loopbanen doorheen onderwijs naar arbeidsmarkt

Steunpunt beleidsrelevant onderzoek

Cel voor Analyse van Onderwijsloopbanen

Dekenstraat 2
B – 3000 Leuven

A linear programming approach to the estimation of pupil flows from enrollment data

G. Van Landeghem & J. Van Damme

Onderzoek in opdracht van de Vlaamse minister van Onderwijs en Vorming,
in het kader van het programma 'Steunpunten voor Beleidsrelevant Onderzoek'

2002

LOA-rapport nr. 6

Voor meer informatie omtrent deze publicatie:

Steunpunt LOA, Unit Onderwijsloopbanen

Auteur(s): G. Van Landeghem & J. Van Damme

Adres: Dekenstraat 2, 3000 Leuven
Tel.: +32 16 32 57 58 of +32 16 32 57 47
Fax.: +32 16 32 58 59
E-mail: Christel.Deno@ped.kuleuven.ac.be

Website: http://www.steunpuntloopbanen.be

A LINEAR PROGRAMMING APPROACH TO THE ESTIMATION OF PUPIL FLOWS FROM ENROLLMENT DATA

G. VAN LANDEGHEM AND J. VAN DAMME

CENTRUM VOOR SECUNDAIR EN HOGER ONDERWIJS

DEPARTEMENT PEDAGOGISCHE WETENSCHAPPEN

KATHOLIEKE UNIVERSITEIT LEUVEN

Part of this research was supported by the Ministerie van de Vlaamse Gemeenschap, Departement Onderwijs, through INES project 9604 and through the Steunpunt Loopbanen in het Onderwijs en de Overgang van School naar Werk.

Correspondence regarding this text can be sent to G. Van Landeghem, K.U.Leuven, Departement Pedagogische Wetenschappen, Centrum voor Secundair en Hoger Onderwijs, Dekenstraat 2, B-3000 Leuven, België. E-mail: Georges.VanLandeghem@ped.kuleuven.ac.be

Abstract

The enrollments of a (sub)population in two consecutive school years in a given educational system restrict the potential values of the array of transitional flows to the area within a polyhedral boundary. This boundary can be explored by means of linear programming, yielding, for example, lower and upper limits for individual flows. Any additional piece of information that can be expressed as a linear constraint on the transitional flows can be incorporated, possibly narrowing the calculated intervals. A detailed example, based on data about the educational systems in the Dutch and French speaking Communities in Belgium, demonstrates how this approach can be successfully implemented in practice. Repetition rates in primary education are derived from enrollment data per grade and year of birth. The present method is compared to earlier attempts to employ linear programming in the estimation of transition flows from enrollments.

A Linear Programming Approach to the Estimation of Pupil Flows from Enrollment Data

Introduction

Enrollments in primary education in Flanders have been counted and published annually for several decades. In contrast, a first tentative official count of the number of repeaters in Flemish primary education—concerning the school year 1996–1997—appeared not until 1998. Indeed, repetition data are still not included in the published annual official statistics on Flemish primary education at present. Therefore, we set about *estimating* the numbers of repeaters, focusing first on the period from 1984–1985 until 1994–1995 (that is: the decade after important reforms in the Flemish educational system). Those estimates (reported by Van Landeghem & Van Damme, 1997) are based on the official enrollment data and on additional information with regard to the structure of the Flemish educational system and the interactions between its constituent parts. Linear programming was employed to combine the information, yielding interval estimates for the numbers of repeaters.

The aim of the present report is to explain our approach of estimating transition rates from enrollments. In the 'Models' section, the Flemish case is used to demonstrate how relevant linear programming problems can be formulated, taking into account the structure of the educational system, the availability of enrollment data, and the inclusion of any additional information concerning pupil flows. Next, in the 'Examples' section, it is explained how those models can be set to work in practice. The section provides illustrations using data from the educational systems in the Dutch and French speaking Communities in Belgium. Finally, in the 'Alternatives' section, two related existing methods of estimating transition rates from enrollments, explained by Crouch (1991), are commented upon from the viewpoint of the present approach.

Models

In this section we construct the basis of a model which serves to estimate the number of repeaters in Flemish regular primary education when enrollments in consecutive school years are known. The primary objective of this section is to introduce our approach of estimating pupil flows from enrollment data by means of linear programming models.

A simple six grades model requiring only enrollments per grade is refined by introducing age by grade enrollments and by taking into account the interactions with special, nursery, and secondary education. Subsequently, additional information concerning the pupil flows within the Flemish educational system is utilized in order to simplify the model.

A Simple Six Grades Model

Conceptual formulation. The Flemish regular primary educational system consists of 6 consecutive grades. The majority of the children born in, say, 1990, have entered the first grade in september 1996 (at age 5 or 6, depending on their month of birth) i.e.: at the beginning of school year 1996–1997, and will leave the sixth grade at the end of june 2002 (the end of school year 2001–2002) i.e.: at age 11 or 12. Having thus stated the main feature of the regular primary educational system in Flanders, we add the following assumptions in order to complete a first model: (1) every pupil is assigned to one single grade in any given school year; (2) no pupil is allowed to skip a grade (for example: to go from grade 1 in 1996–1997 to grade 3 in september 1997); (3) backtracking is not possible (for example: a pupil from grade 2 in 1997–1998 is not allowed to return to grade 1 in september 1998); (4) pupils enter the system exclusively in the first grade; (5) pupils leave the system exclusively from the sixth grade. Note that the assumptions allow pupils to repeat grades. Note also that those assumptions imply that there are no dropouts: any pupil who enters the system is assumed to leave it after six years or more, having completed at least one school year in every grade. With compulsory education until the age of 18 in Flanders, however, the assumption that there are no dropouts from primary education is a reasonable first approximation. This first model can be represented by Figure 1.

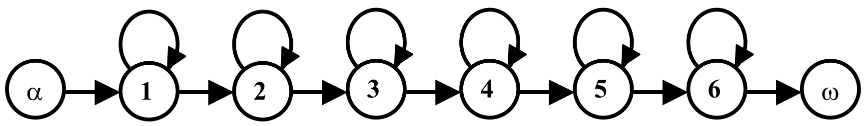

Figure 1. A six grades model.

The numbered nodes in the graph represent the six grades. The node α represents the 'source' (unspecified as yet) of the pupils entering the system for the first time; the node ω represents the 'destination' (equally unspecified) of the pupils leaving the system. The 13 arrows in the graph represent the flows of pupils at the moment of transition (summer of year t) between the end of a school year (school year '(t-1)–t', for example: 1995–1996) and the beginning of the next school year (school year 't–(t+1)', for example: 1996–1997). Specifically, the six loops represent the repeaters in school year t–(t+1) in each grade—a repeater in school year t–(t+1) thus being a pupil who is enrolled in a given grade in both school years (t-1)–t and t–(t+1); the five arrows between grades represent the pupils who move on to a higher grade; the two remaining arrows represent the pupils entering and leaving the system, respectively.

Mathematical formulation. The model implies six links between the flows at moment *t* and the enrollments per grade in school year *t*–(*t*+1), expressing that in each grade the sum of the incoming flows must be equal to the resulting enrollment:

$(i_{1,t})$ $$F_{\alpha,1,t} + F_{1,1,t} = E_{1,t}$$

$(i_{j,t})$ $$F_{j-1,j,t} + F_{j,j,t} = E_{j,t} \quad j = 2,3,4,5,6$$

Symmetrically, the model connects the flows at moment t and the enrollments in school year (*t*-1)–*t*, stating that in each grade the enrollment of the previous school year equals the sum of the outgoing flows:

$(o_{j,t})$ $$F_{j,j+1,t} + F_{j,j,t} = E_{j,t-1} \quad j = 1,2,3,4,5$$

$(o_{6,t})$ $$F_{6,\omega,t} + F_{6,6,t} = E_{6,t-1}$$

Here the *E*'s, with a first index denoting the grade and a second index denoting the school year (by means of the civil year in which the school year starts), represent the enrollments; the *F*'s, with a first index denoting the source of the flow (the external source α or one of the grades), a second index for the destination (one of the grades or the external destination ω), and a third index identifying the moment of transition, represent the flows.

Both the *E*'s and the *F*'s measure the size of groups of pupils sharing the same grade in a given school year (enrollments) or having in common the grades in two consecutive school years (flows). Consequently, the following non-negativity constraints apply:

$(n_{j,t})$ $$E_{j,t} \geq 0 \quad j = 1,2,3,4,5,6 \quad \forall t$$

$(n_{j,k,t})$ $$F_{j,k,t} \geq 0 \quad (j,k) = (\alpha,1),(1,1),(1,2),\ldots,(6,\omega) \quad \forall t$$

Those non-negativity constraints are satisfied automatically when enrollments or flows are measured directly (i.e.: by counting pupils). When, however, predictions of enrollments or estimates of flows are derived from other data, it is necessary to ensure that the constraints are properly incorporated in the algorithm used to produce those results.

Questions answered by means of linear programming. When the enrollment numbers per grade in two consecutive school years (*t*-1)–*t* and *t*–(*t*+1) of Flemish primary education are known, the 12 equations $(i_{1,t})$, ..., $(i_{6,t})$, $(o_{1,t})$, ..., $(o_{6,t})$ together with the 13 additional constraints $(n_{\alpha,1,t})$, $(n_{1,1,t})$, ..., $(n_{6,\omega,t})$ implicitly define the set of (combinations of) flow values that are allowed by our model assumptions. Geometrically, this 'set of feasible solutions' can be pictured as a region in a 13-dimensional space (one dimension for each unknown flow). (Note: we ignore the fact that in reality the flows are integer numbers, and treat the unknown flows as real numbers. As all the relevant flows are much larger than 1, the consequences of this approximation are negligible.) As the equations and inequalities defining this region are linear in the unknown flows, the region is bounded by (hyper)planes.

An explicit question about this set of feasible solutions such as : 'what is the minimal number of repeaters in the second grade allowed by the model assumptions and the given enrollment numbers' or 'what is the maximal flow from grade 3 to grade 4' or 'what is the minimal total number of repeaters', or indeed any question about a set with a polyhedral bound which inquires about a minimum (or maximum) of a *linear* combination of the unknowns (linear 'objective function'), amounts to the formulation of a so-called linear programming (LP) problem (Ravindran, Phillips, & Solberg, 1987, chapter 2).

Most LP problems nowadays are solved by means of a computer implementation of Dantzig's simplex algorithm. This algorithm finds a vertex of the polyhedron and then tracks a path from vertex to vertex, in each step choosing the neighboring vertex corresponding to the optimal rate of change in the objective function, until an optimal vertex (which is guaranteed to exist whenever any solution of the linear programming problem exists) is found.

In general, even a LP problem with only 13 variables and 25 constraints, which is a tiny problem according to LP standards, quickly becomes hard to manage without the help of the simplex algorithm to calculate and find a path along vertices of the set of feasible solutions. In the case of our present model, however, 12 of the 25 constraints are equations. Moreover, the coefficient matrix of this system of equations has rank 12. As a consequence, it is possible to eliminate 12 variables, leaving a LP problem with one variable (and, consequently, a set of feasible solutions with no more than two vertices) and 13 (inequality) constraints. Thus, although general purpose LP software can be applied in order to find, for example, an upper limit for the total number of repeaters in a given school year according to the proposed model, it is not strictly necessary. As we develop the model further, however, the simplex algorithm will quickly be found to be indispensible.

Models per Generation

Enrollment numbers by grade and by year of birth. The tables concerning regular primary education in the Statistical Yearbooks of the Education Department of the Ministry of the Flemish Community (e.g.: Ministerie van Onderwijs, 1984, or Ministerie van de Vlaamse Gemeenschap, departement Onderwijs, 1991) not only supply the enrollments per grade but also the enrollments per grade and year of birth. While the main body of a generation (meaning: all the pupils born in a given year) advances through the educational system en bloc, some of their peers take a lead and, more importantly, a growing number of pupils of the same age lag behind.

Table 1 lists the enrollments by grade and by year of birth in Flanders in 1996–1997 (Ministerie van de Vlaamse Gemeenschap, departement Onderwijs, 1997). It shows that the spread of a generation during its progress through regular primary school exhibits an orderly pattern: in each grade five different generations are present; one subgroup leads by one year with regard to the main group of 'normally progressing' pupils, three subgroups consist of pupils who have fallen behind by one, two and three years respectively. Only 100 pupils (less then 3 in 10000) fall outside this pattern in 1996–1997. The same pattern holds in 1995–1996 and earlier school years (indeed, the Yearbooks for 1983–1984 up to and including 1994–1995 do not mention any enrollments outside this pattern). When Flemish age by grade data are used in the remainder of this report, we will neglect the small number of pupils outside this pattern.

Table 1. Enrollments by Grade and by Year of Birth in the Vlaamse Gemeenschap (Belgium) in 1996–1997

Birth year	Age[a]	Grade 1st	2nd	3rd	4th	5th	6th	Birth year[b]
1992	4–5							*t*-4
1991	5–6	565	4	1				*t*-5
1990	6–7	**65821**	532	5	2	1		*t*-6
1989	7–8	7281	**58550**	545	7	1	2	*t*-7
1988	8–9	309	8295	**55464**	537	9	4	*t*-8
1987	9–10	26	496	8286	**53809**	612	13	*t*-9
1986	10–11	8	42	636	8808	**52726**	587	*t*-10
1985	11–12	4	9	72	797	8675	**50947**	*t*-11
1984	12–13		1	13	89	876	8114	*t*-12
1983	13–14			2	4	44	578	*t*-13
1982	14–15		1	2	3	4	29	*t*-14
1981	15–16							*t*-15

Data source: Ministerie van de Vlaamse Gemeenschap, departement Onderwijs (1997).
[a]Age at end of school year.
[b]*t* = 1996.

Six grades model applied to a single generation. The six grades model developed in the previous subsection can be applied to any subpopulation, e.g.: to boys, girls, an ethnic group, an age group (a generation) etc., as soon as the corresponding enrollment numbers are available. Interestingly, thanks to the pattern exhibited in Table 1, the model simplifies when applied to a single generation.

Consider, for example, the enrollment numbers of the generation born in year *t*-8 (say 1988), in two consecutive school years (*t*-1)–*t* (1995–1996) and *t*–(*t*+1) (1996–1997). According to the pattern of the age by grade data (exemplified in Table 1), the 4th, 5th and 6th grade enrollments of this generation in school year (*t*-1)–*t* are (virtually) zero; in addition, the 5th and 6th grade enrollments in the next school year are zero. In other words, the right hand sides of the equations ($o_{4,t}$), ($o_{5,t}$), ($o_{6,t}$), ($i_{5,t}$) and ($i_{6,t}$) are zero, which implies (when the non-negativity constraints are taken into account) that the flows $F_{4,4,t}$, $F_{4,5,t}$, $F_{5,5,t}$, $F_{5,6,t}$, $F_{6,6,t}$ and $F_{6,\omega,t}$ are zero. Having thus eliminated 6 of the 13 flows and 5 of the 12 equations, we are left with 7 equations for 7 unknown flows which are constrained to be non-negative. One can easily check that the (sparse) coefficient matrix of the 7x7 system has rank 7. Consequently, the system has a single solution which, if it consists of non-negative numbers, constitutes the unique solution of the model (otherwise the model has no solution, i.e.: it conflicts with the data).

A similar argument can be applied to each generation that has a number of pupils in primary education in at least one of both school years, i.e.: the generations with years of birth *t*-5, *t*-6, ..., *t*-15. It is easily verified that in each case except for the age group born in year *t*-10, the simplification yields a square system of equations with full rank, implying that the solution of the model is unique or does not exist.

A condition for the enrollment numbers. When applied to the *t*-10 generation, the mathematical formulation of our model is reduced to a system of 10 equations—namely ($o_{j,t}$), j=1,2,3,4,5 and ($i_{j,t}$), j=2,3,4,5,6—in 9 unknown flows (and the corresponding non-negativity constraints). As the system's coefficient matrix proves to have the maximal rank 9, its solution set is empty or there is a unique solution, depending on whether the 10 by 10 matrix formed by attaching the system's right hand side as an extra column to the coefficient matrix has full rank or not. The latter is true if and only if

$$\text{(c)} \qquad \sum_{\textit{all nodes } j} E_{j,t} = \sum_{\textit{all nodes } j} E_{j,t-1}$$

In this equation 'all nodes *j*' refers to all the nodes of the graph for which enrollments are available. In the present case these are the six numbered nodes representing the grades. As no pupils born in year *t*-10 are enrolled in the 6th grade in school year (*t*-1)–*t*, our model assumptions preclude any outflow of pupils of that generation from primary school. Symmetrically, the fact that the generation *t*-10 is not represented in the 1st grade in school year *t*–(*t*+1) implies that there is no inflow. In other words, for the pupils born in year *t*-10, the system of flows that underlies our model is closed (the nodes α and ω disappear from the corresponding graph). Hence the condition (c), which expresses that the total enrollment of generation *t*-10 pupils does not change during the transition. Given that condition (c) is satisfied, the model for the *t*-10 generation has, again, either a unique solution or no solution at all, depending on whether the solution of the system of equations satisfies the non-negativity constraints.

As we will illustrate at several instances in the further development of our model, condition (c) arises whenever the system of flows underlying the model is effectively closed, regardless of the details of the model. For that reason condition (c) has been formulated in a general way.

Summary. We find that the application of our model to separate generations of Flemish pupils yields a unique solution (or no solution), instead of a range of solutions (which can be quite wide, as will be demonstrated in the Examples section). The merits of what appears to be an important increase of accuracy will be challenged in the next subsection.

Interactions with Special, Nursery and Secondary Education

Flows from and to special education. In contrast with regular primary education, special primary education in Flanders is not structured as a sequence of grades. Consequently, the model developed in the previous subsections inevitably excludes the pupils of special primary education.

According to the 1985–1986 Yearbook, 15986 pupils attended special primary education in Flanders in that school year, i.e.: 3.73 % of the total population of 428614 pupils in primary education (regular plus special). The fraction of pupils in special primary education has expanded steadily since then, reaching 5.53 % (22827 out of 413022 pupils) in 1995–1996 and 5.54 % (23121 out of 417369) in 1996–1997.

This growth of the special education enrollments has been due mainly to an increase of the number of pupils attending so-called ‘type 8’ education: the number doubled in a decade, from 3560 in 1984–1985 to 7447 in 1994–1995. By comparison, the increase of enrollments in ‘type 1’ education (8275 in 1984–1985, 9603 in 1994–1995) and the growth of the total number of pupils attending other forms of special education (4219 in 1984–1985, 5020 in 1994–1995) has been moderate. The latter forms of special education are intended for pupils whose inability to develop in regular primary education or to attain an ‘elementary academic standard’ has a specific and clearly identifiable cause such as a physical disability, a moderate or severe mental disability, a long illness or a personality disorder. Pupils in ‘type 1’ education are expected to be able to attain the previously mentioned ‘elementary academic standard’ with the help of their special education program, despite a mild mental disability. ‘Type 8’ special education is aimed at pupils who do not suffer from any of the problems which point at the need for one of the other forms of special education, but nevertheless—despite their ‘normal abilities’—experience learning difficulties. Pupil flows between regular and special primary education consist mainly of pupils leaving regular education for a ‘type 8’ or ‘type 1’ education and pupils returning to regular primary education from ‘type 8’ (Moenaert, 1991). As a consequence it is likely that not only the fraction of pupils attending special education but also the interactions (pupil flows) between regular and special primary education have increased considerably. Moreover, interactions between regular education and 'type 8' education have probably been intensified by the gradual propagation of a new concept of 'type 8' education since the mid eighties, emphasizing the importance of a timely switch of pupils from regular education to 'type 8' and of a speedy return, often after only one year in special education (Dens, 1992).

It has become increasingly unrealistic, then, to ignore the interaction between regular and special primary education in a model of the regular primary educational system in Flanders. Figure 2 is an extension of Figure 1, taking into account that pupils can go over from any grade in regular primary education to special education (node 's' in the graph) and vice versa.

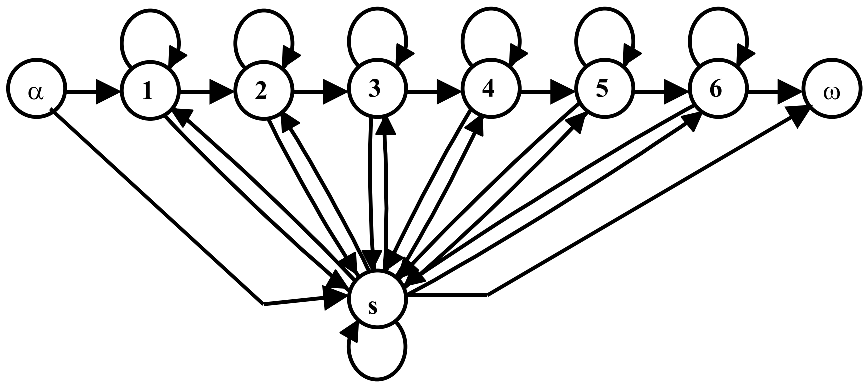

Figure 2. Taking into account special education.

Moreover, the graph reckons with the existence of a direct inflow in special education (i.e.: pupils entering in special education without a previous enrollment in regular primary education) and a direct outflow from special education (pupils leaving special education without returning to or possibly even without ever having been enrolled in regular primary education).

The (non-negative) extra flows need to be incorporated in the twelve equations describing the six grades model:

($i_{1,t}$) $F_{\alpha,1,t} + F_{1,1,t} + F_{s,1,t} = E_{1,t}$

($i_{j,t}$) $F_{j-1,j,t} + F_{j,j,t} + F_{s,j,t} = E_{j,t} \quad j = 2,3,4,5,6$

($o_{j,t}$) $F_{j,j+1,t} + F_{j,j,t} + F_{j,s,t} = E_{j,t-1} \quad j = 1,2,3,4,5$

($o_{6,t}$) $F_{6,\omega,t} + F_{6,6,t} + F_{6,s,t} = E_{6,t-1}$

If enrollments in special education are known (which is true in the case of Flanders), the large increase in the number of unknown flows (Figure 2 contains 15 new flows) is partly made up for by the availability of an equation for incoming flows and its counterpart for the outgoing flows:

$$(i_{s,t}) \quad F_{\alpha,s,t} + F_{s,s,t} + \sum_{j=1}^{6} F_{j,s,t} = E_{s,t}$$

$$(o_{s,t}) \quad F_{s,\omega,t} + F_{s,s,t} + \sum_{j=1}^{6} F_{s,j,t} = E_{s,t-1}$$

Interaction with nursery and secondary education. As the Statistical Yearbooks (e.g.: Ministerie van Onderwijs, 1984, or Ministerie van de Vlaamse Gemeenschap, departement Onderwijs, 1991) also provide the enrollment numbers in nursery and secondary education, it makes sense to ensure that any estimates of flows within or to and from regular primary education are consistent with those data. That is achieved by inserting a node 'n' (for nursery education, both regular and special) and a node 'h' (for regular and special secondary education) in the graph (Figure 3):

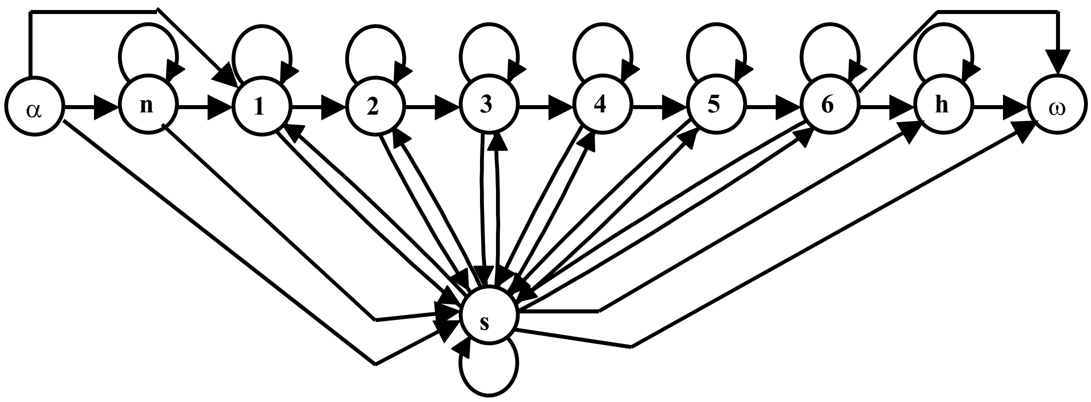

Figure 3. Including interaction with nursery and secondary education.

Although the graph already excludes flows from special primary education back to nursery education and from secondary education back to special primary education ('no backtracking'), the addition of nodes 'n' and 'h' brings eight extra unknown flows in exchange for (only) two extra pairs of equations.

A closed system. Assume, however, that we restrict our study of the transitions of pupils in the summer of year *t* to the set of pupils born in years *t*-5, *t*-6, ..., or *t*-15 (or to any subset thereof). As was illustrated by Table 1, it is fairly accurate to assume that in school year *t*–(*t*+1) only the generations born in the years *t*-5, *t*-6, ..., *t*-14 are represented in regular primary education; similarly, only the generations *t*-6, *t*-7, ..., *t*-15 are present in school year (*t*-1)–*t*. Consequently, the aforementioned restriction leaves all flows (in year *t*) within regular primary education intact, as they consist only of pupils born in years *t*-6, *t*-7, ..., *t*-14. It also leaves intact all flows to and from regular primary education (involving pupils born in years *t*-5, *t*-6, ..., *t*-15).

With a system of nursery education that is attended by virtually all Flemish children and with compulsory education from the age of 6 until 18, it is reasonable to assume in the Flemish context, as a first approximation, that all pupils born in years *t*-5, *t*-6, ..., *t*-15 are enrolled in the educational system at the moment of transition *t*. (In this statement, the term 'educational system' signifies both regular and special nursery, primary and secondary education.) The acceptability of this assumption is illustrated in Table 2. It shows, for example, that before the transition of 1985 (= *t*), 72715 of the children born in 1980 (= *t*-5) are enrolled in the educational system; this number is not much different from the enrollment of the 1980 generation in 1986–1987, which amounts to 71989 pupils, at an age when education is compulsory. Before the transition of 1995, 71821 pupils of the 1980 generation (= born in year *t*-15) are still enrolled in the educational system. (The impact of the small fluctuations in the enrollments of a given generation over the years on our model will be discussed in the Examples section).

Table 2. Enrollment in the Educational System[a] of the Generation born in 1980

School year	Age[b]	Enrollment	School year	Age[b]	Enrollment
1984–1985	4–5	72715	1990–1991	10–11	71681
1985–1986	5–6	72728	1991–1992	11–12	71785
1986–1987	6–7	71989	1992–1993	12–13	70964
1987–1988	7–8	71521	1993–1994	13–14	71702
1988–1989	8–9	71466	1994–1995	14–15	71821
1989–1990	9–10	71523	1995–1996	15–16	70265

Data sources: Ministerie van onderwijs (1985, ..., 1990) and Ministerie van de Vlaamse Gemeenschap, departement Onderwijs, (1991, ..., 1996).
[a]The 'educational system' consists of regular and special nursery, primary and secondary education.
[b]Age at end of school year.

This assumption of a complete enrollment of the generations *t*-5, *t*-6, ..., *t*-15 in the educational system reduces our model to a closed system, eliminating the flows from the source α and the flows towards the destination ω.

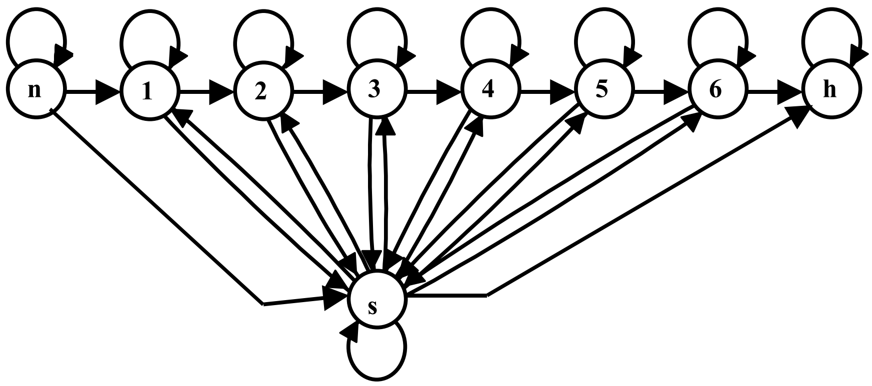

Figure 4. A closed system.

The new graph (see Figure 4), as compared to Figure 2, contains only two extra flows. The four extra equations amount to:

($i_{n,t}$) $F_{n,n,t} = E_{n,t}$

($i_{h,t}$) $F_{6,h,t} + F_{h,h,t} + F_{s,h,t} = E_{h,t}$

($o_{n,t}$) $F_{n,1,t} + F_{n,n,t} + F_{n,s,t} = E_{n,t-1}$

($o_{h,t}$) $F_{h,h,t} = E_{h,t-1}$

The equations ($i_{j,t}$) and ($o_{j,t}$) j=1,2,3,4,5,6,s are updated by substituting 'n' and 'h' for 'α' and 'ω'.

Modeling flows in a single generation. As the Yearbooks provide all the required enrollment numbers (in each grade of regular primary education, in special primary education, in nursery education and in secondary education) per generation, the model can be applied to separate age groups and can be simplified accordingly (see subsection 'models per generation'). The Yearbooks show that, with very few exceptions, which we will ignore here, pupils entering secondary education are at most one year ahead, and pupils leaving nursery education ar at most two years behind. As a consequence, nursery education is only effectively involved in the models for the generations t-8, t-7, ...; and secondary education is only involved in the models for the generations t-11, t-12, ... when flows in year t are studied.

Refined model for pupils born in year t-12 or earlier. The analysis of pupil flows (in the summer of year *t*) in separate age groups allows us to add a further refinement to the model, relaxing the assumption that pupils can leave regular primary education from the sixth grade only: pupils born in year *t*-12 or earlier can move up directly from the fifth grade to secondary education. Those pupils have fallen behind by at least one year (as compared to the majority of the pupils of their generation, who move up from the sixth grade to secondary education in year *t*). They skip the sixth grade of primary education and are enrolled in a branch of the first grade of secondary education that prepares them for a delayed entry in the general track of secondary education or for vocational training. As an example, consider Figure 5, which results when the present model is applied to the pupils born in year *t*-12.

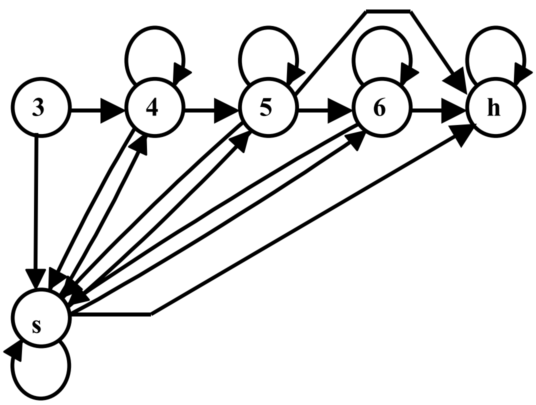

Figure 5. Flows at moment *t* for generation *t*-12.

The new flow $F_{5,h,t}$ is added to the equations ($o_{5,t}$) and ($i_{h,t}$).

Simplifying the Model

The previous subsections have been devoted to the introduction of a six grades model and its subsequent development by refining the model (taking into account special education and relaxing the restriction on the outflow), by including additional data (enrollments in nursery, secondary and special primary education and enrollments per generation), and by pointing out the advantage of generational models.

Rare transitions. In the present subsection we eliminate, in the models per generation, some flows to and from special primary education representing transitions that are rare (or better: that are thought to be so; the available information concerning flows to and from special primary education in Flanders is quite incomplete; a sizeable part of it has been assembled, commented upon and interpreted in section 2.1 of the report by Van Landeghem and Van Damme, 1997: it offers some support to the simplifications introduced in this subsection and in each case does not contradict them). The models are simplified according to the following rules. Loosely speaking, they implement the idea that pupils changing from regular to special primary education or vice versa usually lag behind their peers in regular education, that for pupils older than 12 it is usually more worthwhile to work towards secondary education within their current environment (either regular or special primary education), and that a return to regular primary education is beneficial only for pupils who are not too far behind:

1. Pupils who have just finished a grade of regular primary education with a lead of one year (Table 1) do not move over to special primary education.
2. Pupils who have just finished the 3rd, 4th or 5th grade and who are progressing normally do not move over to special primary education.
3. Pupils born in year *t*-12, *t*-13, ... who have finished the 5th or 6th grade do not move over to special primary education in year *t*.
4. Pupils moving over from special primary education to regular primary education or secondary education do not have a lead.
5. Pupils moving over from special to regular primary education are at least one year and at most two years behind.
6. Pupils born in year *t*-12, *t*-13, ... do not move over from special to regular primary education in year *t*.

Figure 6 shows the resulting graphs of two of the eleven (*t*-5, *t*-6, ..., *t*-15) generational models.

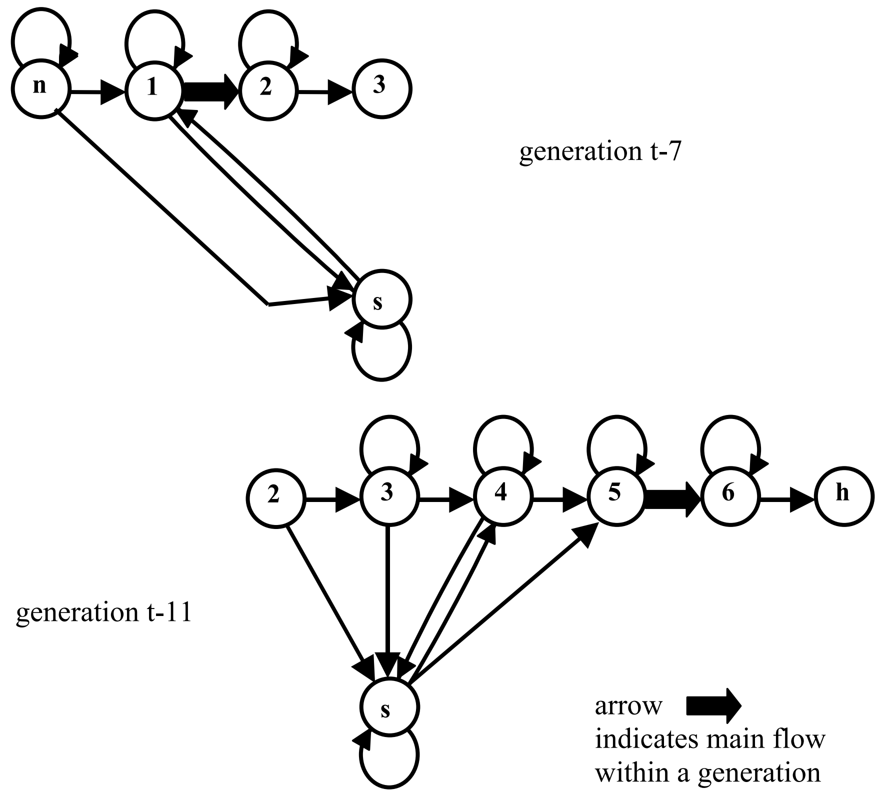

Figure 6. Graphs for generations *t*-7 and *t*-11 after simplification.

Need for LP algorithm. Table 3 lists, for the general case (flows aggregated over age) and for each age group (designated by means of year of birth): the number of equations in the model, the number of flows in the model, the rank of the system of equations and (in the row entitled 'dimension') the difference between the last two numbers, i.e.: the number of unknown flows remaining when a maximal number of flows has been eliminated by means of the available system of equations. Note that the elimination of flows by means of the equations leaves the number of inequalities (non-negativity constraints) unchanged: the inequalities are merely reformulated in terms of the remaining flows. It is clear then that an analytical or graphical solution of the model becomes unwieldy as soon as this remaining dimension is greater than two, which is the case for several age groups (and, of course, for the general model), necessitating the use of a LP algorithm.

Table 3. Characteristics of the Systems of Equations for Flows in the summer of Year *t*

Characteristic	All	*t*-5	*t*-6	*t*-7	*t*-8	*t*-9	*t*-10	*t*-11	*t*-12	*t*-13	*t*-14	*t*-15
Equations	18	3	6	9	10	11	12	12	11	9	7	5
Flows	31	2	5	10	13	13	15	15	13	10	7	4
Rank	17	2	5	8	9	10	11	11	10	8	6	4
Dimension	14	0	0	2	4	3	4	4	3	2	1	0

Note: Column 'all': model for flows aggregated over age. Column '*t-j*': model for generation born in year *t-j*. Rows: number of equations, number of flows (= number of non-negativity constraints), rank of the system, number of 'free' parameters (that is: not subject to equality constraints).

Condition for the enrollment numbers. Note also that in each model listed in Table 3 the rank equals the number of equations minus 1; as a consequence, in each case condition (c) (see 'models per generation' subsection) applies.

Brute force approach versus analytical preparation. Table 3 reflects two ways of solving a model of this kind. On the one hand, one can opt for the 'brute force' method, solving in the case of—for example—the 't-9' age group LP problems with 13 unknown flows and 24 constraints (11 equations and 13 inequalities). On the other hand, one can—in the same example—analytically express 10 flows as a function of the three remaining flows, substitute those expressions in the non-negativity constraints and finally solve LP problems with 3 unknown flows and 13 inequality constraints. And, third, one can also choose for a middle course, analytically eliminating some but not all of the flows. The brute force approach allows a fast implementation of series of LP problems for different age groups (in exchange for some more computer time, which is hardly important with LP problems of this size). Analytical preparation, however,—which is less difficult than it may seem, as the coefficient matrix of the system of equations is quite sparse—may yield interesting insights. In the age group '*t*-12', for example, the elimination of $F_{5,6,t}$ from the equations ($i_{6,t}$) and ($o_{5,t}$) yields the relationship:

$$F_{5,h,t} = E_{5,t-1} - E_{6,t} - F_{5,5,t} + F_{6,6,t}$$

which states that, when the enrollment numbers are known, information about the size of the 'special' flow of pupils going from the fifth grade directly to secondary education might be gained from data about the (difference between the) number of pupils repeating the fifth and the sixth grade.

Examples

Repeaters in Belgium

In order to assess the merit of the model constructed in the previous section, it is used here to estimate the number of repeaters in 1993 in the regular primary educational system of the Vlaamse Gemeenschap (the Dutch speaking Community in Belgium). The necessary data are listed in Tables 4 and 5. In addition, the model is applied to the 1993 transition in primary education in the Communauté française (the French speaking Community in Belgium), using the data in Tables 6 and 7. Thus, it is possible to construct lower and upper limits for the number of repeaters per grade in Belgium in 1993.

Table 4. Enrollments by Position and by Year of Birth in the Vlaamse Gemeenschap (Belgium) in 1992–1993

Birth year	Nurs.	Grade in regular primary education: 1st	2nd	3rd	4th	5th	6th	Spec. prim.	Sec.
1987	66621	553							
1986	2472	**62405**	516					1197	
1985	*129*	7060	**55152**	606				1949	
1984		328	8194	**54361**	541			2705	
1983		28	480	8426	**55632**	560		3132	
1982			47	679	8545	**56187**	649	3578	(5)
1981				81	827	8978	**57444**	3844	644
1980					78	884	8340	3333	*58329*
1979						49	574	842	69998

Note: Nurs. = Nursery education, Spec. prim. = Special primary education, Sec. = Secondary education.
The enrollment numbers for the six grades of regular primary education have been corrected (Ministerie van de Vlaamse Gemeenschap, 1997, private communication) after the publication of the Yearbook (Ministerie van de Vlaamse Gemeenschap, departement Onderwijs, 1993). The corrected data are reported here. The numbers for nursery, special primary and secondary education are as published in the Yearbook. Blank cells indicate zero's. The small enrollment between brackets has been ignored in the calculations. We also ignore the fact that, according to the Yearbook, an unknown number of the 129 pupils assigned to generation 1985 in nursery education (number in italics) may be older than indicated. Similarly, the 58329 pupils on the 1980 row in secondary education comprise a subgroup of 176 pupils (in special secondary education), an unknown number of which may be younger than indicated.

Table 5. Enrollments by Position and by Year of Birth in the Vlaamse Gemeenschap (Belgium) in 1993–1994

Birth year	Nurs.	Grade in regular primary education 1st	2nd	3rd	4th	5th	6th	Spec. prim.	Sec.
1987	2738	**62345**	548					1255	
1986	*101*	7444	**56498**	569				2011	
1985		327	7994	**53250**	625			2779	
1984		65	532	8473	**53414**	571		3297	
1983			54	665	8767	**54895**	573	3575	(4)
1982				86	841	8893	**55557**	3858	641
1981					95	936	8252	3631	*58755*
1980						53	641	816	70192
1979							30	160	71630

Source: Ministerie van de Vlaamse Gemeenschap, departement Onderwijs (1994), with subsequent slight adjustment (Ministerie van de Vlaamse Gemeenschap, 1997, private communication) of the enrollments in regular primary education.
See also notes Table 4.

Table 6. Enrollments by Position and by Year of Birth in the Communauté française (Belgium) in 1992–1993

Birth year	Nurs.	Grade in regular primary education 1st	2nd	3rd	4th	5th	6th	Spec. prim.	Sec.
1987	49329	*1139*	*(3)*	*(1)*				*(21)*	
1986	*2161*	**47261**	1185	(12)	(1)			426	
1985	37	5749	**41242**	1296	(6)	(2)		996	
1984	*(2)*	487	7273	**38887**	1239	(27)	(2)	1472	
1983		44	941	7661	**36559**	1416	(28)	1790	
1982		(16)	100	1541	8577	**36185**	1547	2094	
1981		(4)	(13)	205	2226	9394	**35309**	2319	*1124*
1980			(8)	(26)	324	2783	9187	2251	*36976*
1979			(3)	(8)	(26)	220	2129	923	47366

Source: Communauté française de Belgique, Ministère de l'Education, de la Recherche et de la Formation (1994).
See also notes Table 4.

Table 7. Enrollments by Position and by Year of Birth in the Communauté française (Belgium) in 1993–1994

Birth year	Nurs.	Grade in regular primary education 1st	2nd	3rd	4th	5th	6th	Spec. prim.	Sec.
1987	*2246*	**47891**	1085	(3)				451	
1986	37	5626	**43318**	1081	(5)			1007	
1985	*(4)*	463	7094	**39519**	1129	(10)		1453	
1984		34	965	7792	**37929**	1190	(15)	1836	
1983		(10)	94	1407	8100	**35645**	1335	2068	
1982		(10)	(23)	165	2108	9314	**34952**	2337	*979*
1981		(3)	(6)	(19)	297	2573	9112	2388	*37232*
1980		(1)	(2)	(8)	(43)	223	2172	981	48631
1979					(4)	(13)	61	157	50980

Source: Communauté française de Belgique, Ministère de l'Education, de la Recherche et de la Formation (1995).
See also notes Table 4.

Closed system? In the previous section it was assumed provisionally that all the members of each of the nine generations relevant to our example (years of birth 1979 to 1987) are enrolled in 1992–1993 and in 1993–1994. In actual fact, of course, the equilibrium condition (c) is satisfied only approximately. The 1984 generation in the Vlaamse Gemeenschap, for example, has 223 enrollments less in 1992–1993 than in 1993–1994 (this can be verified by means of Tables 4 and 5). The difference is the net result of migration (between the educational systems within Belgium and between Belgium and foreign countries), counting errors and mortality. When the equilibrium condition is not met exactly, the model has no solution (the set of potential values of the array of flows is empty). The ideal remedy for this problem, which lies outside the scope of the present illustration, would be to acquire information about the migratory flows (if available) and mortality, to incorporate those flows into the model and to correct the counting errors (which is probably infeasible). Instead, we have applied an ad hoc cure by adding either a source (in the case of the 1984 generation in the Vlaamse Gemeenschap, and in most other cases, where the 1992–1993 enrollment is smaller than the 1993–1994 enrollment) or a destination node (when the 1992–1993 enrollment is larger than the 1993–1994 enrollment) to the model. The source (destination) provides (receives) a number of pupils which is exactly equal to the difference between the 1992–1993 and 1993–1994 enrollments. Flows from the source to every node with a non-zero enrollment in 1993–1994 (or, in the models with a destination node, from every node with a non-zero enrollment in 1992–1993 to the destination) are added to the model. No assumptions regarding the relative sizes of those flows are imposed (but their sum is fixed).

Skipping a grade? A basic assumption of the model developed in the previous section is that no pupil is allowed to skip a grade. In the present example we have allowed a slight deviation of this assumption in the case of the 1984 generation in the Vlaamse Gemeenschap. The number of pupils of this generation who are one year ahead in 1993–1994 (571 pupils in the fifth grade, see Table 5) is larger than the leading group in 1992–1993 (541 pupils in the fourth grade, Table 4). The source of 223 pupils that was added to the model in order to satisfy the equilibrium condition (c) is also sufficient to prevent the model from breaking down as a consequence of this increase in the leading group: it can provide the 30 pupils needed to explain the increase. It seems, however, unlikely that pupils migrating into the educational system would be one year ahead (considering problems of adaptation). Therefore we preferred to model the increase in the leading group by building into the model a fixed flow of 28 pupils going from grade three to grade five, skipping the fourth grade. (The two remaining pupils constitute what the source can deliver when its flows are distributed according to the enrollments in the destination nodes.) Similar small adjustments have been made in the models for the 1986, 1985 and 1983 generations in Flanders, with 51 pupils skipping grade two, 18 pupils skipping grade three, and 11 pupils skipping grade five, respectively.

Results. For each of both Communities and for each one of the nine relevant generations, linear programming problems have been set up according to the models of the previous section (with the above adjustments) and solved, twice for each grade, to obtain a lower and upper limit for the number of repeaters. Table 8 reports the results, aggregated over the generations.

Table 8. Percentages of Repeaters in Regular Primary Education in Belgium in 1993

Source	Grade in regular primary education						
	1st	2nd	3rd	4th	5th	6th	Total
Vlaamse Gemeenschap							
Raw upper limit	12.0	13.9	15.6	16.1	16.0	14.7	14.6
Modeled upper limit	9.0	3.9	3.1	3.4	3.2	14.0	6.2
Official	–	–	–	–	–	–	–
Modeled lower limit	5.5	0.0	1.5	1.2	1.2	0.0	1.6
Communauté française							
Raw upper limit	13.4	17.8	21.3	23.7	27.6	27.1	21.6
Modeled upper limit	10.6	6.5	6.9	8.2	10.0	25.0	11.1
Official	6.8	4.7	3.8	4.2	4.7	3.2	4.6
Modeled lower limit	5.7	1.0	2.5	2.1	3.5	0.8	2.6
Belgium							
Raw upper limit	12.6	15.6	18.1	19.4	21.0	19.9	17.7
Modeled upper limit	9.7	5.1	4.8	5.5	6.1	18.7	8.3
Official	–	–	–	–	–	–	–
Modeled lower limit	5.6	0.5	1.9	1.6	2.1	0.3	2.0

The repeaters are the pupils who attend the same grade in both school years 1992–1993 and 1993–1994. In this table, the number of repeaters is expressed as a percentage of the enrollment in 1993–1994. For each of those *percentages of repeaters* the corresponding *repetition rates* (number of repeaters expressed as percentage of enrollment in 1992–1993) can be calculated, when required (via the data in Tables 4-7). In the first six columns, the number of repeaters in a given grade is compared to the enrollment in that grade. In the last column, the total number of repeaters in regular primary education is compared to the total enrollment in regular primary education.
The results in the first third of the Table, concerning the Vlaamse Gemeenschap, are based on the data in Tables 4 and 5. The limits in the middle part, for the Communauté française, are based on Tables 6 and 7. Source of the offical numbers of repeaters for the Communauté française: Communauté française de Belgique, Ministère de l'Education, de la Recherche et de la Formation (1995). The results in the lower third of the Table, for Belgium, were derived by combining the results for the Dutch and French speaking Communities; we ignored the contribution of the Deutschsprachigen Gemeinschaft (German speaking Community), which constitutes less than one percent of the enrollment in regular primary education in Belgium.
A ‘raw upper limit’ is simply the minimum of the two corresponding enrollments (in the two consecutive school years). (The ‘raw lower limit’ is zero.) The modeled lower and upper limits have been calculated by (separately) applying the final model of the Models section to the data of the two Communities.

As is shown in Table 8, official percentages of repeaters per grade in 1993 are available in the case of the Communauté française. Those official percentages lie neatly within the intervals calculated on the basis of our model. In the second, third, fourth and fifth grade, the intervals for the Vlaamse Gemeenschap (for which no official 1993 repeaters data exist) lie below the official data of the Communauté française. In other words: according to our model, there was less repetition (in those grades) in the Vlaamse Gemeenschap than in the Communauté française. Note that in the fifth grade, the intervals calculated for the two Communities do not overlap.

In Table 9, repeaters data (as provided by the UNESCO) for a selection of different countries in 1990 are shown. They serve to compare the 1993 results for Belgium in Table 8, which were obtained by aggregating the results calculated separately for the two main Communities in Belgium. Table 9 appears to indicate that the percentage of repeaters varies substantially between countries, even within Europe. Note, for example, that, apart from a slight deviation in the second grade, the percentages for France lie within the intervals calculated for Belgium, while the percentages for Finland lie below the lower limits for Belgium.

Table 9. Percentages of Repeaters in Primary Education in 1990, UNESCO data

	Grade in primary education						
Population	1st	2nd	3rd	4th	5th	6th	Total
Belgium	10.8	14.5	16.2	17.3	18.9	17.6	15.8
Cameroon	34.5	24.7	32.3	24.6	28.3	30.6	29.7
Finland	0.9	0.6	0.3	0.1	0.1	0.1	0.4
France	7.5	5.6	4.2	3.8	3.8	–	5.0
Indonesia	16.7	11.9	10.3	7.9	6.3	1.3	9.7
Italy	1.1	0.9	0.6	0.6	0.8	–	0.8
Mexico	17.6	11.2	9.1	7.5	5.2	0.9	9.4
Morocco	13.8	12.6	13.8	12.1	9.3	0.0	11.1

The repeaters are the pupils who attend the same grade in the school years 1989–1990 and 1990–1991. The number of repeaters is expressed as a percentage of the enrollment in 1990–1991. The percentages were calculated from the enrollments (table labeled ‘Pupils enrolled by grade and sex’) and numbers of repeaters (‘Repeaters by grade and sex’) available on the UNESCO website (url http://www.unesco.org, september 2001). The percentages of repeaters can also be found in the UNESCO Statistical Yearbooks; see, for example UNESCO (1997), p. 3–149 for the Belgian data. The 1990 data are the most recent repeaters data provided by the UNESCO for Belgium.

Oddly, Table 9 also contains data for Belgium in 1990, although no repeaters percentages for 1990 have been published by the authorities of the Vlaamse Gemeenschap in Belgium. Moreover, as was already remarked by Crouch (1991, p. 267) on the basis of data from the same source for 1983, the repeaters percentages published by the UNESCO for Belgium are strikingly high. They are more in line with the percentages for developing countries such as Indonesia and Cameroon than with the other European data. Also, they are quite outside the 1993 intervals (and similar intervals for 1990) calculated with our model. This anomaly is easily explained, however: the UNESCO data published for Belgium are not the numbers of repeaters but rather the numbers of pupils lagging behind. (The percentages of pupils lagging behind in 1990–1991, as derived from Ministerie van de Vlaamse Gemeenschap, departement Onderwijs, 1991 and Communauté française de Belgique, Ministère de l'Education, de la Recherche et de la Formation, 1992, exactly match the percentages in Table 9.) Note that, in contrast, the UNESCO percentages of repeaters for Finland (Table 9) are correct (Ministry of Education, 1998, and National Board of Education, 1998).

Repeaters in Flanders

The construction of the model underlying the results for the Vlaamse Gemeenschap in Table 8 already requires a fairly detailed knowledge of the Flemish system of primary education. On the other hand, the assumptions on which the model is based are thought to be acceptable to most people who are sufficiently familiar with this educational system. Therefore, we venture that the lower and upper bounds reported in Table 8 for the Vlaamse Gemeenschap are quite sound. In addition, the intervals between those bounds are significantly more narrow than the 'raw' intervals (Table 8), except in the sixth grade. Thus, the example of Table 8 demonstrates that it is possible, by means of linear programming, to elicit useful information about transitional flows from the combination of enrollments (per grade and by age) in two consecutive years with a detailed knowledge of the structure of the educational system.

Adding assumptions. We applied the basic model described above to estimate percentages of repeaters in the Vlaamse Gemeenschap in the school years 1984–1985 to 1994–1995 (see Van Landeghem & Van Damme, 1997). In addition, survey data that document parts of the interaction between regular and special primary education in particular years and some counting data concerning the flows from primary to secondary education in one particular year were modeled and extrapolated (guided by enrollment data available in the Yearbooks), in order to provide additional constraints for each moment of transition. (A detailed description of this work can be found in Van Landeghem & Van Damme, 1997.) For each year (1984 to 1994) the constraints proved consistent with the basic model. (That is: there were feasible solutions left when the constraints were added to the model.) Also, they turned out to cut down significantly the area of feasible solutions, narrowing the intervals for the numbers of repeaters. Table 10 shows partial results, for 1993.

Table 10. Percentages of Repeaters in Regular Primary Education in the Vlaamse Gemeenschap (Belgium) in 1993

Source	Grade in regular primary education						Total
	1st	2nd	3rd	4th	5th	6th	
Raw upper limit	12.0	13.9	15.6	16.1	16.0	14.7	14.6
Modeled upper limit (A)	9.0	3.9	3.1	3.4	3.2	14.0	6.2
Modeled upper limit (B)	8.6	3.9	3.1	2.5	3.2	14.0	5.9
Modeled upper limit (C)	8.6	3.8	2.9	2.1	2.7	13.7	5.7
Modeled upper limit (D)	8.6	3.8	2.9	2.1	2.7	6.9	4.5
Official	–	–	–	–	–	–	–
Modeled lower limit (D)	7.8	2.9	2.5	2.0	2.7	0.4	3.1
Modeled lower limit (C)	7.8	2.9	2.5	2.0	1.6	0.0	2.9
Modeled lower limit (B)	7.6	2.4	1.7	1.4	1.4	0.0	2.5
Modeled lower limit (A)	5.5	0.0	1.5	1.2	1.2	0.0	1.6

The repeaters are the pupils who attend the same grade in both school years 1992–1993 and 1993–1994. The number of repeaters is expressed as a percentage of the enrollment in 1993–1994. The modeled lower and upper limits (A) have been calculated by applying the final model of the Models section to the data in Tables 4 and 5. Limits (B) were obtained by adding constraints involving flows between nursery, regular primary and secondary education on the one hand and special primary education on the other hand. The narrower limits (C) result from additionally fixing each individual flow from the external source into the system (or from the system to the external destination node), instead of merely keeping the total flow fixed. The interval for the sixth grade is narrowed considerably—limits (D)—by adding partial information about the flows between primary and secondary education. (A detailed description of the constraints can be found in Van Landeghem & Van Damme, 1997).

Official repeater counts in Flanders. A new pupil based system for the acquisition of enrollment data was introduced gradually during the nineties in Flanders. It provides, theoretically, the necessary information to count the number of repeaters by grade and by age in primary education. This innovation was first embodied in a publication about repeaters in 1996–1997 (Ministerie van de Vlaamse Gemeenschap, departement Onderwijs, 1998). Table 11 compares estimates based on our linear programming approach to official data about the 1996 transition.

Table 11. Percentages of Repeaters in Regular Primary Education in the Vlaamse Gemeenschap (Belgium) in 1996

Source	Grade in regular primary education						Total
	1st	2nd	3rd	4th	5th	6th	
Raw upper limit	11.1	13.8	14.6	16.1	16.3	15.6	14.5
Modeled upper limit	7.5	3.9	2.6	2.7	2.6	14.6	5.6
Official	5.4	2.7	1.6	1.3	1.2	0.7	2.3
Modeled lower limit	4.0	0.0	1.4	1.1	1.1	0.0	1.3

The repeaters are the pupils who attend the same grade in both school years 1995–1996 and 1996–1997. The number of repeaters is expressed as a percentage of the enrollment in 1996–1997. The modeled lower and upper limits have been calculated by applying the final model of the Models section to enrollment data from the Ministerie van de Vlaamse Gemeenschap, departement Onderwijs (1996, 1997). Source of the official data: Ministerie van de Vlaamse Gemeenschap, departement Onderwijs (1998).

The survey data on which the narrower interval estimates (B, C, D) in Table 10 are based can probably not be extrapolated safely to the 1996 transition. In addition, the official information about the 1996 transition (which was based on the 88% of the pupils for which the information about the position before 1996 was deemed sufficiently reliable) exhibits a number of inconsistencies that may be due to teething troubles in the new enrollment data acquisition system. Therefore, we have not attempted to drive the comparison beyond what is reported in Table 11.

Alternatives

In this section two existing methods of estimating pupil flows from enrollment data are discussed, namely: Schiefelbein's method (as described by Crouch, 1991) and the method advanced by Crouch (1991). Their approaches are described here in a way that is aimed at eliciting their close connection with *and* differences from the method presented in the previous sections.

The model of the regular primary educational system in Flanders that is summarized by Figure 2 serves as the starting point of this description. The removal of the flows from special to regular primary education (and the assumption that the enrollments in special primary education are not available) yields the model depicted in Figure 7.

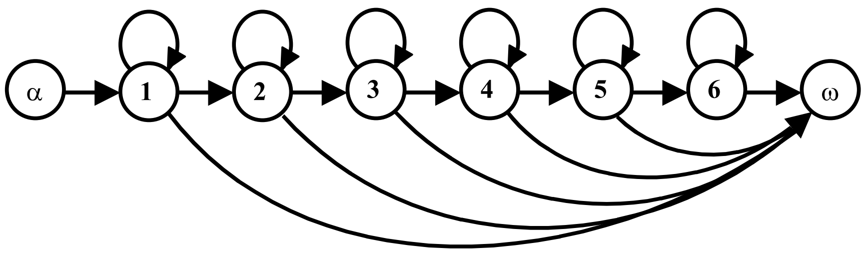

Figure 7. Graph corresponding to Crouch's (1991) model.

The graph in Figure 7 is the primary educational systems' outline that is at the center of Crouch's (1991) article. It takes into account explicitly the possibility of prematurely dropping out of the system, which is typical of the situation in developing countries. As the flows from special to regular primary education in Flanders are known to be much smaller than the flows from regular to special primary education (see Van Landeghem & Van Damme, 1997), Figure 7 also still represents a reasonable approximation of the Flemish primary educational system. (When the graph is interpreted in the latter way, the pupils flowing from the grades 1 to 5 towards the destination node ω are not dropping out of the educational system but entering special education.)

When the enrollments in the six grades in two consecutive years are known, the graph in Figure 7 gives rise to a system of 12 equations and 18 non-negativity constraints for 18 unknown flows. Following Crouch (1991), we will assume that, apart from the enrollments, the number P_t of children that are eligible for entry in primary education at time t is known, and add a (19th) constraint to the system:

$$(c_{\alpha,1,t}) \qquad F_{\alpha,1,t} \le P_t .$$

The system of 12 equations has full rank. Imitating Crouch (1991), we use six equations (namely ($o_{6,t}$) and ($i_{j,t}$), j = 2, 3, 4, 5, 6) to express the flows $F_{1,2,t}$, $F_{2,3,t}$, $F_{3,4,t}$, $F_{4,5,t}$, $F_{5,6,t}$, and $F_{6,\omega,t}$ in terms of the enrollments and the other flows. After their substitution in the remaining equations and inequalities, those six expressions can be set apart. We are then left with a system of six equations and 19 constraints for 12 flows. It is written down below in the formulation used by Crouch (1991), that is: expressed in terms of repetition rates, dropout rates, and an entrance rate:

$$(i_{1,t}) \qquad E_{1,t} = e_t P_t + r_{1,t} E_{1,t-1}$$

$$(o_{j-1,t}) \qquad E_{j,t} = r_{j,t} E_{j,t-1} + (1 - r_{j-1,t} - d_{j-1,t}) E_{j-1,t-1} \qquad j = 2, 3, 4, 5, 6$$

$$(n_{j,j,t}) \qquad r_{j,t} \ge 0 \qquad j = 1, 2, 3, 4, 5, 6$$

$$(n_{j,\omega,t}) \qquad d_{j,t} \ge 0 \qquad j = 1, 2, 3, 4, 5$$

$$(n_{\alpha,1,t}) \qquad e_t \ge 0$$

$$(n_{j-1,j,t}) \qquad r_{j,t} \le E_{j,t} / E_{j,t-1} \qquad j = 2, 3, 4, 5, 6$$

$$(n_{6,\omega,t}) \qquad r_{6,t} \le 1$$

$$(c_{\alpha,1,t}) \qquad e_t \le 1 .$$

with the unknown repetition rates (r's), dropout rates (d's) and entrance rate defined by:

$$r_{j,t} = F_{j,j,t} / E_{j,t-1} \qquad j = 1, 2, 3, 4, 5, 6$$

$$d_{j,t} = F_{j,\omega,t} / E_{j,t-1} \qquad j = 1, 2, 3, 4, 5$$

$$e_t = F_{\alpha,1,t} / P_t .$$

Table 13 illustrates (by means of the 1992 and 1993 transitions in the Communauté française, using the data in Table 12, which are aggregated over age) that this system of equations and constraints does not necessarily provide much information about the repetition rates: the modeled upper and lower limits are hardly narrower than the raw limits.

Table 12. Enrollments and Number of Children Eligible for Entry in Primary Education in the Communauté française (Belgium)

		Grade in regular primary education					
Sex	Eligible	1st	2nd	3rd	4th	5th	6th
				1990–1991			
Girls	23655	25884	23927	24643	24932	25191	23257
Boys	25547	27249	25457	25756	25663	25984	23759
				1991–1992			
Girls	24304	25768	24759	23726	24525	25167	23896
Boys	26703	27518	25517	25312	25455	25879	24275
				1992–1993			
Girls	24732	26460	24734	24484	23724	24607	24063
Boys	26797	28241	26034	25153	25238	25431	24224
				1993–1994			
Girls	–	26720	25667	24384	24513	23816	23610
Boys	–	28424	26921	25611	25102	25152	24037

Source: Communauté française de Belgique, Ministère de l'Education, de la Recherche et de la Formation (1992, 1993, 1994, 1995). Some of the column totals in Tables 6 and 7 are smaller than the corresponding enrollments listed here, because Tables 6 and 7 do not show all generations.

The number of children eligible for entry in 1991 (1992, 1993) is defined as the 1990–1991 (1991–1992, 1992–1993) enrollment in ordinary or special preschool education of children born in 1985 (1986, 1987). (This definition of the number of children eligible for entry is not beyond discussion, but that does not affect the methodological point we make in this text.)

Table 13. Repetition Rates (in percentages) in the Communauté française (Belgium)

Source	Year	Sex	Grade in regular primary education					
			1st	2nd	3rd	4th	5th	6th
Raw upper limit	1992	G + B	100.0	100.0	100.0	98.0	98.0	100.0
Raw upper limit	1993	G + B	100.0	100.0	100.0	100.0	97.9	98.7
Modeled upper limit (Fig. 7)	1992	G + B	98.0	98.8	100.0	98.0	98.0	100.0
Modeled upper limit (Fig. 7)	1993	G + B	94.8	98.0	98.7	100.0	97.9	98.7
Modeled lower limit (Fig. 7)	1992	G + B	6.9	2.3	1.1	0.9	1.0	0.0
Modeled lower limit (Fig. 7)	1993	G + B	6.6	3.0	1.5	1.4	1.4	0.0
Official	1992	G + B	7.4	4.7	4.1	4.5	5.0	3.2
Official	1993	G + B	6.8	4.9	3.8	4.2	4.6	3.1
Schiefelbein (0.9)	1993	G + B	7.3	4.3	3.3	3.7	4.2	1.2
Schiefelbein (0.6)	1993	G + B	4.5	2.5	2.6	4.4	7.4	4.3
Crouch	1992	Girls	8.4	4.7	3.8	21.4	22.1	28.7
Crouch	1992	Boys	46.6	44.4	43.3	42.8	43.1	39.9
Crouch	1992	G + B	27.0	23.7	23.0	22.4	23.0	18.6
Crouch	1993	Girls	8.4	5.7	4.8	5.1	5.3	11.8
Crouch	1993	Boys	13.0	9.0	7.7	7.5	27.4	62.9
Crouch	1993	G + B	6.9	3.3	2.1	2.1	2.1	29.4

The repeaters in 1992 (1993) are the pupils who attend the same grade in both school years 1991–1992 and 1992–1993 (1992–1993 and 1993–1994)). The number of repeaters is expressed as a percentage of the enrollment in 1991–1992 (1992–1993).
The modeled upper and lower limits (lines 3 to 6) are defined by the model of six equations and 19 constraints for twelve flows based on Figure 7 (see text). Together with the raw upper limits (lines 1 and 2), they are based on the data in Table 12 (after aggregation over sex).
The official repetition rates have been calculated from the enrollments and numbers of repeaters in: Communauté française de Belgique, Ministère de l'Education, de la Recherche et de la Formation (1994, 1995).
Schiefelbein's method (see text) was applied to the age by grade data in Tables 6 and 7. The ratio of the number of repeaters to the number of pupils not promoted was set equal to 0.9 in the first and to 0.6 in the second estimate. (Backward calculation for the generations 1984 and younger, forward calculation for the generations 1983 and older, see text.)
Crouch's method was applied to the data in Table 12, three times for each of the two years (first girls, then boys, then aggregated enrollments).

As explained in the Models and Examples sections of this text, there seems to be only one way to improve those estimates of the repetition rates, namely: to acquire more data, insert them into the model, and check that they narrow the intervals of allowed values for the repetition rates. The next subsections describe the methods applied by Schiefelbein and by Crouch to improve the estimates.

Schiefelbein's Method

The approach of Schiefelbein that is described by Crouch (1991) does indeed draw in a large amount of data to make the most of the system of equations and constraints based on Figure 7. First, it requires age by grade enrollment numbers. Secondly, this method requires additional inputs that are sufficient (when all the data are exact) to leave a unique solution. It is, for example (and as suggested by Crouch's account on pp. 261-262), sufficient to provide (for every generation!) one of the repetition rates and the five ratios

$$f_{g,j,t} = \frac{r_{g,j,t}}{r_{g,j,t} + d_{g,j,t}} \qquad j = 1, 2, 3, 4, 5$$

The index g (denoting the birth year) has been added to indicate explicitly that the rates and ratios refer to a particular generation. Here $f_{g,j,t}$ is the ratio of the number of pupils of generation g who start to repeat grade j at time t, to the number of pupils of generation g who do not go up from grade j to grade j+1 at time t. Specifically, when $r_{g,k,t}$ is given, the repetition rates for higher grades are uniquely determined by

$$r_{g,j+1,t} = \left(\frac{r_{g,j,t}}{f_{g,j,t}} - 1 \right) \frac{E_{g,j,t-1}}{E_{g,j+1,t-1}} + \frac{E_{g,j+1,t}}{E_{g,j+1,t-1}}$$
$$j = k, k+1, \ldots, 5$$

and the rates for lower grades are determined by:

$$r_{g,j,t} = f_{g,j,t} \left(r_{g,j+1,t} \frac{E_{g,j+1,t-1}}{E_{g,j,t-1}} + 1 - \frac{E_{g,j+1,t}}{E_{g,j,t-1}} \right)$$
$$j = 1, 2, \ldots, k-1.$$

In order to soften those data requirements (which, except for a transformation, amount to demanding the repetition rates themselves directly) one may, for example, approximate all the ratios $f_{g,j,t}$ (for all generations g and grades j) by a single ratio f. When this or some less radical simplification is adopted, or when some of the other data are not exact, it is possible that the rates calculated by means of the formulas above lie outside the region of feasible solutions. In the approach of Schiefelbein described by Crouch, ad hoc measures are suggested to force the estimates to respect (most of) the inequality constraints.

As an example, we have implemented two versions of this approach. First, the 'forward algorithm' starts with a given rate $r_{g,1,t}$ that satisfies

$$0 \le r_{g,1,t} \le \frac{E_{g,1,t}}{E_{g,1,t-1}}.$$

The repetition rates for the next five grades are calculated according to the following algorithm:

For $j = 1, 2, 3, 4$:

$$r_{g,j+1,t} = \max\left\{0, \min\left\{\left(\frac{r_{g,j,t}}{f} - 1\right)\frac{E_{g,j,t-1}}{E_{g,j+1,t-1}} + \frac{E_{g,j+1,t}}{E_{g,j+1,t-1}}, \frac{E_{g,j+1,t}}{E_{g,j+1,t-1}}\right\}\right\}$$

$$r_{g,6,t} = \max\left\{0, \min\left\{\left(\frac{r_{g,5,t}}{f} - 1\right)\frac{E_{g,5,t-1}}{E_{g,6,t-1}} + \frac{E_{g,6,t}}{E_{g,6,t-1}}, \frac{E_{g,6,t}}{E_{g,6,t-1}}, 1\right\}\right\}.$$

Secondly, the 'backward algorithm' starts with a given rate $r_{g,6,t}$ that satisfies

$$0 \leq r_{g,6,t} \leq \min\left\{\frac{E_{6,t}}{E_{6,t-1}}, 1\right\}.$$

The repetition rates for the five previous grades are calculated according to the following algorithm:

For $j = 5, 4, 3, 2, 1$:

$$r_{g,j,t} = \max\left\{0, \min\left\{f\left(r_{g,j+1,t}\frac{E_{g,j+1,t-1}}{E_{g,j,t-1}} + 1 - \frac{E_{g,j+1,t}}{E_{g,j,t-1}}\right), \frac{E_{g,j,t}}{E_{g,j,t-1}}\right\}\right\}.$$

Both algorithms force the estimate to satisfy the constraints ($n_{g,1,1,t}$), ..., ($n_{g,6,6,t}$), ($n_{g,\alpha,1,t}$), ($n_{g,1,2,t}$), ..., ($n_{g,5,6,t}$), and ($n_{g,6,\omega,t}$).

The above implementation of Schiefelbein's approach works well in the case of the 1993 repetition rates in the Communauté française, as is shown in Table 13 (where the official 1993 rates are also reported). The reported estimates have been obtained by aggregating the output rates of the above algorithms over the generations, producing estimated rates per grade.

(According to Crouch, 1991, this may smooth out irregularities in the age by grade estimates. Although the approach of Schiefelbein explained by Crouch makes use of age by grade enrollments, its ultimate aim is to produce an accurate estimate of the repetition rate in each grade. It does not claim to yield accurate estimates of the age by grade rates.) The backward algorithm was applied for the generations 1984 and younger, the forward calculation for the generations 1983 and older. This strategy exploits the structure of the Belgian age by grade enrollment data (see Models section, especially Table 1), which is particularly favourable for Schiefelbein's method: the starting rate $r_{g,6,t}$ (backward calculation, younger pupils) or $r_{g,1,t}$ (forward calculation, older pupils) is known to be exactly zero. The availability of accurate starting values seems to make for estimates of the right magnitude. In this example, the use of a single ratio f does not seem to have adverse effects on the estimated rates (except, perhaps in the sixth grade). The contrast between the $f = 0.9$ rates and the (much less accurate) $f = 0.6$ rates in Table 13 indicates that it is necessary to have, apart from the age by grade enrollments, sufficiently accurate information about f to obtain useful estimates of the repetition rates.

The method of Schiefelbein described by Crouch (1991) always yields a solution. This can be seen as a strength, as it spares the user the disappointment of being left without a result, but it is also a weakness: even in the case of conflicting data, (possibly very inaccurate) estimates of the repetition rates are produced anyhow. In contrast, our approach (see Models and Examples sections) is more prudent, in that it acknowledges that the set of feasible solutions may be empty and in that case alerts the user to conflicts in the data. Secondly, our approach is more flexible: it allows the user to enter the data gradually, in order of reliability. The interval estimates indicate when sufficient accuracy has been achieved. In that way it may be possible, in some instances, to avoid using the less reliable data, whereas Schiefelbein's method (Crouch, 1991) always requires a full set of inputs (however inaccurate) that pin-points a single solution.

Crouch's Method

Crouch has attempted to use the system of equations and constraints based on Figure 7 without the need for enrollments by age and grade, because such data are not available in many countries (Crouch, 1991, p. 262). Instead, Crouch proposes to add information by requiring the enrollments (and the number of children eligible for entry) of an additional school year. In the case of the estimation of the 1993 repetition rates in the Communauté française, for example, the enrollments of 1991–1992, 1992–1993, and 1993–1994 (available in Table 12) are required. As explained above, the 13 givens for 1992–1993 and 1993–1994 (12 enrollments and the number eligible for entry) act as the coefficients in a system (call it the 1993 system) of six equations and 19 constraints in the 12 unknown rates for the 1993 transition. The joint data for 1991–1992 and 1992–1993 are the coefficients of the (otherwise identical) 1992 system of six equations and 19 constraints in the 12 unknown 1992 rates. Next, Crouch couples the two systems by adding the assumption that the rates in the consecutive years are equal (in the example: each 1992 rate is equal to its 1993 counterpart). After the elimination of the 1992 rates by means of this equality assumption, the overall system has 12 equations (mutually independent, apart from exceptional circumstances) for 12 unknown rates. Therefore it yields a unique solution or no solution at all (depending on the constraints). Indeed, no solution is found in the above case of the Communauté française, despite the fact that the 1992 and 1993 regions of feasible solutions are large (at least in certain directions, see modeled upper and lower limits in Table 13) and close together (as the enrollments do not vary much between consecutive years). This situation can be graphically illustrated in the case of a single grade model (Figure 8).

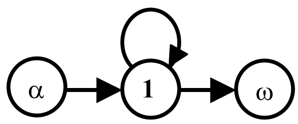

Figure 8. A single grade model.

The overall system of equations of Crouch's approach, when applied to the graph of Figure 8, has two variables: $r_{1,t}$ and e_t. Figure 9 shows the corresponding t = 1992 and t = 1993 sets of feasible solutions—truncated lines in the ($r_{1,t}$, e_t) plane—for girls in the Communauté française.

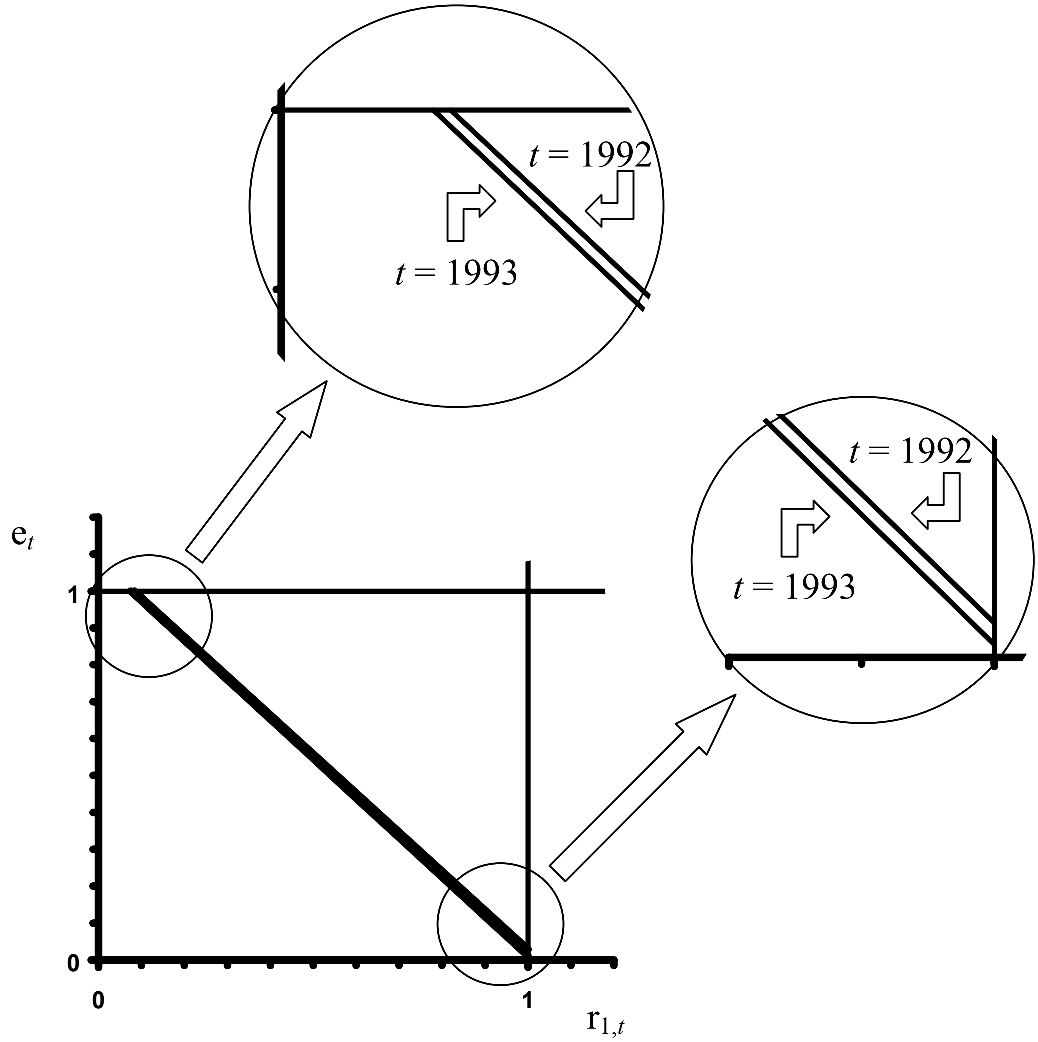

Figure 9. Sets of feasible solutions according to the single grade model for t = 1992 and t = 1993 for girls in the Communauté française (data in Table 12).

To ensure the existence of a solution, Crouch relaxes the 12 equations minimally, that is: just enough to allow a solution that satisfies the constraints. (In Figure 9, in the case of a single grade system, this means that the two truncated lines representing the 1992 and 1993 sets of feasible solutions are thickened until they touch in one point.) This relaxation can be implemented in several ways. Crouch's implementation allows the quantities E*, which are defined as

$$E^*_{1,t} = e_t P_t + r_{1,t} E_{1,t-1}$$

$$E^*_{1,t-1} = e_t P_{t-1} + r_{1,t} E_{1,t-2}$$

$$E^*_{j,t} = r_{j,t} E_{j,t-1} + (1 - r_{j-1,t} - d_{j-1,t}) E_{j-1,t-1}$$

$$E^*_{j,t-1} = r_{j,t} E_{j,t-2} + (1 - r_{j-1,t} - d_{j-1,t}) E_{j-1,t-2}$$

$j = 2, 3, 4, 5, 6,$

and can be interpreted as predictions of the enrollments, to deviate minimally from the corresponding enrollments $E_{1,t}$, $E_{1,t\text{-}1}$, …, $E_{6,t\text{-}1}$. Specifically, Crouch's model amounts to the minimization of the function

$$\sum_{\substack{j=1,2,\ldots,6 \\ h=t,t-1}} E^*_{j,h}$$

of the twelve variables e_t, $r_{1,t}$, …, $r_{6,t}$, $d_{1,t}$, …, $d_{5,t}$, subject to the 36 constraints of the overall system. The first 12 constraints

$$E^*_{j,h} \geq E_{j,h} \quad j = 1, 2, \ldots, 6 \qquad h = t, t-1$$

essentially serve as a simple way to avoid the need for the absolute value operator in the objective function to express a minimization of a sum of prediction errors (see Crouch, 1991, p. 263). Next, there are the 12 non-negativity constraints for the variables and the upper limit ($c_{\alpha,1,t}$) for the entrance rate. Finally, the constraints

($n_{j\text{-}1,j,t}$) $\quad r_{j,t} \leq E_{j,t} / E_{j,t-1} \quad j = 2, 3, 4, 5, 6$

($n_{j\text{-}1,j,t\text{-}1}$) $\quad r_{j,t} \leq E_{j,t-1} / E_{j,t-2} \quad j = 2, 3, 4, 5, 6$

($n_{6,\omega,t}$) $\quad r_{6,t} \leq 1$

are not mentioned explicitly by Crouch (1991), but follow logically from the construction of the model.

Table 13 shows some estimates obtained by means of Crouch's model. Those estimates are erratic. The cause of this unstability is visible—in the case of a single grade model—in Figure 9: a slight change in the enrollment data may shift the point of closest proximity between the two sets of feasible solutions from the upper left corner to the lower right. In other words: a very small change in the data may cause a very large change in Crouch's estimates. In technical language, this means that Crouch has defined his estimates as the solutions of an ill-conditioned problem.

In short: Crouch's approach is flawed, because either the enrollments of consecutive years are markedly different (indicative of rapid changes in the educational system and/or its environment), in which case the assumption of nearly equal repetition rates in consecutive years becomes untenable; or the enrollments are quite similar, a situation in which his model degenerates to an ill-conditioned problem, yielding estimates that are oversensitive to details of the data.

Conclusions

The enrollments of a (sub)population in two consecutive school years in an educational system with a given structure restrict the potential values of the array of transitional flows to the area within a polyhedral boundary. Any additional piece of information that can be expressed as a linear constraint on the transitional flows may be incorporated, possibly restricting further the set of feasible solutions. The boundary of this set can be explored by solving linear programming problems. In this report we have demonstrated that this approach yields useful interval estimates of repetition rates by combining publicly available data (enrollments, structural information and some additional information about particular transitions) about primary education in Belgium. Moreover, we have shown that this method has two important advantages: it is both flexible and prudent. Note also that the computational resources required are quite limited: the numerical results reported here have been obtained by implementing and solving linear programming problems using the 'Solver add-in' of the Microsoft Excel 97 software.

Our method is more flexible than alternative methods that require sufficient constraints to pin-point a unique solution. It allows a gradual addition of information. The merit of the introduction of a particular piece of information can be assessed by looking at its impact on the (narrowing of the) interval estimates. This impact can be weighed against what the researcher knows about the reliability of this information. Conversely, it is possible to experiment with artificial constraints in order to discover what kind of information is required to obtain interval estimates with a given narrowness.

The proposed method of estimating repetition rates from enrollments is prudent in the sense that it is geared to detect both deficits and inconsistencies in the available data. When a particular set of data hardly contains any information concerning a particular transition, our approach makes this plain by reporting a very wide interval. In addition, the smallest inconsistency in the data causes the set of feasible solutions to be empty, forcing the user to explicitly investigate and remedy such a problem. Alternative methods that are designed to yield a solution no matter how uninformative or inconsistent the data, are obviously easier to apply. But their use is, in our opinion, not to be recommended.

Our experience with the estimation of Belgian repetition rates suggests that useful results can not be obtained without enrollments by age and grade or without a detailed knowledge of the educational system at hand. Consequently, when an attempt is made nonetheless to estimate repetition rates from enrollments by grade (and not by age), we strongly advise to use a method that is able to signal an information deficit. When it turns out that—as expected—the enrollments by grade contain scarcely any information with regard to the repetition rates, it is useless or worse to try to remedy this by algorithmic tricks. Finally, we are sceptical about attempts to estimate transition rates from enrollment data alone, without bringing in detailed local knowledge of the educational system in question.

References

COMMUNAUTÉ FRANÇAISE DE BELGIQUE, MINISTÈRE DE L'EDUCATION, DE LA RECHERCHE ET DE LA FORMATION (1992) [(1993), (1994), (1995)]. *Annuaire Statistique 1990-1991* [*1991–1992, 1992–1993, 1993–1994*]. Bruxelles: Communauté française de Belgique, Ministère de l'Education, de la Recherche et de la Formation, Régime linguistique français, Service des Statistiques.

Crouch, L. A. (1991). A simplified linear programming approach to the estimation of enrollment transition rates: Estimating rates with minimal data availability. *Economics of Education Review, 10* (3), 259-269.

Dens, G. (1992). Buitengewoon en geïntegreerd: Tien jaar integratie van type-8 leerlingen in het gewoon onderwijs na één jaar buitengewoon onderwijs. *Pedagogische Periodiek, 18*, 235-237.

MINISTERIE VAN ONDERWIJS (1984) [(1985), ..., (1990)]. *Statistisch Jaarboek van het onderwijs 1983–1984* [*1984–1985, ..., 1989–1990*]. Brussel: Ministerie van Onderwijs, bestuur Gemeenschappelijke Diensten, dienst Onderwijsstatistieken.

MINISTERIE VAN DE VLAAMSE GEMEENSCHAP (1997). Private communication 24 February 1997. Brussel: Ministerie van de Vlaamse Gemeenschap, departement Onderwijs, afdeling Begroting en Gegevensbeheer.

MINISTERIE VAN DE VLAAMSE GEMEENSCHAP, DEPARTEMENT ONDERWIJS (1991) [(1992), ..., (1997)]. *Statistisch Jaarboek van het Vlaams onderwijs 1990–1991* [*1991–1992, ..., 1996–1997*]. Brussel: Ministerie van de Vlaamse Gemeenschap, departement Onderwijs, afdeling Begroting en Gegevensbeheer.

MINISTERIE VAN DE VLAAMSE GEMEENSCHAP, DEPARTEMENT ONDERWIJS (1998). *Zittenblijven en schoolse vertraging in het Vlaams onderwijs: Een kwantitatieve analyse: 1996–1997.* Brussel: Ministerie van de Vlaamse Gemeenschap, departement Onderwijs.

MINISTRY OF EDUCATION (1998). Private communication 15 October 1998. Finland: Ministry of Education, Department for Education and Science Policy, General Education Division.

Moenaert, H. (1991). *Intelligentie en schoolse problemen: Een onderzoek bij zeven- tot tienjarigen*. Leuven: K.U.Leuven, Faculteit der Psychologie en Pedagogische Wetenschappen, Afdeling Psychodiagnostiek en Psychologische Begeleiding, unpublished doctoral dissertation.

NATIONAL BOARD OF EDUCATION (1998). Private communication 14 October 1998. Finland: National Board of Education, Databases in Education.

Ravindran, A., Phillips, D., & Solberg, J. (1987). *Operations research: Principles and practice* (2nd ed.). New York: Wiley.

UNESCO (1997). *Statistical Yearbook*. Paris: UNESCO.

Van Landeghem, G., & Van Damme, J. (1997). *Een schatting van het aantal zittenblijvers in het Vlaams lager onderwijs in de schooljaren 1984–1985 t.e.m. 1994–1995*: *Eindverslag van INES project 9604*. Leuven: K.U.Leuven, LIVO, unpublished report.