THE MODELLING OF THE TRANSMISSION OF HIV INFECTION IN SELECTED EUROPEAN COUNTRIES: A MARKOW CHAIN APPROACH

STANISŁAW MACIEJ KOT

GDANSK UNIVERSITY OF TECHNOLOGY,

Department of Economics and Business Management, Narutowicza St., 11/12, 80-952 Gdansk, Poland e-mail: skot@zie.pg.gda.pl

JOLANTA WOJCIECHOWSKA

University of Gdansk, Department of Pedagogy, Krzywoustego St., 19, 80-952 Gdansk, Poland e-mail: pedjw@univ.gda.pl

ABSTRACT

In the paper, Markov chains are applied to modelling of the transmission of the HIV infections between the populations of: homo/bisexual men, injection drug users, and heterosexual people. These groups will be called "the states" and denoted by s_1 , s_2 , and s_3 respectively. At the moment t, new infections are mathematically described by the three state discrete random variable X_t , which takes the state s_j with the probability $p_j(t) = P(X_t = s_j)$, $j = 1,2,3$. The set of these three probabilities, i.e., $\{p_1(t), p_2(t), \text{ and } p_3(t)\}$ describes the structure of new infections at the moment t. The sequence of such structures over time t reflects the evolution of the HIV epidemic. Let $p_{ij} = P(X_t = s_j | X_{t-1} = s_i)$, i,j = 1,2,3, denote the conditional probability that the HIV-positive person who belongs to the ith population at t-1 will infect the person from the *j*th population at t. Particularly, the conditional probability p_{ii} describes the transmission of HIV inside the same *i*th population. The 3x3 matrix $P = [p_{ij}]$ of these nine conditional probabilities defines the transition of HIV between (or inside of) these populations completely.

In the paper, the method of the estimation of unknown transition probabilities p_{ii} is proposed when the observations of the structures $\{p_1(t), p_2(t), p_3(t)\}\$ at $t = 1, 2, \ldots, T$ are available only. In empirical example, those probabilities are estimated for four European countries: Germany, United Kingdom, Sweden, and Poland. The half yearly data on new HIV infections are used for the period 2000-2003.

1. Introduction

In this paper, we adopt Markov-chain methods to develop a simply mathematical model of HIV epidemic. Three populations of HIV-positive persons are chosen, according to the modes of the HIV transmission: homo/bisexual men, injecting drug users, and heterosexual women and men. These populations will be called "the states", and denoted by 1 ('HOMO/BI'), 2 ('IDU'), and 3 ('HETERO') respectively. The proposed model describes the transition of HIV between these states. It means that we treat the HIV epidemic as a selfadopting system which dynamic is controlled only by the probabilities of the transition of HIV from one state to another.

The estimation of unknown transition probabilities is the main difficulty when the Markov chain is to be fitted to sample data. Micro-data' are usually used for this purpose. The term "micro" means information about the transition of HIV on individual person's level, i.e. the transition of HIV from person A to person B. The 'ideal' micro-data could be collected from reports which contain at least four elements: t_A and t_B - the dates of infection of person A and B respectively, and the labels s_i , s_j of the populations (HOMO/BI, IDU, HETERO) to which person A and B belongs respectively, $i, j = 1, 2, 3$. The micro-data would contain *n* such records for *n* diagnosed persons B. For established time unit Δ ('step'), we may select the records for which the distance t_B – t_A is close to Δ , and calculate the numbers n_{ij} of the HIV passes from the population s_i to s_j . If n_i denotes the total number of such passes from s_i , the ratio: n_{ij}/n_i will be the estimate of the transition probability from s_i to s_j in one step.

However, the typical surveillance reports base only on two elements: t_B and s_i , of micro-data described above, i.e. on information about the date of infection of diagnosed person B and the population of the person A who has infected B. The surveillance reports present aggregate data ('macro-data') in the form of aggregate numbers of new HIV diagnoses (of persons B) which are classified according to the route of infection s_i. Therefore, the estimates of the transition probabilities cannot be obtained from the macro-data in the way presented above when micro-data are available.

In this paper, we propose an alternative method of the estimation of unknown transition probabilities, when only macro-data are in our disposal. The proposed method is applied to modelling of the HIV epidemic in four European countries: Germany, United

Kingdom, Sweden, and Poland. Each of these countries represents different 'model' of the routes of HIV transmission.

The rest of the paper is organised as follows. In Section 2, the mathematical model of HIV epidemic is presented. Section 3 contains the description of statistical data. In Section 4, the results of estimation are presented. Section 5 offers concluding remarks.

2. Mathematical model of the HIV epidemic

Let ${X_t}_{t=1,2}$ denote a discrete random variable which takes K values (states) s_i at time t with probability $P(X_t = s_i) = p_i(t)$, where $i = 1, 2, ..., K$, $t = 1, 2, ...$. The sequence $\{X_t\}_{t=1,2,...}$ will form a finite Markov chain if the conditional probability:

$$
P(X_t = s_j | X_{t-1} = s_i, X_{t-2} = s_i, ...)
$$

satisfies the following requirement:

$$
P(X_t = s_j | X_{t-1} = s_i, X_{t-2} = s_i, \ldots) = P(X_t = s_j | X_{t-1} = s_i) = p_{ij}(t) \tag{1}
$$

where $i,j,l = 1,2,...,K$, $t = 1,2,...$ The conditional probability $p_{ii}(t)$ is called the *transition probability* from state s_i to state s_j in one step. The term 'step' means the assumed time unit.

Moreover, if the transition probabilities are independent of time t, i.e., if $p_{ii}(t) = p_{ii}$, for every $i, j = 1, 2, \ldots, K$, the Markov chain is called *'time homogeneous'* or *stationary*. The KxK matrix $P = [p_{ij}]$ is called the transition matrix. This matrix describes the stationary finite Markov chain completely. Stationarity means the lack of any external influences; the whole dynamics of any Markov chain-type system is controlled solely by the probabilities of the transition between states.

Let states s_1 , s_2 , and s_3 denote the following populations of people respectively: 1. homo/bisexual men ('HOMO/BI'), 2. injecting drug users ('IDU'), and 3. heterosexual men and women ('HETERO'). The HIV epidemic will occur if the virus HIV transits from one person to another, either between these populations or within one of them. In mathematical terms, the HIV epidemic consists in the transition of HIV between states s_1 , s_2 , and s_3 over time.

We propose the modelling of the HIV epidemic by the stationary Markov chain with the states s_1 , s_2 , and s_3 described above. The following symbols will be used in the paper:

 $p_i(t)$ – the unconditional probability that a HIV-positive person belongs to the *i*th group at time t, $j = 1,2,3$. The vector $p(t) = [p_1(t), p_2(t), p_3(t)]$ reflects the 'structure' of infected populations at t;

 $p_{ij} = P(X_t = s_i | X_{t-1} = s_i)$ – the conditional probability that the HIV-positive person who belongs to the *i*th population at t -*l* will infect at t the person who belongs to the *j*th population. Particularly, p_{ii} denotes the conditional probability of HIV infection between two persons who belong to the same ith group.

The transmission of HIV between three states in our epidemic model is displayed in fig.1.

(insert fig.1 here)

For the stationary Markov chains, the following equations hold:

$$
p_j(t) = \sum_{i=1}^{K} p_i(t-1)p_{ij}, j=1,2,...,K
$$
 (2)

or in matrix notation:

$$
p(t) = P p(t-1) \tag{3}
$$

The equation (2) (or (3)) is crucial for the proposed method of the estimation of the transition matrix P . Let us notice that the equation (2) resembles the regression model with dependent variable $p_i(t)$, independent variables $p_1(t-1)$, $p_2(t-1)$, …, $p_k(t-1)$, and the K conditional probabilities p_{1i} , p_{2i} , p_{Ki} as the coefficients of regression.

If $y_i(t)$, $y_i(t-1)$, $y_k(t-1)$ denote the observations of the unconditional probabilities $p_j(t), p_1(t-1), p_2(t-1), ..., p_K(t-1)$ respectively, the regression model (2) will take the following form:

$$
y_j(t) = \sum_{i=1}^{K} y_i(t-1)p_{ij} + Z_j, j=1,2,...,K
$$
 (4)

where Z_j is a disturbing term with zero mean and variance σ^2 [Miller, 1952].

Miller estimated unknown transition probabilities p_{ij} by least squares method (LSM) for every $j = 1, 2, ..., K$ separately.

In general, Miller's approach does not provide correct estimates of the transition probabilities. It is due to the fact that these probabilities should satisfy the following constraints:

$$
\sum_{j=1}^{K} p_{ij} = 1, i = 1, 2, ..., K
$$
 (5)

and:

$$
0 \le p_{ij} \le 1, \quad i,j = 1,2, \dots, K \tag{6}
$$

Lee et al (1970) show that LSM estimators of the unknown probabilities p_{ij} always satisfy the condition (5) but the condition (6) may be violated, in general. Hence, cited authors develop some alternative estimators which fulfil both constraints (5) and (6).

The estimates of the transition probabilities may be used for the calculation of the Bartholomew's mobility index I_B :

$$
I_B = \sum_{i=1}^{K} \sum_{j=1}^{K} y_j(1) p_{ij} |i - j|
$$
 (7)

This index summarises the mobility of HIV between populations. It will take zero if there is no mobility between states, i.e. when the transition matrix P is diagonal. The index (7) assigns the weights in the form of the distance $|i - j|$ to each transition probability p_{ij} . The greater this distance, the higher the weight. Therefore, the greater mobility between the states of a Markov chain, the greater the value of the index I_B .

3. Statistical data

The source data are in the form of the numbers of HIV infections newly diagnosed in homo/bisexual men (state 1), injecting drug users (state 2), and in persons infected through heterosexual contact. The data come from EuroHIV reports: "HIV/AIDS Surveillance in Europe' 2000-2003. Four countries were selected: Germany, United Kingdom, Sweden, and Poland.

Reporting of HIV diagnoses has become a key surveillance instrument to monitor the HIV epidemic. It has progressively replaced AIDS surveillance which, since 1996 with the introduction and wide spread use of HAART therapy, has become less reflective of the HIV underlying trends.[EuroHIV, 2004, p.5].

In our analysis, 6 month period is used as a time unit t . For extraction of half-yearly data we have used mid-year reports and end-year reports. The mid-year reports provide data on the cases of newly diagnosed HIV infections by 30 June. The subtraction of these data from yearly data provides data for the second half of a year. EuroHiv reports do not allow for recovering the time units shorter than 6 month.

For every time unit t, the proportions (fractions) $y_i(t)$ of persons in the jth population (state) are calculated, $j = 1,2,3$. These fractions are presented in Table 1 and 2.

(insert Table 1 and 2 here)

Each raw of these tables describes the structure of the HIV epidemic, with respect to the mode of HIV transmission, at a half year. The last element of every raw is the total number of infections (sample size).

The trends of the HIV epidemic in selected countries are plotted in fig.2.

(insert fig.2 here)

It is seen that the total number of reported HIH diagnoses has been rising in United Kingdom since 2000. Certain stabilisation of the HIV epidemic is observed in remaining countries.

Fig.3 displays the changes in the structure of HIV epidemic in Germany over time.

(insert fig.3 here)

One can see in fig. 3 that the fractions of HIV infections newly diagnosed in the populations of heterosexual men and heterosexual persons were slowly increasing in years 2000-2001, whereas the fractions in injecting drug users were decreasing continually over that period. The fractions in the heterosexual population started to decrease in the first half of the year 2002 what combining with constant trend of the fractions in the IDU population resulted in increasing trend of the fractions in the HOMO/BI population.

The next figure (4) shows the trends of the fractions of HIV infections newly diagnosed in United Kingdom.

(insert fig.4 here)

It is seen in fig. 4 that the increasing trend of fractions in the HETERO population is accompanied with the decreasing trend in the HOMO/BI population in years 2000-2001. It seems to be a continuation of those trends from the past. Yearly data (not quoted here) indicated on the year 1996 as a starting point of those trends. The fractions of new HIV infections in the HOBO/BI population outweighed the fractions in the HETERO in years 1996-1998. A reverse picture was observed since 1999 [Wojciechowska, 2004, p. 53]. Data from other countries indicate that the rise in diagnoses of heterosexual infections is largely due to an increasing number of cases among persons originating from countries with generalised HIV epidemics: from 30% of heterosexual infections in 1998 to 53% in 2003, over 90% of which were in emigrants from sub-Saharan Africa [EuroHIV, 2004, p.5].

In fig. 4 it is also seen that the rates of the growth of the fractions of both heterosexual and homo/bisexual infections have diminished markedly since 2002. The fractions of new HIV infections in the population of injecting drug users in UK persisted on very law level with tiny tendency to decrease during whole period.

The structure of HIV epidemic in Sweden is presented in fig.5.

(insert fig.5 here)

We see that the picture of Swedish HIV epidemic is similar to the British one. However, the fractions of HIV infections in injecting drug users and heterosexual persons in Sweden are higher than those in UK.

The trends of fractions of HIV infections in populations under discussion in Poland are presented in fig. 6.

(insert fig.6 here)

The picture of he structure of Polish HIV epidemic is typical for central and eastern Europe: very high proportion of HIV infections among injecting drug users and law proportions for homo/bisexual and heterosexual persons.

4. Empirical results.

The results of estimation of transition probabilities are presented in Table 3.

(insert table 3 here)

The estimates of probabilities on main diagonal, i.e. p_{11} , p_{22} , and p_{33} in each of the transition matrices presented in table 3 show the scale of 'intra-group infection" or "self-infection", i.e. the transition of HIV among persons from the same group. The elements outside the main diagonal reflect the inter-group transmission of HIV. We will interpret the obtained results for every country separately.

Germany. The first raw of the estimated transition matrix shows that homo/bisexual person can only infect the person from the same group. In fact, the first state of underlying Markov chain is absorbing, which means that there is no return from this state to another one. This absorbing state can be reached only from state 3. It means that HOMO/BI is a selfinfecting population, i.e. it do not expose other groups to the risk of infection. However, bisexual persons are exposed to the risk of infection by heterosexual persons.

The second raw of the transition matrix informs that injecting drug uses can infect either other IDU (with probability 0.41) or heterosexual persons (probability= 0.59). There is not the way of transmission of HIV from IDU to HOMO/BE.

Third raw describe the transmission of HIV from heterosexual persons to people from other groups. There is 3% chance of infection someone from HOMO/BI by a person from HETERO; 12% chance of infection of IDU, and 84% chance of infection inside HETERO group.

United Kingdom. Transmission of HIV from HOMO/BI to other groups is more likely in UK than in Germany. Although the probability p_{11} of self-infection, i.e. between persons from HOMO/BI is still very high (0.76) the probabilities of infection of other groups are greater than zero: 0.07 for IDU and 0.16 for HETERO.

The second raw of the transition matrix for UK reveals interesting phenomenon: injecting drug users cannot infect themselves but they can infect heterosexual persons with probability one This results suggest that injecting drug users in UK are better protected against HIV transmission from themselves than those in Germany. Perhaps exchange needle programs work better in UK than in Germany.

Heterosexual persons can infect a bisexual person with probability 0.12. The selfinfection among heterosexual persons is high probable ($p_{33} = 0.88$).

Sweden. In this country, the transition of HIV from homo/bi sexual man to other groups is more likely than in Germany and UK. The probability of intra-group infection among homo/bisexual men is equal to 0.48. Homo/bisexual person can infect an injection drug user with probability 0.3, and a heterosexual person with probability 0.22.

The second raw of the transition matrix shows that the probability of the HIV transmission between injecting drug users is very small (0.06). Much greater than that is the probability of the HIV transmission from injecting drug users to heterosexual persons (0.94). The small value of the probability p_{22} on the main diagonal of the transition matrix may indicate that injecting drug users enjoy good-working protective programs in Sweden.

The flow of HIV between homo/bisexual persons and heterosexual ones seems to be symmetric in Sweden: the probability of infection of a bisexual man by a heterosexual women equals 0.24. There is only 2% of chances that an injecting drug user will be infected by a heterosexual person. The scale of intra-group infection among heterosexual persons is very large ($p_{33} = 0.74$).

Poland. The transition matrix P reveals that the mechanism of the HIV infection in Poland differs slightly from that in previously analysed countries. Homo/bisexual man cannot infect another man from the same populations. However, they can transmit HIV to another populations, e.g., to injecting drug users with probability 0.49 and to heterosexual persons with probability 0.51 .

The transfer of HIV from injecting drug users to other populations occurs with small probability: 0.09 to homo/bisexual men and 0.05 heterosexual persons. Very large value of the probability $p_{22} = 0.86$ reveals weakness of exchange needle programs or the methadone programs in Poland.

The third raw of the transition matrix for Poland shows that heterosexual persons transmit HIV to injecting drug users exclusively.

The Bartholomew's index I_B ranks four selected countries in ascending order. The HIV epidemic seems to be the most 'stable' in Germany and the least stable in Poland.

4. Concluding remarks.

The proposed method of the estimation of the transition probabilities of HIV between three distinguished populations or within each of them seems to be a useful tool in depicting important features of the mechanism of the HIV epidemic. Those probabilities appear in many stochastic models [see, among others, Tan and Xiang, 1999, Desai, 2004, Tan, 2005].

However, our Markov chain model of the HIV epidemic bases on two strong assumptions: 1) discrete time domain, and 2) homogeneity (stationarity) of the Markov chain. These assumptions need some comments.

One may argue that the continuous time Markov processes provide better models of the HIV epidemic than our discrete time Markov process (Markov chain). When applying the continuous time Markov processes we have to estimate the matrix of the intensity of the transition between states instead of the transition matrix. It requires micro-data which are recorded at very small time units in order to estimate all necessary characteristics. Even though a continuous time Markov chain is chosen as a correct model of the HIV epidemic, the estimation technique requires the discretisation of time, i.e. the choice of discrete increments ∆t. It leads eventually to the discrete-time Markov processes (Markov chains) at the estimation stage. If so, our procedure of estimation might be applied.

Stationarity (homogeneity) of a Markov chain means the absence of any extraneous impacts on the system. Then the transition matrix P is constant over time. This assumption can be checked by the step-wise estimation procedure. For a sequence of structures $v(t)$ recorded at t=1,2,…T, we may chose an initial time period of the length of m, and estimate the transition matrix, say P_1 . Then we may estimate another transition matrix, say P_2 , using m observations from $t = 2$ to $t = m+1$, i.e. rejecting first observation and adding observation next to m. Repeating this procedure we can obtain the sequence of the transition matrices P_1 , P_2 , $..., P_{T-m+1}$. If these matrices are constant over time, the homogeneous Markov chain could be accepted as a model of the HIV epidemic. Otherwise, we may regress these matrices on some exogenous variables and obtain the model of the non-stationary Markov chain.

In the both problems described above, long time series of the observed structures are necessary. Until now, the half yearly data seem to be the smallest time units which can be recovered from surveillance reports offered by EuroHIV or by other organisations. It means

that we have to wait some years until the HIV databases will cover longer time interval than the recent one.

Bibliography

- Desai K. (2004) 'Challenges of Modelling the HIV Epidemic in the United States", DIMACS, Rutgers University.
- EuroHIV (2000-2003) 'HIV/AIDS Surveillance in Europe', Mid-year and End-year reports., Institut de veille sanitaire, Saint-Maurice
- Kemeny J.G., Snell J.L. (1960) 'Finite Markov Chains', Princeton Univ. Press, Pronceton.
- Lee T.C., Judge G.G., and Zellner A. (1970) 'Estimating the Parameters of the Markov Probability Model from Aggregate Time Series Data', North-Holland Publ. Co., Amsterdam.
- Miller G.A. (1952) 'Finite Markov Processes in Psychology', Psychometrica, No. 17, pp. 149-167.
- Wojciechowska J. (2004) 'Stigma and Discrimination as the Third Phase of AIDS', Gdansk University Press, Gdansk (in Polish).
- Tan W.Y. (2005) 'Stochastic Modeling of AIDS Epidemiology and HIV Pathogenesis', World Scientific Pub. Co.
- Tan Q.Y., Xiang Z. (1999) 'A State Space Model for the HIV Epidemic in Homosexual Populations and some Applications', Mathematical Biosciences, Vol. 152, Issue 1, pp. 29-61.

TABLES

Year/ Part	Germany				United Kingdom			
	HOMO /BI	IDU	HETERO	Sample size	HOMO /BI	IDU	HETERO	Sample size
$2000/1$ st	0.45545	0.13036	0.41419	606	0.46005	0.02405	0.51590	1289
$2000/2^{\rm nd}$	0.45522	0.12313	0.42164	804	0.43420	0.03988	0.52592	2006
$2001/1^{st}$	0.46516	0.10451	0.43033	488	0.40867	0.02927	0.56206	1845
$2001/2^{nd}$	0.46978	0.08636	0.44387	579	0.34682	0.02801	0.62517	2249
$2002/1$ st	0.48824	0.10097	0.41079	723	0.36602	0.02715	0.60683	2284
$2002/2^{nd}$	0.49872	0.07801	0.42327	782	0.34578	0.02387	0.63035	3268
$2003/1$ st	0.51573	0.07250	0.41176	731	0.31970	0.02097	0.65933	2909
$2003/2^{nd}$	0.54414	0.09841	0.35745	691	0.34123	0.01579	0.64297	3039

Table 1. HIV infections newly diagnosed in homo/bisexual men (HOMO/BI), injecting drug users (IDU), and in persons infected through heterosexual contact, in Germany and United Kingdom, 2000-2003 (half-yearly data)

Source: Authors' calculations using data from 'HIV/AIDS Surveillance in Europe', end-year and mid-year reports, 2000 – 2003.

Table 2. HIV infections newly diagnosed in homo/bisexual men (HOMO/BI), injecting drug users (IDU), and in persons infected through heterosexual contact, in Sweden andPoland, 2000-2003 (half-yearly data)

Year/ Part	Sweden				Poland			
	HOMO /BI	IDU	HETERO	Sample size	HOMO /BI	IDU	HETERO	Sample size
$2000/1$ st	0.31868	0.08791	0.59341	91	0.05759	0.87958	0.06283	191
$2000/2^{nd}$	0.36667	0.06667	0.56667	120	0.09615	0.79327	0.11058	208
$2001/1^{st}$	0.29286	0.17857	0.52857	140	0.08861	0.82911	0.08228	158
$2001/2^{nd}$	0.23853	0.11927	0.64220	109	0.06135	0.85276	0.08589	163
$2002/1$ st	0.28302	0.13208	0.58491	106	0.09848	0.86364	0.03788	132
$2002/2^{nd}$	0.26531	0.11565	0.61905	147	0.15789	0.68421	0.15789	95
$2003/1^{st}$	0.23333	0.07500	0.69167	120	0.09917	0.80165	0.09917	121
$2003/2^{nd}$	0.27869	0.09290	0.62842	183	0.02985	0.85821	0.11194	134

Source: Authors' calculations using data from 'HIV/AIDS Surveillance in Europe', end-year and mid-year reports, 2000 – 2003.

Table 3. The estimated probabilities of the transition of HIV between the populations (states) of: 1. Homo/bisexual men (HOMO/BI), 2. Injecting drug users (IDU), 3. Heterosexual persons (HETERO) for selected European countries in years 2000-2003 (half-yearly data)

Source: Authors' calculations using data from Table 1 and 2.

FIGURES

Fig.1. The transition of HIV between groups.

Fig.2. The total number of HIV infections newly diagnosed in years 2000-2003 by country (half-yearly data).

Fig.3. The fractions of HIV infections newly diagnosed in Germany in years 2000-2003 by the mode of transmission. (half-yearly data)

Fig.4. The fractions of HIV infections newly diagnosed in UK in years 2000-2003 by the mode of transmission (half-yearly data)

Fig.5. The fractions of HIV infections newly diagnosed in Sweden in years 2000-2003 by the mode of transmission (half-yearly data)

