Working Paper

Use of the Response Surface Method for the Quantification of a Sexual Behavior Model

Carina van Vliet

WP-95-110
October 1995

IIASA
International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 807 □ Fax: +43 2236 71313 □ E-Mail: info@iiasa.ac.at

Use of the Response Surface Method for the Quantification of a Sexual Behavior Model

Carina van Vliet

WP-95-110
October 1995

IIASA

International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria

Telephone: +43 2236 807 □ Fax: +43 2236 71313 □ E-Mail: info@iiasa.ac.at

Abstract

A microsimulation model of sexual behavior is developed, which simulates individual life histories of individuals in a hypothetical community. Heterogenity of the community is taken into account. The model has the form of a discrete event system in which changes in the characteristics of the individuals result from random events that occur at random times. The model is quantified against real world data by a response surface method.

Key words: simulation, sexual behavior, discrete event system, identification, response surface method

1 The STDSIM simulation model for decision support in STD control

1.1 General introduction to STDSIM

In developing countries Sexually Transmitted Diseases (STDs) are a major cause of acute and chronic ill-health especially in women and neonates. The last years a lot of efforts have been made to combat HIV/AIDS, the most prominent STD. Because of the assumed interaction between the transmission of HIV and other (classical) STDs, nowadays the interest in controlling classical STDs is rising. Control measures for HIV and classical STDs include a wide range of options like promotion of safe sex, sexual health education in schools, safe blood transfusion, case detection and treatment, improving management of AIDS and STD control.

The Commission of the European Communities has requested the Department of Public Health of the Erasmus University of Rotterdam to develop a microsimulation model to support the Commission services and health planners in developing countries in planning of STD interventions. This model will have a general design in order to be useful in different contexts.

Aspects that are taken into account in this STDSIM model relate to demography (birth, death and migration), sexual behavior (e.g. number of sexual partners, prostitution), medical aspects (e.g. transmission risks, natural history of infections) and health care aspects (e.g. percentage of infected persons that seek care, cure rates). Some parameters of the model, for instance related to sexual behavior and organization of health services, will depend on local circumstances. The model will initially be applied to decision support in STD control in Nairobi/Kenya, and will include the following 5 infections: gonorrhea, chlamydia, syphilis, chancroid and HIV.

In this microsimulation model life histories of a large number (e.g. 5,000-10,000) of hypothetical individuals are simulated by means of a computer program. Together, these simulated individuals constitute a hypothetical community. In the computer program, individuals are represented by a number of characteristics. Some of these characteristics remain constant (e.g. sex and date of birth), while others change during the simulation (e.g. infection and disease status, number of sexual partners). Changes in characteristics are the results of events which occur during life history. Examples of events are: starting or ending of relationships, new infections and treatment. Most of the events are stochastic and are simulated by generating random variables from probability distributions. Comparison with field data occurs on aggregated level and in the same format as field observations: e.g. patterns of sexual behavior and prevalences of infection and morbidity.

1.2 Quantification of STDSIM

An important step in the modelling process is model quantification: functional relationships and probability distributions have to be specified and parameters have to be quantified. The model quantification can be checked by comparing the outcomes of the complete model or submodels with empirical data with respect to sexual behavior, prevalences and/or morbidity. In this phase uncertain parameters can be estimated. This is done by seeking for quantifications of parameters that minimize the gap between the model outcome and empirical data. Because of the stochastic components in the model, the outcome of microsimulation models is subject to random variation which hampers conventional optimization. Furthermore, the time required for a single run of the model is considerable, since populations have to be simulated as a whole due to interactions between individuals. The development of efficient algorithms for estimating parameters that minimize the gap between the model and real world data is therefore essential.

One of the main parts of the STDSIM model is the sexual behavior submodel. In this submodel sexual behavior of individuals is modelled explicitly. This means that for every individual in our model decisions are made about when he/she will start having sexual contacts, with whom (another simulated individual) he/she will have sexual relationships of what duration, and at what moment of time. Just as with other aspects in the model, such decisions are made by drawing from probability distributions.

It has been shown that assumptions about sexual behavior have an extremely large impact on the simulated spread and consequences of HIV, and likely the same will hold for other STDs (UN/-WHO, 1989). Therefore, proper specification of the sexual behavior submodel is very important. However, quantitative data on sexual activity, especially in developing countries, is rather limited, and this makes it difficult to specify correct probability distributions. Fortunately, for Kenya some quantitative information on sexual behavior is available, and this can be used to specify and validate the assumptions and quantifications in the sexual behavior submodel.

In the next chapter an overview is given of the data that are available on sexual behavior in Kenya. Chapter 3 contains a detailed description of the sexual behavior submodel, while in Chapter 4 the Response Surface method that will be used for estimation of parameters is presented. Chapter 5 shows the results of these estimation exercises for the sexual behavior submodel. The conclusions of this working-paper are stated in Chapter 6.

2 Data for the sexual behavior submodel

The National Council for Population and Development of Kenya in cooperation with the Central Bureau of Statistics regularly performs a demographic and health survey (KDHS) in Kenya. In this survey also questions about sexual behavior are asked. Data from this survey will be used for the quantification of the sexual behavior submodel. Therefore, in this section some data from the most recently performed survey (1993) are presented. The presented data relate to whole Kenya. In most tables a distinction is made according to the background variable age, while in one table according to marital status. On some topics the results for Nairobi are also presented separately. However, it was not possible to make in the data for Nairobi a distinction according to background variables.

Start of sexual activity

Table 2.1 shows for two age-cohorts percentages of females and males who have already experienced first sexual intercourse at certain ages.

Table 2.1: Age at first sexual intercourse: accumulated percentage of women (n=4148) and men (n=2336) who had already experienced first sexual intercourse by exact age 15, 18, 20, 22 and 25 according to current age

	Females					Males				
current age	15	18	20	22	25	15	18	20	22	25
20-24	17.7	57.7	79.4	NA	NA	27.7	72.8	89.1	NA	NA
25-29	17.9	59.0	81.4	92.2	97.3	26.6	68.7	89.5	94.1	98.3

NA: not applicable

On average women report to start sexual activity at higher ages than men. For females in Kenya the current median age at first intercourse is 17.3 years, while for males it is 16.3 years.

Number of recent sexual partners

The most important data on sexual behavior that are available are probably the number of recent sex partners. Table 2.2 shows per age-class the number of sexual partners reported in the six months preceding the survey.

Table 2.2: Percent distribution of women (n=7540) and men (n=2336) by number of sexual partners in the six months preceding the survey, according to age-class

Age	Females				Males			
	None	1 partn.	2 partn.	3+ partn.	None	1 partn.	2 partn.	3+ partn.
15-19	65.7	29.3	3.6	1.3	-	-	-	-
20-24	26.4	69.2	3.4	1.0	15.8	39.3	20.8	24.0
25-29	14.1	82.1	2.6	1.2	7.8	60.3	14.0	18.0
30-34	11.9	84.8	1.8	1.4	6.0	67.3	13.7	13.0
35-39	17.5	79.4	1.6	1.4	3.8	65.9	21.4	8.9
40-44	19.1	78.8	1.9	0.2	5.6	69.9	18.6	5.8
45-49	26.1	70.2	2.4	1.1	10.2	64.6	15.6	9.7
50-54	-	-	-	-	8.4	64.7	18.2	8.6
Overall	29.9	66.2	2.7	1.2	8.7	59.6	17.5	14.2
Nairobi	24.3	67.6	4.9	3.3	9.4	61.8	12.4	16.5

Table 2.2 shows that males report strikingly more sexual partners in the six months before the survey than females. This gender difference can partly be attributed to response bias. In Kenya it is socially more acceptable for males than for females to have premarital and extramarital sexual contacts, which might have led to underreporting in females. Furthermore, it might be the case that women with three or more recent sexual partners have a much higher average number of sexual partners than men in this activity group, for example if this group of women consists partly or mainly of prostitutes. Prostitution is quite widespread in cities in Sub Saharan Africa and plays a crucial role in the spread of STDs (Over and Piot, 1992).

As can be seen in Table 2.3, numbers of reported sexual partners do not only differ by age, but also strongly by marital status. The term "married" refers to individuals who are in official or unofficial, for instance traditional, unions. Approximately 10% of the married males and 20% of the married females are in polygynous unions (in which a male has more than one wife).

Table 2.3: Percent distribution of women (n=7540) and men (n=2336) by number of sexual partners in the six months preceding the survey, according to type of union

Type of union	Females				Males			
	None	1 partn.	2 partn.	3+ partn.	None	1 partn.	2 partn.	3+ partn.
Unmarried	64.5	28.6	4.9	2.0	20.1	35.9	17.5	26.5
Monogyn. Married	7.2	91.0	1.3	0.6	4.3	75.9	11.0	8.6
Polygynous Married	12.1	84.8	2.0	1.1	2.2	16.3	67.3	14.2

Table 2.3 indicates, amongst others, that for males the relation between marital status and reported number of recent sexual partners is weaker than for females.

Both in Table 2.2 and Table 2.3 it is not specified whether relationships with reported sexual partners were long-term (for example marriage), short-term (for example up to 3 months) or only for one night. Such information would be valuable, because in the presence of concurrent relationships there is not a one-to-one relationship between number of partners in a certain time interval and rate of partner change, and it makes quite a difference for the spread of STDs whether someone has every 6 months 2 new sexual partners, or has the same 2 sexual partners for years.

Number of lifetime sexual partners

Next to the reported number of recent partners, the reported number of lifetime sexual partners, as indicated in Table 2.4, can give extra information on the rate of partner change. In interpreting such data one must be aware of the possible presence of cohort-effects; it is likely that sexual behavior has changed during the latest decades, and this is reflected in lifetime number of partners of older age-groups. IN Table 2.4 numbers of missing answers are explicitly mentioned, because these might be biased towards individuals with a lot of sexual partners.

Table 2.4: Percent distribution of women (n=7540) and men (n=2336) by number of sexual partners in their life, according to age-class

Age	Females						Males					
	None	1	2-3	4-5	6+	Miss.	None	1	2-3	4-5	6+	Miss.
15-19	54.0	22.1	17.1	4.6	1.1	1.1	-	-	-	-	-	-
20-24	10.4	32.3	38.9	13.2	3.5	1.6	5.5	7.1	18.4	15.6	46.9	6.4
25-29	1.3	36.2	41.0	13.9	5.4	2.2	1.3	4.3	10.4	15.9	61.9	6.2
30-34	0.8	34.6	39.4	15.5	6.6	3.2	0.3	3.4	10.0	13.4	66.0	6.8
35-39	0.6	41.8	37.7	13.0	5.0	1.8	0.0	4.4	8.7	10.5	69.8	6.6
40-44	0.2	41.3	39.3	11.1	5.3	2.8	0.5	3.5	11.4	14.3	61.8	8.6
45-49	1.1	52.2	31.7	9.9	4.1	1.0	0.5	1.3	8.1	9.7	72.1	8.3
50-54	-	-	-	-	-	-	1.2	6.1	8.9	13.9	65.6	4.3
Overall	15.3	33.8	33.7	11.3	4.1	1.9	1.7	4.5	11.7	13.7	61.5	6.8
Nairobi	13.4	26.7	34.1	13.9	9.0	3.0	2.3	4.7	11.7	11.7	59.6	9.9

Table 2.4 shows that the difference between males and females in reported number of lifetime sexual partners are even higher than in reported numbers of recent partners, with a modus at 6+ lifetime sexual partners for males and 1-3 lifetime sexual partners for females. Unfortunately, for the category of 6+ lifetime sexual partners (to which the majority of males belong) no more refined data are available at present.

Age difference between spouses

In modelling patterns of sexual behavior of a population, age differences between partners are important. Individuals from different age-classes do not form pairs randomly. There is a general tendency that males prefer younger females, and this can influence the impact of STD transmission tremendously, see also Anderson et al (1992). The KDHS gives information about age difference between spouses. In Figure 2.1 this variable is shown per age-class of the female partner. Again, not only official marriages are taken into account, but also consensual unions.

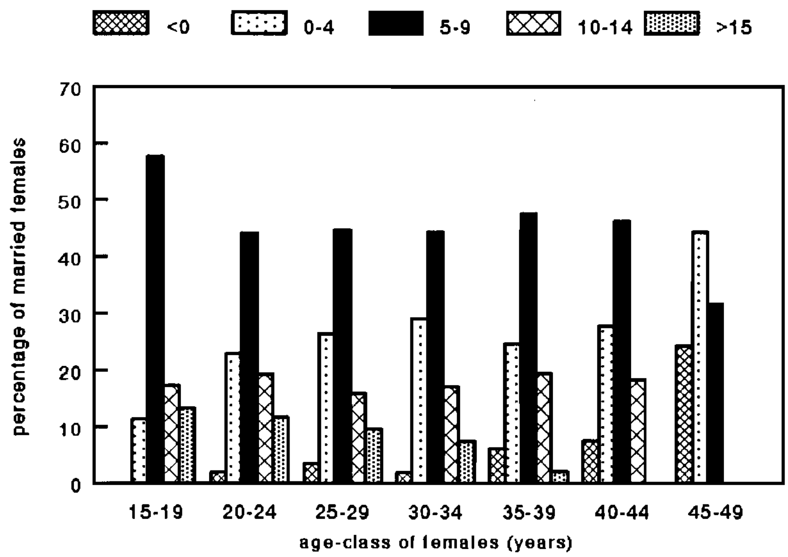

Fig. 2.1: Age difference between spouses: husband's age - wife's age (n=1265)

Overall, husbands are a median of 7 years older than their wives. This difference is higher for younger than older wives, presumably in part due to polygyny, in which men take second wives who are considerably younger than themselves. The median age differences between husbands and their second+ wives is 14.8 years. Furthermore, death removes the largest age differences between spouses; older wives do not often have much older husbands, because most of the much older husbands have already died. For other types of sexual relationships the age differences between partners might differ from the figures for married couples. However, information on this topic is lacking.

3 Description of sexual behavior submodel

3.1 Introduction

In this section a formal description of the sexual behavior model is given. Before we look in detail at the specification of this submodel, we will first remark that the simulation approach in STDSIM is event-driven. This means that for every individual in the model population it is determined which events take when place. During a simulation run one jumps through continuous time from event to event with variable time steps, in such a way that all events of all individuals take place in the appropriate order. At the moment that an event happens, characteristics of the individual(s) involved are updated and the times that new events happen are determined.

The events that are most important in the sexual behavior submodel are the starting and ending of sexual relationships. In the model, these sexual relationships can be initiated both by males and by females, and there is a possibility of concurrent relationships. Homosexuality, which is marginal in Kenya, is not considered in the model. Individuals that have a relationship with each other are linked in the computer model, because events happening to one partner can have impact on the other partner. For instance, if a man is infected with an STD after visiting a prostitute, his wife or girlfriend can later on also become infected. It is assumed that individuals in the model population only have relationships with each other. Marriages will not be explicitly taken into account, because of definition problems and extra parameters involved with the explicit inclusion of marriages. The model population is open in the sense that birth, death and migration occur.

In the partner finding process the concepts of *availability* and *active searching* are crucial. Every new pair consists out of one person who has been searching actively for a partner and a person of the opposite sex who has been available for a new relationship. For every individual in the model population it is at birth decided at which age he/she will become available for a first sexual relationship. If it takes a long time before an available individual is selected by someone of the opposite sex, he/she will go to search actively for a sexual relationship him/herself at a certain moment. At this moment of active searching it is decided in which age-class a partner is searched for, based on a certain preference structure. One of the available individuals of the opposite sex in the appropriate age class is then randomly selected as a sex partner. Later on we may also take into account partner mixing on basis of other factors than age, for instance social status, religion and ethnical groups. Finally, when a relationship starts, it is decided how long this relation will last, and when both partners become available for a new relationship.

In the sexual behavior submodel not only the process of starting and ending sexual relationships is important, but also commercial sex (prostitution). Some women are labelled as commercial sex worker (CSW) during a certain period in their life, which indicates that they are available for single random commercial contacts. To regulate visiting of CSWs, at every new relationship initiated by males it is determined whether this relationship will only be a visit to a CSW or a 'normal' (non-commercial) relationship. We assume that after a CSW visit males are immediately available for another relationship. However, as usual, it will take some time before CSW visitors are going to search actively for another relationship themselves. Contacts with prostitutes do not lead to links between the persons involved, because of the short duration of contact.

Initially, we will not consider special groups with respect to sexual activity apart from CSWs, because reliable data on this topic is totally lacking. However, we are aware of the tremendous impact that heterogeneity in sexual behavior, both between subgroups and between individuals within subgroups, can have on the spread of STDs, see also Anderson and May(1992).

In the following, the assumptions related to starting sexual activity, mixing, duration of relationships, availability and moment of active searching and prostitution are discussed in greater depth and in a more formal way. Note that in the list of parameters a distinction has been made between fixed parameters and free parameters. The former are relatively easy to estimate, or thought to be of relatively minor importance, while the latter are very uncertain and of major importance. We will focus our quantification efforts to the free parameters.

3.2 Formal model description

Indices

s = sex (0 males, 1 females)
i = cumulative number of sex partners of individual
j = age class
k = individual
r = type of relationship (0 casual, 1 steady)

Variables indicating the timing of events

$t\_birth_{s,k}$ = time of birth of individual k of sex s
$t\_death_{s,k}$ = time of death of individual k of sex s
$t\_avail_{s,k,i}$ = time that individual k of sex s is available for his i-th cumulative sex partner
$t\_search_{s,k,i}$ = time that individual k of sex s is going to search actively for i-th cumulative partner
$t\_end_{s,k,i}$ = time that i-th cumulative relationship of individual k of sex s ends
$t\_agech_{k,j}$ = time that individual k changes from age class j to j+1
$t\_cswstart_{k}$ = time that female k becomes CSW
t_cswend_{k} = time that female k ends work as CSW

Stochastic variables regulating the timing of events

$Lifetime_{s,k}$ = duration of lifetime of individual k of sex s
$Durrel_{s,k,i,s',k',i'}$ = duration of relationship i of individual k of sex s with individual k' of sex s'
$Untilav_{s,k,i}$ = time until availability for the i-th partner of individual k of sex s
$Untilse_{s,k,i}$ = time until searching for i-th cumulative partner of individual k of sex s
$Untilcsw_{k}$ = time between starting sexual availability and becoming CSW
$Durcsw_{k}$ = duration of working period as CSW
$Com_{k,i}$ = indicator whether i-th relationship of male k is commercial or not
$Potcsw_{k}$ = indicator whether female k will become CSW
$I_{k,k',i,i'}$ = indicator whether individual k and k' are engaged in a relationship, that is the i-th relationship for individual k and the i'-th relationship for individual k'

Derived variables

$nrrel_{s,k}(t)$ = current number of relationships of individual k of sex s at time t

$cumrel_{s,k}(t)$ = cumulative number of relationships of individual k of sex s at time t

$pref_{s,k}(t)$ = age group in which individual k from sex s searches a partner at time t

Sets

$TOTSET_s(t)$ = set of all individuals of sex s

$AVAILSET_{s,j}(t)$= set of available individuals of sex s and age class j

CSWSET(t) = set of CSWs at time t

Fixed parameters

α_j = length of age-class j

$\pi_{s,j,j'}$ = preference matrix: an individual of sex s and age j searches a partner in age class j' with probability $\pi_{s,j,j'}$

j0 = minimum age group in which males can start a steady relationship

Δ = delay if no suitable partners are available

ϕ = probability that a female becomes a CSW

γ = mean of the distribution of the time between starting sexual activity and becoming CSW

ρ = mean of the distribution of the working period duration of CSWs

θ_s, k_s = scale and shape parameter of the distribution of becoming available for the first sexual relationship of an individual of sex s

x_s = minimum age for becoming available for first sexual relationships for sex s

Free parameters

$\mu_{1,s}$ = mean of the distribution of the time until becoming available for individuals of sex s who are engaged in at least one relationship

μ_2 = mean of the distribution of the time until becoming available after the end of the last relationship of an individual

λ = mean of the distribution of the time between becoming available and starting searching actively for a partner

β_r = mean of the distribution of the duration of casual and steady relationships

δ = probability that a relationship initiated by a male is actually a CSW contact

Time indicator

t_n = time of n^{th} event in the simulation

t_n^+ = time just after the n-th event

Transition equations

$t_n = \min_{s,k,i,j}(t\_birth_{s,k}, t\_death_{s,k}, t\_avail_{s,k,i}, t\_search_{s,k,i}, t\_end_{s,k,i}, t\_agech_{k,j}, t\_cswstart_k, t\_cswend_k) > t_{n-1}$

If t_n = $t\_birth_{s,k}$

- $TOTSET_s(t_n) = TOTSET_s(t_{n-1}) \cup \{k\}$
- $t\_death_{s,k} = t_n + Lifetime_{s,k}$
- $t\_agech_{k,1} = t_n + \alpha_0$
- $t\_avail_{s,k,1} = t_n + Untilav_{s,k,1}$
- if $s = 1 \wedge Potcsw_k = 1 \rightarrow t\_cswstart_k = t_n + Untilav_{s,k,1} + Untilcsw_k$
- $nrrel_{s,k}(t_n) = 0$
- $cumrel_{s,k}(t_n) = 0$

At birth of an individual the model population is expanded. For the new individual the time of transition to the second age class and the first time of becoming available for a sexual relationship are determined. For a potential CSWs, the starting time of her CSW career is determined. Furthermore, the number of current and cumulative relationships are initialized at 0.

if t_n = $t\_agech_{k,j}$

- $t\_agech_{k,j+1} = t_n + \alpha_j$
- if $k \in AVAILSET_{s,j}(t_n) \rightarrow$
  - $AVAILSET_{s,j}(t_n) = AVAILSET_{s,j}(t_{n-1}) \setminus \{k\}$
  - $AVAILSET_{s,j+1}(t_n) = AVAILSET_{s,j+1}(t_{n-1}) \cup \{k\}$

At the time of transition to another age class, the moment of transition to the next age class is determined. If an individual is available for a sexual relationship at this moment, he/she is removed from the set of available individuals of the old age class and added to the set of available persons of the new age class.

if t_n = $t\_avail_{s,k,i}$

- $t\_search_{s,k,i} = t_n + Untilse_{s,k,i}$
- $AVAILSET_{s,j}(t_n) = AVAILSET_{s,j}(t_{n-1}) \cup \{k\}$, for j such that $t\_agech_{k,j-1} < t_n < t\_agech_{k,j}$

At the time that an individual becomes available for a new sexual relationship, it is determined when he/she is going to search actively for a new sexual relationship. Furthermore, the individual is added to the set of available persons of his age class.

if t_n = $t\_search_{s,k,i}$

- if $s = 0 \wedge Com_{k,i} = 1 \rightarrow$
  - if $|CSWSET| = 0 \rightarrow t\_search_{s,k,i} = t_n + \Delta$

    else
    - $nrrel_{s,k}(t_n) = nrrel_{s,k}(t_{n-1}^+) + 1$
    - $nrrel_{s,k}(t_n^+) = nrrel_{s,k}(t_n) - 1$
    - $cumrel_{s,k}(t_n) = cumrel_{s,k}(t_n) + 1$
    - $nrrel_{s',k'}(t_n) = nrrel_{s',k'}(t_{n-1}^+) + 1$, for k' such that $I_{k,k',i,i'} = 1$
    - $nrrel_{s',k'}(t_n^+) = nrrel_{s',k'}(t_n) - 1$, for k' such that $I_{k,k',i,i'} = 1$
    - $cumrel_{s',k'}(t_n) = cumrel_{s',k'}(t_n) + 1$, for k' such that $I_{k,k',i,i'} = 1$
    - $t\_search_{s,k,i+1} = t_n + Untilse_{s,k,i+1}$

· else

· if $\Sigma_{j': \pi\_s,j,j' \neq 0}$ $|AVAILSET_{s',j'}(t_n)| = 0 \rightarrow t\_search_{s,k,i} = t_n + \Delta$

else
- $nrrel_{s,k}(t_n) = nrrel_{s,k}(t_{n-1}^{+}) + 1$
- $cumrel_{s,k}(t_n) = cumrel_{s,k}(t_{n-1}) + 1$
- $nrrel_{s',k'}(t_n) = nrrel_{s',k'}(t_{n-1}^{+}) + 1$, for k' such that $I_{k,k',i,i'} = 1$
- $cumrel_{s',k'}(t_n) = cumrel_{s',k'}(t_{n-1}) + 1$, for k' such that $I_{k,k',i,i'} = 1$
- $t\_end_{s,k,i} = t\_end_{s',k',i'} = \min(t_n + Durrel_{s,k,i,s',k',i'}, t\_death_{s,k}, t\_death_{s',k'})$, if $I_{k,k',i,i'}=1$
- $t\_avail_{s,k,i} = t_n + Untilav_{s,k,i}$
- $t\_avail_{s',k',i'} = t_n + Untilav_{s',k',i'}$, for k' such that $I_{k,k',i,i'} = 1$
- $AVAILSET_{s,j}(t_n) = AVAILSET_{s,j}(t_{n-1}) \setminus \{k\}$
- $AVAILSET_{s',j'}(t_n) = AVAILSET_{s',j'}(t_{n-1}) \setminus \{k'\}$

At the moment of active searching for a relationship, it is for male searchers determined whether they want to have a commercial or a non-commercial relationship. All relationships initiated by females are by assumption non-commercial in nature.

If the new relationship will be commercial, randomly a female partner is chosen from the set of CSWs. In case there is no CSW in the whole population, the active searching is delayed. The duration of a commercial relationship is taken infinitesimal small, and according to this the current and cumulative number of relationships of the male and CSW involved are updated. The CSW visitor is assumed to be immediately available for a new sexual relationship. Therefore, the timing of searching for a new sexual relationship is determined while the set of available persons does not have to be updated.

If the new relationship will be non-commercial it is checked whether there are persons of the opposite sex available in the age-groups that are positively preferred by the searcher. If yes, first the age group in which a partner is searched for is determined, and then the partner itself is selected. If no, the time of active searching is delayed. If a relationship can be established, the current and cumulative number of relationships are updated for both partners. Also the time of the ending of the relationship is determined by drawing a relationship duration. However, the end of the relationship is advanced if one of the partners will die earlier. Furthermore, it is for both partners determined when they become available for a new sexual relationship. Finally, both partners are removed from the current set of available persons.

if $t_n = t\_end_{s,k,i}$

- $nrrel_{s,k}(t_n) = nrrel_{s,k}(t_{n-1}^{+}) - 1$
- $nrrel_{s',k'}(t_n) = nrrel_{s',k'}(t_{n-1}^{+}) - 1$, for k' such that $I_{k,k',i,i'} = 1$
- if $nrrel_{s,k}(t_n) = 0 \wedge k \notin \bigcup_j AVAILSET_{s,j}(t_n) \rightarrow$
  $t\_avail_{s,k,c} = t_n + Untilav_{s,k,c}$, with $c = cumrel_{s,k}(t_n) + 1$
- if $nrrel_{s',k'}(t_n) = 0 \wedge k' \notin \bigcup_j AVAILSET_{s',j}(t_n) \rightarrow$
  $t\_avail_{s',k',c'} = t_n + Untilav_{s',k',c'}$, with $c' = cumrel_{s',k'}(t_n) + 1$

At the end of a sexual relationship the number of current relationships of both partners involved are updated. If the number of current relationships of one (or both) of the partners involved is equal to zero after the ending of this relationship and he/she is not yet available, a new moment of becoming available is determined. Because of the used distribution functions this will lead normally to an advancement in becoming available.

if t_n = $t\_death_{s,k}$

- $TOTSET_s(t_n) = TOTSET_s(t_{n-1}) \setminus \{k\}$
- if $k \in AVAILSET_{s,j}(t_{n-1}) \rightarrow AVAILSET_{s,j}(t_n) = AVAILSET_{s,j}(t_{n-1}) \setminus \{k\}$
- if $k \in CSWSET(t_{n-1}) \rightarrow CSWSET(t_n) = CSWSET(t_{n-1}) \setminus \{k\}$

At the death of an individual, he/she is removed from the model population set. If the dying individual was CSW or available for a new relationship, the set of CSWs, respectively available persons, is also updated.

if t_n = $t\_cswstart_k$

- $CSWSET(t_n) = CSWSET(t_{n-1}) \cup \{k\}$
- $t\_cswend_k = t_n + Durcsw_k$
- if $k \in AVAILSET_{1,j}(t_{n-1}) \rightarrow AVAILSET_{1,j}(t_n) = AVAILSET_{1,j}(t_{n-1}) \setminus \{k\}$

At the starting time of a CSW career, the set of CSWs is updated and the end of this CSW career is determined. Furthermore, new CSW are removed from the set of available persons. Note that CSWs will in first instance be described as a homogeneous group. However, later on it can be taken into account that in reality this group is non-homogeneous with large differences with regards to number of clients and types of clients.

if t_n = t_cswend_k

- $CSWSET(t) = CSWSET(t_{n-1}) \setminus \{k\}$
- $t\_avail_{s,k,i} = t_n + Untilav_{s,k,i}$

At the end of a CSW career, the CSW is removed from the set of CSW. Furthermore, it will be determined when she will become available for a new (non-commercial) sexual relationship.

Distributions of the stochastic variables regulating the timing of events

- $Lifetime_{s,k}$ distributed according to the Kenyan life tables for males and females used for United Nations demographic projections. The processes of birth and death are regulated outside the sexual behavior submodel. Overall, the model population size is growing with about 1.5% a year.

- $Untilav_{s,k,1} - x_s \sim Weibull(\theta_s, k_s)$
  $Untilav_{s,k,i} \sim Exponential(\mu_{1,s})$ if $i \geq 1 \land nrrel_{s,k}(t_n) \geq 1$
  $Untilav_{s,k,i} \sim Exponential(\mu_2)$ if $i \geq 1 \land nrrel_{s,k}(t_n) = 0$

The time until availability for a first sexual relationship is determined at birth. From a certain minimum age x_s on, the first time until availability follows a Weibull distribution which differs between males and females. The time until availability for following relationships is determined at the beginning of a new relationship, and is exponentially distributed. This time also differs between males ($\mu_{1,0}$) and females ($\mu_{1,1}$), because in Table 2.3 it is shown that the association between marital status and number of recent partners is much stronger for females than for males.

Initially, a dependency between duration of relationship and time until availability is not directly implemented. However, our hypothesis is that individuals who have no relationship are more available for a new relationship than individuals involved in at least one relationship. Therefore, if at the end of a relationship no other relationships are left for an individual, the time until availability will be advanced by drawing out of an exponential distribution with $\mu_2 < \mu_{1,s}$.

Possible extensions are the inclusion of heterogeneity between individuals, or dependencies with duration and/or number of relationships.

- $\text{Untilse}_{s,k,i} \sim \text{Exponential}(\lambda)$

The time until active searching for a new relationship is determined at the moment that availability starts, and is exponentially distributed. Note that initially we do not take into account interdependencies with age, sex, or number of current relationships.

- $\text{Durrel}_{s,k,i,s',k',i'} \sim \text{Exponential}(\beta_0)$, if $t_n < t\_agech_{k,j0} \wedge s = 0$, or if $t_n < t\_agech_{k',j0} \wedge s' = 0$
  $\text{Durrel}_{s,k,i,s',k',i'} \sim \text{Exponential}(\beta_1)$, otherwise

Initially we thought about using only one distribution for the duration of relationships, but with this assumption it was very hard to get results with the STDSIM model that resemble the data. Therefore a distinction has been made between steady and casual relationships. In this it is assumed that relationships involving males under a certain age are casual, and other (non-commercial) relationships are assumed to be steady. Both for steady and casual relationships we have used exponential distributions.

Further extensions might be including dependencies between relationship duration and age at the starting of relationships, and including dependencies between relationship duration and the presence of other (long term) relationships.

- $P[\text{Com}_{k,i}=1] = \delta$

The probability that a relationship initiated by a male is actual a CSW contact is determined at the moment of active searching, and is assumed to be fixed for all males.

- if $\text{Com}_{k,i} = 0 \rightarrow$

$$P[pref_{s,k}(t_n) = j'] = \frac{\pi_{s,j,j'}}{\sum_{j':AVAILSET_{s,j'}(t_n)\neq\emptyset} \pi_{s,j,j'}} \qquad if\ AVAILSET_{s,j'}(t_n)\neq\emptyset$$

$$P[I_{k,k',i,i'}=1] = \frac{1}{|AVAILSET_{s',j'}(t_n)|} \quad if\ k' \varepsilon AVAILSET_{s',j'}(t_n) \wedge i'=cumrel_{s',k'}(t_n)+1 \wedge j'=pref_{s,k}(t_n)$$

The probability that a (non-commercial) partner is searched in a certain age-group is equal to the initial preference structure, conditional on the availability of partners in the possible age groups. Note that this probability is equivalent to the probability that results from a procedure in which, based on the initial preference structure, age-groups are drawn until an age group is found in which potential partners are available. One individual of the appropriate age group is uniformly randomly selected as a partner.

- if $Com_{k,i} = 1 \rightarrow$

$$P[I_{k,k',i,i'} = 1] = \frac{1}{|CSWSET(t_n)|} \quad if\ k' \ \varepsilon\ CSWSET \wedge i' = cumrel_{1,k'}(t_n) + 1$$

The probability that a certain CSW is chosen by a male searching for a commercial contact is equal to the inverse of the number of CSWs.

- $P[Potcsw_k = 1] = \phi$

The probability that a female becomes CSW is determined at birth and is fixed for all females.

- $Untilcsw_k \sim Exponential(\gamma)$

The time between starting sexual activity and becoming CSW is determined at birth, and is exponentially distributed.

- $Durcsw_k \sim Exponential(\rho)$

The duration of the working period of a CSW is determined at the starting of the CSW career, and is exponentially distributed.

Used probability density functions

$$f(x) \sim Weibull(\theta,k) : f(x) = \frac{k}{\theta^k} \cdot x^{k-1} \cdot e^{-(\frac{x}{\theta})^k} , \ mean = \theta\Gamma(1 + \frac{1}{k})$$

$$f(y) \sim Exponential(\beta) : f(y) = \frac{1}{\beta} \cdot e^{-\frac{y}{\beta}} , \ mean = \beta$$

4 Method for quantification of the sexual behavior submodel

4.1 Introduction

Our aim is to find a quantification for the model such that the sexual behavior of the hypothetical individuals in the simulation model closely resembles the actual sexual behavior in Nairobi, as indicated in Tables 2.1, 2.2, 2.3 and 2.4 and Figure 2.1. We will focus our parameter estimation efforts on getting a good fit with respect to the number of recent sexual partners and number of lifetime partners reported (Table 2.2 & 2.4), by minimizing the following expression with respect to parameter set p:

$$F(p) = \sum_h \sum_s \sum_i \sum_j w_{h,s,i,j}(\mathrm{E}[model_{h,s,i,j}(p)] - data_{h,s,i,j})^2 \qquad (4.1)$$

in which:
- h = index over different time horizons of reported data (6 months, lifetime)
- s = index over the sexes
- i = index over age classes
- j = index over levels of sexual activity of individuals (model outcome)
- $w_{h,s,i,j}$ = weight for sex s, age class i and activity class j and time horizon h
- $model_{h,s,i,j}(p)$= model outcome in terms of the number of individuals of sex s and age class i that have activity level j at a time horizon h for a given parameter set p
- $data_{h,s,i,j}$ = number of individuals of sex s and age class i that have activity level j for time horizon h, if the probability of this follows the figures in respectively Table 2.2 and Table 2.4

For the weight function $w_{h,s,i,j}$ we will use $1/data_{h,s,i,j}$. For any given horizon, sex and age class the probability of falling into a certain sexual activity class is multinomially distributed, hence it can be shown that:

$$\sum_j^{N_j} \frac{(model_{h,s,i,j}(p) - data_{h,s,i,j})^2}{data_{h,s,i,j}} \sim \chi^2(N_j - 1) \qquad (4.2)$$

With this statistic (4.2) one is able to test whether for a given horizon, sex and age class the distribution of numbers of partners is equal to the distribution in the data set. Unfortunately, extension of this statistic by summing over all the horizons, sexes and age classes will not deliver another Chi-squared distributed goodness of fit statistic. Only the sum of independent Chi-squared variables is also Chi-squared distributed, and it is very unlikely that the outcomes for different horizons, sexes and age classes are independent. Because of the interdependencies mentioned, also other statistical measurements for goodness-of-fit are hard to obtain.

To determine the term $\mathrm{E}[model_{h,s,i,j}]$ for a given parameter set p a complete run of the sexual behavior submodel has to be performed, which takes considerable running time. Because of the (stochastic) microsimulation technique used such a model run produces not the exact value of $\mathrm{E}[model_{h,s,i,j}(p)]$, but only an estimation. The strategy of simply repeating the model experiment for the same set of parameters p until the noise has almost disappeared will cost too much running time, especially if in a optimization procedure a whole range of parameter sets has to be checked. Free et al (1987) showed that one of the few effective methods for optimization of computationally expensive noisy functions is the Response Surface Method (RSM).

4.2 Response Surface Method

In the Response Surface Method (RSM) one tries to reduce the number of evaluations (in our case: running of the sexual behavior submodel) by approximating the original function F(p) by a simple function over a subregion of the total domain. After optimization of this approximating function, the search for the optimum of the original function is continued near this optimum by constructing and minimizing a new approximating function. This process continues (in principle) until convergence to the actual optimum occurs.

For determination of the approximation function the original function F(p) is estimated at several design points (in our case: combinations of values for the parameters that have to be estimated) and the coefficients of the approximating function are estimated by applying a least squares method to the estimations of F(p). In this least squares estimation two types of errors play a role: 1) systematic, or bias, errors that indicate the difference between the original function F(p) and the approximating function; 2) random error caused by the fact that a model run does not deliver the exact value of F(p), but only an estimation.

In all publications about applications of the RSM we have seen, the impact of random errors is of minor importance or even totally absent, see for instance Toropov (1989), Vanderplaats (1989) and Toropov et al (1993). Because of the stochastic nature of our microsimulation model, the impact of random errors is considerable and might outweigh the systematic errors. Hence, it is unlikely that RSM implementation strategies from literature are suitable for our application, and we will have to find our own implementation strategies. In the following, several decisions that have to be made for the actual implementation of the RSM are discussed.

Approximation region

We will not fit an approximation function to the function F(p) over the whole permitted domain (in our case: for six parameters all positive numbers, and for the δ-parameter all values in [0,1]), because this would probably require a very complex function. For a simple function the systematic error between F(p) and the approximation function will be too large. In general, the smaller the size of the approximation region, the smaller the systematic error will be. Therefore we will limit the size of the approximation region. However, in our application the size of this approximation region should also not be too small, because otherwise the impact of the variables on the function value F(p) is totally hidden by random noise.

Ideally, the approximation region should contain the optimum. However, because the actual location of the optimum is unknown, this is not a useful requirement. Instead, it is important that the approximation region can move in the different steps of the optimization process. The approximation region will act as a trust region: the design points are not only chosen in this region, but the approximation function is also optimized within the bounds of the approximation region. After optimization a new approximation region is defined around the minimum of the approximation function and the procedure is repeated.

In Figure 4.1 the procedure is graphically illustrated for a two-parameter case. Point 1 indicates the center of the initial approximation region, which is indicated by a box. In this example, the other design points are located on the border of the approximation region. The function outcomes in the design points are used to fit an approximation function, which has a minimum within the chosen approximation region in the point indicated with a star. As can be seen in Figure 4.1, around this point 2 a new approximation region is established.

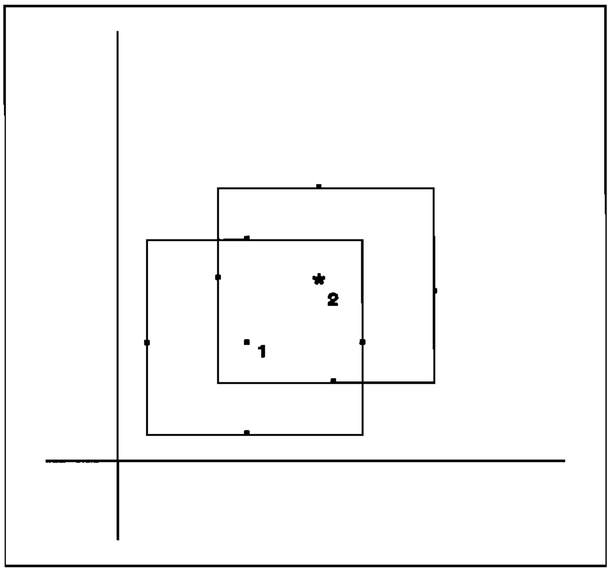

Fig. 4.1: Graphical representation of the Response Surface Method

It is hard to determine beforehand whether the size of this approximation region should stay constant or decrease in time according to certain rules. Therefore, we will initially fix the size of the approximation region, and later on pay more attention to the topic of the size of the approximation region.

Design points

Another point of consideration is how many design points are used and how they are located in the approximation region. Straightforwardly, the number of, linearly independent, design points should be at least equal to the number of coefficients that have to be estimated. Using more design points is of course possible and increases the accuracy of the estimated coefficients. On the other hand this also increases the running time per design. It is hard to make a clear decision on the number of design points needed beforehand, but because of the presence of noise the number of design points must not be too low.

With respect to the location of design points, it is desirable that the design is orthogonal, because this improves the reliability of the least squares estimation. Examples of orthogonal designs are two-level factorial designs (also called cubes), with terms of the form $(x_1, x_2,..x_n) = (\pm a, \pm a,....,\pm a)$, or star designs with $(x_1, x_2,x_3,...x_n) = (\pm b, 0, 0,..,0), (0, \pm b, 0,...,0),...., (0, 0, 0,..,\pm b)$, where the center point of the design is (0,0,..,0). These cube and star designs can be supplemented with one or more center points. For a careful analysis of optimal location of design points we refer to Box and Draper (1986).

Initially, we have used in our RSM application a combination of a factorial and a star design, while we evaluated the center point four times. Furthermore, we have reused design points of earlier approximation iterations if they are located within the approximation region of a new iteration. In Figure 4.1 this would mean that in the second approximation iteration we would reuse the center point and the 2 points located on the upper, respectively right bound of the initial approximation region.

Scaling and transformation

Scaling around the center point of the design has briefly been mentioned in the previous section. This scaling does not only relate to translation of the center point to position (0,0,...0), but also to multiplication such that for every variable the borders of the approximation region are taken as -1 and 1. Scaling is extremely important, because this guarantees orthogonality in the estimation procedure, which enhances the precision of estimates. Furthermore, the optimum found in an iteration might depend on scaling. For instance, this is the case if one would estimate a first order polynomial and apply a steepest descent method to determine the location of the optimum.

Apart from scaling of the variables, also transformation of the variables or the function F(p) might be useful. We have transformed the function F(p) by taking the square root. The minimum of F(p) and the square root of F(p) is the same, and intuitively we thought that it would be more likely that parameters would influence $\sqrt{F(p)}$ in a polynomial way than F(p). Later on we will test whether this transformation is effective.

Class of approximation functions

To construct an approximating function one first has to define a class of functions that will be considered. A main criterium for the choice of the class of approximating functions is that this class of functions can be optimized relatively easily; it is not useful to replace a difficult optimization problem by another difficult optimization function. In literature often the class of first or second order polynomials are considered as approximating function. However, generally for the class of second order polynomials convexity, and thereby optimality, is difficult to prove. This is especially the case when the number of variables is considerable. The subclass of second order separable polynomials does not have this drawback. As will be shown in the following subsection, one can easily determine the optimum of a separable second order polynomial with bounded variables analytically. Another advantage of separable second order polynomials is that the minimum number of designs points needed is linear in the number of variables taken into account, instead of quadratic as for a full second order polynomial; for a separable second order polynomial of k variables 2k+1 coefficients have to be estimated, instead of $\frac{1}{2}k^2+1\frac{1}{2}k+1$ for a full second-order polynomial, so only 2k+1 linearly independent design points are minimally needed.

Because of the reasons mentioned above we have decided to use the class of separable second order polynomials for approximation of the original function.

Method for optimization of the approximation function

As already indicated, optimization problems for separable second order polynomials with bounded variables can easily be optimized analytically. To prove this, note that:

$$\begin{aligned} &Min \;\; a_0 + \sum_{j=1}^{N} a_{1,j} \cdot X_j + a_{2,j} \cdot X_j^2 \qquad s.t. \; l_j \le X_j \le u_j \;\forall j \;\Leftrightarrow \\ &Min \;\forall j : a_{1,j} \cdot X_j + a_{2,j} \cdot X_j^2 \qquad s.t \; l_j \le X_j \le u_j \end{aligned} \tag{4.3}$$

With the following optima of the subproblems for a single variable X_j:

$$
\begin{aligned}
&\text{if } a_{2,j} > 0 \wedge l_j \le -\frac{a_{1,j}}{2a_{2,j}} \le u_j && \Rightarrow \quad X_j^* = -\frac{a_{1,j}}{2a_{2,j}} \\
&\qquad\qquad \wedge -\frac{a_{1,j}}{2a_{2,j}} < l_j && \Rightarrow \quad X_j^* = l_j \\
&\qquad\qquad \wedge -\frac{a_{1,j}}{2a_{2,j}} > u_j && \Rightarrow \quad X_j^* = u_j \\
&\text{if } a_{2,j} = 0 \wedge \quad a_{1,j} > 0 && \Rightarrow \quad X_j^* = l_j \\
&\qquad\qquad \wedge \quad a_{1,j} < 0 && \Rightarrow \quad X_j^* = u_j \\
&\qquad\qquad \wedge \quad a_{1,j} = 0 && \Rightarrow \quad X_j^* = [l_j, u_j] \\
&\text{if } a_{2,j} < 0 \wedge -\frac{a_{1,j}}{2a_{2,j}} > \frac{l_j + u_j}{2} && \Rightarrow \quad X_j^* = l_j \\
&\qquad\qquad \wedge -\frac{a_{1,j}}{2a_{2,j}} < \frac{l_j + u_j}{2} && \Rightarrow \quad X_j^* = u_j \\
&\qquad\qquad \wedge -\frac{a_{1,j}}{2a_{2,j}} = \frac{l_j + u_j}{2} && \Rightarrow \quad x_j^* = \{l_j, u_j\}
\end{aligned}
\tag{4.4}
$$

Stopping rules

Because initially we did not have a clear idea about a good stopping criterium for the RSM, we have decided to start with the simple rule that after a certain number of function calls, corresponding with a number of designs completed, the algorithms stopped. Later on we will look at more sophisticated stopping rules.

5 Results

5.1 Introduction

In this section the results of the parameter estimation process of the sexual behavior submodel are presented. However, before the actual parameter estimation process has been performed, several preparatory steps have been undertaken, which are also described in this section. The first preparatory step has been the generating of a test data set by running the model several times with all parameters set on a fixed level. The advantage of a generated test data set in comparison with a real data set is that the parameters by which the data set is generated are known. This makes it, amongst others, possible to examine to what extent a parameter estimation procedure is able to estimate good values. In Section 5.2 the generating of the test data set is discussed in greater depth.

As can be seen in Section 5.3, the next preparatory step has been the performance of a sensitivity analysis at which the test data set has been used. The aim of this sensitivity analysis is to identify how the model reacts to changes in parameters, which enhances the understanding of the model. As a further preparatory step, the test data set has been used to test the performance of the Response Surface Method. According to the results of these tests the implementation of the RSM has been adapted. In Section 5.4 the testing of the RSM is described. Finally, in Section 5.5 the application of the RSM for the actual parameter estimation of the sexual behavior submodel is described.

In all analyses in this section the fixed parameters of the sexual behavior submodel have been given the following values:

- length of age-classes: $\alpha_0 = 15$; $\alpha_j = 5$, j=1,2,..7; $\alpha_8 = 40$; which implies that we have got the following age classes: 0-15, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, over 50 years.
- minimum age-group in which males can start a steady relationship: j0 = 3, corresponding with age 25.
- preference matrix $\pi_{sjj'}$: based on the data of Figure 2.1, a preliminary preference table has been established. For simplicity, this table does not depend on the age at which a relationship is started. Note that for some age-classes and sexes not all preferred younger and/or older age classes of partners exist (for instance, for males of age class 2, there are no partners of age class -1 and -2). These non-existing age classes are treated the same way as classes without available persons.

Table 5.1: Age difference between partners at the start of a relationship

Female 1 age class older	Partners of same age class	Female 1 age class younger	Female 2 age classes younger	Female 3 age classes younger	Female 4 age classes younger
5%	15%	40%	20%	15%	5%

- delay if no suitable partners are available $\Delta = 0.5$ years (arbitrarily)
- probability that a female becomes a CSW $\phi = 0.05$
- mean time between starting sexual activity and becoming CSW $\gamma = 2$ years
- mean duration of working period as CSW $\rho = 8$ years
- parameters of the distribution of availability for first sexual relationship for respectively males and females: $\theta_0 = 4.3$, $k_0 = 1.7$, $x_0 = 12.5$; $\theta_1 = 4.5$, $k_1 = 1.2$, $x_1 = 13.5$; these figures are estimates based on Table 2.1.

A problem with the data in Table 2.4, which indicates the number of lifetime sexual partners and will be used to estimate parameters, is that the size of the missing answer category is quite extensive. We have, arbitrarily, assumed that the answers on this question of non-respondents do not systematically differ from the answers of respondents, and redistributed the non-response according to this assumption over the different sexual activity classes.

5.2 Generating the test data set

As indicated in the introduction, we have generated a test data set that, amongst others, can be used for sensitivity analysis and to test whether the RSM works effectively and efficiently at our application. It will be clear that results of a sensitivity analysis and of RSM testing are only relevant if the test data set and the real data set are more or less similar, at least in a qualitative sense. Therefore, this test data set has not been generated by running the model for a randomly chosen set of parameters. Instead, we have applied a first implementation of the RSM to the actual data set to determine reasonable values for the free parameters, and the attained values of the free parameters have been used to generate the test data set.

First adaptation: inclusion of significance criteria in the RSM

One adaptation to the RSM was already made in this phase of finding reasonable parameter values to generate a test data set, namely the inclusion of a criterium on the significance of the parameters. If for a certain variable X_j both the values of the coefficients $a_{1,j}$ and $a_{2,j}$ are 'very close' to 0, the optimal value X_j^* can get any value, because the estimated coefficients $\hat{a}_{1,j}$ and $\hat{a}_{2,j}$ can by chance be positive or negative and $\hat{a}_{1,j}/\hat{a}_{2,j}$ can both have a very small or very large absolute value. Therefore, we include the criterium that a new design is only moved to the minimum of the approximation function if at least one of the estimated coefficients has an absolute t-value of 1.3, corresponding with a very loose significance level of 20%. This criterium prevented indeed boun-c-ing of values of variable, while it was loose enough to have movements into the right directions.

If only $a_{2,j}$ is very close to 0, the term $\hat{a}_{1,j}/\hat{a}_{2,j}$ will in general be large. However, this is no problem because both for positive and negative values of $\hat{a}_{2,j}$ the movements will always be in the same direction as in the case that $a_{2,j}$ is exactly 0. Furthermore, the step size will be limited by X_j's upper and lower bounds of the approximation region, just as in the case that $a_{2,j}$ is exactly 0. If only $a_{1,j}$ is very close to 0, and $a_{2,j} > 0$ the optimal value of X_j will be close to zero so there can not be large movements into wrong directions. However, if $a_{1,j}$ is very close to 0 and $a_{2,j} < 0$, quite arbitrarily the optimum will be on the lowerbound or the upperbound of the approximation region. Unfortunately, we have neglected to implement significance criteria for this case.

Results of generating test data set

The initial implementation of the RSM has been run from different starting points to find good parameters to generate a test data set. Independent of the starting point, it was quite hard to get real close to the data represented in the Tables 2.2 and Table 2.4. One of the main reasons for this is that with the model the large gap in answers between males and females in these tables is hard to generate. Furthermore, the model assumptions are crude on some points, for instance by ignoring a direct effect of age on sexual behavior. In particular, these (over)simplifications make it difficult to get a good fit with respect to both the number of recent and lifetime sexual partners. Another explanation for the difficulties in getting real good results might be that the RSM implementation used is not able to deal with our parameter estimation problem, and needs further adaptations.

With the initial implementation of RSM the following estimates of the free parameters have been found:

- mean time until becoming available for males in a relationship $\mu_{1,0} = 26$
- mean time until becoming available for females in a relationship $\mu_{1,1} = 43$
- mean time until becoming available at the end of last relationship $\mu_2 = 1.5$
- mean time between becoming available and going to search actively $\lambda = 0.8$
- mean duration of casual relationships $\beta_0 = 4.5$
- mean duration of steady relationships $\beta_1 = 20$
- probability that a relationship initiated by a male is in fact a CSW contact $\delta = 0.95$

With these parameters 100 STDSIM runs of 200 year have been made, starting with a start population of 500 individuals distributed over the different age classes. In every run tables of the same format as Table 2.2 and Table 2.4 are generated, and the final test data set is generated by taking the average over the tables of the 100 runs, see Table 5.2 and Table 5.3.

Table 5.2: Percent distribution of women an men by number of sexual partners in the six months preceding the model survey, according to age-class; results of 100 runs of 200 years, end population over 100 runs: 287392 women and 287982 men

Age	Females				Males			
	None	1 partn.	2 partn.	3+ partn.	None	1 partn.	2 partn.	3+ partn.
15-19	60.5	37.2	1.1	1.2	66.1	27.9	5.0	1.1
20-24	31.8	63.0	2.8	2.4	38.1	53.0	7.3	1.6
25-29	22.5	71.5	4.1	2.0	26.2	64.7	7.7	1.4
30-34	16.0	76.1	6.3	1.6	14.8	75.1	8.9	1.2
35-39	12.1	78.7	8.1	1.2	7.8	77.2	13.3	1.6
40-44	10.4	78.5	9.9	1.1	5.4	75.5	16.7	2.3
45-49	11.4	76.4	11.1	1.2	4.5	71.7	19.9	3.8
50-54	14.1	74.2	10.8	0.9	5.1	69.6	20.8	4.4
Overall*	26.1	66.8	5.6	1.6	16.7	68.8	12.5	2.1

* *To guarantee comparability with the overall figures in Table 2.2, for males the age-group 15-19, and for females the age-group 50-54 are excluded from the overall category.*

Table 5.3: Percent distribution of women and men by number of sexual partners in their life, according to age-class; results of 100 runs of 200 years with an end population over 100 runs of 287392 women and 287982 men

Age	Females					Males				
	None	1	2-3	4-5	6+	None	1	2-3	4-5	6+
15-19	53.9	37.0	7.7	0.1	1.3	47.5	13.8	20.1	11.0	7.5
20-24	15.7	42.8	36.6	1.2	3.7	6.5	8.4	20.3	21.0	43.8
25-29	3.8	25.8	57.6	8.2	4.6	0.3	2.8	10.2	14.9	71.8
30-34	0.8	13.2	60.5	19.4	6.0	0.0	0.6	5.3	10.1	84.0
35-39	0.2	7.2	53.4	31.2	8.0	0.0	0.2	3.4	8.3	88.1
40-44	0.0	4.2	45.1	38.5	12.1	0.0	0.0	2.2	6.8	90.9
45-49	0.0	2.8	37.2	43.7	16.4	0.0	0.0	1.5	5.9	92.6
50-54	0.0	1.5	30.0	46.4	22.1	0.0	0.0	1.1	5.0	93.9
Overall*	13.1	21.5	41.3	17.5	6.6	1.3	2.1	7.4	11.3	77.9

**To guarantee comparability with the overall figures in Table 2.4, for males the age-group 15-19, and for females the age-group 50-54 are excluded from the overall category.*

To check to what extent the results in Table 5.2 and 5.3 depend on the choice of the simulation horizon we have generated similar tables for the average of 100 runs with a time horizon of 25, 50, 75, 100, 125, 150 and 175 years. For the 3 shortest horizons the results with respect to number of lifetime partners are too low, especially for the older part of the population. This can be explained by noting that the initial population starts without a sexual history. Furthermore, the demographic structure of the population slightly differs for different choices of the simulation horizon, and this influences sexual activity because of the mixing function used.

However, the results in terms of numbers of partners of runs of 100, 125, 150 and 150 and 175 years do not systematically deviate from the results of runs of 200 years. This can be verified by looking at the appendix tables A.1 and A.2, which show the average outcomes of 100 runs with a simulation horizon of 100 years. Note that the deviances between runs are larger for shorter simulation horizons, because of smaller population sizes (e.g. runs of 100 year are a factor 4 smaller than runs of 200 years).

If one compares Table 5.2 with Table 2.2, one sees that for older females the group with 2 recent partners is too large, while for males, especially for young males, the numbers of recent partners are too low. Comparison of lifetime number of partners in the real and test data set shows that in Table 5.3 both for older males and females the number of lifetime partners are too high.

Despite these discrepancies between the real data set and the model outcome, to our opinion this data set is still good enough to test the impact of different parameters in the model and improve the implementation of the RSM. Hence, we have decided to wait with making adjustments to the sexual behavior submodel STDSIM and/or the optimization procedure, and have first made a further analysis of the model and the optimization procedure using the test data set of Table 5.2 and Table 5.3.

5.3 Sensitivity Analysis

To test how the model reacts to changes in parameter values, we have performed a sensitivity analysis. In this analysis we have changed the parameter values of the test data set one by one with first -50% of the original value, and then by +50%, the latter except for parameter δ which has to be in the interval [0, 1]. For all 13 parameter sets we have made 100 runs of 100 years. Note that this simulation run is shorter than the run used for generating the test data set to reduce running time. As mentioned above, a horizon of 100 years does not produce systemic bias.

In Table 5.4 the results of the sensitivity analysis are summarized. For every parameter is separately for males and females indicated what the main effect is on the number of recent partners and number of lifetime partners, compared with the test data set. For this indication we have used '+' and '-' signs, where a '+' indicates that increasing the value of the parameter leads to an increase in the number of partners for a given horizon and sex, and a '-' indicates a decrease in number of partners. If the effect is negligible small, a 0 is used. The number of signs varies to indicate the strength of the effect, see the legend of Table 5.4 for exact criteria. The percentage mentioned in this legend refers to the shift between activity classes for the age groups that are most affected in case of a 50% increase. Furthermore, in Table 5.4 it is also indicated which age groups are affected most by a change in a specific parameter.

Table 5.4: Results of sensitivity analysis

Parameter	Number of recent partners		Number of lifetime partners	
	Females	Males	Females	Males
$\mu_{1,0}$	– primarily effect on 25+	–– primarily effect on 30+	– primarily effect on 25+	––– pronounced effect in 30+
$\mu_{1,1}$	–– primarily effect on 30+	– primarily effect on 20-35	– Overall	0
μ_2	–– Overall	–– Overall	–– Overall	– Overall
λ	–– primarily effect on 15-30	––– primarily effect on 15-30	–– Overall	–––– primarily effect on 15-35
β_0	+ primarily effect on 15-25	++ primarily effect on 20-30	––– Overall	–– primarily effect on 25-40
β_1	+ primarily effect on 35+	+ primarily effect on 35+	–– primarily effect on 35+	– primarily effect on 35+
δ	––– primarily effect on 15-30	––/++ primarily effect on 15-40	––– Overall	++++ Overall

+	highest pos. impact on an age-class 1%-3%	–	highest neg. impact on an age-class 1%-3%
++	highest pos. impact on an age-class 3%-6%	––	highest neg. impact on an age-class 3%-6%
+++	highest pos. impact on an age-class 6%-10%	–––	highest neg. impact on an age-class 6%-10%
++++	highest pos. impact on an age-class > 10%	––––	highest neg. impact on an age-class > 10%

Discussion sensitivity analysis

An increase in the parameter $\mu_{1,0}$ means that males within a relationship become later available for another relationship. Therefore, the number of males with more than one relationship at a time decreases. Generally, this has a negative impact on both number of lifetime and recent partners of males and females. However, not indicated in Table 5.4 is that for young males a higher value of $\mu_{1,0}$ leads to a decrease of the number with 0 partners. An explanation for this is that males without a relationship have less competition of males with a relationship. Furthermore, the occurrence of this effect in especially young males possibly relates to the fact that old males are relatively more preferred by females; when $\mu_{1,0}$ increases, older males are less available and young males are selected instead.

The impact of an increase in the parameter $\mu_{1,1}$, by which females within a relationship become later available for another relationship, is quite analogous to the impact of $\mu_{1,0}$. Generally, we see the same negative impact on number of partners. However, the effect on lifetime number of partners of males is negligible. An explanation is that this figure is dominantly determined by CSW contacts. Hence changes in parameters of 'normal' women have relatively less impact. Not indicated in Table 5.4 is that increasing $\mu_{1,1}$ leads to a decrease in the number of females above 35 years with 0 partners. This not so popular age-group benefits from the fact that females of a more attractive age-group with a relationship are less available.

An increase in the parameter μ_2 delays the time of becoming available after the end of a person's last relationship. In all age groups we see at increasing μ_2 a rising number of persons with 0 recent partners: these will be most persons whose relationship have ended but who are not yet available for another relationship. The slower 'turnover' of relationships also affects both for males and females the total lifetime number of partners. For males this effect is less pronounced than for females, because males still end up with high number of lifetime partners anyway.

An increase in the parameter λ leads to a longer time between becoming available and starting active searching, which has a negative impact on number of partners. For males an increase in λ has a much stronger impact than an increase in μ_2. The main reason for this is that males who have visited a CSW are immediately available for a new relationship but have to wait some time before they will search actively for a new relationship themselves. So for these CSW visitors the time until their next relationship is directly influenced by λ but not by μ_2, and there is only the indirect effect of females who become later available and are also going to search for a male partner somewhat later. For females changes in λ have about the same impact as changes in μ_2, although for λ younger females are more affected than older females. A possible explanation for this age-effect is that females who are searching for relatively scarce older males can more easily find a partner at higher values of λ. For older females this positive effect may partly outweigh the negative effect of going to search later.

An increase in the parameter β_0 leads to an average longer duration of casual relationships. With respect to number of recent partners, changes in this parameter affect primarily younger age-groups. This is not surprising, because casual relationships start only if the male partner involved is under 25 years. The positive impact of increasing β_0 on recent partners is caused by the fact that persons are longer in a relationship, and therefore the number of persons with 0 partners decreases. The effect that longer relationships lead to a lower rate of partner change is not noticeable at number of recent partners, because the basic duration of relationships and time until availability and searching are so large that not many persons will have more than 1 non-commercial relationship in a half year. However, on lifetime number of partners the effect of lower rate of partner change is noticeable, especially for females.

An increase in the parameter β_1 leads to an average longer duration of steady relationships, which lowers the number of persons with 0 recent partners. Changes in this parameter primarily affect older age groups, because individuals that start a steady relationship are older than persons that start a casual relationship. Furthermore, the basic average duration of steady relationships is very long which also increases the age at which changes in these parameters are noticeable. Note that the effect of changes in β_1 on number of lifetime partners are smaller than the effects of changes in β_0. A possible explanation for this is, that steady relationships last so long, that even if the average duration changes with 50%, persons cannot have more than 2 or 3 steady relationships in their life. Hence changes in number of steady partners have only limited effects.

For evaluating the effect of an increase in the parameter δ, which indicates the proportion of relationships initiated by males that are in fact CSW contacts, we can only look at the difference between δ=0.475 and δ=0.95. Of all variables this variable has the largest impact on number of partners of males and females. Looking first at males, we see that a higher value of δ leads to a decrease in the number of males with 1 recent partner, while there is an increase in the number of males with 0 and 2 or 3+ partners. The explanation for this is that when males have more commercial contacts, the rate of partner change is higher which leads to a rise in the 2 and 3+ category. On the other hand, it is also more likely that some males have no sexual relationship at all in a half year, because the time until searching actively is quite long. The effect of increasing δ on lifetime number of partners of males is definitively positive, because after a CSW contact males are immediately available for another contact, while this is not the case after the start of a non-commercial relationship. For females, the effect of increasing δ is negative for both number of recent and lifetime partners. If males visit CSWs more often, it takes longer before they are going to search for non-commercial contacts. This lowers the number of partners of females who are not involved in commercial sex. The increase of sexual activity of CSWs is not reflected in Table 5.4; all CSWs are already in the highest activity level, and the size of the group of CSWs is not influenced by the demand for their services.

Note that no parameter has such effects that changes in a single parameter diminishes the gap between the test data set and the real data set with respect to both recent and lifetime number of partners of males and females. This is even not the case for δ, that is almost on it's natural boundary of 1. Increasing δ leads among others also to an increase in the number of lifetime partners of males, and this is undesirable.

5.4 Testing the Response Surface Method

As indicated in Section 5.1 the test data set has also been used to ascertain whether the RSM works effectively and efficiently at our application, and how RSM can be implemented best. In the first step of the testing procedure we have looked at the one parameter estimation case, in which all parameters except one are fixed at the value used to generate the test data set. The only free parameter is given a very low or a very high start value, and we can then evaluated whether the RSM approximately finds the parameter value used to create the test data set. This one parameter case is subsequently performed for all parameters.

According to the results of the one parameter estimation case, we have adapted the RSM. Next we have turned to testing of the RSM in the multi parameter estimation case, in which we try to estimate all 7 free parameters simultaneously. The results of these multi parameter estimation case are also evaluated, after which we have decided whether further adaptations of the RSM are necessary.

Initial implementation of the Response Surface Method for the one parameter estimation case

We have started the testing phase for the one parameter case with the following implementation of the RSM:

- fixed approximation regions
- scaling of the design points between -1 and 1
- transformation of F(p) to $\sqrt{F(p)}$
- designs of 8 points consisting of 4 center points (i.e. 4 evaluations at position 0), and 4 other points at position -1, $-1/\sqrt{2}$, 1 and $1/\sqrt{2}$
- approximation functions from the class of separable polynomial functions function
- runs of 200 years with the STDSIM model, with a starting population of 500 and an end population of approximately 5000
- if both estimated coefficients of X_j and X_j^2 have a t-value in the interval [-1.3, 1.3] the center of the next design does not move, but stays on the same location
- stopping after the performance of 15 designs (120 function calls)

Furthermore, the design is translated if one of the parameters would violate his natural restrictions (greater than 0 for first 6 parameters, being in the interval [0,1] for δ)

Second adaptation: fixing demography

Unfortunately, at first the RSM was not able to solve the one parameter estimation case by finding subsequently the right parameter values that have been used to generate the test data set. One reason for this is the tremendous variance in the value of F(p).

Careful analysis of the results showed that on average the value of F(p) was higher for larger population sizes. At first this seemed strange; normally, Chi-squared goodness-of-fit statistics are independent of population sizes. However, this is only the case if the model is properly specified, such that E[$\text{model}_{h,s,i,j}$(p)]/(number of persons of sex s and age i) is multinomially distributed with a mean of $\text{data}_{h,s,i,j}$/(number of persons of sex s and age i). However, if the model specification is incorrect in the sense that the average model outcome is unequal to the distribution of the data set one gets that the larger the population size, the higher the value of F(p) will be on average. This does not only causes extra variance, but also leads to movements into the direction of parameter values that by chance have been accompanied with a small population size. In our application the final population size varies considerable, differences between runs of 10%-20% are quite common, and therefore this bias is quite large.

To get rid of changes in population size, we have used in our simulation model a fixed series of random numbers for the demographic processes, while the random numbers for the sexual behavior processes still vary. After this adaptation the effectiveness of the RSM improved. A risk of fixing demography in the actual parameter estimation process is that the parameters which are found in the optimization procedure fit well for a specific series of random numbers, and not for other series of random numbers. However, because in this phase of the quantification process we focus on testing the RSM and not on determining the actual parameters this is not yet problematic.

Results of testing initial implementation RSM for the one parameter estimation case

After the above mentioned adjustment to the simulation model, we have applied the RSM again to test whether the method can subsequently find the parameters values used to generate the test data set. In Table 5.5 the results of these tests of the RSM are indicated. In the first column the original value of the parameter that is freed is indicated. In the second and third column the results of the RSM procedure for respectively a much too low and much too high starting point of the freed parameter are represented. Because the parameter δ is almost on it's natural upperbound of 1, we have for this parameter not started a search from a too high starting point. Note that the mentioned sizes of the initial approximation region determines, amongst others, how fast one can move through the parameter space.

Table 5.5: Results of testing initial implementation RSM for the one parameter estimation case

Parameter	Low starting point			High starting point		
	Start value	Initial region	End value	Start value	Initial region	End value
$\mu_{1,0}^*$: 26	5	1-9	29.2[†]	45	41-49	24.9
$\mu_{1,1}^*$: 43	15	11-19	36.5[†]	75	71-79	76[†]
μ_2^*: 1.5	0.5	0.3-0.7	1.49	3	2.8-3.2	1.41
λ^*: 0.8	0.25	0.05-0.45	0.84	2	1.8-2.2	0.84
β_0^*: 4.5	1.5	1.1-1.9	4.3	8	7.6-8.4	6.7[†]
β_1^*: 20	5	2-8	18.7	36	32-40	25.6[†]
δ^*: 0.95	0.6	0.55-0.65	0.94	-	-	-

[†]*End parameter value more than 10% away from the actual parameter value*

The figures in Table 5.5 show that, for the given design and length of a run, it is very hard to find the exact value of a parameter with the RSM. However, in 8 of 13 cases it was possible to get within 10% of the actual optimal value, even while the starting points of the RSM were far away from the actual optimum. For sensitive parameters, like for instance λ and δ, the RSM procedure works best because the impact of noise is relatively small. The parameter $\mu_{1,1}$ that has not got much impact on the simulation outcome, is the most difficult to estimate. If the starting value of this parameter is too high, the function $\sqrt{F}(\mu_{1,1})$ is so flat and the impact of noise is so large that no steps into the right direction are made at all.

The impact of noise can be reduced by generating more design points or making longer and/or bigger runs, but this will cost also more running time. A solution which does not increases the running time and works quite well, is enlarging the size of the approximation region. For the parameters $\mu_{1,0}$ and $\mu_{1,1}$ the effectiveness increases for larger approximation regions.

In general, it appeared to be harder to find the optimal value in cases that the initial starting value was taken too high than in cases that it was taken too low. The explanation for this is that in our application for most parameters the impact of absolute changes in parameter values is relatively larger at lower parameter values. For instance the effect of decrease or increase in $\mu_{1,0}$ of 4 is much larger when the basic value is 5 than when the basic value is 45.

In this specific case, the only exception on the rule of better performance for lower starting points has been parameter $\mu_{1,0}$. Analysis of the results in different parameter estimation steps showed that the value of this parameter has been long close to the optimal value. However, just before the end of the search due to random noise a switch has been made to a too high parameter value.

Third adaptation: flexible size of approximation region

One way to reduce the impact of such random fluctuations in the end stage of the RSM is to use smaller approximation regions. However, this might hamper the effectiveness of the RSM in it's initial phase; for some parameters the chosen approximation regions were already too small. The sketched dilemma can be resolved by making the size of the approximation region flexible: start with a relatively large approximation region, and decrease the size of the approximation region during the optimization procedure according to certain criteria. Such a strategy has been implemented in Toropov (1989) and Toropov et al (1993).

One of the criteria used in these papers is that the size of the approximation region is reduced after the detection of an interior optimum. However, because of the relatively large impact of noise in our application quite regularly an interior optimum is found by chance. Direct implementation of the mentioned rule for decreasing the size of the approximation region leads in our application to a too fast decreasing size of the approximation region. To prevent this, the size of the approximation region is only reduced when two times in succession an interior optimum has been found. At the implementation of this rule we have doubled the size of the initial approximation region.

Fourth adaptation: stopping rule

The stopping rule applied in the initial implementation of the RSM, simply stopping after a fixed number of runs, is very crude and can be quite inefficient. The combinations of parameter and starting point for which the RSM worked quite well had reached after about 8 designs already a good parameter value. However, simply stopping after a smaller number of iterations would have made the results for the other parameter/starting point combinations even worse.

In Toropov (1989) and Toropov et al (1993) the stopping rule is used that after a certain number of reductions of the approximation region size the RSM is stopped. In our application this rule guarantees that at the moment of stopping 2 interior optima have subsequently been found in a relatively small region. This means that the minimum of the approximation function found is relatively stable. To prevent an infinite running time in case the model does not converge, we also force the model to stop if after a certain number of runs the stopping criterium is still not met.

Results of testing adapted RSM for the one parameter estimation case

After the implementation of these adaptations we have run the RSM model again to test whether the results improved. Note that because the initial approximation regions are larger, some low starting points have moved a bit to stay within the natural boundaries. Furthermore we let some other parameters start somewhat further away from the optimum. Also the low starting point of parameter $\mu_{1,1}$ has been increased a bit, because for small values of $\mu_{1,1}$ the number of relationships is so large that the model runs out of memory. In Table 5.6 the results of the adapted RSM are represented in same way as in Table 5.5.

Table 5.6: Results of testing adapted implementation RSM by freeing parameters one by one

Parameter	Low starting point			High starting point		
	Start value	Initial region	End value	Start value	Initial region	End value
$\mu_{1,0}^*$: 26	9	1-17	23.4	50	42-58	25.0
$\mu_{1,1}^*$: 43	16	8-24	44.3	75	67-83	$59^{\dagger}$
μ_2^*: 1.5	0.5	0.1-0.9	1.58	4.5	3.7-5.3	1.62
λ^*: 0.8	0.5	0.1-0.9	0.86	3	2.6-3.4	0.82
β_0^*: 4.5	1.0	0.2-1.8	$3.9^{\dagger}$	8	7.2-8.8	4.2
β_1^*: 20	8	2-14	19.6	35	29-41	20.7
δ^*: 0.95	0.5	0.4-0.6	0.96	-	-	-

†*End parameter value more than 10% away from the actual parameter value*

If we compare the results of the adapted implementation of the RSM in Table 5.6 with the initial results in Table 5.5 we see some improvements. In 11 of the 13 cases it was possible to get as close as 10% of the original value, while at the same time the running time decreased with approximately 40%. Only for the low starting point of β_0 and the high starting point of $\mu_{1,1}$ the RSM is unable to get close to the optimal value.

Testing fit between original function and approximation function

To test the adequacy of the approximation function in the different phases of the RSM, we have used the Q^2-statistic presented in Box and Draper (1987):

$$Q^2 = \frac{\sum_i (F(p_i) - \bar{F})^2/k}{\sigma^2}$$

with: p_i = i-th design point

$F(p_i)$ = True value of $F(p_i)$ in the i-th design point (see Formula 4.1 for definition)

$\bar{F}$ = Mean of $F(p_i)$ over all design points

k = number of parameters that are to be estimated in the approximation function

σ^2 = variance of $F(p_i)$

An estimator of this statistic Q^2 is given by:

$$\hat{Q}^2 = \frac{\sum_i [g(p_i) - \bar{g}]^2/k}{\sum_i [f(p_i) - g(p_i)]^2/(n-k-1)} - 1 \sim F_{k,n-k-1} - 1$$

with: $f(p_i)$ = estimation of $F(p_i)$, based on the model outcome in the i-th design point

$g(p_i)$ = value of the estimated approximating function in the i-th design point

$\bar{g}$ = mean of $g(p_i)$ over all design points

n = number of design points

To test the statistical significance of the estimated parameters $Q^2 + 1$ should be greater than $F_{k,n-k-1}(1-\alpha)$. In our case with 3 parameters and 8 design points $F_{k,n-k-1}(0.95) = 6.59$. Box and Draper indicate that to have an adequately estimated response function, the F-value should be minimally about 10 times higher than $F_{k,n-k-1}(1-\alpha)$. Calculation of the Q^2 statistic for all the regression analyses that have been performed in the different iterations of the test runs of the RSM, show that this requirement is hardly ever met. Still, most of the times the movements made based on the estimated regression equation were in the right direction. A general conclusion might be that the RSM does not need a perfect fit to work fairly well.

Adaptation of the transformation of F(p)

The Q^2-statistic can, amongst others, be used to test different specifications and transformations of the response functions. To ascertain whether the used square root transformation on F(p) worked out well, we have calculated the Q^2-statistics in all iterations for the original function F(p), $\sqrt{F(p)}$, and log(F(p)). The outcome was that the values of Q^2 were not qualitatively different for the three possibilities, although most of the times the square-root transformation resulted in slightly higher values of the Q^2-statistic, while most of the times the log-transformation outperformed both other possibilities. Remarkably, the directions and step sizes of movement based on the performed regression analysis showed hardly no difference for the three transformations. Still, we decided that it might be better to use a log-transformation on F(p) than a square root transformation. Because the impact of the differences in the movements for the three transformations were only limited, we decided not to recalculate all our earlier analyses, but use the log-transformation only from this point on.

Final test: simultaneous estimation of all parameters

Until now, the tests of the RSM only dealt with the estimation of one parameter at a time, while the other parameters were kept fixed at their original value. To test whether the RSM is also able to find parameters in a multi-parameter optimization problem, we have released all 7 parameters simultaneously and evaluated whether the RSM was able to find all parameters values that were used to generate the test data set. For all parameters we have used starting points far from the optimal values; some parameters have been given a high starting value and others a low starting value. This was done in such a way that compared to the optimal parameter values some starting values of parameters led to increased numbers of partners, while others led to a decreased numbers of partners.

For this final test we have used a composite design consisting of a double (i.e. 2 evaluations at every point) star with length 1, a double star with length $1/\sqrt{2}$ and supplemented this with 8 center points. The total design consisted of 64 points. This large design was considered to reduce the impact of noise. Note that a star design is very useful for estimating separable polynomials, because one can easily measure the effects of changes in parameters relating to one variable X_j. To estimate the two coefficients relating to a variable X_j, 16 design points (including 8 center points) are available.

In all the runs a simulation horizon of 200 years has been used. The stopping criterium we have used was the same as in the adapted one-by-one case, with a maximum of 1280 runs (20 iterations). It appeared that the program stopped on the latter criterium and not on the criterium of number of design decreases. In Table 5.7 the results of the final test are indicated.

Table 5.7: Results of simultaneous estimation using adapted RSM implementation

Parameter	Start value	Initial region	End value
$\mu_{1,0}^*$: 26	9	1-17	30.2†
$\mu_{1,1}^*$: 43	75	67-83	61.3†
μ_2^*: 1.5	0.5	0.1-0.9	1.51
λ^*: 0.8	3.0	2.6-3.4	0.82
β_0^*: 4.5	1.0	0.2-1.8	4.3
β_1^*: 20	35	29-41	18.1
δ^*: 0.95	0.5	0.4-0.6	0.93

†*End parameter value more than 10% away from the actual parameter value*

One can see in Table 5.7 that after the completion of 20 designs, 5 out of 7 parameters have come within 10% of the real value. As before, there are problems with the estimation of parameter $\mu_{1,1}$. Furthermore, also parameter $\mu_{1,0}$ is too high. Looking at the path through the parameter space in the different iterations of the RSM, we see that in the first iterations the very sensitive parameters δ and λ moved fast to their optimal value. After δ and λ have found their optimal value more or less, the less sensitive parameters μ_2, β_0 and β_1 started to move to their optimal value. After 10 designs also the parameters μ_2, β_0 and β_1 have found their position, and only the last parameters $\mu_{1,0}$ and $\mu_{1,1}$ did not have a good value yet. However, in the last 10 design rounds the RSM did not manage to move these parameters to the right direction.

In the appendix Table A.3 and A.4 indicate which distribution of numbers of partners results with the parameters found in the final test of the RSM. If we compare the results of table A.3 and A.4 with the test data set (Table 5.2 and Table 5.3), we see that for males and females above 30 the numbers of recent partners are somewhat to low (maximal difference between activity classes ±4%), while there are also small differences for males and females in lifetime numbers of partners (maximal differences between activity classes of ±2%). These differences are not very large, but unfortunately systematic.

5.5 Results parameter estimation with real data

Although the RSM was not able to find in the tests the exact parameter values again, to our opinion the results are good enough to use the current implementation of the RSM for the actual parameter estimation problem. In the actual parameter estimation we compare the model outcome at different parameter values with the real data set instead of the test data set. The estimation is started from the parameters that have been used to generate the test data set. As mentioned in Section 5.2 these parameters have been found by a preliminary application of the RSM to the real data set, and hence they are supposed to be at least of the right order of magnitude. To estimate the parameters we have used the same implementation of the RSM as in the simultaneous parameter estimation for the test data set. The only difference is that to limit running time we stop the algorithm already after 10 designs instead of after 20. Note that in the final tests all major steps have been made in the first 10 iterations. In Table 5.8 the results of the parameter estimation process are indicated:

Table 5.8: Results of parameter estimation based on the real data set

Parameter	Start value	Initial region	End value
$\mu_{1,0}$	26	18-34	50.1
$\mu_{1,1}$	43	67-83	79.3
μ_2	1.5	1.1-1.9	1.35
λ	0.8	0.4-1.2	0.63
β_0	4.5	3.7-5.3	6.6
β_1	20	14-26	40.6
δ	0.95	0.91-0.99	0.955

Table 5.8 shows that the estimation procedure results in major changes in the parameters $\mu_{1,0}$, $\mu_{1,1}$, β_0 and β_1. These are parameters for which the model is not very sensitive (see also Table 5.4) and which were relatively difficult to estimate in the testing procedures. A possible explanation of the finding that changes especially occur in these parameters is that for the more sensitive parameters the values found by the preliminary implementation of the RSM were already quite good and for the less sensitive parameters not. Table 5.9 and Table 5.10 indicate which distribution of sexual

Table 5.9: Percent distribution of women and men by number of sexual partners in the six months preceding the model survey, according to age-class; results of 100 runs of 200 years with parameters found in estimation of real data set; end population after 100 runs 239900 women and 245700 men

Age	Females				Males			
	None	1 partn.	2 partn.	3+ partn.	None	1 partn.	2 partn.	3+ partn.
15-19	57.5	40.4	0.7	1.4	61.3	30.6	6.3	1.8
20-24	25.8	69.3	2.5	2.4	31.8	56.9	8.8	2.6
25-29	15.4	78.6	3.8	2.1	20.0	70.7	7.3	2.0
30-34	10.3	83.4	5.0	1.4	9.7	82.4	7.1	0.8
35-39	7.3	84.7	6.9	1.1	5.3	85.2	8.9	0.7
40-44	7.1	84.4	7.7	0.8	3.8	82.5	12.7	1.1
45-49	7.3	82.9	9.0	0.8	2.6	80.9	14.8	1.6
50-54	9.8	80.5	9.1	0.7	3.1	78.2	16.5	2.2
Overall*	20.5	73.3	4.7	1.5	12.4	75.6	10.4	1.6

* *To guarantee comparability with the overall figures in Table 2.2, for males the age-group 15-19, and for females the age-group 50-54 are excluded from the overall category.*

Table 5.10: Percent distribution of women and men by number of sexual partners in their life, according to age-class; results of 100 runs of 200 years with parameters found in estimation of real data set, end population after 100 runs 239900 women and 245700 men

Age	Females					Males				
	None	1	2-3	4-5	6+	None	1	2-3	4-5	6+
15-19	53.2	39.1	6.1	0.1	1.5	45.3	12.1	18.2	11.8	12.5
20-24	14.5	49.3	31.7	0.8	3.7	6.3	7.3	16.4	16.9	53.1
25-29	3.6	34.5	52.6	4.6	4.7	0.2	2.9	8.5	11.2	77.2
30-34	0.7	21.0	61.9	11.2	5.3	0.0	1.4	5.6	8.7	84.3
35-39	0.1	13.0	62.3	18.6	5.9	0.0	0.5	4.0	7.1	88.4
40-44	0.0	9.2	60.2	24.3	6.4	0.0	0.2	2.6	5.0	92.3
45-49	0.0	6.6	55.6	30.2	7.7	0.0	0.1	2.4	5.3	92.3
50-54	0.0	4.6	50.5	34.8	10.2	0.0	0.0	1.6	4.3	94.0
Overall*	12.2	26.4	45.2	11.5	4.8	1.2	2.1	6.6	8.9	81.2

* *To guarantee comparability with the overall figures in Table 2.4, for males the age-group 15-19, and for females the age-group 50-54 are excluded from the overall category.*

activity results when the STDSIM model is run with the estimated parameters.

Comparison of Table 5.9 and Table 2.2 shows that the model simulates too many 30+ females having more than 1 recent partner, and too few with 0 recent partners. For males, the percentage with 3 or more recent partners is much too low. Looking at lifetime numbers of partners, Table 5.10 shows quite reasonable results for females, while for males we see that numbers of lifetime partners are much too high.

It appears to be difficult to identify model adaptations that could improve the fit between model outcome and the observed data. For instance, the too high numbers of lifetime partners in males indicate that the rate of partner change is too high. But if the rate of partner change is lowered, we run into trouble with respect to numbers of recent partners, unless more males are in concurrent relationships. However, if more males are in concurrent relationships, there must be female (non-prostitute) partners, which increases the numbers of both recent and lifetime numbers of partners of females to a too high level. The difference between the reported number of partners of males and females seems to be too large to be reproduced with our model. Of course, it might be possible that inclusion of extra parameters, for instance representing heterogeneity in sexual behavior for different age classes and/or between individuals could improve the results somewhat, but it is questionable whether this could fill the gap in answers between males and females. It is more likely that because of reporting bias it will never be possible to find a model specification that perfectly fits to the data.

If we compare Table 5.9 and Table 5.10, generated by the parameters found in the estimation procedure, with the results of Table 5.2 and Table 5.3, generated by the start values of the parameters in the estimation procedure, at first sight the results have not improved that much. Note that this comparison can only be made with caution, because Table 5.9 and Table 5.10 are generated with a fixed seed for demographic processes, and Table 5.2 and Table 5.3 are not. Despite this observation, the value of the function F(p) has decreased with approximately 30% in the estimation process.

The estimation procedure for the test data set has also been started from other starting points, and the resulting optimum was qualitatively the same. Results of these estimation procedures are not further presented in this working-paper.

6 Discussion

6.1 Summarizing conclusions

This working-paper has described the first steps in the process of quantifying a sexual behavior submodel that constitutes the backbone of the stochastic microsimulation model STDSIM. The general outline and aim of the overall model STDSIM has been presented in Chapter 1. In general, the difficulty of quantification of a model largely relates to the level of detail in the model that is to be quantified. Very detailed models are most of the times hard to quantify, especially when data are limited. However, very crude models will often not generate results that resemble real world processes, and hence such models are not useful for decision support. In model development a trade-off has to be made between statistical validity, requesting a parsimonious model, and face-validity, requesting a detailed model.

The data available for the estimation of the parameters of the sexual behavior submodel has been presented in Chapter 2. At first sight the available data might look sufficient. However, it is not possible to derive estimates for all parameters from these data immediately, because we only have aggregated cross sectional data, while our microsimulation model deals with evolvements in time at an individual level. Furthermore, we are faced with the more general problem that reported data, especially on sexual behavior, might be far from reliable. However, because better sources of quantitative information on sexual behavior are not available, we have decided to act in the process of model quantification as if we trust the available data.

In Chapter 3 the basic structure of the sexual behavior submodel has been presented. We have tried to design a model structure that is relatively simple, but still detailed enough such that the most important processes are included. In the model we have identified 7 key parameters which have to be estimated in the quantification process. In a sensitivity analysis presented in section 5.2 we have investigated how the model outcomes react to changes in these key parameters.

To ascertain how closely our model resembles real world data we have defined in Section 4.1 a function that measures the distance between the model outcome and the data. In the parameter estimation procedure we will try to minimize this function. Based on experiences in literature for similar problems, we have used the Response Surface Method to solve this minimization problem. In implementing the RSM many decisions have to be made regarding e.g. experimental design, approximation region, type of approximation function etcetera. These topics have been mentioned in Section 4.2, and have been further discussed in Chapter 5.

The model STDSIM has been used to generate a test data set, in order to ascertain whether the RSM is suitable for our application. According to the results of these test we have modified our initial implementation of the RSM. In Chapter 5 it is shown that the adapted RSM is most of the times able to find approximately the right parameters in the test data set, both in estimation of single parameters as in simultaneous estimation of several parameters. However, it appeared to be hard to find parameters that result in model outcomes which are close to the real world data of Chapter 2. This can be explained by inconsistency in data of males and females, but maybe the basic model is also still too crude with respect to some assumptions. Further research to throw more light on this point would be very valuable.

6.2 Further work

The further work in constructing and quantifying the sexual behavior submodel can be divided into three main parts: improvement of the model, improvement of the optimization method used for parameter estimation and collection of more data of better quality. Apart from these three main points of further work, it might be useful to study first the simulated life-histories on an individual level, to see whether the evolvement of numbers of partners in time of individuals seems 'reasonable'. Even with a good fit on an aggregated level, the results on an individual level might be counter-intuitive.

Possible improvements of the sexual behavior submodel are the inclusion of more heterogeneity in sexual behavior between individuals, and the differentiation of sexual behavior according to age. At the moment that such adaptations and extensions are implemented in the model it has to be carefully ascertained what the impact is on the model outcome, and whether the fit between the model outcome and the data is significantly improved by adding more complexity to the model. Another point of attention is that in the model presented migration is not taken into account yet. In Nairobi, and also in other cities in developing countries, migration flows are considerable. Because in these migration flows males are often strongly overrepresented, migration leads to an unbalanced population structure which also affects sexual behavior. Hence, inclusion of migration can lead to marked changes in the specification of the sexual behavior submodel.

Although the current implementation of the Response Surface Method performs reasonably well, considerable improvement might still be possible. Before the RSM is further adapted, it might be worthwhile to first compare the current implementation of the RSM with other optimization methods, like Stochastic Approximation, Genetic Algorithm or Simulated Annealing. Depending on the results of these comparisons it has to be decided whether it is worthwhile to make further improvements to the RSM. Some points that might deserve attention in improving the implementation of the RSM are:

- Design: until now the experimental design has been quite arbitrarily chosen. However, there are lot of theoretical considerations, like unbiasedness of coefficients, minimization of variance of the error of the approximation (see Box and Draper, 1987) that can be looked upon. It has to be tested whether design structures from literature are suitable for our application. Furthermore, it might be worthwhile to change the design in the various iterations of the RSM. For instance, one can implement a decision rule that when the fit between the original function and approximating polynomial is bad, extra points will be added to the design.
- Improvement of stopping rules: in the adapted implementation of the RSM we have used a stopping rule relating to the size of the approximation region. Unfortunately this did not work out well, because in all cases the algorithm stopped on the artificial stopping rule of a maximum on the number of iterations. Therefore, other stopping criteria have to be considered.
- Estimation procedure of the approximation function: until now we have applied the ordinary least squares method for estimating the coefficients of the approximating function. The OLS method assumes that the errors are normally distributed. However, we have not tested this assumption yet. If the assumption of normally distributed errors does not hold other estimation techniques can be more valid. For instance, Box and Draper (1987) mention a weighing technique that increases the robustness of the estimator for heavy-tailed error distributions.

- Transformation of variables: until now we have only paid attention to the question whether transformation of the function F(p) would be useful. However, it might be the case that transformations of the explanatory variables (i.e. parameters in the simulation model) would be much more valuable.

In the discussion of the results of the application of the RSM to the 'real' data set (Section 5.5) it is mentioned that the discrepancy between the answers of males and females is so large, that it is unlikely to be true. Consistent data is surely needed, but will be hard to get because the phenomenon of inconsistency in the answers of males and females is present in sex surveys in all countries. Ideally, we would like to have data about the evolvement of number of sexual partners in time on an individual level, because such data closely relate to the mechanisms in our model. New studies would have to be conducted to gather this data, but it does not seem reasonable to expect results of new studies in the short run. A more attainable goal might be to include in current surveys questions relating to number of new partners in different time intervals (one month, 6 months, 2 years). Such questions could shed more light on the topic whether high numbers of partners in a certain time interval indicate a high rate of partner change or a high number of concurrent relationships.

Although there is still a lot of work to be done before we will have a specification and quantification of the sexual behavior submodel that describes the sexual behavior of individuals in Nairobi, the research described in this working-paper is the first important step in this direction. To our opinion, the formal description of the model specification enhances discussion of the used assumptions with experts from a broad range of fields. Furthermore, the RSM showed to be a fairly good method for estimating parameters in the type of model under study, and the first model experiments helped to identify what kind of data is needed to get a reliable model that can be used for decision support in the control of STDs.

Acknowledgements

This research was partially supported by the Netherlands Organization for Scientific Research (NWO). The author would like to thank Prof. A. Ruszczyński and Prof. G. Pflug for their helpful discussions and comments during her stay at IIASA, Dr C.P.B. van der Ploeg and Dr. G.J. van Oortmarssen for their careful reading of the manuscript, and S.J. De Vlas MSc for all his efforts in the coding of the STDSIM model and for providing an initial implementation of the RSM.

Bibliography

Anderson RM, May RM. Infectious diseases of humans: dynamics and control. Oxford: Oxford University Press, 1992.

Anderson RM, May RM, Ng TW, Rowley JT. Aged-dependent choice of sexual partners and the transmission dynamics of HIV in Sub-Saharan Africa. Philosophical Transactions of the Royal Society of London Biological sciences, vol 336, 135-155 (1992)

Box GEP, Draper NR. Empirical model building and response surfaces. New York [etc]: John Wiley & Sons, 1987.

Free JW, Parkinson AR, Bryce GR, Balling RJ. Approximation of computationally expensive and noisy functions for constrained nonlinear optimizations. Transactions of the American Society of Mechanical Engineers, vol 109, 528-532 (1987)

National Council for Population and Development/ Central bureau of Statistics, Kenya. Kenya Demographic and Health Survey 1993. Calverton, Maryland: NCPD, CBS and MI, 1994

Over M, Piot P. HIV Infection and Sexually Transmitted Diseases. In: Disease control priorities in developing countries, eds D.T. Jaminson et al (455-527). New York: Oxford University Press for the World Bank, 1992.

Toropov VV. Simulation approach to structural optimization. Structural Optimization 1, 37-46 (1989).

Toropov VV, Filatov AA, Polynkin AA. Multiparameter structural optimization using FEM and multi-point explicit approximations. Structural Optimization 6, 7-14 (1993)

United Nations/World Health Organization. The AIDS epidemic and its demographic consequences. Proceedings of the UN/WHO Workshop on modelling the demographic impact of the AIDS epidemic in pattern II countries: progress to date and policies for the future. New York, 13-15 december 1989.

Vanderplaats GN. Effective use of numerical optimization in structural design. Finite elements in analysis and design 6, 97-112 (1989)

Appendix

Table A.1: Percent distribution of women and men by number of sexual partners in the six months preceding the model survey, according to age-class; results of 100 runs of 100 years with an end population over 100 runs of 77636 women and 77099 men

Age	Females				Males			
	None	1 partn.	2 partn.	3+ partn.	None	1 partn.	2 partn.	3+ partn.
15-19	58.8	38.7	1.2	1.4	65.3	28.2	5.2	1.3
20-24	32.7	61.4	3.3	2.7	37.2	54.4	6.8	1.5
25-29	22.4	71.5	4.2	1.9	24.2	66.7	7.9	1.3
30-34	16.8	76.3	5.4	1.5	16.6	73.5	8.9	1.0
35-39	12.7	77.8	8.0	1.5	8.4	77.7	12.3	1.7
40-44	11.4	77.6	10.0	1.0	6.4	74.2	16.8	2.6
45-49	12.3	76.5	10.4	0.8	4.9	72.1	19.0	4.0
50-54	14.3	74.3	10.4	1.0	5.3	71.6	19.3	3.8
Overall*	26.4	66.6	5.5	1.6	16.8	69.1	12.1	2.1

** To guarantee comparability with the overall figures in Table 2.2, for males the age-group 15-19, and for females the age-group 50-54 are excluded from the overall category.*

Table A.2: Percent distribution of women and men by number of sexual partners in their life, according to age-class; results of 100 runs of 100 years with an end population over 100 runs of 77636 women and 77099 men

Age	Females					Males				
	None	1	2-3	4-5	6+	None	1	2-3	4-5	6+
15-19	52.3	38.5	7.7	0.1	1.5	46.9	14.3	20.3	11.1	7.4
20-24	16.5	41.6	37.0	1.2	3.9	6.7	9.3	20.3	20.6	43.2
25-29	3.8	25.4	58.1	8.0	4.8	0.2	2.8	10.4	15.5	71.1
30-34	0.8	13.6	60.2	19.5	5.9	0.0	0.8	5.2	11.2	82.8
35-39	0.2	6.5	53.6	31.9	7.9	0.0	0.3	3.6	8.9	87.3
40-44	0.0	4.2	45.0	39.1	11.6	0.0	0.0	2.4	6.9	90.7
45-49	0.0	2.4	36.3	44.5	16.8	0.0	0.0	1.8	5.5	92.7
50-54	0.0	1.8	29.3	47.2	21.7	0.0	0.0	1.3	5.4	93.3
Overall*	13.1	21.4	41.3	17.8	6.3	1.3	2.4	7.5	11.5	77.3

**To guarantee comparability with the overall figures in Table 2.4, for males the age-group 15-19, and for females the age-group 50-54 are excluded from the overall category.*

Table A.3: Percent distribution of women and men by number of sexual partners in the six months preceding the model survey, according to age-class; results of 100 runs of 200 years with parameters found in estimation of 7 parameters of test data set, end population after 100 runs 239900 women and 245700 men

Age	Females				Males			
	None	1 partn.	2 partn.	3+ partn.	None	1 partn.	2 partn.	3+ partn.
15-19	60.4	37.5	0.8	1.4	66.2	27.9	4.8	1.1
20-24	32.3	63.2	2.2	2.4	40.2	51.2	7.2	1.4
25-29	21.8	72.9	3.1	2.2	32.0	59.5	7.2	1.3
30-34	15.5	78.8	4.5	1.2	22.4	69.4	7.1	1.1
35-39	11.7	81.0	6.2	1.1	8.9	80.0	10.0	1.1
40-44	11.0	80.7	7.6	0.7	5.9	78.2	14.3	1.7
45-49	11.1	80.0	8.2	0.8	5.1	75.0	17.6	2.3
50-54	12.9	78.5	8.1	0.5	5.5	74.3	17.2	2.9
Overall*	25.3	69.0	4.3	1.4	19.1	68.4	10.9	1.6

* *To guarantee comparability with the overall figures in Table 2.2, for males the age-group 15-19, and for females the age-group 50-54 are excluded from the overall category.*

Table A.4: Percent distribution of women and men by number of sexual partners in their life, according to age-class; results of 100 runs of 200 years with parameters found in estimation of 7 parameters of test data set, end population after 100 runs 239900 women and 245700 men

Age	Females					Males				
	None	1	2-3	4-5	6+	None	1	2-3	4-5	6+
15-19	54.0	36.8	7.5	0.1	1.6	47.5	14.0	20.4	10.8	7.2
20-24	15.2	42.7	37.2	1.4	3.5	6.7	9.1	21.2	21.5	41.6
25-29	3.8	26.2	56.9	8.2	5.0	0.3	2.6	9.8	14.8	72.5
30-34	0.7	13.7	60.7	19.2	5.7	0.0	0.7	4.7	9.2	85.3
35-39	0.1	8.1	54.3	29.6	7.9	0.0	0.2	2.2	6.0	91.7
40-44	0.0	5.1	47.5	37.0	10.4	0.0	0.0	1.2	3.6	95.2
45-49	0.0	2.9	39.5	42.2	15.3	0.0	0.0	0.9	3.3	95.8
50-54	0.0	1.8	31.7	45.0	21.5	0.0	0.0	0.5	2.3	97.3
Overall*	12.5	21.1	42.2	17.7	6.5	1.2	2.2	6.7	9.6	80.2

* *To guarantee comparability with the overall figures in Table 2.4, for males the age-group 15-19, and for females the age-group 50-54 are excluded from the overall category.*