A MATHEMATICAL PROGRAMMING APPROACH TO APPLICANT SELECTION FOR A DEGREE PROGRAM BASED ON GENDER, REGIONAL ORIGIN AND SOCIOECONOMIC CRITERIA

Abstract

In 2007, the Department of Industrial Engineering at the University of Chile inaugurated a Master's degree in globalization management, in alliance with a major Chilean mining company. The new program is aimed at helping meet the challenges currently facing the country in the development of human and social capital through the training of young professionals. This paper describes the use of mathematical programming models in the program's applicant selection procedure for the first three entering classes, subject to equity criteria on gender, regional origin and socioeconomic background. The models generated robust solutions in a matter of minutes, an achievement practically impossible with manual methods. The success of this application demonstrates how Mathematical Programming and Operations Research can make a contribution to social policy, in this case by generating a list of candidates that best fits the admission profile of a university degree program incorporating equity considerations. The mathematical tool developed also added transparency to the selection process.

Keywords: Equity, Integer Programming, Operations Research, Robust Selection.

1. Introduction

In 2007, the University of Chile inaugurated a Master's degree program in globalization management whose mission is to provide an education of excellence in business administration to young Chilean professionals. It is run by an alliance of the Department of Industrial Engineering, a unit of the University's Faculty of Physical and Mathematical Sciences, and one of Chile's largest mining companies.

More specifically, the goal of the program is to address the challenges currently facing the country in the development of human and social capital through the training of professionals from a wide spectrum of socioeconomic backgrounds who have the potential to perform effectively in globalized businesses. One of its key aspects is that all those who are admitted are eligible for a grant allowing them to study full time. The 18-month program includes courses given in Chile as well as internships abroad at universities in countries such as Australia, China, Canada and the United States. Applicants must meet a series of requirements regarding age, educational background and work experience.

For the first entering class (2007) the program directors set the total number of admitted students at 53, reducing the figure to 51 for the second class (2008) and 47 for the third (2009). It was also decided to apply equity or "positive discrimination" selection criteria based on gender, region of origin and socioeconomic background. This policy reflects another central objective of the program which is to ensure genuine equality of opportunity and initiate a reversal of Chile's traditional concentration of highly trained human resources on men from the Santiago (capital) region in the top income quintile. It was thus ruled that in 2007 at least 30% of total admissions would be women, 60% would come from non-Santiago regions and 80% would belong to the lower four income quintiles. In 2008 program organizers chose to prioritize slightly the applicants' general qualifications over the positive discrimination criteria, lowering the above percentages to 30%, 55% and 70%, respectively. For 2009 they were returned to the previous year's values of 30%, 60% and 80%, and a 30% minimum non-engineer requirement was added to boost diversity after data for the first two years indicated the successful entrants were almost exclusively engineers.

The more than 600 applicants who met the minimum requirements in each year entered the first stage of the selection process, in which they were each assigned a certain number of points based on their academic and work backgrounds. Of these, some 500 in each year advanced to the second stage, which involved a series of aptitude tests in various fields of

knowledge and a psychometric evaluation. The results were combined with the first stage point total to arrive at a new score, on the basis of which some 160 applicants in each year progressed to the third stage. These were given a psychological evaluation and, in 2008 and 2009, an English test the need for which had become apparent over the course of the first year. Those who passed this stage, numbering 87 in 2007, 83 in 2008 and 86 in 2009, formed the short list from which the final group of admitted applicants would be selected along with a 20-candidate waiting list in case any of that group declined the admission offer. In 2009, the psychological evaluation was used not only to narrow down the number of applicants but to define a series of scenarios for the final group selection instead of just the one used previously.

The method of evaluating the scores and minimum conditions just described for arriving at the short list was defined by the program organizers and will not be discussed in this paper. What will occupy us here is how the selection process used mathematical methods to identify the final group of candidates in a way that reflected the program organizers' desire to choose those with the best qualifications profile while ensuring that advantages of gender, regional origin and socioeconomic background would not be decisive. The precise identification of lower quintiles and non-Santiago region status (based on place of birth or secondary school completion) was also decided by program officials. During the 2007 selection process, two different quintile definitions were employed until the last moment. The first one defined the top quintile as applicants who had attended private secondary schools while the second one added to the first the condition that at least one of their parents were members of traditionally high-income professions (medicine, engineering, law, economics). The latter definition, which was more restrictive and therefore broadened the base of the lower quintiles, was the one finally settled on and was maintained in 2008 and 2009.

The objective of the present analysis will thus be to show how integer linear programming (ILP) models were used to select the final group of applicants from the short list who best fit the qualifications profile of the Master's program while satisfying the equity constraint minima. The goal was to obtain a definitive solution that was robust in the sense that it would not vary greatly with small variations in the admission criteria. Achieving this with a manual procedure in a reasonable time period would have been practically impossible, which is what originally prompted the resort to mathematical methods. Also, the ILP algorithm we employed brought much greater transparency to the selection process. In this sense our study demonstrates the potential of Operations Research to make a contribution to social policy, and more particularly to strengthen equality of opportunity in graduate level education. The programming tool we will describe was used in all three of the program's selection processes held so far.

The use of Management and Operations Research techniques in selection processes is associated in the existing literature with various applications of the Analytic Hierarchy Process (AHP) [5]. Examples include a number of articles in the fields of health [1], education [3] and business management [4]. In [1] the author reports a study by a pharmaceutical company to determine which projects to implement in the search for new cancer drugs. The findings demonstrate the advantages of using AHP in the decision-making process, particularly when various criteria are employed and even more so when the criteria are subjective. In [3], AHP is used to choose new instructors at a higher educational institution. The selection process generated wastes no resources (such as selection committee time), considers all criteria and applies a procedure that is fair to all participants. Finally, [4] presents another AHP application that finds the best way of selecting suppliers for a manufacturing firm based on both quantitative and qualitative factors.

To the best of our knowledge, the use of integer linear programming techniques in a selection process is a novel feature in the present work. The idea of combining the results of multiple scenarios to arrive at a final robust decision is taken from [2], where the authors apply a mathematical model to create a set of winning offers in an combinatorial auction for the supply of school meals throughout Chile.

In the remainder of this paper, Section 2 describes three different mathematical models utilized in the selection process, Section 3 develops a selection algorithm for combining them to obtain a more robust solution, Section 4 sets out the results and Section 5 presents our

conclusions. Tables containing the final results of the three selection processes (2007, 2008 and 2009) and a specific example of the selection algorithm for the 2008 process are given in the appendices.

2. Mathematical models

Three different mathematical models were developed for the selection process, each of which incorporates a different selection criterion. The first model maximizes the sum of the scores assigned to the selected applicants, the second one minimizes the sum of their rankings and the third minimizes the ranking of the last candidate selected. In all three cases, applicants must satisfy the gender, lower income quintile and non-Santiago region criteria and, for 2009, the new criterion regarding professional background. In what follows we first set out the notation, decision variables and constraints common to all of the models and then describe the specifics of each one individually.

Notation

Let N be the number of persons to be admitted, K the set of short-listed applicants, M the set of all female applicants, R the set of all non-Santiago region applicants, Q the set of lower income quintile applicants and P the set of non-engineers. Also, p_i is the score of applicant *i* (without loss of generality we may assume that the scores are ordered from high to low).

Decision variables

$$x_i = \begin{cases} 1 \; if \quad applicant \quad i \quad is \quad selected \\ 0 \quad\quad otherwise \end{cases}$$

Constraints

1. Total number of applicants to be selected is predetermined by program organizers (the value of N was 53 in 2007, 51 in 2008 and 47 in 2009).

$$\sum_{i \in K} x_i = N$$

2. At least m% of the selected applicants must be women. The value of m utilized in all 3 years was 30.

$$\sum_{i \in M} x_i \geq \frac{m}{100} \cdot N$$

3. At least r% of the selected applicants must be from non-Santiago regions. The values of r utilized were 60 in 2007 and 2009, and 55 in 2008.

$$\sum_{i \in R} x_i \geq \frac{r}{100} \cdot N$$

4. At least q% of the selected applicants are from the bottom four income quintiles. The values of q utilized were 80 in 2007 and 2009, and 70 in 2008.

$$\sum_{i \in Q} x_i \geq \frac{q}{100} \cdot N$$

5. At least p% of the selected applicants must not be engineers. The value of p utilized was 30 (this constraint was added for 2009).

$$\sum_{i \in P} x_i \geq p/100 \cdot N$$

We now describe the objective functions of each model and, in the case of the third one, an additional decision variable and constraint.

2.1. Model 1

The objective is to maximize the sum of the selected applicants' scores. The idea behind this is to find a global optimum score.

Objective Function

$$\max \sum_{i \in K} x_i \cdot p_i$$

2.2 Model 2

The idea behind this model is similar to that for Model 1, the difference being that here we consider the candidates' ranking order rather than their scores. The objective is therefore to minimize the sum of the selected applicants' rankings.

Objective Function

$$\min \sum_{i \in K} i \cdot x_i$$

In the event of tied scores between two applicants, a better ranking is attributed to the applicant who satisfies a greater number of the equity and professional background characteristics the program seeks to favor (women, lower income quintiles, non-Santiago regions of origin, non-engineers). If the tie persists, the ranking is defined randomly but the details are recorded, and if the applicants in question are among those admitted in the final selection stage or placed on the waiting list, the organizers make the final decision based on a qualitative criterion they consider appropriate.

2.3 Model 3

This model aims to provide a sort of guarantee regarding the whole set of selected applicants by imposing the condition that the last one chosen has the best ranking possible. The objective is thus to minimize the ranking of the last selected applicant. The model contains an additional decision variable *y* (positive real) not appearing in the other two whose value is greater than or equal to the ranking of all the selected applicants, and once minimized will be the ranking of the last chosen candidate. The model also incorporates an extra constraint that requires the new variable to be greater than or equal to the position on the (ordered) list of all the selected applicants. The objective function value will minimize the sum of this variable's value and that of the objective function value of Model 2, the latter multiplied by a very small number. This is done so that given two set of candidates with a tie in the ranking of the last chosen candidate, the set of best ranked applicants is selected. Clearly, the second term of the sum will not effect the result if the two set of candidates have different ranking of the last chosen applicant.

Decision variable

y: the relative position greater than or equal to selected applicants.

Constraint

$$i \cdot x_i \leq y \qquad \forall i$$

Objective Function:

$$\min\left\{ y + \left(0.0002 \cdot \sum_{i \in K} i \cdot x_i \right) \right\}$$

From these models we could easily construct examples in which each one identified a different group of admitted applicants. This is confirmed by the computational results set out below in Section 4, which show that the admissions list cutoff varies slightly between the models. To achieve a more robust solution, therefore, we will employ all 3 models instead of just 1 in a procedure that combines their results using a selection algorithm described in the next section.

3. Selection Algorithm

The procedure defined by the selection algorithm runs the models a set number of times before combining the best solutions so generated to produce a single final solution. This number of runs is a parameter chosen by the user. Here, we used the three best solutions of each model, Run 1 yielding the best solution, Run 2 the second best, and Run 3 the third best. The second best solution is obtained by adding a constraint to the models that renders the best solution infeasible; the third-best solution is derived by similarly eliminating also the second-best one. If there exist unique best, second best and third best solutions, the applicants in each of them are assigned the coefficients 1, 0.6 and 0.3, respectively. These values are then summed across all three solutions (*i.e.*, runs) and models for each applicant. If, for example, an applicant is selected in Run 1 of models 1 and 2, Run 2 of models 1 and 2 and Run 3 of Model 1, he or she is assigned a general weighting coefficient of 3.5. Finally, this value is then multiplied by the person's point total to arrive at a new score.

More specifically, the steps in the algorithm are as follows:

1. First Selection: The applicants appearing in the optimal solution (Run 1) of all three models are identified. These candidates are immediately admitted to the program. If the three models return the same optimal solution, the admissions list is complete and the algorithm jumps to step 5 to identify the waiting list. If not, it goes to step 2.

2. New Score: The coefficients described above are now calculated for each applicant not selected in Step 1 and then multiplied by their respective point totals to generate new scores.

3. Second Selection: The composition of the admissions decided in the First Selection in terms of candidates from the three equity categories (women, non-Santiago regions and lower income quintiles) is evaluated to determine how many more of each category are needed to meet the required percentage minima. Model 2 is then run using the scores obtained in Step 2 with constraints that ensure, first, that it selects at least the number of candidates required to make up these minima, and second, that those so selected equal the number lacking in the First Selection to satisfy the program total *N*. The algorithm then checks whether the optimal solution is unique. If it is not, the algorithm proceeds to Step 4; if it is, the applicants in the solution are selected, thus completing the admissions list, and the algorithm jumps to Step 5.

4. Third Selection: The sum of the scores of each of the solutions found in Step 3 is calculated (in other words, Model 1 is applied). The group with the highest point total

completes the list of admitted applicants. If two or more solutions are still tied, all of the alternatives are presented to the program organizers for a final decision.

5. Waiting List: If the number of applicants who appear in any of the nine runs but are not admitted is greater than 20, the top 20 scorers among them are placed on the waiting list. If the number is less than 20, all of them are placed on the waiting list and the additional applicants needed to complete it are chosen from the best scorers (before the weighting) among those who were not selected in the best solutions of any of the models.

The identification of the waiting list candidates does not take into account the equity criteria. If, however, any of the admitted applicants later decline to enter the program, they are replaced with the highest scorers among the waiting list applicants in such a way that the equity category minima are met.

The selection algorithm guarantees the robustness of the final solution in the sense that the applicants admitted to the program will have all figured in various of the best solutions of each model. This clearly reveals the value of using mathematical programming models, for it would be practically impossible to obtain such results in relatively few minutes using manual methods. Furthermore, our programming tool gives the process a high level of transparency. To illustrate how the algorithm actually functions, its application to the 2008 selection process is set out in Appendix 2.

4. Results

In the selection processes for the three years, the models were used in the first two stages only to make sure there were enough applicants from the three equity categories. They therefore acted simply as a support tool for deciding which candidates would advance to the next stage, and the selection algorithm was not used at all. The results we will present in this section relate to the final stage of the three processes when the models are employed complete with selection algorithm. Note that the admissions list thus arrived at was adopted as the definitive one by the program organizers in all three years.

We begin with the selection process for 2007, and then examine the process for 2008 and 2009.

4.1 Selection process for 2007

The objective function results for the 3 best solutions of each model with the 2 different income quintile definitions used by the program organizers in 2007 are shown in Table 1.

Model	Best solution	O.F. value (1st quintile definition)	O.F. value (2nd quintile definition)	Bound for the O.F.
1	1	3334.0798	3392.6797	3399.414
1	2	3334.0717	3392.4	
1	3	3333.6946	3391.9229	
2	1	1792	1470	1431
2	2	1795	1473	
2	3	1795	1476	
3	1	71.3702	64.294	53.2862
3	2	71.3714	64.2946	
3	3	71.3716	64.2952	

Table 1: Objective function values under both income quintile definitions in the 2007 selection process.

As can be observed, the second quintile definition leads to superior objective function values with all three models. This is so because under this definition, the set of persons in quintiles other than the top one is larger. Also, note that an upper bound for the objective function in Model 1, which would be obtained if the only constraint were the selection of the top 53 applicants to fill the program without any equity criteria restrictions, is 3399.414.

The corresponding theoretical lower bounds for models 2 and 3 are 1431 and 53.2862, respectively (the decimals in the latter figure are used to break a tie if two solutions select the same candidate as the last admitted applicant). These are the values the function would have if the best 53 scorers were selected, that is, if there were no equity constraints. Again, the constraints have a major impact on the objective function values when the first quintile concept is applied.

Another interesting point is that under the second quintile definition the selection algorithm jumps directly from Step 1 to Step 5, meaning that the admitted applicants are the same with all three models. If the first definition is employed, however, the algorithm must execute Step 3 before going to Step 5. This is so because the models' best solutions coincide on only 48 (of a possible 53) selected candidates. Such cases illustrate the importance of designing a transparent process that combines the results of the different models and completes the admissions list by applying the equity constraints. This process also ensures the robustness of the final solution. In this case, for example, the 5 candidates selected to complete the list of 53 must appear in various of each model's best solutions (something that would be impossible to check in a reasonable amount of time with a manual method).

As regards the waiting list, under the first quintile definition there were 9 applicants who appeared in the run solutions but were not admitted. To complete the list, therefore, 11 more candidates had to be chosen from among those who did not appear in any solution. Under the second definition there were only 3 who figured in the solutions but were not admitted, leaving 17 additional waiting list applicants to be selected. These data indicate that under the first definition, 62 applicants appear in the nine runs while under the second, the number falls to 56. In other words, with the second definition the models coincide to a high degree not only in their best solutions but the second and third best ones as well. Indeed, the second best solutions coincide perfectly as do the third best for models 1 and 3, the latter solutions differing on only one candidate from Model 2. In the end, the program organizers opted for the second quintile definition (using it again in 2008 and 2009) so as to improve the academic quality of the set of chosen candidates. In the rest of this study the second quintile definition will therefore be applied exclusively.

4.2 Selection process for 2008

The objective function results for the 3 best solutions of each model in the 2008 selection process are shown in Table 2.

Model	Best solution	O.F. value	Bound for the O.F.
1	1	3322.65	3325.4
1	2	3322.6	
1	3	3322.5	
2	1	1351	1326
2	2	1353	
2	3	1353	
3	1	55.2726	51.2662
3	2	55.273	
3	3	55.2734	

Table 2: Objective function values in the 2008 selection process.

The best possible value for Model 1 is 3325.4, the sum of the scores for the top 51 candidates (the number admitted in 2008) with no other constraints applied.

As for Model 2, with the reduction of admissions from 53 to 51, the objective function value assuming no other constraints is 1326 while for Model 3 it is 51.2662.

Thus, in 2008 the optima for the three models are closer to their ideal values than in 2007. This result may be attributable to the fact that, as mentioned earlier, the socioeconomic level and regional origin constraints in 2008 were less restrictive than the year before.

When the selection algorithm was run, the best solutions of the three models coincided on 49 (of a possible 51) applicants in Step 1. The algorithm thus had to execute Step 3 before jumping to Step 5. The details are given in Appendix 2.

Regarding the waiting list applicants, since 4 of them figured in a run solution the total number appearing in any of the 9 runs was 55.

As regards the runs themselves, note that Model 2 generates two second best solutions, and both Model 1 and Model 2 yield the same optimal solution which differs from Model 3 in 2 applicants.

4.3 Selection process for 2009

For the 2009 process the program organizers decided to broaden the criteria used for deriving candidates' scores. With this in mind, two changes were implemented. The first was to define another pair of weighting factors that placed slightly more importance on work history and slightly less on academic background. Each applicant therefore had two initial point totals in 2009, one based on the newly defined weights and the other on the original weights used in 2007/2008.

The second change was that applicants who passed the psychological test, the last stage in defining the short list, were grouped into three aptitude categories: more qualified (I), qualified (II) and less qualified (III). Their point totals were then multiplied by a coefficient corresponding to the category they were assigned to. Three sets of such coefficients were defined: 1, 0.95, 0.9; 1, 0.97, 0.92; and 1, 0.92, 0.85.

The combined effect of the changes resulted in the creation of six final-score totals, denoted scenarios. Thus, Scenario 1 combined the 2007/2008 point-total weights with the first set of coefficients, Scenario 2 combined them with the second set, and Scenario 3 with the third set. Scenarios 4, 5 and 6 were formed analogously by combining the 2009 point-total weights with the first, second and third coefficient sets, respectively.

For each of the 6 scenarios, the same selection procedure used in 2007 and 2008 was applied to determine which of the short-listed applicants would be admitted. The admissions list was completed by maximizing the number of scenarios in which an applicant was selected subject to the equity and no-engineer constraints. Candidates chosen in any scenario who were not in the final group were placed on the waiting list, which was filled with the 20 top scorers among those who were not selected in any scenario. For this purpose the point totals used were the averages over the six scenarios.

The objective function results for the best solutions of each model and each scenario in the 2009 selection process are shown in Table 3.

Scenario	Model	O.F. value	Bound for the O.F.

E1	1	2846.306	2930.316
E1	2	1556	1128
E1	3	75.3276	47.2256
E2	1	2880.602	2964.523
E2	2	1559	1128
E2	3	74.3256	47.2256
E3	1	2793.659	2875.891
E3	2	1512	1128
E3	3	75.3090	47.2256
E4	1	2831.390	2927.985
E4	2	1610	1128
E4	3	74.3262	47.2256
E5	1	2865.769	2961.709
E5	2	1617	1128
E5	3	74.3270	47.2256
E6	1	2778.286	2872.713
E6	2	1569	1128
E6	3	74.3182	47.2256

Table 3: Objective function values in the 2009 selection process

Note that since the applicants' scores vary from one scenario to the next, the upper bound for Model 1 also varies by scenario.

In models 2 and 3, however, the scores do not affect the objective function result and their lower bounds therefore do not vary by scenario, the respective values being 1128 and 47.2256.

The data in the table reveal that the differences between the ideal theoretical values and the actual values are greater than for the previous years. This implies that the positive discrimination and no-engineer constraints impact more strongly on the results. The divergence in the case of Model 1 is widest for Scenario 4, while for models 2 and 3 the gap is greatest in Scenarios 5 and 1, respectively.

Under all six scenarios the selection algorithm had to be employed to complete the admissions list of 47 candidates, and did so successfully. The highest number of selected applicants for which the results of all three models coincided was 44 (of a possible 47), observed in scenarios 3 and 6, while the lowest number was 41, generated by scenario 2.

To elect the 47 admitted applicants we checked which ones were repeated in all six final scenario solutions. A total of 43 appeared in all of the solutions and the remaining four candidates were those selected in the greatest number of scenarios subject to the organizers' constraints. This procedure led to a tie between three groups of four. In each of them there was one applicant that appeared in five final scenario solutions, another one in four solutions and two in three solutions. The group finally chosen was the one whose point total weighted by the six scenarios was the highest.

Execution time did not exceed 5 seconds for any of the 9 model runs. The entire procedure for each of 2007 and 2008 was completed in about 20 minutes while for 2009 it required slightly more than 90 minutes due to the multiple scenarios involved. The model solutions were generated using CPLEX 10.0 on a computer equipped with a 2.0 GHz Pentium IV processor and 1 GB of RAM.

5. Conclusions

In the first part of this section we present various sensitivity analyses in order to determine the effects on the results of the various equity criteria constraints. The significance of this step was explained by Ms. Lysette Henríquez, Executive Director of the Master's program during the two first selection processes, in the following manner: "A key aspect of the model's application is the sensitivity analyses conducted during the decision processes. Visualizing a solution given a set of constraints is fundamental and practically impossible to do manually, but perhaps even more important is being able to vary the program parameters within a reasonable margin or make minor modifications to the objective function to examine other interesting elements of the program. A key factor is the ability to appreciate how robust is the presence of certain applicants in the solution, that is, whether or not they appear systematically in the final solution. Having this information allows the decision-makers to feel more certain they are making the right admission choices."

The first of these *a posteriori* sensitivity analyses investigates how many of those admitted to the program would not have been without the application of the equity criteria. The results show that 4 of the 53 admitted candidates in 2007 (7.5%), 2 of the 51 in 2008 (3.9%) and 14 of the 47 in 2009 (29.8%) would not have been accepted without this positive discrimination. The decline in the number for 2008 reflects that fact that the percentage minima for each equity category were reduced slightly for the non-Santiago region and income quintile criteria. Even though the percentage changes relative to the admissions based solely on ranking are small, the fact that the process involves decisions that impact the applicants' personal and professional futures makes it imperative the criteria adopted are backed by a fair and transparent mechanism such as the one we have developed.

The positive discrimination criteria applied in 2009 impacted strongly on the admissions list, revealing the need for greater use of mathematical models to ensure a fair admissions process. It should be stressed that this goes further than the simple fact that six scenarios were run for 2009 instead of just one. If, for example, there had only been one scenario in 2009 (scenario 1), the number of applicants rejected due to the constraints would also have been 14 (with similar results for the other 5 scenarios). This may be due to the addition of the new (non-engineer) constraint but could also be attributed to the random differences in the makeup of the original applicants from one year to the next.

An analysis of the results for the 2007 process shows that the admitted applicant numbers exactly fulfilled the minima required by the female and non-Santiago region equity criteria but not the minima for the lower income quintiles. Analyzing the 2008 results, we find that the admitted applicants exactly equaled the non-Santiago region and lower income quintile minima while the female category minimum was exceeded by 1. Similarly, in 2009 admission numbers equaled the required minima in all categories save women, where the amount was one higher than the lower bound.

In the case of the 2007 process, if the female admissions minimum is eliminated, one fewer women and one more man would be selected. If there is no non-Santiago region minimum, two more Santiago candidates would be admitted. This indicates that removing one of the equity constraints while maintaining the others has no major effect on the final solution.

Turning to the 2008 process, if the regional origin constraint is eliminated the algorithm terminates upon completing Step 1 (implying the three models give the same best solution) after selecting the exact minima for female and lower income quintile admissions and one fewer non-Santiago applicant than the required minimum. If the income quintile constraint (the other active restriction in 2008) is excluded, the algorithm again terminates once Step 1 has been executed after selecting the exact minimum numbers of female and non-Santiago candidates. On the other hand, the percentage of lower income quintile candidates is 64.7% (three fewer lower income quintile applicants than the minimum required when this constraint is included).

As for the 2009 process, if the 6 scenarios are run without the minimum non-engineer restriction we get a final result in which only 23.4% of the selected group are non-engineers as

opposed to the 30% imposed by the constraint. If we eliminate the gender constraint in all scenarios, upon combining the solutions we end up with a list of selected applicants that also contains only 23.4% females instead of 30% if the constraint is applied. Removal of the income quintile constraint results in only 57.44% of admitted candidates from the lower quintiles. And finally, if the region of origin constraint is removed, the final solution contains 55.32% from the non-Santiago region, relatively close to the constraint requirement of 60%.

An aspect of the selection process that greatly interested the program organizers was the impact of the psychological evaluation on candidate selection. This test eliminated 42% of the applicants who had made it through to the third stage in 2007. The figure for 2008 was considerably lower at 19.41%, although in this process the test was applied after the English language test which did not exist the previous year. In 2009 the psychological evaluation, which partitioned those who passed into three aptitude categories, was also administered after the English test and eliminated 19.62% of the candidates. It therefore had a major impact on applicant selection; indeed, had the evaluation not been given, 13 successful candidates (24.52% of the total) in 2007, 11 (21.56%) in 2008 and 12 (25,53%) in 2009 would have been denied admission in favor of others who failed it.

Another general observation of interest is that if the psychological evaluation is excluded, the Model 1 result improves whereas the other models' results may either improve or deteriorate. This is because the test reduces the number of applicants. Model 1 could not therefore produce a better solution than the one it generated without the test because with the same point totals the list of applicants is a subset of the original one. For models 2 and 3, however, a shortened list does not *a priori* affect the sum of the applicants' rankings or the ranking of the last admitted applicant because the relative positions of those who remain on the list may change in the new scenario. Hence, the impact on the objective function value in both Model 2 and Model 3 will depend entirely on which applicants are eliminated by the test.

According to Ms. Henríquez, "these *a posteriori* analyses reveal the consequences of applying certain restrictions and enable us to make program policy decisions with full awareness of their impacts." She added that "in short, the contribution of the model has been fundamental to ensuring transparency of decisions involving the award of a grant of some US$ 75,000 per student for a program that received more than 800 applications from around the country in 2008 and required an investment for the first three years of close to 12 million dollars. This is particularly significant considering that the program's purpose is to stimulate the creation of a meritocracy."

The 2009 admission process incorporated a series of different scenarios to increase the robustness of the final decision. As explained by Patricio Meller, Academic Director of the program since its inception, "having alternative models and scenarios has been fundamental to our ability to select more suitable candidates. No model is perfect, so with various scenarios the risk of rejecting a good applicant is lower. That's the real advantage of selecting those who appear in the final solution in the majority of scenarios."

As a general conclusion of the study we would emphasize its demonstration of the Operations Research and Mathematical Programming ability to contribute to social policy issues, and in particular the usefulness of these techniques in identifying the applicants to a degree program who best fit the desired profile in terms of equity criteria based on regional origin, socioeconomic background and gender.

It is still too early to conduct a full analysis of the impact of the program. Meller, however, offers this preliminary verdict: "The first graduating class has already completed the program and the 53 students all did extremely well. For students coming from lower-income families or regions distant from Santiago, the program can significantly change their life paths. This is the sort of impact we hope to achieve with positive discrimination. On this point we're convinced that the mathematical models we applied enabled us to choose the most appropriate applicants among those who met the constraints imposed. We cannot imagine decision-making for future admission processes without the support of these tools."

We conclude by underlining the fact that finding robust solutions to the admissions problem in a matter of minutes using manual techniques would have been simply impossible. The mathematical tools developed for this task also had the added advantage of bringing transparency to the selection process.

Acknowledgements: To Lysette Henríquez Amestoy, Patricio Meller, Gastón Held, and the executive and academic personnel at the MGG, with whom it was a joy to work. To Andrés Weintraub and Kenneth Rivkin for their interesting comments which helped to improve this study. And to the anonymous referees for their valuable suggestions that helped us improve this work. The first author was partially supported by FONDECYT Project No. 1080286 and the Millenium Complex Engineering Systems Institute.

References

[1] F. Chan, "Interactive selection model for supplier selection process: an analytical hierarchy process approach", International Journal of Production Research 41 (15) (2003), 3549 – 3579.

[2] R. Epstein, L. Henríquez, J. Catalán, G. Weintraub and C. Martínez, "A Combinatorial Auction Improves School Meals in Chile", Interfaces 32 (2002), 1-14.

[3] J. Grandzol, "Improving the faculty selection process in higher education: a case for the analytic hierarchy process", IR Applications 6 (2005), 1-13.

[4] M. Ross and R. Nydick, "Selection of licensing candidates in the pharmaceutical industry: an application of the analytic hierarchy process", Journal of Health Care Marketing 12 (2) (1992), 60-65.

[5] T. Saaty, "The Analytic Hierarchy Process" (1980), McGraw Hill.

Appendix 1: Final results of selection processes, 2007, 2008 and 2009.

Table 4 displays the final results for the 2007 selection process, with the applicants arranged by descending order of scores. As indicated by the "selected" column, applicants in positions 44, 45, 50 and 52 are the 4 of the top-ranked 53 who are not among the 53 candidates on the definitive admissions list (*i.e.*, after application of the equity criteria), while those in positions 54, 55, 57 and 64 are their replacements.

Ranking	Score	Gender	Non-Santiago region	Lower income quintiles (Definition 2)	Selected
1	77.3967	Male	Yes	Yes	Yes
2	74.6663	Male	No	Yes	Yes
3	73.1412	Female	No	Yes	Yes
4	70.9622	Male	No	Yes	Yes
5	70.2533	Male	Yes	Yes	Yes
6	70.0854	Male	Yes	Yes	Yes
7	68.9846	Male	No	Yes	Yes
8	68.3338	Male	Yes	Yes	Yes
9	68.2611	Male	No	Yes	Yes
10	68.2314	Female	Yes	Yes	Yes
11	67.7061	Male	No	No	Yes
12	67.5873	Male	Yes	Yes	Yes
13	67.4197	Male	Yes	Yes	Yes
14	67.3683	Female	No	Yes	Yes

15	67.336	Male	Yes	Yes	Yes
16	67.148	Male	No	No	Yes
17	65.7685	Male	No	Yes	Yes
18	65.3751	Female	Yes	Yes	Yes
19	65.0443	Male	No	Yes	Yes
20	64.495	Female	No	Yes	Yes
21	64.2388	Female	Yes	Yes	Yes
22	63.8693	Male	No	Yes	Yes
23	63.4154	Male	No	Yes	Yes
24	63.3793	Male	No	Yes	Yes
25	63.0156	Male	Yes	Yes	Yes
26	62.8584	Male	No	Yes	Yes
27	62.7446	Male	No	Yes	Yes
28	62.6285	Male	Yes	Yes	Yes
29	62.2127	Male	Yes	Yes	Yes
30	62.1483	Male	Yes	Yes	Yes
31	62.1481	Male	Yes	Yes	Yes
32	62.0838	Female	Yes	Yes	Yes
33	62.0832	Male	No	Yes	Yes
34	62.0342	Male	No	Yes	Yes
35	61.9296	Female	No	Yes	Yes
36	61.7264	Male	Yes	Yes	Yes
37	61.4523	Male	Yes	No	Yes
38	61.1665	Male	Yes	Yes	Yes
39	60.8406	Male	No	Yes	Yes
40	60.8082	Male	Yes	Yes	Yes
41	60.6791	Male	No	Yes	Yes
42	60.6263	Male	Yes	Yes	Yes
43	60.4522	Male	Yes	No	Yes
44	60.3994	Male	No	Yes	No
45	60.0838	Male	No	Yes	No
46	59.9693	Female	Yes	Yes	Yes
47	59.8533	Male	Yes	No	Yes
48	59.4615	Female	Yes	Yes	Yes
49	59.4572	Female	Yes	Yes	Yes
50	59.4336	Male	No	Yes	No
51	59.2239	Female	Yes	Yes	Yes
52	58.7673	Male	No	Yes	No
53	58.659	Female	Yes	Yes	Yes
54	58.6414	Female	Yes	Yes	Yes
55	58.5681	Male	Yes	Yes	Yes
56	57.7766	Male	No	Yes	No
57	57.7504	Female	Yes	Yes	Yes
58	57.5946	Male	No	Yes	No
59	57.5842	Male	No	Yes	No
60	57.5556	Male	No	Yes	No
61	57.5294	Male	Yes	Yes	No
62	57.3431	Male	Yes	Yes	No
63	57.2793	Male	No	No	No
64	56.9899	Female	Yes	Yes	Yes
65	56.5452	Male	Yes	Yes	No
66	56.5313	Female	No	Yes	No

67	56.4444	Male	Yes	Yes	No
68	56.421	Male	Yes	No	No
69	56.2399	Male	No	Yes	No
70	56.1681	Male	No	Yes	No
71	55.9509	Female	No	Yes	No
72	55.821	Male	Yes	Yes	No
73	55.6551	Male	No	Yes	No
74	55.4727	Female	No	Yes	No
75	55.4646	Female	No	Yes	No
76	55.4358	Male	Yes	Yes	No
77	55.2457	Male	No	Yes	No
78	55.2024	Male	No	Yes	No
79	55.0265	Female	No	Yes	No
80	55.0064	Male	No	Yes	No
81	55.0007	Male	Yes	Yes	No
82	54.9741	Female	No	Yes	No
83	54.9328	Male	Yes	Yes	No
84	54.59	Male	No	No	No
85	54.4713	Male	Yes	Yes	No
86	54.3639	Female	Yes	No	No
87	54.1758	Male	No	Yes	No

Table 4: Final results of 2007 selection process.

Table 5 displays the final results for the 2008 selection process. As with the preceding table, the applicants are arranged by descending order of scores. The "selected" column indicates that those ranked in positions 39 and 51 are the 2 candidates of the top-ranked 51 who are not among the 51 on the definitive admissions list (*i.e.*, after application of the equity criteria), while those in positions 53 and 59 are their replacements.

Ranking	**Score**	**Gender**	**Non-Santiago region**	**Lower income quintiles**	**Selected**
1	74.80	Female	Yes	No	Yes
2	74.45	Male	Yes	Yes	Yes
3	73.60	Male	No	No	Yes
4	70.85	Male	Yes	Yes	Yes
5	69.45	Female	Yes	Yes	Yes
6	69.10	Male	Yes	Yes	Yes
7	69.10	Male	No	Yes	Yes
8	68.15	Male	Yes	Yes	Yes
9	68.15	Female	No	Yes	Yes
10	68.00	Male	No	Yes	Yes
11	67.65	Male	No	No	Yes
12	67.40	Male	Yes	No	Yes
13	67.35	Male	Yes	Yes	Yes
14	67.35	Male	Yes	Yes	Yes
15	67.05	Male	Yes	No	Yes
16	66.90	Male	Yes	Yes	Yes
17	66.30	Male	Yes	Yes	Yes
18	66.10	Male	No	Yes	Yes

19	66.10	Male	No	No	Yes
20	65.55	Male	No	No	Yes
21	65.40	Female	No	Yes	Yes
22	65.30	Female	Yes	Yes	Yes
23	65.00	Male	No	Yes	Yes
24	64.50	Female	No	Yes	Yes
25	64.40	Male	No	Yes	Yes
26	64.05	Male	No	No	Yes
27	63.90	Male	Yes	No	Yes
28	63.90	Male	No	No	Yes
29	63.85	Male	No	Yes	Yes
30	63.85	Male	Yes	No	Yes
31	63.60	Male	Yes	Yes	Yes
32	63.60	Female	No	No	Yes
33	63.55	Male	Yes	Yes	Yes
34	63.25	Female	No	No	Yes
35	63.00	Male	Yes	Yes	Yes
36	62.65	Female	Yes	Yes	Yes
37	62.65	Female	No	No	Yes
38	62.50	Female	Yes	No	Yes
39	62.50	Male	No	No	No
40	62.45	Female	No	Yes	Yes
41	62.35	Male	No	Yes	Yes
42	62.25	Female	No	Yes	Yes
43	62.20	Male	Yes	Yes	Yes
44	62.15	Male	Yes	Yes	Yes
45	61.70	Female	Yes	Yes	Yes
46	61.60	Female	Yes	Yes	Yes
47	61.45	Male	No	Yes	Yes
48	61.40	Male	Yes	Yes	Yes
49	61.40	Female	No	Yes	Yes
50	60.80	Male	Yes	Yes	Yes
51	60.80	Male	Yes	No	No
52	60.75	Female	Yes	No	No
53	60.60	Male	Yes	Yes	Yes
54	60.60	Female	No	No	No
55	60.40	Male	No	Yes	No
56	60.40	Male	Yes	No	No
57	60.35	Male	No	No	No
58	60.10	Male	No	Yes	No
59	59.95	Female	Yes	Yes	Yes
60	59.90	Male	Yes	Yes	No
61	59.90	Male	No	Yes	No
62	59.70	Male	Yes	No	No
63	59.50	Male	No	Yes	No
64	59.50	Male	No	No	No
65	59.45	Female	Yes	Yes	No
66	59.45	Female	No	No	No

67	59.25	Female	No	Yes	No
68	59.00	Male	No	No	No
69	58.95	Male	No	No	No
70	58.80	Male	No	Yes	No
71	58.75	Male	Yes	Yes	No
72	58.75	Female	No	Yes	No
73	58.35	Male	Yes	No	No
74	57.90	Male	Yes	Yes	No
75	57.90	Male	No	Yes	No
76	57.65	Female	No	Yes	No
77	57.60	Male	No	No	No
78	57.55	Female	Yes	No	No
79	57.30	Female	Yes	Yes	No
80	57.00	Male	Yes	Yes	No
81	56.95	Male	No	Yes	No
82	56.90	Male	Yes	Yes	No
83	56.75	Male	No	Yes	No

Table 5: Final results of 2008 selection process.

Table 6 displays the final results for the 2009 selection process. As with the preceding table, the applicants are arranged by descending order of scores. In this case, however, the score is the average of the 6 scenarios. The “is an engineer” column reflects the profession constraint, which did not exist in 2007 and 2008. The rightmost column gives the number of scenarios in which an applicant is selected in the final solution. The “selected” column indicates the 14 candidates who were among the best 47 scorers but who were replaced in the final admissions list by 14 others who were not originally among those 47. Note, for example, that the 8th ranked applicant is not on the final list due to his status as a male engineer from the Santiago region in the top income quintile.

Ranking	Weighted scores	Gender	Non-Santiago region	Lower income quintiles	Is an engineer	Selected	No. of scenarios
1	72.13	Male	No	Yes	True	Yes	6
2	70.14	Female	Yes	No	True	Yes	6
3	68.97	Male	No	Yes	True	Yes	6
4	68.68	Male	Yes	No	True	Yes	6
5	67.51	Male	No	No	True	Yes	6
6	67.18	Male	Yes	Yes	True	Yes	6
7	66.98	Male	Yes	No	True	Yes	4
8	66.26	Male	Yes	No	True	No	2
9	65.57	Male	No	Yes	True	Yes	6
10	65.46	Female	Yes	Yes	False	Yes	6
11	65.27	Male	No	Yes	True	Yes	6

12	64.74	Male	Yes	Yes	False	Yes	6
13	64.49	Male	No	No	True	Yes	6
14	64.15	Male	No	Yes	True	Yes	6
15	63.54	Female	No	Yes	True	Yes	6
16	63.22	Male	No	No	True	Yes	6
17	62.96	Male	Yes	Yes	False	Yes	6
18	62.86	Male	Yes	Yes	True	Yes	6
19	62.60	Male	Yes	Yes	False	Yes	6
20	61.35	Male	Yes	No	True	No	
21	61.30	Female	Yes	No	True	Yes	5
22	61.26	Male	No	No	True	No	1
23	61.21	Male	Yes	Yes	False	Yes	6
24	61.11	Male	No	No	True	No	2
25	60.66	Male	No	No	True	No	
26	60.62	Male	No	Yes	False	Yes	6
27	60.56	Male	No	Yes	True	Yes	6
28	60.47	Male	Yes	No	True	No	
29	60.41	Male	No	No	True	No	
30	60.24	Male	Yes	No	True	No	
31	59.96	Male	Yes	No	True	No	
32	59.75	Male	No	No	True	No	
33	59.63	Male	Yes	No	False	No	1
34	59.61	Male	No	Yes	True	Yes	6
35	59.53	Male	Yes	No	True	No	
36	59.39	Male	Yes	Yes	True	Yes	6
37	59.35	Male	No	Yes	True	Yes	6
38	59.14	Male	No	Yes	True	Yes	6
39	59.13	Male	Yes	Yes	True	Yes	6
40	59.10	Male	Yes	Yes	True	Yes	6
41	58.66	Male	Yes	No	True	No	
42	58.37	Female	Yes	Yes	False	Yes	6
43	58.15	Male	No	No	False	Yes	6
44	57.64	Female	Yes	No	True	No	
45	57.37	Female	No	Yes	True	Yes	6
46	57.26	Female	No	No	True	Yes	3
47	56.97	Male	No	Yes	False	Yes	6
48	56.88	Male	Yes	No	True	No	
49	56.14	Male	No	Yes	False	Yes	6
50	56.11	Male	No	Yes	True	Yes	6
51	56.07	Female	Yes	No	True	No	

52	55.71	Male	Yes	No	True	No	
53	55.53	Female	Yes	Yes	False	Yes	6
54	55.42	Female	No	No	True	No	
55	55.22	Female	No	Yes	False	Yes	6
56	54.99	Male	Yes	Yes	True	No	
57	54.90	Female	No	Yes	True	Yes	6
58	54.62	Male	Yes	Yes	True	No	
59	54.53	Female	No	Yes	True	Yes	6
60	54.49	Female	Yes	No	True	No	
61	53.96	Female	No	Yes	True	Yes	6
62	53.63	Male	No	Yes	True	Yes	3
63	53.43	Female	No	Yes	True	Yes	6
64	53.29	Male	No	Yes	True	No	3
65	53.22	Male	No	Yes	True	No	
66	53.16	Male	No	No	True	No	
67	53.14	Male	No	No	True	No	
68	53.12	Male	No	Yes	True	No	
69	53.11	Female	Yes	Yes	True	Yes	6
70	53.08	Female	No	No	True	No	
71	53.07	Male	No	Yes	False	Yes	6
72	52.84	Female	No	Yes	True	Yes	6
73	52.25	Male	No	Yes	True	No	
74	51.54	Female	No	No	True	No	
75	51.01	Male	No	Yes	False	Yes	6
76	50.81	Male	Yes	Yes	True	No	
77	50.36	Female	Yes	No	False	No	
78	50.16	Male	Yes	No	False	No	
79	49.94	Male	Yes	Yes	True	No	
80	49.38	Male	Yes	No	True	No	
81	49.03	Male	No	No	True	No	
82	48.98	Female	Yes	Yes	False	Yes	6
83	48.17	Male	Yes	Yes	True	No	
84	47.60	Male	Yes	Yes	True	No	
85	47.40	Male	Yes	No	True	No	
86	47.04	Male	Yes	No	True	No	

Table 6: Final results of 2009 selection process.

Appendix 2: Application of Selection Algorithm to 2008 Process

In what follows we illustrate the functioning of the selection algorithm as it was applied to the 2008 selection process. Table 7 contains all of the applicants appearing in at least one best solution of any model, indicating which of the 9 solutions they were selected by. The listing is ordered by the candidates' personal ID numbers shown in the leftmost column, which were assigned to ensure anonymity.

Note that Model 2 has two second best solutions, denoted 2a and 2b. In case of a tie between any of the solutions of a given model, the technique for assigning the general weighting coefficient is generalized. Thus, if (as in this case) there is one best solution and two second-best solutions, the best one is assigned a value of 1 as usual while the two second best ones are each assigned one-half the sum of the second- and third-best coefficients ((0.6 + 0.3) / 2 = 0.45). The condition that the sum of the coefficients equals 1.9 is thereby maintained.

Table 8 shows the successful applicants' progress through the stages of the selection process, indicating in the various columns whether they were immediately selected for admission in Step 1, advanced to the following steps (in which case their weighting coefficients are also given), selected for admission in Step 3, or placed on the waiting list.

In both tables, if applicant *i* possesses attribute *j*, component *i,j* in the table is marked with an X.

Applicant ID	BS1 Mod 1	BS2 Mod 1	BS3 Mod 1	BS1 Mod 2	BS2a Mod 2	BS2a Mod 2	BS1 Mod 3	BS2 Mod 3	BS3 Mod 3
13			X		X				
21	X	X	X	X	X	X	X	X	X
42	X	X	X	X	X	X	X	X	X
49	X	X	X	X	X	X	X	X	X
62	X	X	X	X	X	X	X	X	X
66	X	X	X	X	X	X	X	X	X
139	X	X	X	X	X	X	X	X	X
169	X	X	X	X	X	X	X	X	X
175	X	X	X	X	X	X	X	X	X
176							X	X	X
198	X	X	X	X	X	X	X	X	X
241	X	X	X	X	X	X	X	X	X
249							X	X	X
250	X	X	X	X	X	X	X	X	X
258	X	X	X	X	X	X	X	X	X
261	X	X	X	X	X	X	X	X	X
290	X	X	X	X	X	X	X	X	X
291	X	X	X	X	X	X	X	X	X
302	X	X	X	X	X	X	X	X	X
314	X	X	X	X	X	X	X	X	X
315	X	X	X	X	X	X	X	X	X
325	X	X	X	X	X	X	X		X
371	X	X	X	X	X	X	X	X	X
372	X	X	X	X	X	X	X	X	X
382	X	X	X	X	X	X	X	X	X
392	X	X	X	X	X	X	X	X	X
398	X	X	X	X	X	X	X	X	X
402	X	X	X	X	X	X	X	X	X
413	X	X	X	X	X	X	X	X	X

444	X	X	X	X	X	X	X	X	X
456	X	X	X	X	X	X	X	X	X
469	X	X	X	X	X	X	X	X	
485	X	X	X	X	X	X	X	X	X
499	X	X	X	X	X	X	X	X	X
510	X	X	X	X	X	X	X	X	X
517	X	X	X	X	X	X	X	X	X
531	X	X	X	X	X	X	X	X	X
538	X	X	X	X	X	X	X	X	X
544	X	X	X	X	X	X	X	X	X
548	X	X	X	X	X	X	X	X	X
567	X		X	X	X				
577	X	X	X	X	X	X	X	X	X
593	X	X	X	X	X	X	X	X	X
635	X	X	X	X	X	X	X	X	X
647	X	X		X		X		X	X
663	X	X	X	X	X	X	X	X	X
669	X	X	X	X	X	X	X	X	X
710	X	X	X	X	X	X	X	X	X
756	X	X	X	X	X	X	X	X	X
757	X	X	X	X	X	X	X	X	X
784	X	X	X	X	X	X	X	X	X
808	X	X	X	X	X	X	X	X	X
818		X				X			
868	X	X	X	X	X	X	X	X	X
882	X	X	X	X	X	X	X	X	X

Table 7: Applicants selected in best solutions, by model and solution (2008 selection process). BS: best solution; MOD: model.

Applicant ID	**Selected (Step 1)**	**Applicants advancing to following steps (weighting)**	**Selected (Step 3)**	**Waiting list (Step 5)**
13		X (0.75)		X
21	X			
35				X
42	X			
49	X			
62	X			
66	X			
116				X*
139	X			
144				X
169	X			
175	X			
176		X (1.9)		X
198	X			
208				X
228				X
241	X			

249		X (1.9)		X
250	X			
258	X			
261	X			
290	X			
291	X			
297				X
302	X			
314	X			
315	X			
325	X			
371	X			
372	X			
382	X			
387				X
392	X			
398	X			
400				X
402	X			
407				X
412				X
413	X			
444	X			
456	X			
459				X
469	X			
485	X			
499	X			
510	X			
517	X			
531	X			
538	X			
544	X			
548	X			
567		X (2.75)	X	
577	X			
593	X			
613				X*
628				X
635	X			
647		X (3.95)	X	
658				X
663	X			
669	X			
710	X			
729				X
756	X			
757	X			
758				X

762				X
784	X			
808	X			
818		X (1.05)		X
868	X			
882	X			

Table 8: Applicants by progress through selection algorithm (2008 selection process).