myriad variables involved and the lack of certainty concerning various indicators. The estimate here does, however, locate the mean risk of transmission at the population level as being very low and thus serves to set the corresponding variable in the model in relation to empirically observed probabilities. The value 3.2% was chosen as the upper limit to correspond with the estimate from a similar study cited by the authors which produced estimates ranging from 0.8-3.2. The range of 0-3.2% thus encompasses the full range of probabilities from both studies, including confidence intervals.

[23] For the cases in which information on membership in the three disadvantaged groups is available, 10.9% are in none of the three groups, 46.2% were in at least one of the groups, 34.1% in two of the groups, and 8.8% in all three. Thus, the quality of belonging to one, two, or three of the disadvantaged groups is assigned to the sex worker agents randomly so that this distribution is reached.

on a scale of 0-100, with 0 representing maximum stability and 100 maximum instability. This baseline level of instability is then multiplied by the disadvantage score and also by the discrimination score (scaled as a percentage). These two values are then added to the baseline instability score to produce the final instability score for the sex worker. Thus, the level of discrimination and disadvantage have the effect of exacerbating the level of instability already present.

The baseline probability of engaging in unprotected sex is set randomly between 0-100%. This score is then multiplied by the instability value. The result is added to the baseline unprotected sex probability to produce the final unsafe sex score for the sex worker. The level of instability thus has a direct bearing on the readiness to engage in unprotected sex.

We thus have a hierarchy of nested variables bearing on the situation of each sex worker (Figure 8) in which the following equations depict the interrationship of key variables:

Psychosocial Instability = Baseline Instability + (Baseline Instability * Disadvantage)
+ (Discrimination * Baseline Instability)

Unsafe Sex Probability = Baseline Unsafe Sex Probability + (Psychosocial Instability
* Baseline Unsafe Sex Probability

The amount of time in the scene is randomly set between 1 "month" and 5 "years" so that at the population level an approximately normal distribution around the mean of 2.5 years results. This is based on estimates of the projects working in this area.[24] What constitutes "month" and "year" in the model will be discussed below.

In the case of the clients, each agent has values for:

- **HIV status**
- a probability of engaging in **unprotected sex**
- the amount to **time** he will spend on the scene
- and in the case of those who are HIV positive, a level of **infectiousness**

[24] In a special inquiry as part of the model-building process, the author asked the five projects in Germany working with male street prostitutes to estimate how long the sex workers and their clients are active in the scene. Four of the projects responded. The estimate across projects was 1 month to 15 years, with an estimated mean in each city of approximately 2.5 years.

Given the lack of information about the clients, simple randomizations were used to generate values. HIV status was initially randomly distributed to reflect the prevalence in the general German population. The probability of engaging in unprotected sex is randomly set 0-100%. The level of infectiousness for HIV positive clients is randomly set between 0-3.2%. The amount of time in the scene was set for 1 "month" to 20 "years," with a population mean of 10 "years," based on project estimates.

The model thus begins with a heterogeneous population of 234 sex workers and 209 clients. Some are more likely than others to have unprotected sex. Some are positive and some are not, with some agents who are positive being more infectious than others. Among the sex workers, some are very stable and are unlikely to engage in unprotected sex while others are more disadvantaged and more likely to undergo risk for infection. At the population level the distribution of the various variables reflects empirical findings, where possible.

At this point the conscientious reader is probably irritated by the way in which the numerous variables have been operationalized. We went from a broad albeit plausible conceptual explanation of HIV transmission among sex workers to a quantification scheme for which there is limited empirical support. That certain variables can be operationalized on a scale of 0-100% may be acceptable to many, but when it comes to a variable such as relative disadvantage, the chosen increments are admittedly pulled out of the proverbial hat. It is important to remember that we are involved here in a *gedankenexperiment*—an exercise in "what if"—in order to set our ideas in action. There are no claims to an exact depiction of the "real" world. Most important for the operationalization of the variables is the creation of a nested hierarchy of causal levels, with each variable being assigned an adequate magnitude of influence within the appropriate level. Different quantification schemes would likely produce different outcomes. It would be interesting to explore this model further by trying different forms of operationalizing the variables in order to see, for example, in what way the observed trends vary. The scheme proposed here should be viewed as a starting point, being one variation among many.

5.3.3 Running the Model

After the start button is pressed, the following events happen during each cycle:

- Each agent moves one space in a random direction.
- If two sex workers or a sex worker and a client are both on the same space, this constitutes a contact.
- When a contact takes place, there is a 0-100% chance that sex will occur.
- If sex occurs and one of the partners is HIV positive, a transmission of the virus takes place if (1) unprotected sex occurs and (2) the level of infectiousness is high enough.
- Unprotected sex occurs at the mean probability of the two partners' unprotected sex scores.
- If unprotected sex occurs, the transmission of the virus takes place at the probability of the infected partner's infectiousness.

If the migration switch is set to "off," the above continues until all are infected. If the switch is set to "on," sex worker and client agents leave the scene as soon as their pre-set time expires. As agents leave, new ones come to take their place such that the population of sex workers and clients varies approximately normally around the initial population values. The new agents coming on the scene are created to have the various individual properties described above, assigned these properties on a random basis but within the pre-programmed population parameters.

For example: A sex worker and a client have contact. There is a 0-100% chance they will have sex. If they have sex, the type of sex is determined by the mean probability of their combined unprotected sex values. That is, if the sex worker has unprotected sex 100% of the time and the client 50% of the time, the chance of their having unprotected sex during the encounter is 75%. If the unprotected sex occurs and one is positive, the chance of transmission of the virus is determined by the infectiousness score of the latter (which is set at 0-3.2%).

Although the various traits are assigned randomly at the time of each agent's creation, associations between variables may develop over time. For example, between being HIV positive and having a high level of psychosocial need, given the relationship between need and risk-taking programmed into the model.

5.4 Questions of Interest

Several questions arose over the course of running the model. These can be divided into two categories: Model Exploration and Prevention Experiments.

Model Exploration included all questions aimed at understanding the basic functioning of the model over time. These were:

- How might time expressed in the model relate to time in the real world?
- How does the risk for infection differ among sex worker and client agents?
- What happens when there is no migration in and out of the scene?
- Under what conditions is a 15% prevalence among sex workers sustained?
- What is the association between HIV status and the following variables for sex workers: probability of unprotected sex; time on the scene; level of instability and disadvantage?

The Prevention Experiments were aimed specifically at influencing the model in such a way as to minimize the prevalence of HIV among sex workers over time. The question was:

- Which parameters can be changed so as to provide the most impact over time on HIV prevalence among sex workers?

Before proceeding it is important to note the time constraints impinging on the running of models. The decision was made to extend all runs at least 3000 cycles in order to take into account 8.2 "years," and thus provide for a considerable amount of agent migration. Even on a fast computer a set of runs can take a considerable amount time to complete. Particularly in experiments involving a range of parameter values, the amount of runs which could be conducted was constrained by the time available. However, the behavior of the model as a whole could be observed over the course of several hundred runs, the subset of data reported here consistent with this overall behavior.

5.5 Exploring the Model

Here we will explore the behavior of the model, as defined above, as well as some assumptions about the initial parameter values. This exploration will be described by answering a series of questions posed about the model. Here all baseline parameter values (discrimination, baseline instability, baseline probability of unprotected sex, and infectiousness) are placed under no constraints, thus varying within their full range, as described above.

5.5.1 Time

How might time expressed in the model relate to time in the real world? As explained earlier, during each cycle of the simulation the agents move about, have contact with each other and, if the probabilities coincide, pass on infection. An interesting initial question to ask is to what extent each cycle may correspond to a unit of actual time. One marker of time based on the behavior being modeled here would be the occurrence of sex between the agents. Unfortunately, there is no data available regarding how often sex takes place either between sex workers and clients or among sex workers. Over several runs of the model we find a mean of about 25 sexual encounters per cycle, with a range of about 10-40. Given the amount of agents being over 400 at any given time (approximately 234 sex workers and 209 clients), it is not unreasonable to assume that a cycle may be equivalent to a day in the prostitution scene. This assumption was made in order to code the time each agent spends in the scene, based on available estimates (see above). The range of values for sex worker agents is thus 30-1825 cycles (to reflect the mean of 2.5 years and a range of 1 month to five years) and for client agents 30-7300 cycles (to reflect the mean of 10 years and a range of 1 month to 20 years). A "month" in the model is thus 30 cycles and a "year" 365 cycles.

5.5.2 Risk for Infection

How does the risk for infection differ among sex worker and client agents? As the parameters described above would suggest, the mean risk for infection among sex workers in the model is generally greater than that for clients. Also, the assumption that sexual contact takes place between sex workers but not between clients exposes sex workers to more contact with potentially infected partners. Finally, the exacerbating effects of psychosocial instability among the sex workers on their readiness to perform unprotected sex results in higher probabilities of this behavior. At the start of a typical run, the mean values for unprotected sex for sex workers is about 70% and that for clients about 50%. On the other hand, client agents are on the average active for a much longer time on the scene, which can mean more potential exposures.

5.5.3 Migration

What happens when there is no migration in and out of the scene? Figure 12 shows the growth in prevalence for both sex worker and client agents as well as for the entire population when there is no migration in and out of the scene. As would be expected, all agents eventually become infected, with the growing prevalence taking the form of a logistic curve[25]. New infections occur slowing, then rise steeply, and level out at a lower level as the population approaches saturation.

It is important to note how slowly HIV spreads among the agents in the model. Even in this scenario providing the best conditions for the spread of the disease within the given parameters, a total of 13,000 cycles (or 35.6 "years") were necessary to generate the above curves. In that time the prevalence among sex workers reached 100%, that among clients 96.2%, and for the population as a whole 98%. This amount of time vastly exceeds the maximum time in the scene for both sex workers (1825 cycles or 5 "years") and clients (7300 cycles or 20 "years"), showing that saturation would not be possible under more realistic time constraints.

Why does HIV spread so slowly in the model? The answer lies in the various probabilities. We have a 5% chance of two sex workers or a sex worker and a client coming in contact (a mean of 25 contacts per cycle, with a total of 443 moves per cycle). When contact takes place, there is a mean 50% chance that sex will take place. If sex takes place, there is an approximate mean of 60% that the sex will be unsafe (average of the client and sex worker agent means, see above). At the outset there is a 15% chance that a sex worker is positive. The mean infectiousness of agents carrying HIV is 1.6%. We therefore have a mean probability of transmission of about 0.004% per contact at the start of the model. Even with an HIV prevalence of 75% among the sex workers, the initial mean risk would only be at 0.02%. Of course, the mean risk increases as more sex workers and clients become infected, but such a process takes a considerable amount of time, given the various probabilities.

5.5.4 Sustaining the Prevalence Among Sex Workers

Under what conditions is a 15% prevalence among sex workers sustained? We now begin exploring the model with the migration switch turned "on," meaning that the population of sex worker and client agents is allowed to fluctuate within the ranges stated above. At each cycle, the agents whose time is up leave and a random number of others come on to the scene for the first time.

As stated above, the 15% prevalence of HIV among sex workers is based on collected data and project estimates in several cities. Unfortunately, we do not know how long the prevalence has been at this level or if there have been fluctuations over time. We can assume, however, that there is some stability to this figure, given project reports. A realistic model would then need to set up the conditions in which a 15% prevalence among sex workers could be generated and maintained.

Experiment 1

We begin with the most simple assumption: The transmission of the virus takes place only among members of the scene, where the prevalence level is already at 15% for sex workers. At the outset, 15% of the sex worker agents are set as positive, the client agents are all negative (based on the prevalence in the general population of 0.05%), and all new agents coming onto the scene are negative. A sample of ten runs of 3,000 cycles (8.2 "years") shows not only considerable *quantitative* differences but also strong *qualitative* similarities in the prevalence curves (Figure 13). The *quantitative* differences are the result of the various random functions in the model, which shows the considerable effect of such stochastic influences. The prevalence levels between the models differ considerably over time. For example, in one run, the prevalence for both groups and for the sample as a whole settles into a pattern of 1-2%. On the other end of the spectrum, there is the run where the prevalence among clients rises to about 10%, that among sex workers drops to a lower level and fluctuates between 4-6%. Most importantly in terms of the question posed here, no run manages to maintain the level of infection among sex workers at 15%. Qualitatively, however, we see commonalities between the runs: There is a general tendency for the prevalence among sex workers to drop steadily, that among clients to rise steadily, and for both to settle into a more stable pattern later in the run.

Experiment 2

In the next experiment to attain a stable level of 15% prevalence among sex workers we again assume an existing 15% prevalence among sex workers and that

[25] SPSS has difficulty creating graphics and running corresponding tests (like the curve estimation here) for very large files. Therefore, a random subset of 50% of the total cases was used.

both clients and sex workers only become infected in the prostitution scene. Here we add, however, an existing prevalence of 5% among clients. It seems reasonable to assume that a certain percentage of clients are infected, given the amount of transmission which takes place under the assumptions of the model. The question is: Can we find an existing prevalence level at which the 15% prevalence among sex workers can be maintained? A look at Figure 14 reveals a qualitatively similar pattern as before, and a continued lack of sustainable prevalence among sex workers at the 15% level.

Experiment 3

In a third attempt to sustain the level of a 15% prevalence among sex workers, we will assume as above an existing 15% prevalence among sex workers and that both clients and sex workers only become infected in the prostitution scene. Here, however, an existing prevalence of 15% among clients will be modeled. Qualitatively, we see a strong resemblance to the tails of the curves in the first two experiments (Figure 15). In all three attempts, once the prevalence between clients and sex workers has become equal, there is a descent in the number of infections among sex workers until a lower level is reached, at which point the curve stabilizes. A quantitative difference between this and the previous two experiments is that the prevalence among client agents is sustained at a higher level. And as above, the 15% level among sex workers cannot be maintained.

Experiment 4

It is not known whether sex workers in Germany become infected predominantly in the course of the sex work or through other sexual contacts (either within or outside of the scene). Here as in other countries, the focus of prevention has been on preventing infections between client and sex worker. International research indicates, however, that sex workers are probably more likely to perform unprotected sex with partners other than their clients (see literature summary in Wright 2000b). We have thus far accounted for such contacts in the model by allowing for sexual relations between the sex worker agents. This, however, is not sufficient to sustain the 15% prevalence among the sex workers. In this experiment we assume that infections for sex workers can occur outside of the prostitution scene, as well. We will set the initial HIV prevalence among sex worker agents at 15%, that among client agents at 0, and that among new sex workers coming onto the scene at 10%. The latter change incorporates the idea of infections from the outside. Can this constellation of variables sustain the prevalence level among the sex workers?

In the sample runs (Figure 16a) we see for the first 3000 cycles a sustained 15% prevalence level for the sex workers, but considerable quantitative differences in terms of the actual prevalence reached for client agents. There is a general qualitative tendency toward the prevalence for clients rising until it meets that of the sex workers, a trend which was found in the above experiments in which the clients began at a lower level of prevalence than the sex workers. At 3000 cycles we are left, however, with the question of how the curves will develop. To what degree will the rising prevalence among client agents pull the sex worker prevalence beyond a mean of 15%? In the last five runs, an additional 2000 cycles were completed (Figures 16b). There may, indeed, be a tendency for the sex worker prevalence to accompany the client prevalence on a slow but steady climb. There is a higher mean and range for the last 2000 cycles in the five runs.

Experiment 5

The results of the experiments thus far suggest that a prevalence of 15% among the sex worker agents is, indeed, sustainable if (1) infection from outside the scene is provided for and (2) the level of infection among client agents is close to but does not exceed that of the sex workers. Let's test this hypothesis by conducting a series of runs in which the initial prevalence among sex workers is 15%, that among clients is 10%, and the chance that new sex workers entering the scene are HIV positive is 10%. In the sample runs (Figure 17a) we see after the first 3000 cycles a range of prevalence values from 10-25%, a mean of 17.2% (range of means for all runs: 15.6-19.1%). For those runs for which we have data for 5000 cycles (Figure 17b), we see a range of 10-27 with a mean of 19.7% (range of means for all runs: 16.1-20.0%). We see that by accounting for infection among sex workers outside of the scene as well as a comparable initial level of infection among clients, we are able to create a more stable prevalence among sex workers. Further experimenting with a range of values reveals that an HIV prevalence among clients of 7% and that among incoming sex workers of 9% provides us with the best combination in order to sustain a mean prevalence level among sex workers of near 15% over time. Averaged over several runs, these values result in a baseline mean prevalence of 14.3% among sex worker agents with an average range of 9.8 - 18.4%. All following model variations will be compared to these baseline values.

5.5.5 The Association Between HIV Status and Other Sex Worker Variables

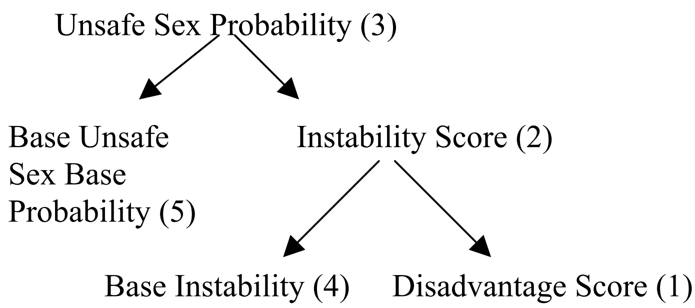
What is the association between HIV status and the following variables: probability of unprotected sex; time on the scene; levels of instability and disadvantage (for sex workers)? At the beginning of a model run, HIV status and the various other sex worker characteristics are randomly distributed among the agents. This results in both HIV positive and HIV negative agents having a comparable mean score on all variables at the beginning of each run, as averaged over several runs. All variables listed here have a range of 0-100%, except for the disadvantage score which ranges from 0-75% and scene time which is by definition 0 at the beginning of a run. The approximate average scores are:

Variable	Average Score
Disadvantage Score	34
Instability Score	80
Unsafe Sex Probability	70
Base Instability	50
Base Unsafe Sex Probability	50
Scene Time	0

Based on how we have constructed our model, one would expect that over time HIV positive agents would show elevated scores on these variables, at least on the variable unsafe sex probability which is most proximal to the infection event. Averaged over several runs (Table 1) we see, in fact, that all scores related to disadvantage, unsafe sex, and instability are higher for positive than for negative agents. The differences are small, however, ranging from 0.1-2.1 points. Of course, any time trends related to individual agents are diluted by migration into and out of the scene, new agents being assigned values on these variables based on random functions independent of HIV status. We see that the range of means for HIV positives tends to be greater and also tends to include more values in the upper range. There is an interesting pattern in the ranking of the variables when comparing the difference between mean positive and negative scores. From greatest to least difference, the variables are:

Variable	Difference
1. Disadvantage Score	2.1
2. Instability Score	1.0
3. Unsafe Sex Probability	0.6
4. Base Instability	0.3
5. Base Unsafe Sex Probability	0.1

This ranking takes on meaning when we recall the nested hierarchy of the variables. The unsafe sex probability is composed of the base unsafe sex probability and the instability score. The instability score is composed of the base instability and the disadvantage score. A picture makes this relationship clearer (the numbers in parentheses are the rankings from the above table):



The two base scores show the least difference; the two variables which combine with the base scores to give the true probability of unsafe sex show the most difference. And of these two, it is the most distal variable (disadvantage) which is the most different for positive and negative agents. Also, we see a mathematical relationship between the variables: The difference between the instability scores (1.2) is approximately the average of the differences for base instability (0.3) and the disadvantage score (2.1). And the difference in the unsafe sex probabilities (0.6) is approximately the average of the differences for the base unsafe sex probability (0.1) and the instability score (1.0). These relationships can be explained if we recall how the model is constructed. The baseline values for unsafe sex and instability are decided through a random function which sets values from 1-100. The result is a near normal distribution of scores centering around the value 50. The disadvantage score was assigned differently, reflecting the proportion of the sex workers in Düssedorf with one, two, or three markers of disadvantage. In this way, three subgroups were built into the sex worker population which differ from the general population. Therefore, of the three agent-specific variables which determine the true unsafe sex probability (baseline unsafe sex, baseline instability, and disadvantage), only disadvantage has a unique distribution pattern at the population level. Although disadvantage is a distal factor in the above hierarchy, the unique differences it ascribes to subgroups of sex worker agents are apparent. These differences are mitigated, however, by the additional layers of the other factors over the course of calculating the true unsafe sex probability for each agent.

The differences in scene time of infected and uninfected agents constitutes a more robust finding, being present in every run of the model. The infected agents are clearly more likely to have been active in the scene longer than agents who are uninfected. More time on the scene means a greater probability of risk contacts during which a transmission can occur.

Here it is important to note that several aspects of the model cannot be measured in terms of aggregate scores and their association with HIV status. Discrimination is a systems level variable which affects all agents, exacerbating sex worker's true level of instability in proportion to their baseline level of instability. Also, the infectiousness of the partners with whom the agents have sexual contact is not measurable. Finally, the probability of unsafe sex of the sexual partners is not included, and thus the other half of the equation determining the actual probability of unsafe sex in these encounters is not present in the aggregate scores.

5.6 Prevention Experiments

The prevention experiments conducted in the second half of the modeling process were aimed at answering the question: Which parameters can be changed so as to provide the most impact over time on the HIV prevalence among sex workers? Reflecting actual prevention practice to this point, the focus was on making changes in the factors affecting sex workers with the goal of reducing the HIV prevalence among this population. Interventions for clients of sex workers were considered in a final step. The order of the experiments proceeded as follows:

Testing the Effects of Each Level—The four variables *infectiousness*, the baseline probability of *unsafe sex*, baseline *instability*, and *discrimination* were each manipulated individually to test their effects on the HIV prevalence among sex workers. The variables were tested in this order, moving from the most proximal cause of infection (the infectiousness itself) to the more distal cause (discrimination).

Combining Adjacent Levels—The four variables were then grouped in pairs of adjacent levels to test combined effects: infectiousness and the baseline probability of unsafe sex; the baseline probability of unsafe sex and baseline instability; baseline instability and discrimination.

Combining all Levels—All four levels were combined at two different intensities to measure the effects of intervention at all levels of the system.

Testing a Realistic Scenario—A particular configuration of intensity at all four levels was tested, in an attempt to reflect what may be a more common scenario in the actual practice of HIV prevention for sex workers.

Testing the Effects of Reducing Clients' Readiness for Unsafe Sex—The effects of adding interventions to reduce the unsafe sex probability on the part of clients in addition to interventions for sex workers was tested by reducing the probability of unsafe sex among clients in addition to the interventions which had already been considered.

As we saw above in the initial exploration of the model, the problem in terms of HIV infection is not only the high prevalence, but also the instability of the prevalence curve over time. Ideally, we would like not only to reduce the mean level of infection among sex workers, but also to change the overall dynamic of transmission. This would result in the prevalence curve becoming more even ("flattening") and, if possible, falling over time as a result of intervention. One measure of such an effect which will be applied here is the relative linearity of the various prevalence curves. Thus, in considering the effects of the various intervention we will be looking at the following over time:

- The mean prevalence
- The range of the prevalence (minimum, maximum and spread)
- The shape of the prevalence curve

For all intervention experiments, the migration switch is "on". It is also important to note that the interventions reduce the maximum level of a particular variable, but continue to allow for a heterogeneous (and approximately normal) distribution of values among the agents. This implies that interventions cannot reduce the risk for all members of a population to some stated goal, but that the average risk in the overall population can be reduced, thus affecting the overall disease burden (measured here in terms of prevalence). Although this is commessurate with basic public health principles, it is often not the assumption in the practice and evaluation of disease prevention (see the often cited discussion of this issue by Geoffrey Rose 1992).

5.6.1 Intervention 1: Reducing Infectiousness

In this first experiment we reduce the initial maximum probability of transmission from 3.2% to 1.6% and then to 0 (which would correspond to the 100%, 50%, and 0% levels of the variable, respectively). This would represent medical or behavioral interventions resulting in a reduced likelihood that the virus is transmitted (for example, but reducing viral load or modifying certain forms of unsafe sex). This level is most proximal to the transmission of the virus.

In Table 2 the results of ten sample runs are presented. Here, we see a clear trend in the reduction of HIV prevalence, as represented by falling minimum, maximum, and mean prevalences. Also, the prevalence curves become increasingly relatively more linear as the level of infectiousness is reduced, as evidenced by the linearity scores[26]. Interestingly, the lowest mean score of 10.9% (when the infectiousness of all sex workers is 0) is indeed a considerable improvement on the model baseline of 14.3, but is far from approaching zero. Here it is useful to remember that HIV positive sex workers are also infected outside of the scene, entering at the level of 9% infected. Also, the level of client agent infectivity remains at 3.2%, thus constituting

[26] The linearity scores show the average *relative* linearity for the stated variable values as related to the linearity of all other combinations *included in the same run*. Therefore, the scores between the tables are not comparable. For each run, the linear fit of all variable combinations in the table are caculated (raw fitness scores). Then the lowest raw score in the run is subtracted from each of the other raw fitness scores, thus producing the scores shown in the tables. Through this normalization procedure, the lowest raw score is thus set to zero. In tables with multiple variables, linearity scores are not show for baseline values of each variable because these were calculated separate from the intervention experiments. The same holds true for the baseline values for the model as a whole.

an ongoing level of infectiousness in the system which is independent of that for sex workers.

5.6.2 Intervention 2: Reducing the Baseline Probability of Unsafe Sex

We recall from the above description of the model that each sex worker has a baseline readiness to engage in unsafe sex. This baseline score is combined with the level of instability to produce the true unsafe sex probability. By reducing the baseline probability of unsafe sex we are simulating the primary goal of most prevention interventions in the "real" world, namely: influencing the psychological factors at the individual level which affect risk-taking. In our experiment we reduced the maximum base probability of unsafe sex by increments of 0.25 (100, 0.75. 0.50, 0.25., 0) (Table 3). If we compare the 0, 0.50, and 100 levels (corresponding to the levels in Intervention 1), we see a clear reduction at the 0.50 level but with no further gains at the 0 level. The difference between the highest and lowest scores is only 1.4 as compared to 3.4 for Intervention 1. The same pattern holds true for the linearity scores, with no clear settling of the curve under the 0.50 level. Therefore, there is no appreciable gain in reducing the maximal probability from 100-75%, nor from 50% to 0. The lowest average mean score overall is 12.9%.

5.6.3 Intervention 3: Reducing the Level of Baseline Instability

The model specifies that a baseline psychosocial instability is intensified proportionally by discrimination and level of disadvantage. This total disability score affects, in turn, the probability of engaging in unsafe sex. The intervention tested here would represent all attempts to address the pre-existing psychosocial problems of the sex workers irrespective of problems in the prostitution scene (e.g., family problems, unemployment, homelessness, alcohol and drug problems, etc.). As in Intervention 2, the variable was incrementally reduced by 0.25 to produce the levels: 100, 0.75. 0.50, 0.25., 0 (Table 4). We see that nothing short of eliminating this variable all-together (reducing its level to 0) appears to produce changes in the prevalence scores. There is, however, no clear trend of increasingly linearity. Whereas a 50% reduction in the baseline unsafe sex score produced a noticeable difference, changes in this variable which is one step further out on the causal hierarchy are only evident at the level of 0. The lowest average mean score for the intervention is 13.1%.

5.6.4 Intervention 4: Reducing the Level of Discrimination

The systems level variable of discrimination was modeled to be one factor (along with level of disadvantage) which impinges on the overall level of psychosocial instability experienced by the sex workers. As such, this variable is the most distal in terms of its affect on HIV transmission. The intervention would represent all efforts to improve the working conditions and social status of prostitutes so as to minimize work-related stressors. The levels of intervention were 0, 0.25, 0.50, 0.75 and 100. In looking at the mean prevalence (Table 5) we see no clear trend based on level of discrimination. There is also no qualitative difference between the curves to be observed. We have apparently reached a level on the causal hierarchy at which intervention is without effect.

5.6.5 Intervention 5: Infectiousness and Unsafe Sex

Here we combine all levels of the two variables *infectiousness* and *baseline unsafe sex* to see what can happen when simultaneous reductions are made in these two factors most near to the HIV transmission event. In the individual experiments above we saw that these two variables were the most effective singular interventions in reducing prevalence. The lowest mean in the combined intervention (10.4%) is a half point lower than lowest mean (10.9%) produced by reducing the level of infectiousness to 0 only. In general, the combined effect of the two variables is greater than each of the variables alone in terms of prevalence level, but with no clear pattern of increase in the relative linearity of the curves being observed (Table 6, Figure 22). There is a stagnation regarding improved prevalence when the level of infectiousness is 0. At this level, changes in the maximum baseline probability of unsafe sex do not make a clear difference. This can be explained when we consider the hierarchical structure of the model. If none of the sex workers are infectious, then it does not matter how often they have unsafe sex; the virus cannot be transmitted. Another interesting observation is that the 25% and 75% levels of the unsafe sex variable take on meaning when combined with lower levels of infectiousness. Whereas, the 75% level differed little from 100% and the levels of 25% and 50% were much like 0 when only the baseline unsafe sex variable was manipulated; here we see gains when adding the effects of reducing infectiousness. For example, even if the maximum baseline probability of unsafe safe remains at 75%, a 50% drop in infectiousness results in a mean prevalence of 12.3% (2 points lower than with no intervention at all). The lowest mean prevalence in this combined intervention is 10.4.

5.6.6 Intervention 6: Unsafe Sex and Instability

In this intervention we reduce the baseline probability of unsafe safe while reducing the baseline instability score. It is the baseline instability which is augmented by the level of discrimination to produce the true instability score. This true instability score is combined with level of disadvantage to adjust the baseline probability of unsafe sex, producing the true probability of unsafe sex for each sex worker agent. As the level of disadvantage cannot be manipulated (that is, membership in one of three most disadvantaged groups cannot be switched), this combined intervention represents the factors in the model amenable to change which most directly impinge on the unsafe sex probability.

Here we see fortuitous results through the combined effects of the two interventions which, however, stagnate in a large area of intersecting values (Table 7). In general, intervening at both levels results in a lower prevalence than intervening at one level alone. In many cases, however, the difference is not large and the effect does not necessarily intensify as the percentage of the two variables decrease. There is no clear evidence for a settling of the curves as interventions are applied. The lowest mean prevalence in this combined intervention is 12.1, higher than that for Intervention 5.

5.6.7 Intervention 7: Instability and Discrimination

The baseline level of instability is augmented by the level of discrimination then combined with the level of disadvantage to produce the true instability score. The instability score, in turn, is used to calculate the true probability of unsafe sex. Combining discrimination and baseline instability thus amounts to manipulating the most distal variables in the model. As we see in Table 8, there is no noticeable trend regarding prevalence when combining these two interventions.

5.6.8 Intervention 8: Combining all Interventions

Here we combine reductions in all of the variables considered thus far: First at the 75% level then at the 50% level. The infectiousness variable was set at the 50% level (1.6%) for both rounds. The mean prevalence scores at both levels only differ by a half point, but they rival the lowest prevalence levels achieved in prior interventions (accept where infectiousness was set to 0) (Table 9). More interesting, however, is the family of curves which is produced. For the first time we have signs that the transmission dynamic is deteriorating. In the majority of runs in this sample we see either a strong falling pattern, settling at a lower level, or varying around or just below the original mean (Figures 18 a,b).

5.6.9 Intervention 9: Combining all Interventions in a Realistic Scenario

In this final experiment we combine all variables in a scenario based on what may be more realistic goals in practice. In a last step, the maximum baseline probability of unsafe sex on the part of client agents is reduced to 50%. The scenerio presented here claims to be more realistic in terms of the relative weight of the changes in the various variables. Although the projects working with sex workers attempt to intervene at all levels, the focus is on addressing the causes of psychosocial instability, followed by influencing safer sex behavior and providing medical care. The larger dynamic of discrimination only changes slowly thus cannot be ameliorated to a large degree in the shorter term.

We see here (Table 10) a low mean prevalence (11.4-12.0) which does not vary much even with the client intervention. This prevalence is lower than of all previous models, except where infectiousness was set to 0. Even more significant, however, is a robust trend found in all runs showing a falling curve (Figure 19).

5.7 What Have We Learned?

> *Representation of the world, like the world itself, is the work of men; they describe it from their own point of view which they confuse with the absolute truth.*
>
> Simone de Beauvoir (1972)

Because of the unusual nature of social simulation and of its still uncommon application to social problems, the critical reader may at this point question the relevance of the above exercise: "But what does this have to do with the real life transmission of HIV?" Like all models, the one presented here is nothing more—and nothing less—than a metaphor for the phenomenon we are trying to describe. If it is a good metaphor, we are able to gain new perspectives on the problem. If the metaphor does not work, the resultant observations are obvious or trivial—or both! Only when a model is confused with reality itself does a researcher demand "absolute proof" of its validity (see Chapter 4). Here we will summarize the questions raised and the insights gained based on the simulation exercise. These will be presented as a series of propositions for further thought, focusing on the structure and dynamics of HIV transmission and its prevention.

5.7.1 Little is Known about the Dynamic Process

One of the primary benefits of social simulation is the challenge it poses to researchers and practitioners to clarify theory and to identify deficits in knowledge (cf. Müller 2000, Gilbert & Troitzsch 1999). It is a daunting task to look at a phenomenon as a whole by considering various levels of causality simultaneously and to set these in relationship to one another. We are forced to move beyond narrowly defined areas of expertise to take in "the big picture." The fact that HIV prevention for male sex workers in Germany is based on the concept of *structural prevention* (see Chapter 3) greatly facilitated the model-building process. However, in defining and operationalizing the various factors, we realize how imprecisely this concept is applied. It is one thing to promote the necessity of both behavioral and social levels of intervention (*Verhaltens- und Verhältnisprävention*), but quite another to describe the important factors on each level and how the levels are related. For example, the model-building process raises the following questions:

- How exactly do the various factors influence each other (directly or indirectly?) and to what degree? How are the various levels of the model interrelated?
- Where do male prostitutes become infected: inside or outside of the scene? And from which partners are they likely to receive an infection (clients, other sex workers, others)?
- How often does potential exposure to HIV infection take place (e.g., how many sexual contacts occur, how often are they unsafe, etc.)?
- What is the prevalence of HIV among the clients of sex workers?
- To what extent do the modeled probabilities reflect those in the "real" world?
- To what degree does HIV prevalence vary within the prostitution scene over time?

5.7.2 HIV Prevalence is Inherently Unstable

It is commonly assumed that an infectious disease is either on the increase, on the decrease, or is stable in any particular population. Our model suggests something else: HIV prevalence may be naturally erratic, as has been observed in the case of other epidemics (Philippe & Mansi 1998). In a dynamic system in which migration occurs, it may not be possible to ascertain a prevalence trend over a longer period of time. Simple assumptions about the level of HIV infection among sex workers and client agents were not sufficient to sustain a 15% prevalence. Only by accounting for infections from outside of the scene could we approach maintaining this level. But even then, the level of infection among sex workers was highly irregular, the runs producing a diverse family of curves showing peaks and valleys at various times. These results suggest the following:

- HIV prevalence may be highly variable, the amount of variability being an indicator of the levels of the various parameters and their relationship to each other at a given point in time. That is, with more information about the various parameters and their interrelationships and given more experience with the model, we may be able to infer changes in parameter values based on the shape and stability of the prevalence curve over time.
- The prostitution scene is not a closed system of HIV transmission, but is maintained by infections occurring outside.
- Not only sex workers, but also clients are transmitting the virus within the prostitution scene.
- A rise or fall in HIV prevalence among sex workers can occur *without any changes in the probability of unsafe behavior or in any other parameter*. Fluctuations can occur as part of the natural history of the complex system.

5.7.3 Distal Social Variables Can be Instrumental

It is common practice in HIV prevention research (and in epidemiological research in general) to use aggregate scores of at-risk or infected individuals to analyze the factors contributing to disease spread, whereby the usual focus is on proximal biological and behavioral causes of disease. In our model, the immediate behavioral cause of viral transmission (unsafe sex) was constructed as the pinnacle of a nested hierarchy of variables including psychosocial instability and social disadvantage.

Over time we were, indeed, able to observe slightly higher unsafe sex scores among infected as opposed to uninfected sex worker agents, which in turn were based on higher disadvantage and instability scores. The difference in disadvantage scores, the most distal variable in the hierarchy, were most pronounced. This would suggest that higher disadvantage scores would be a stronger predictor of being HIV positive in our model than the unsafe sex scores.

This finding recalls a central observation of social epidemiology; namely, the stratification of disease by social class, a trend which has continued to be documented throughout the twentieth century, in spite of an overall improvement in living conditions in the industrialized world (Mielck 1994; Adler, et al. 1994). Poverty (marked social disadvantage) is such a robust predictor of so many health problems, that it can be considered a *fundamental cause* of disease in general (Link & Phelan 1995). The stratification of HIV by class was discussed in Chapter 2. Our model shows how such a trend can be reproduced as the result of nested hierarchies. The observations concerning the effects of disadvantage also reflect the results of empirical studies as reported in Chapter 4 on multi-level analysis. As in the real life examples, our model demonstrates how a distal social cause may be a key component in the disease dynamic.

This being said, the model cautions that changes in distal factors do not result in a significant impact on disease prevalence until they are combined with interventions on lower levels of the nested causal hierarchy. Including causes more removed from the transmission event eases the pressure on more immediate causes, thus lowering the necessary target levels for such interventions.

5.7.4 More Time Means More Potential Exposure

In our model we found a particularly robust trend that HIV positive sex worker agents are active for a longer time on the scene than our HIV negative sex workers. Although the amount of the time on the scene can be measured at the individual level, is can best be viewed as an expression of the effects of time within the larger dynamic rather than a quality of specific individuals. The more time on the scene, the more exposures can take place. The robustness of the finding could lead us to conclude that we should reduce the amount of time individual sex workers are involved in prostitution. This would be a very ambitious goal which would essentially mean the end of the prostitution scene as it is currently structured. Less ambitious and thus more efficient would be reducing the level of exposure in the system as a whole by orchestrating multi-level interventions. The model suggests that this can be done in such a way as to lead to a long-term trend of falling prevalence levels, presumably do to an ever decreasing number of exposures.

5.7.5 Effective Prevention Means Focusing on the Larger Dynamic

Prevention is primarily about influencing the larger dynamic, changing specific factors is secondary. A typical uni-level approach to prevention may focus on the maximum baseline probability of unsafe behavior while ignoring other levels of causality. In this simulation, reducing this probability resulted only in a 1.4% decrease in mean prevalence from baseline to 12.9%, with effects tapering off at the 50% level. When combined with a reduction in infectiousness or a reduction in baseline instability we were able to attain prevalence levels lower than by manipulating any of the variables alone. Only by combining all levels of intervention and then by introducing interventions for clients could we attain both low mean levels of prevalence *and* a deterioration in the system dynamic driving new infections. Thus, even the perfect intervention which managed to reduce the readiness of sex workers to engage in unsafe sex to 0 is not as effective as simultaneous interventions at various levels with more modest goals. The lead question for prevention then needs to be: How can we most efficiently intervene at several levels in order to affect the overall dynamic of the epidemic?

5.7.6 Not Only Effectiveness But Also Efficiency of Interventions Is Important

Interventions at proximal levels are more effective, but including other levels is more efficient. As discussed in Chapter 4, one of the central controversies surrounding social epidemiology from a biomedical perspective is the inclusion of distal causes of disease (such as poverty), challenging an exclusive focus on proximal causes (such as specific pathogens). This controversy has carried over to the field of prevention in that the focus is on disease status and risk behaviors, and not on social context. In this model, we see that reducing the infectiousness of HIV to 0 is, indeed, the most effective singular intervention. Given that this has not yet been attainable, efforts have focused on changing individuals' readiness to engage in unsafe behavior. As a singular intervention, this is the next best thing to reducing infectiousness. It also represents the next level within the causal hierarchy. Many factors impinge directly on whether or not unsafe sex occurs, however, including those modeled here: psychosocial instability and the readiness of the sexual partner to engage in unprotected sex. By including more than one level, less ambitious goals for each intervention can not only produce better short terms results but also improve the transmission dynamics in the system as a whole in the longer term.

5.7.7 Migration Must Be Considered

It is common practice in HIV prevention research (and in epidemiological research in general) to focus on aggregate scores of at-risk or infected individuals to analyze the factors contributing to disease spread. If, however, persons come and go in a population and they are exposed to disease in other situations outside of the particular scenario under study (here the prostitution scene), real markers of vulnerability may be more difficult to identify in cross-sectional data. The aggregation of individual scores dilutes trends for those who have been exposed over a longer period to the particular scene dynamic by combining their scores with those of persons who are fresh on the scene.

Another important and more obvious impact of migration is the changing prevalence levels independent of transmission patterns taken place within the scene itself. This was most striking in the case of the level of infectiousness of sex workers being set at 0; infections from outside of the scene (as well as from infected clients) continued to maintain prevalence levels.

5.7.8 Focus on Reducing Average Risk, Not Absolute Risk in Each Case

The structure of the model itself necessitated focusing on reducing average levels of risk for the population as a whole, rather than reducing the absolute risk of each individual. Although this principle is recognized as a cornerstone of public health practice, interventions for HIV prevention at the community level are rarely designed and implemented under this aspect. The interventions in the model are not focused on the high risk minority, but rather on shifting the curve for the majority. Even slight shifts in the risk curve for the entire population result in large returns in terms of prevention (cf. Rose 1992). Conceptualizing prevention in this way implies setting more modest and thus, perhaps, more realistic goals, particularly for disadvantaged populations. When several levels are combined do such modest goals appear to bear the most fruit.

5.8 Conclusion: Thinking in "Layers"

The core idea presented within the modeling exercise carried out in this chapter is the conceptualization of HIV transmission and its prevention as a system composed of a hierarchy of nested levels. This provides an epistemological alternative to causal explanations which may consider multiple causal factors, but which propose linear relationships between these factors and which ignore dynamic processes. The observations made on the basis of this exercise are, more generally stated, observations of how a multi-level system can function and how interventions can be conceptualized to best influence that system. Once this shift in thinking takes place, several of the observations are less than surprising, particularly the fact that changing parameters at several levels is more effective than changing those at one level only. As with all models, the more parameters we can control, the more we are able to steer the hypothesized outcomes. More remarkable are the particular interrelationships between levels as well as the dynamic system characteristics which the simulation model suggests. Thinking of HIV transmission and prevention in this way brings new challenges, both methodological and epistemological, but it also brings us a step further from static explanations into a world of interactions more strongly reflective of practice experience and of the broader social epidemiological discourse.