Optimization modeling for resource allocation in the Chilean public educational system *

Jorge Amaya† Dominique Peeters‡ Paula Uribe§ Juan P. Valenzuela¶

August 27, 2013

Abstract

This paper presents an economic decision model for planning in public education, considering the satisfaction of the total demand for education in a municipality and the annual cost involved in the system. Given a set of established schools in a municipality, the aim of the model is to find the optimal supply in terms of which schools should be kept open or closed, which grades (or year levels) should be made available and the number of classes that should be provided within each such grade to meet student demand. The model includes specific information about curriculum requirements in terms of the number of hours per year per subject to be provided for each grade level. The objective function essentially seeks to minimize the total fixed and variable costs to schools, which mainly relate to human resources (principal and direction, teaching, administrative and support staff). We propose an integer optimization model for solving this problem and we apply it to two large municipalities in Chile.

Key words: Education planning, Optimization models.

*Supported by CONICYT, under projects: BASAL CMM-Univ. de Chile, CONICYT-PIA project CIE-05 and FONDEF VRI-V10P0021.

†Centro de Modelamiento Matemático and Departamento de Ingeniería Matemática, Universidad de Chile

‡Center for Operations Research and Econometrics, Université catholique de Louvain

§Centro de Modelamiento Matemático, Universidad de Chile

¶Centro de Investigación Avanzada en Educación and Departamento de Economía, Universidad de Chile

1 Introduction

In Chile, one of the main challenges for increasing competitiveness and advancing development of the country is to improve the standard and quality of education, particularly in the public education system. The public sector represents around 40% of total primary and secondary education infrastructure and, in general, caters to the poorest segment of the young population.

The Chilean system of elementary and high-school education can be decomposed into three well-identified sectors: the public sector, managed by local authorities (municipalities); the private subsidized system, managed by private sponsors with partial funding from the state; and the private non subsidized sector, managed by private enterprise which derives economic benefit from educational activity. Also some of them are managed by non profit institutions such as religious organizations. In Chile, there is almost no restriction on the establishment of private schools in any geographic region and the private subsidized system is distinguished by a specific agreement between a private sponsor and the national Ministry of Education. Some of these schools are non-profit and are allowed to impose relatively moderate fees on parents and a complementary subsidy which is allocated by the state. All three systems are competitive, seeking to attract the best students, but the public sector is completely free and open, in the sense that all candidates must be admitted and if there is no place available at a given public school in the district, then the authority must look for places at another school in the public system, even outside the municipality. In general, families choose schools in the vicinity of their homes, especially for primary school. In addition, there is no gender segregation, most schools being mixed, and racial balance is not an issue in the country.

The Ministry of Education only provides general regulations for managing the Chilean educational system. These are in the form of a minimal standard for the teaching curriculum which schools are required to meet or exceed. In the particular case of the public sector, the state, throughout federal, regional and local levels, also provides economic resources, but the rest of the responsibility falls on the local authorities. These resources are often limited, which is especially severe, because (according to the law) the public system must satisfy current demand and essentially no financial contribution may be asked of students' families, at least among primary school students. Moreover, exceptionally, when some public high schools use charges to families, they are very small comparing those fees charged by private subsidized schools ([14]). In this context, the optimal use of public resources becomes a critical concern in each municipality.

During the last ten years, household income in the country has experienced a systematic increase. This has caused a fall in demand for public education as higher incomes have brought previously unaffordable private subsidized sector schooling within the reach of many more families. Also demographic changes in the whole country are reducing total population at schooling age, affecting both public and private schooling services. Both sources are producing an annual reduction between 1%-2% of public schools market share on national educational system ([4]). Management of public resources is also strictly limited by legal regulations, particularly in terms of human resources. For instance, teachers having tenure positions must receive higher wages but transfers from central government are not high enough to afford these larger costs. This represents a real challenge for the public sector: Municipalities must propose an attractive program to students and their families, while at the same time using public funds in an efficient way.

The problem we address in this paper can be described as follows: Given the set of existing public schools and the possible new locations that could be opened in a municipality, and knowing the expected evolution of the demand for the next school years, the aim is to decide how to structure the supply of student places and classes. This means that the model must decide, for a given time horizon, which schools must be kept open, the grade levels to be offered by each school and the number of parallel classes within each grade. The main constraints are essentially the satisfaction of demand, infrastructure availability (number of classrooms in each school) and compliance with minimal teaching curriculum requirements, expressed in terms of the number of teaching hours per week for each subject. This last variable is important because it determines the human resources that need to be allocated, which represents a very important part of the total budget of the municipality.

The location of schools can have a great economic and social impact, particularly in countries having centralized public education systems. The criteria for the location, size or structuring of schools can vary due to factors ranging from spatial or demographic considerations to social integration and administrative or economic constraints. Different authors have proposed various models, depending on particular state regulations and main objectives.

In 1993, Church and Murray ([5]) proposed a model to address the problem of school consolidation. Their work is about the assignment of students to schools with constraints that take into account racial issues and school balancing. In the situation we study here, we select schools to remain in operation and decide how many classes of a given grade will be made available, looking to meet demand for student places. Another difference is that we also consider the balancing of classes within each 'cycle' for each school, but not the balance between schools as they do in their paper (in the Chilean system, primary education is partitioned into two so-called 'cycles' which will be described later). Church and Murray also consider a fixed number of schools to operate, while in our case, this decision is made by the model. Also, our objective function considers the cost of operating a school, unlike in their work where the monetary cost of the system is fixed, and they endeavor to use resources in a balanced way.

In [2], the authors use the simulated annealing heuristic to tackle this kind of problem and report good behavior of the strategy in real applications. Also, reference [7] deals with the problem of school redistricting in which certain, stipulated blocks of the region must be assigned to establishments, combining the use of a binary decision model with a geographic information system, and obtaining encouraging practical results in real-world case studies. The problem of assigning students to schools, considering both capacity and racial balance, and minimizing the total distance from houses to schools, is addressed in [12]. The resulting integer linear problem is solved by solving a sequence of network instances.

A simple binary optimization model is used in [8] to assign students to subjects at university level, but in that case the criteria are essentially the preferences of the students. In [10], the authors present a partial static competitive equilibrium logit model specified for a given scenario of policies, yielding the expected equilibrium locations, prices of schools, and students' school choices. The demand-supply equilibrium is studied and a fixed-point algorithm is used to find a unique solution. For the districting problem, in [6], the authors proposed a decision support system, which uses distances between the student population and schools as the main criteria for assignment.

A dynamic multi-period model is proposed in [1] for the evolution of school networks, allowing size reduction and/or expansion within certain pre-defined facility size limits. This generates a relatively complex mixed linear programming program, whose variables represent assignment, location and capacity decisions. The objective is to minimize the total discounted cost. By using a variant of the *p-median* model, in [13], the authors propose to maximize the accessibility of students to schools (through minimization of total distances) and they apply the resulting facility location model to real case studies.

In [9], the authors deal with the existence of contradictory interests of all parties (administrators, teachers, parents, students, etc.). They propose using multi-criteria models for spatial allocation of educational resources, and give some critical comparison of the different multi-criteria techniques that can be used: Pareto optimality, data envelopment analysis, parametric programming and goal programming methods.

In [11], the authors focus on the location problem, for the real case of two counties in the state of Rio de Janeiro, using heuristic procedures to solve the *p-median* problem. The criterion is to minimize the expected sum of overall home-school distances of students. In [3], the authors address the location and sizing of rural schools in Chile (many of them are multigrade school, having only one teacher for all students), essentially considering accessibility and costs, including transportation costs. They do not consider the internal balance in cycles, neither the academic requirements, in terms of minimal number of teaching hours for each topic. They also impose budget limits for construction/closing, operations and transportation. On the contrary, our work is oriented to relatively large municipalities, which are in general located in urban areas.

This paper is organized as follows: In Section 2 we give the main principles for the planning model and in Section 3 we define the variables, parameters and equations describing the optimization model. Section 4 is devoted to the application of the model to two real case studies and, in Section 5, we give some general conclusions and ideas for further research.

2 The conceptual model

We begin by describing the main constraints of the system, most of which stem from legal and structural regulations. First, the total demand for each grade must be met because the public sector cannot deny enrollment to students. Second, the minimum number of hours per week of each subject is essentially imposed by the academic program established by the state authority. Furthermore, the specific academic plan of a school may fine-tune this number. Third, the maximum number of students in each class is limited by law, but a lower limit could be imposed depending on the education level, on the infrastructure available at the school and on preferences of local authorities related to improve average quality for a particular school or the whole local public educational system. Quality of education is not explicitly included as an endogenous factor in our model, but this issue can be addressed by tuning the classmate size and the number of teaching hours of different topics.

Our purpose is to design an optimization model suggesting which schools should remain open in a given area and, for each of these schools, the forecast number of students at each level, the cycles that are to be set up, the number of classes for each grade that are to be created, how many teaching hours should be allowed, etc. The objective is to provide a given level of service at minimal cost. Actually, we have to cope with an existing situation and to consider the possibility of closing/consolidating schools. Theoretically, one could merge any set of schools to optimize resources. Practically, this is not always possible due to geographic restrictions, population distribution and quality heterogeneities, among other factors. In order to implicitly take those restrictions into account, the model considers a partition (or zonification) of the geographic area of interest, each zone containing a certain number of schools that can be opened, closed or merged without restrictions, based on location criteria and satisfying the local demand. In our two particular applications those partitions are predefined by the decision-maker and at least one school must be opened at each zone.

The Chilean education system is made up of four cycles as follows:

i) Pre-school: 2 years (children being 4 and 5 years of age),
ii) Primary1: 4 years,
iii) Primary2: 4 years, and
iv) High school: 4 years.

The inputs to our model are essentially the resources, school capacities, costs and demands (estimated according the historic evolution and the and the pre-registration process). This means the existing number of classrooms (infrastructure availability), unitary costs of teaching and other personnel (variable), other fixed and variable costs of operating each school, capacities (how many students can be admitted to each class), and the total capacity of the school in each grade. The model considers the evolution of the demand is fixed and known along the planning horizon (this can be valid for short and medium term), but we do not enforce satisfaction of expected demand for each school and grade.

We specify here the main assumptions of our model:

i) All data are annual, but the scope of the proposed model is designed to decide on short time operation (one to five years).

ii) The total demand is known for each grade and geographic zone. In this model, a specific school can satisfy demand coming from different zones, but we penalized this fact in the objective function.

iii) If the model decides that a given school should be kept open, at least one cycle must be made available, but if two non consecutive cycles are to be provided, then any intermediate cycles must also be provided (for example, if the first primary cycle and high school cycle are opened, then the second primary cycle must also be available in order to offer continuity to students completing the first primary cycle).

iv) Due to infrastructure availability we also impose an upper limit on the number of parallel classes within every grade. If a grade is to be made available, then the entire corresponding cycle must also be available.

v) The model also requires that the total number of classes in a given school be restricted to the available infrastructure.

vi) Each school must have a coherent distribution of students and grades. This means that the number of parallel classes in the same grade must be as similar as possible for every grade throughout the same cycle. Besides to relax this constraint does not change the equilibriums in the model. Note that we impose balancing inside a school and not between schools.

vii) Concerning educational requirements, it is necessary to comply with the total number of teaching hours per week for every grade and subject. This is dictated by national regulations but schools can impose additional teaching hours according to their specific academic program.

viii) It is required that the number of students per classroom be smaller than a specified limit, and this limit being fixed according to legal, pedagogic and infrastructure criteria. We note here that the classmate size is usually related to quality of education, but we do not explicitly include quality issues in our formulation.

3 The deterministic mathematical model

The deterministic model seeks to minimize the total cost, satisfying the structural and legal regulations of the educational system, uses the notion of student allocation to schools for every grade by including the corresponding costs of those allocations in the objective function. We suppose that the demand is known for every grade and zone along the time horizon.

Let us define the following variables (that characterize a decision):

x_{ik}^t : Number of classes in grade $i = 1, \ldots, N_G$ at school $k = 1, \ldots, N_S$, period $t = 1, \ldots, T$
y_{ik}^t : Binary, 1 if grade i is available at school k, period t; 0 otherwise
z_k^t : Binary, 1 if school k is open at period t; 0 otherwise
u_{jk}^t : Binary, 1 if cycle $j = 1, \ldots, N_C$ is to be provided by school k at period t; 0 otherwise
v_{mik}^t : Total hours per year assigned to subject $m = 1, \ldots, N_M$, grade i, school k, period t
w_k^t : Maximum difference between the number of classes of two consecutive grades, school k, period t
a_{ihk}^t : Number of children in grade i from zone h who are allocated to school k, period t.

We also define the parameters (problem data),

W_k^t : Fixed annual cost of school k, period t
Q_{mi}^t : Teaching cost per hour for subject m, grade i, period t
C_{ik} : Capacity (maximum number of places per class) of grade i, school k
T_k^t : Total number of available classrooms at school k, period t (capacity of the school)
H_{mik} : Number of hours of subject m required per year in grade i at school k
R_{ik} : Maximum number of classes in grade i at school k
r_{ik} : Minimum number of classes in grade i at school k
$\{C_j\}$: Partition of grades into cycles (e.g. pre-school, primary, high school)
$\{Z_h\}$: Partition of the geographic area where schools are located into zones.
D_{ih}^t : Demand for grade i in zone $h = 1, \ldots, N_Z$, period t
P_k^t : Penalty coefficient for heterogeneity within school k, periode t
E_0 : Set of existing schools at period 0
N_0 : Set of non existing schools at period 0 (possible new locations)
ρ_{ihk}^t : Penalty associated with a child in grade i, zone h, allocated to school k, at period t
K_{kt} : Cost of closing an existing school k at period t
K'_{kt} : Cost of opening a new school k at period t (to use a new location).

The objective function seeks to minimize total fixed and variable school operating costs, mainly related to human resources (teachers, administrative and support staff). The variable a_{ihk}^t gives information about the optimal allocation, but only represents the estimated supply, not a real decision. To enable the calculation of allocation costs, we must define the corresponding parameters ρ_{ihk}^t representing the cost of assignment of children to schools, but in fact it only gives an estimate of the total cost of the system since student assignment to schools is not mandatory and parents have the final decision about where to send their children. Under rational assumptions, it is expected that families would tend to choose the nearest schools, and we represent this fact by using $\rho_{ihk}^t = 0$, for $k \in Z_h$.

The mathematical model can be written as:

$$(P) \qquad \min \sum_{mikt} Q_{mi}^t v_{mik}^t + \sum_{kt} W_k^t z_k^t + \sum_{kt} P_k^t w_k^t + \sum_{ihkt} \rho_{ihk}^t a_{ihk}^t + \sum_{k \in E_0,\ t} K_{kt}(z_k^t - z_k^{t+1}) + \sum_{k \in N_0,\ t} K'_{kt}(z_k^{t+1} - z_k^t)$$

$$\begin{aligned}
u_{jk}^t \leq z_k^t \leq \sum_{j'} u_{j'k}^t \quad & \forall j,k,t && (1)\\
\sum_{k \in Z_h} z_k^t \geq 1 \quad & \forall h,t && (2)\\
r_{ik} y_{ik}^t \leq x_{ik}^t \leq R_{ik} y_{ik}^t \quad & \forall i,k,t && (3)\\
u_{jk}^t = y_{ik}^t \quad & \forall i,j,k,t \text{ such that } i \in C_j && (4)\\
\sum_i x_{ik}^t \leq T_k^t \quad & \forall k,t && (5)\\
v_{mik}^t \geq H_{mik} x_{ik}^t \quad & \forall m,i,k,t && (6)\\
u_{jk}^t + u_{j'k}^t - 1 \leq u_{lk}^t \quad & \forall k,t,\ 1 \leq j \leq j'-2 \leq N_C - 2,\ j+1 \leq l \leq j'-1 && (7)\\
-w_k^t \leq x_{ik}^t - x_{(i+1)k}^t \leq w_k^t \quad & \forall k,t,\ 1 \leq i \leq N_G - 1 && (8)\\
\sum_h a_{ihk}^t \leq C_{ik} x_{ik}^t \quad & \forall i,k,t && (9)\\
\sum_k a_{ihk}^t \geq D_{ih}^t \quad & \forall i,h,t && (10)\\
z_k^t \geq z_k^{t+1} \quad & \forall k \in E_0, t && (11)\\
z_k^t \leq z_k^{t+1} \quad & \forall k \in N_0, t && (12)\\
x_{ik}^t, v_{mik}^t, a_{ihk}^t \geq 0, \text{ integer} \quad & \forall m,i,k,h,t &&\\
u_{jk}^t, y_{ik}^t, z_k^t, \text{ binary} \quad & \forall i,j,k,t &&\\
w_k^t \geq 0 \quad & \forall k,t &&
\end{aligned}$$

Constraint (1) says that a school must be open when at least one cycle is open at that school. It also means that when no cycles are created at a school, then the school will not be open. In this constraint, the right inequality $z_k^t \leq \sum_{j'} u_{j'k}^t$ could be redundant, because variable z_k^t has a positive coefficient in the objective function, so it tends to zero when $u_{jk}^t = 0$.

Constraint (2) impose that at least one school must be open at each zone and period.

Constraint (3) says that if a given grade is opened at school k, then it must satisfy bounds on the number of classes in that grade (infrastructure availability).

Constraint (4) means that if a cycle is to be available at a school, then all grades belonging to that cycle must be provided and if one grade is to be taught, then the entire cycle must be taught.

Note that (1), (3) and (4) together imply that when a school is closed, all grades within it are closed (the x variable for each grade is zero), so the constraint is satisfied. Also, there is no incentive to change the variable value just to achieve a more balanced solution, because if the variable changes to a value greater than zero, the cost of the solution increases (because the whole school must be opened if at least one grade is to be offered).

Constraint (5) gives an upper bound on the number of classes to be taught at each school.

Constraint (6) says that the number of teaching hours of the school must satisfy the requirements given by the academic program of the school, for each subject.

Constraint (7) imposes a coherent structure, in the sense that, if two non-consecutive cycles are offered, then the intermediate cycles must also be offered.

Constraint (8) says that the number of classes in a given cycle must be quite smooth. This allows for a coherent school, in the sense that the number of parallel classes in a given grade does not differ with the number of classes in the following grade.

Constraints (9) and (10) require that the total capacity of the system in a given zone or area be sufficient to satisfy demand for every grade.

Constraints (11) and (12) mean regularity of the solution: if an existing school is closed then it remains closed for the subsequent periods, and if a new school is opened, then it also remains opened along the planning period.

Remark 3.1 *The variables, parameters and (in)equations of this mathematical model have been explicitly described to make clear the structure and readability of the problem. The model could be simplified if, for example, we eliminate some variables y_{ik}^t by using equality (4), but this is performed in the numerical resolution (preprocessing phase).*

Remark 3.2 *Note that* $\bigcup_{h=1}^{N_Z} Z_h = \{1, \ldots, N_S\}$ *and* $Z_h \bigcap Z_{h'} = \emptyset$ *for* $h \neq h'$. *In fact, the zones* Z_h *can correspond to different municipalities or specific areas inside them.*

Remark 3.3 *Note that* $E_0 \bigcup N_0 = \{1, \ldots, N_S\}$ *and* $E_0 \bigcap N_0 = \emptyset$.

Remark 3.4 *The costs depend in general on the period* t. *This could take the form of a a fixed value weighted by a decreasing factor* ε^t, *where* $0 < \varepsilon < 1$. *For example* $W_k^t = W_k \varepsilon^t$.

Remark 3.5 *At the optimal solution, the real variables* w_k^t *take integer values, because in Constraint (8) the term* $x_{ik}^t - x_{(i+1)k}^t$ *is integer and positivity of* P_k^t *permits to set* w_k^t *at an integer values.*

Remark 3.6 *The coefficient* ρ_{ihk}^t *accounts for several aspects. Firstly, it may include the private or individual cost (to families), for example, transportation from area* h *to school* k *(which can be supposed to be zero for* $k \in Z_h$*), the social perception of this travel by the families living in area* h*, the perception of the value of the travel time, the age of children in grade* i*, etc. Secondly, it may include the marginal cost to school* k *for admitting pupils from other zones. In this context, the objective function is a weighted additive combination of three criteria: direct cost of education, indirect cost due to heterogeneity within cycles and private costs to the families of students coming from other zones. Evidently, this aspect could be handled by using multi-criteria modeling and appropriate algorithms.*

Finally, the model permits to handle situations in which demand is overestimated, for example, as a result of pre-enrollment processes. We can ease this requirement by changing Constraint (10) into a relaxed inequality,

$$\lambda_{ih}^t D_{ih}^t \leq \sum_k a_{ihk}^t \quad \forall i, h, t.$$

This relaxation allows that only a certain given proportion $\lambda_{ih}^t \in]0,1]$ of the (local) demand in each zone be met. To encourage demand to be satisfied, we should add to the objective function a term of the form $\sum_{iht} \tau_{ih}^t (D_{ih}^t - \sum_{k \in Z_h} a_{ihk}^t)$, where τ_{ih}^t represents the penalty for not admitting a pupil in grade i from zone h to a school in his own zone. Note that the term $\sum_{ih} \tau_{ih}^t D_{ih}^t$ is a constant, but we keep it in the last expression to highlight the fact that failing to satisfy demand comes with a cost.

4 Case studies

In this section, we present two case studies, corresponding to two municipalities in Chile (designated here Municipality 1 and Municipality 2). Both have typically middle and upper-middle class population. A large share of students from these families are enrolled in private subsidized schools, while public schools accommodate essentially poor and low-middle class students. The decision-makers decided to divide a priori each of both municipalities into eight zones, with different demands per grade. This choice could indeed be left to the model, but we defer this approach to future study. The Municipality 1 system has 15 preexisting schools and Municipality 2 has 24 ones, with estimated demand of 10,320 and 19,946 students respectively.

The cost of a teaching hour is supposed to be a constant, but in general this value should depend on the subject or topic and the specific teacher conditions. Even if the scope of the general model is strategic, in this application we only consider one year horizon, due to the lack of valid information about the demand evolution along the time. We use costs (expressed in US dollars) for the 2010 academic year, while the demand corresponds essentially to the number of students applying for enrollment at each school at the end of 2010. That is to say, the demand is the expected student population for 2011.

The annual fixed cost corresponds to the purchase and maintenance of infrastructure (buildings, equipment, etc.) and the total cost includes the fixed cost plus the wages of academic and non-academic staff of the school. This includes the cost of teaching hours. Rational planning based on the application of our model should essentially reduce the human resource cost (variable component) as well as the fixed cost due to establishments that are closed. The cost of closing and opening a school are explicitly included in the model through the coefficients K_{kt} and K'_{kt}.

For both municipalities, we solved the mathematical model in two scenarios. Scenario 1 is defined by the maximal number of students in every class, fixed at 35. In Scenario 2 we use a bound of 35 students per class at pre-school level and 45 for the primary and high school levels. For each scenario, we consider two instances: $\lambda_{ih}^t = 1$ (forcing the satisfaction of the total demand) and $\lambda_{ih}^t = 0.5$ (relaxing this constraint to allow as little as 50% of the demand to be satisfied).

In these simulations we used some realistic values of the parameters:

- C_{ik}: Maximum number of children per class. We consider two scenarios, already defined in the previous paragraph. Despite 45 pupils is a legal constraint, the municipalities could opt for a lower maximum to reach better quality outcomes. The computational results could be dramatically different depending on this value, but here we use the legal bound and a tighter value, knowing that currently average class in urban public primary schools is around 31 students ([4]).
- R_{ik}: Maximum number of parallel classes in a grade, varies from 0 to 4 in Municipality 1 and 0 to 3 in Municipality 2. Setting to 0 means that the corresponding grade is forced not open in the school.
- r_{ik}: Minimum number of classes, set to 0 everywhere.
- D_{ih}^t: Demand in zone h for grade i is detailed per school in Annex, Tables 1 and 2.
- H_{mik}: The number of teaching hours/week of subject m required for grade i at school k; this varies from 0 to 9 in Municipality 1 and from 0 to 8 in Municipality 2. This is not constant, because it depends on each subject and on the particular pedagogic choices made by each school. In Chile, the law only establishes minimal requirements for each subject.
- Q_{mi}^t: The unit cost per hour of subject m in grade i is constant at \$10.9.
- P_k^t: The cost of heterogeneity at school k, period t; we use \$1 to represent this weight. Another possibility is to limit the number of parallel classes by an upper bound fixed ahead of time, but in that case there is a risk of infeasibility.
- ρ_{ihk}^t: The annual cost of education per student in grade i, zone h at school k, period t is fixed at a constant value of \$1,163. The same value is used for the penalty coefficient τ_{ih}^t.

In the following subsections we show the main data and the results for our two case studies. The application of this mathematical model produces a significant reduction in total costs to both municipalities. These savings come from the optimal configuration of classes (satisfying legal and institutional regulations) and from the optimal allocation of teaching resources (personnel) to the teaching hours needed.

These instances correspond to medium size mixed linear programming problems, which can be solved by most existing commercial and academic software. All instances were run using CPLEX-9.1 on an IBM-iDataplex with 2 processors.

4.1 Main data

Tables 1 and 2 show the main data for each case study. The first column contains an identifier for the school. The second column is the current number of available classrooms in each school. In fact we consider here the classrooms as indistinguishable, but it is well known that, in general, some classrooms are adapted for pre-school or primary-school pupils while others are allowed to be used by high-school students. The next column is the identifier of one of the eight zones to which the schools belong. The next two columns contain the total fix cost (infrastructure, general services, etc.) and the total annual cost, which includes the variable cost, namely teachers' salaries. The last column shows the total expected demand for the year 2011, but the calculations use the detailed distribution of the demand for every grade and school.

4.2 Simulations results

The results are reported in Table 3. First, we observe that it is not possible to satisfy the demand in case of Scenario 1 (capacity equal to 35 pupils per class in all grades), but in the relaxed cases the model find solutions leaving 3% and 5% of the demand unsatisfied, respectively in the two municipalities.

For comparison, we use the base case, which corresponds to the observed costs in year 2010. The total monthly hours, 81,496 for Municipality 1 and 107,308 for Municipality 2 in this base case, can be decomposed into two components. The first is related to the "classroom hours", that is, the number of hours dedicated to teaching, and corresponds to 75,336 and 96,972 monthly hours, respectively. The second part represents the hours dedicated to supporting the educational labor (for example, the hours of the principal are included in this figure), and corresponds to 6,160 and 10,336 monthly hours, respectively. Also it includes current costs for goods and services related with education, salaries to municipal staff who manage educational services and conditioned transfers from central and regional governments to finance specific programs to support public education.

After the optimization, the monthly "classroom hours" decrease by up to 47% for Municipality 1 and by up to 34% for Municipality 2 (depending on the scenario), which can be interpreted as the current percentage of overstaffing levels. This decrease results in a total annual savings of up to around \$4,500,000–\$5,500,000. In fact, if we calculate the average of current daily teaching hours for both cases, the result is between 9 and 12, which does not reflect reality while, after optimization, this value decreases to 7 in all cases; this is consistent with an ordinary academic curriculum in any school. Regarding the number of classes in the base case, 308 are needed in Municipality 1 and 548 in Municipality 2 (maybe not really needed, because of eventual overestimation). After the optimization, this number can decrease up to 13% for Municipality 1 and to 14% for Municipality 2. Note that, although the optimization efficiency varies from 25% to 47% in terms of teaching hours in both municipalities, the percentage of improvement in terms of number of classes is much smaller. This suggests that for these cases, the most significant inefficiency is found in the allocation of human resources. It is consistent with trajectories among municipalities in charge of public education, which have kept constant the total teacher hours hired during last decade rather than a systematic reduction in total enrollment in the same period.

The enormous difference between base case and optimized cases can be explained by the overestimation of the required teaching staff (but we include here the salaries of the principal and other non-teaching staff of the school).

5 Conclusions and further remarks

The model presented here permits the planner to estimate the infrastructure and human resources needed in order to satisfy demand for the municipal school system and the corresponding costs of education at each school and for each grade. At the same time, the model gives the configuration of each school in terms of the size of the student population and the number of classes in each grade. The most significant extra costs of base case from optimal alternative come from the allocation of human resources, but it is sensitive to definition of maximum classmates size, which can also be simulated with the general model introduced here. Some open questions for further research arise from this work.

i) The model provides a general framework for handling real instances of economic planning of the education system. It gives an indication of the total number of hours that are necessary to satisfy student demand, but not how many teachers are necessary to fulfill the teaching demands of each school. Determining the optimal number of teachers is a non trivial function of the academic structure of the school.

ii) We did not consider the quality outcomes as a constraint in our simulations. The choice of variables to represent this factor in the mathematical model remains an open question. The classmate size and the number of teaching hours in our model could be interpreted as indirect quality indicators.

iii) In this paper, we have not considered the random behavior of demand, which is inherent to future demographic and economic evolution, which is specially true of Chile. This needs a new model together with algorithms for model resolution using stochastic or robust optimization techniques.

iv) Closing a school may have a large impact, hence the benefits or costs of this decision must be closely examined. First, significant public infrastructure is disaffected and the use of this spared resource must be envisaged. Next, it may alter deeply a neighborhood, for instance a deprecated area of a city or a village. Incorporating these in a model is a real challenge.

v) The current model only deals with the public sector, without considering the other two sectors of the Chilean education system, which are very dynamic and competitive. The challenge is to study how these sectors influence the demand for public education, in order to propose more general education models.

vi) An open question is how to take advantage of the staircase structure of the optimization model. In fact, we have not discussed the algorithmic aspects of solving the model, since the real instances we solved were only of moderate size, but in a more general case it may be necessary to propose decomposition or heuristic methodologies. Larger instances could appear because the evolution of demand over several years could generate a much larger number of variables and constraints.

   This problem has the general synthetic form:

$$
\begin{array}{rlll}
(P') \ \min & c_1^\top x_1 + \cdots + c_T^\top x_T & & \\
& E_1 x_1 + \cdots + E_T x_T & \leq & 0 \qquad (13) \\
& A x_1 \qquad\qquad & \leq & b_1 \qquad (14) \\
& \qquad \ddots & & \\
& \qquad\qquad A x_T & \leq & b_T \qquad (15) \\
& B x_1 \qquad\qquad & = & 0 \qquad (16) \\
& \qquad \ddots & & \\
& \qquad\qquad B x_T & = & 0 \qquad (17) \\
& x_1, \ldots, x_T & \in & \{0,\,1\}^{n_1} \times \mathbb{N}^{n_2} \times \mathbb{R}_+^{n_3} \qquad (18)
\end{array}
$$

   Each vector $x_t,\ t = 1, \ldots, T$ is defined as follows,

$$x_t = (y_{ik}^t, u_{jk}^t, z_k^t, x_{ik}^t, v_{mik}^t, a_{ihk}^t, w_k^t),$$

   where $(y_{ik}^t, u_{jk}^t, z_k^t)$ are binary, $(x_{ik}^t, v_{mik}^t, a_{ihk}^t)$ are integer and (w_k^t) are real positive variables.

Constraints (14)-(15) represents the requirements for every period, and Constraint (13) links the periods between them (see Constraints (11)-(12), the only set of constraints that involves different periods).

The corresponding matrix $[E_1, \ldots, E_T]$ is very sparse, containing only two nonzero coefficients 1 and -1 in each row. Moreover, it only involves the variables z_k^t. Constraints (16)-(17) correspond to equations (4).

In a more general case, having several planning period, the structure of the optimization problem described here should play a crucial role in terms of resolution time.

Acknowledgements. The authors would like to acknowledge the data provided by two Chilean municipalities and the very useful suggestions and remarks from the referees.

References

[1] Antunes, A. and D. Peeters. A dynamic optimization model for school network planning. *Socio-Economic Planning Sciences*, 34 (2000) 101-120.

[2] Antunes, A. and D. Peeters. On solving complex multi-period location models using simulated annealing. *European Journal of Operational Research* , 130 (2001) 190-201.

[3] Araya F., R. Dell, P. Donoso , V. Marianov, F. Martinez, A. Weintraub. Optimizing location and size of rural schools in Chile. *International Transactions in Operational Research*, 19 (2012) 695-710.

[4] Bellei, C., P. González and J.P. Valenzuela. Fortalecer la educación pública: un desafío de interés nacional. In: Ecos de la Revolución Pingüina (Bellei, C., D. Contreras and J.P. Valenzuela, eds). UNICEF and CIAE-Universidad de Chile. Pehuén Editores (2010) 225-256.

[5] Church, R. L. and A. T. Murray. Modeling school utilization and consolidation. *Journal of Urban Planning and Development (ASCE)*, 119 (1993) 23-38.

[6] Ferland, J. and G. Gunette. Decision support system for the school disctricting problem. *Operations Research*, 38 (1990) 15-21.

[7] Caro, F., T. Shirabe, M. Guignard and A. Weintraub. School redistricting: embedding GIS tools with integer programming. *Journal of the Operational Research Society*, 55 (2004) 836-849.

[8] Geiger, M.J. and W. Wenger. On the assignment students to topics: a Variable Neighborhood Search approach. *Socio-Economic Planning Sciences*, 44 (2010) 25-34.

[9] Malczewski, J. and M. Jackson. Multicriteria spatial allocation of educational resources: an overview. *Socio-Economic Planning Sciences*, 34 (2000) 219-235.

[10] Martínez F., L. Tamblay, A. Weintraub. School locations and vacancies: a constrained logit equilibrium model.*Environment and Planning A* 43(8) (2011) 1853-1874.

[11] Pizzolato, N. D. and H. B. Fraga da Silva. The location of public schools: evaluation of practical experiences. *International Transactions in Operational Research*, 4 (1997) 13-22.

[12] Schoepfle, O.B. and R.L. Church. A fast, network-based, hybrid heuristic for the assignment of students to schools. *Journal of the Operational Research Society*, 40 (1989) 1029-1040.

[13] Teixeira, J., A. Antunes and D. Peeters. An optimization-based study on the redeployment of a secondary school network. *Environment and Planning B: Planning and Design*, 34 (2007) 296-315.

[14] Valenzuela, J. P., C. Bellei and D. De los Ríos. Segregación escolar en Chile. In: Martinic, S. and Elacqua, G. (Eds). Fin de Ciclo? Cambios en la gobernanza del sistema educativo. OREALC-UNESCO. Pontificia Universidad Católica de Chile, (2010) 209-229.

APPENDIX: Tables

School Id Municipality 1	Available classrooms T_k^t	Zone Z_h	Annual fixed cost W_k^t	Annual total cost	Effective annual demand
1	20	8	\$ 340,158	\$ 948,428	723
2	18	8	\$ 159,858	\$ 717,439	439
3	38	6	\$ 391,602	\$ 1,534,137	992
4	32	2	\$ 365,143	\$ 1,371,324	929
5	6	3	\$ 48,444	\$ 247,652	70
6	29	1	\$ 384,239	\$ 1,276,876	834
7	18	4	\$ 315,956	\$ 897,361	545
8	38	7	\$ 593,062	\$ 1,734,583	1,192
9	55	6	\$ 834,686	\$ 2,421,765	1,885
10	20	8	\$ 195,865	\$ 628,751	429
11	16	7	\$ 201,030	\$ 437,242	331
12	23	7	\$ 319,585	\$ 925,828	641
13	20	5	\$ 243,611	\$ 809,302	507
14	14	7	\$ 71,888	\$ 222,942	89
15	22	7	\$ 323,172	\$ 942,087	714
Total	369		\$ 4,788,296	\$ 15,115,718	10,320

Table 1: Main data for Municipality 1

School Id Municipality 2	Available classrooms T_k^t	Zone Z_h	Annual fixed cost W_k^t	Annual total cost	Effective annual demand
1	32	6	$ 236,366	$ 938,919	1,054
2	27	6	$ 218,565	$ 1,080,281	843
3	33	6	$ 307,573	$ 1,176,893	1,300
4	23	2	$ 182,962	$ 650,316	793
5	23	2	$ 200,763	$ 751,755	860
6	17	2	$ 147,358	$ 597,985	608
7	19	2	$ 154,281	$ 625,691	603
8	20	2	$ 163,182	$ 649,291	614
9	18	2	$ 154,281	$ 600,346	584
10	14	6	$ 120,656	$ 502,346	467
11	20	4	$ 165,160	$ 580,811	622
12	16	2	$ 170,105	$ 608,060	707
13	22	2	$ 194,829	$ 729,094	905
14	18	3	$ 180,984	$ 680,272	777
15	24	1	$ 205,708	$ 763,290	929
16	14	6	$ 124,612	$ 620,352	492
17	37	3	$ 308,562	$ 1,109,958	1,361
18	25	2	$ 225,488	$ 854,034	833
19	39	2	$ 330,320	$ 1,148,444	1,563
20	28	2	$ 248,234	$ 922,908	1,092
21	27	3	$ 234,388	$ 820,356	1,004
22	40	5	$ 356,033	$ 1,441,796	1,667
23	9	7	$ 89,008	$ 297,848	184
24	6	8	$ 53,405	$ 220,173	84
Total	551		$ 4,772,822	$ 18,371,218	19,946

Table 2: Main data for Municipality 2

Municipality 1	Base case	Scenario 1 (P)	Scenario 1 Relax. (P)	Scenario 2 (P)	Scenario 2 Relax. (P)
Demand	10,320	10,320	10,320	10,320	10,320
Satisfied demand	10,320	INFEASIBLE	9,991	10,320	10,175
Numb. of schools	15		15	14	14
Total cost	$ 15,115,718		$ 10,583,595	$ 10,114,303	$ 9,542,702
Annual saving			$ 4,532,123	$ 5,001,415	$ 5,573,016
% reduction/base case			30%	33%	37%
Tot. teach. hours/month	75,336		45,732	42,596	40,040
% reduction/base case			39%	43%	47%
Total classes	308		305	284	268
% reduction/base case			1%	8%	13%
Municipality 2	Base case	Scenario 1 (P)	Scenario 1 Relax. (P)	Scenario 2 (P)	Scenario 2 Relax. (P)
Demand	19,946	19,946	19,946	19,946	19,946
Satisfied demand	19,946	INFEASIBLE	19,014	19,946	19,849
Numb. of schools	24		24	23	24
Total cost	$ 18,371,218		$ 13,990,656	$ 12,991,598	$ 12,835,449
Annual saving			$ 4,380,562	$ 5,379,620	$ 5,535,769
% reduction/base case			24%	29%	30%
Tot. teach. hours/month	96,972		72,740	66,156	63,624
% reduction/base case			25%	32%	34%
Total classes	548		531	490	474
% reduction/base case			3%	11%	14%

Table 3: Results for Municipality 1 and Municipality 2.