

# Simulation of Career Development in the European Commission

Stephan Onggo, Michael Pidd, Didier Soopramanien, Dave Worthington

Department of Management Science, Lancaster University Management School, Lancaster University,  
Bailrigg, Lancaster LA1 4YX, United Kingdom {s.onggo@lancaster.ac.uk, m.pidd@lancaster.ac.uk,  
d.soopramanien@lancaster.ac.uk, d.worthington@lancaster.ac.uk}

The European Commission (the Commission) employs more than 22,000 officials who provide administrative services to the European Union. In 2003, the Commission introduced a performance appraisal and promotion system based on points that the officials earn each year. In 2006, the Commission realized that the system needed to be revised. To support the review process, the Commission invited tenders for a project to develop simulation models that it could use to project the future performance of the existing system. A team from Lancaster University won the bid and subsequently worked closely with Commission officials to develop a new system. In 2009, the stakeholders in the Commission's performance appraisal and promotion system agreed to implement the improved system. The simulation model is unusual in the field of manpower planning because it models the consequences of appraisal-system rules. It uses novel, accurate, and efficient sampling techniques that are based on regression models of the underlying relationships in the data. The model was a crucial part of renegotiating the appraisal and promotion system and implementing a new system.

*Key words:* simulation; application; statistical data analysis; manpower planning; government.

*History:* This paper was refereed. Published online in *Articles in Advance* April 7, 2010.

---

## The European Commission and Its Officials

The European Commission (the Commission) is the executive branch of the European Union, which currently has 27 member states. The Commission, which is the administrative body, employs more than 22,000 officials and is organized into departments known as Directorates General (DGs). The Commission is a supranational body, separate from any national government and charged with operating across the European Union. It proposes legislation, which the Council of the European Union and the European Parliament can accept (or not accept), and is responsible for ensuring the implementation of this legislation by member governments or through pan-European bodies. Most Commission functions are based in Brussels, Belgium. More details on the Commission and its operations can be found on its website (<http://ec.europa.eu>).

The Commission is a large public sector body and is legally obliged to treat its employees equally. Therefore, it uses an annual performance appraisal system and a round of promotions to encourage performance

excellence by officials; it offers more rapid promotion to officials who perform particularly well. A promoted official moves to a higher grade and receives a salary increase. The rates at which officials can be promoted are enshrined in the Commission's staff regulations (European Commission 2008), which specify mean promotion speeds (time to transit a grade) for most grades.

An official who is promoted is not obliged to change jobs or functions. Unlike a system in which a promotion is linked to successfully applying for a higher post, most changes of jobs within the Commission are independent of promotion and have no impact on the official's grade. However, some posts—mainly high-level or management posts—may be held only by officials in specific grades.

## The Appraisal and Promotion System

Under a system introduced in 2003, each official is awarded *merit points* and *priority points* annually. Although other points can also be awarded, they apply only in a few special cases.

Each official's line manager conducts an annual performance appraisal during which the official is awarded between 0 and 20 merit points that are then accumulated in the official's *promotion rucksack*. An award of between 10 and 13 points indicates satisfactory performance; 10 points is barely acceptable. Merit points above 14 indicate very good or outstanding performance. The scoring system assumes that the Commission employs officials of above-mean ability compared with the general population. Hence, the distribution of merit points across each DG is expected to have a mean value of approximately 14.

At the promotion round that follows a performance appraisal, priority points are also added to the official's rucksack. Of these, the most significant are DG priority points, which the relevant Director General awards using a scale of 0 to 10. Whereas the merit points are a result of an appraisal process that recognizes meritorious performance, the priority points are awarded based on three main criteria—merit (i.e., a link to the appraisal system), level of responsibilities, and use of languages while executing duties. Each Director General receives a quota of 2.5 points per official in the DG and distributes the points within each grade to the officials in that grade.

Officials with sufficient points in their rucksacks will be promoted, with any excess points carried over to the following year. A promotion *threshold* defines how many points are needed for an official to be promoted to the next grade. The threshold values vary by grade and depend on the available budget each year. Thus, each year, between 0 and 30 points (i.e., the minimum and maximum number of merit and priority points, respectively) are added to each official's promotion rucksack. If the system was in steady state, a promotion threshold could be set that would allow a mean promotion speed to be achieved. Thus, if the mean points added to each rucksack were 17, setting a grade's promotion threshold at 51 would ensure a mean promotion rate of one promotion per official per three years.

## The Study

In 2006, Siim Kallas, the vice president who is responsible for staff policy in the Commission, decided to review the appraisal and promotion system and identified several shortcomings. First, he discovered that

appraisal drift, an inflation of the target average of merit points, followed an earlier decision to increase the merit point average from 14 to 14.5. Second, the distribution around the average had narrowed over time; most officials were awarded between 14 and 16 merit points, which allowed only small differentiations between officials. Third, the criteria for awarding priority points were not always perceived as clear and transparent. Fourth, predicting promotion speeds was difficult because the promotion thresholds were not constant across the years.

To support the review process, the Commission invited tenders for a project to develop simulation models that it could use to project the future performance of the existing system and to show how alternative systems would be likely to behave. The simulation models would be used by appropriately trained officials in the unit responsible for the appraisal and promotion system.

The new system had to meet the following criteria:

- (1) The mean promotion speed (time for an official to transit a grade) should be three years in each grade.
- (2) The points awarded to provide the basis for the promotion system should properly discriminate between different performance levels, with high-flyers awarded significantly more points than average or below-average performers.
- (3) The promotion costs should be within the Commission's staffing budget limits.

The models developed in this project became a crucial part of renegotiating the appraisal and promotion system and its implementation in 2009. It allowed various replacement systems to be tested and convinced stakeholders that the new performance system would be better than the existing system.

The Lancaster University tender to undertake the work was accepted. Project governance was overseen by a Commission liaison group consisting of representatives from a range of DGs and chaired by a senior official who was subsequently promoted to Director General. The day-to-day work on the project was managed by a Commission working group that had frequent contact with the Lancaster team.

## Outline of the Remainder of Our Paper

We began our work in late 2006 and conducted it over five months during which the Lancaster

University project team worked closely with its Commission clients. Activities included an initial briefing meeting, analysis of the Commission's large appraisal database, building of a conceptual model, meetings to discuss progress and ideas, statistical modeling, simulation model building and validation, scenario analyses, and presentation and delivery of a final model. The project was an interwoven mixture of technical and nontechnical components. The remainder of this paper is structured around the technical components of the work—overall model development and use, statistical modeling of the existing system, and simulating options for a revised system. We describe the interactions between the team and client groups, which were triggered by and made use of the modeling work as it developed, with the technical aspects. This paper concludes with reflections on the project.

## Model Development and Use

### Intended Model Use: To Support Thinking

In any simulation study, the analysts and clients must clearly understand and agree on its intended use. The Commission officials who liaised with the Lancaster team were highly trained technically and understood the difference between scenario exploration and prediction. One approach to discussing model use is to distinguish between models that are intended to automate decision making and those that are intended to support thinking on how a particular system operates. Models of the latter type are best considered as tools for thinking (Pidd 2003).

It is a parody of reality to argue that there are only two ways in which models can be used to support decision making. Hence, treating these two approaches as archetypes is best, and a discussion between the Lancaster team and Commission officials about those archetypes was extremely valuable. Both modes of model use assume that models will be successively refined until they fit the purpose; however, major differences exist (Pidd 2010). As a result of this discussion, the officials agreed wholeheartedly that the simulations would not be intended to produce predictions but to demonstrate the likely consequences of particular scenarios. In doing so, they would allow performance comparisons between

the existing system and possible options for change in a new system. That is, although the simulations might be subject to bias because of various modeling assumptions, this bias would affect all scenarios. Hence, all involved agreed that the simulations were intended to support Commission officials in carefully considering their options for improved performance assessment and a revised promotion system.

### Data Cleansing

The appraisal and promotion system is intended to apply to all 22,000 Commission officials. However, the seemingly smooth face of bureaucracy can conceal many surprising variations. At any given time, some officials are seconded (i.e., temporarily assigned) to other institutions or national government departments, are on sabbatical leave, or have no history of appraisal and promotion because they have recently joined the Commission. Hence, it quickly became clear that any statistical modeling should not be based on all 22,000 officials. Thus, the available personnel records (which were made anonymous to the Lancaster team) had to be reduced to include only *standard* officials (i.e., full-time officials who had neither joined nor left the Commission or been seconded elsewhere during the period 2004–2006) using a multistage process. In addition, like most databases, the Commission's personnel system contained some records with anomalous values. We removed these from the data set and reduced its size to only 15,000 officials. The Lancaster team and Commission officials agreed that this would be the basis for the modeling and simulation.

### Joiners and Leavers

Although the records of officials who joined or left the Commission during 2004–2006 were not in the standard data set, the simulation models must represent these joiners and leavers. Hence, the models allow users to inject new officials (joiners) into the system at specified grades (if this is appropriate). Once such joiners are injected into a model, their progress is based on the progress of typical standard officials; however, they behave differently in their first (probationary) year. Likewise, the simulation models allow a number of simulated officials to leave the Commission each year (leavers) at rates to be specified by

the user. That is, the numbers of joiners and rates of leaving are parameters that the user specifies as part of a simulated scenario and are subject to careful scrutiny by Commission officials to ensure that they are realistic.

In the scenarios we discuss below, the Commission officials specified the number of joiners. Their estimates were based on projected recruitment plans and the knowledge that in the first year of the simulation (2007), Romania and Bulgaria would join the European Union. Hence, they specified the number of joiners in the first and second years (2007 and 2008) at about 1,000 per year and specified a much lower number in subsequent years. They assumed that future leaving rates would follow recent patterns in which turnover was very low (approximately 3 percent in higher grades and close to zero in lower grades).

### The Excel Application

The simulation models had to be in a form that trained Commission officials could use. Because the appraisal and promotion systems are based on annual rounds, a time-sliced simulation based on annual updates was appropriate rather than a next-event simulation. In essence, the models would use actual records from the Commission's personnel database; the record for each official would occupy a row in a worksheet and the simulations would apply appropriate rules and sampling procedures to generate simulated futures. In turn, this suggested that the models could be developed as Excel workbooks with linked worksheets. The code necessary to run the simulations could be developed in Visual Basic for Applications (VBA), allowing the models to be presented as executables within the user interface that Excel provides. This method is widely used for business modeling (Powell and Baker 2004, Powell and Batt 2008, Ragsdale 2007, Winston 2004); in addition, the software is familiar to users.

Figure 1 shows the organization of the main components of the Excel application. On starting the application, the user is presented with an interface that allows that user to examine the standard data set, specify which promotion rules and statistical models to use in a simulation, and specify other parameters such as joiners, leaving rates, run length, and number of replications. When a run has been completed,

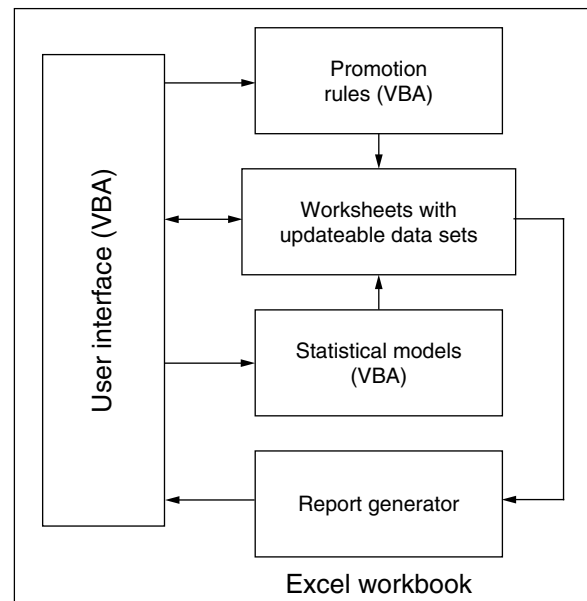


Figure 1: This figure shows the internal organization of the simulation models.

a report generator provides summary statistics and graphs; a user who is so inclined can access the full data set to conduct more detailed investigations.

Although the problem fitted naturally into Excel, we had some concerns about the memory and speed requirements of running the simulation with a large number of simulated officials. We could estimate the memory requirements in advance because we knew that we would need to represent 15,000 officials with approximately 10 variables for each of 10 years, giving a requirement for approximately  $15,000 \times 10 \times 10 = 1.5$  million cells of information. All other memory requirements (e.g., lookup tables and performance measures) were negligible. Therefore, there was a risk that an Excel implementation might run too slowly to allow proper investigation of relevant scenarios. Hence, although the models use Excel as a vehicle, much of the computation is conducted using VBA, reading to and writing from Excel tables as appropriate.

### Actual Model Use

Initially, we developed a series of very simple models, discussed each with the clients, and modified or discarded models as appropriate. However, it is important to realize that during this process both

the Lancaster team and Commission officials were learning about the operation of the existing system and about options for a replacement system. Hence, viewing this work as the development of a model by a technical team that then presented the model to users to do whatever they chose with it would be a mistake. The model building was a cooperative effort in which all involved learned about what was feasible and sensible.

At an initial meeting, one Commission official said, “Well, I think what you’ve taken on is impossible.” Moreover, at some stages in the modeling work, the Lancaster team agreed with this! However, as the statistical analysis became more refined and as officials realized what might be possible, a model of the existing appraisal and promotion system (which they referred to as the “as-is” model) gradually emerged and was refined until all agreed that it was suitable—that it showed what was likely to happen to standard officials if the existing system remained in place for an additional 10 years.

Only when all agreed that the existing system had been modelled satisfactorily did they turn their attention to modeling possible replacement systems; Commission officials realized that they needed to refine their initially loose specifications. In this refinement and the consequent attempts to model alternative systems, it became clear that some features were unnecessary. Hence, the attempt to model new systems led to a helpful, parsimonious view of the main features of such a system. Based on this, we developed the final version of a model of the possible replacement system, tested it, and handed it over to the Commission officials for their detailed evaluations of likely performance. The officials used the simulation results in negotiations with stakeholders in the Commission, including trades unions, to gain agreement on the necessary changes.

### Example Model Output

In the Excel model, the points awarded to each official and the consequent promotions are recorded over the 10-year simulation period. Therefore, extracting statistics (e.g., number of officials, number promoted, number leaving, number joining, and time to promotion) and summarizing these statistics by grade, by year, or over a number of years is simple.

For example, a major requirement for any new appraisal system is that it must be capable of achieving a greater range of promotion speeds than the existing system. It was accepted within the Commission that year-on-year variations might lead to some deviations from the target and that promotion speeds must vary by grade (in line with the legal requirements set out in the staff regulations). Hence, results that summarize repeated simulations of the 10-year period (Figure 2) provided a basis for comparing alternative appraisal systems. Figure 2 shows an example of promotion speeds for the following four appraisal systems for officials in an entry grade.

- AS-IS represents the existing system with 0–20 merit marks awarded, an intended mean of 14.5 in each grade, and 0–10 DG priority points given. This clearly results in a narrow range of promotion speeds for entry-grade officials.

- ADMIN represents an appraisal system proposed by the Commission clients. It results in a much wider range of promotion speeds; however, it retains a mean speed of three years. It uses five performance boxes to award promotion points (see Table 1).

- ALLIANCE represents an appraisal system proposed by a group of trades unions that represent officials. This system is no better than the AS-IS system in terms of promotion speeds for the entry grade. It is radically different from the AS-IS and ADMIN

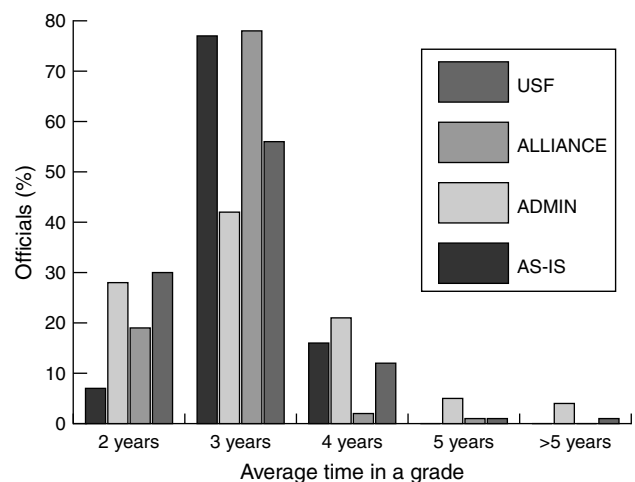


Figure 2: This figure shows a comparison of promotions speeds of entry-grade officials for four appraisal systems.

Performance description	Staff in a specific grade who are placed in this box (%)	Promotion point range
Exceptional performance	10	10–12
Excellent performance	15	7–9
Very good performance	65	4–6
Adequate performance	8	1–3
Performance needing improvement	2	0

**Table 1:** The table shows an example of performance boxes.

systems—it has no performance boxes and it awards one, two, or three points to each official each year; most officials would be awarded two points each year.

- USF represents an appraisal system proposed by another group of trades unions; it results in a range of promotion speeds that lie between the AS-IS and ADMIN systems for entry-grade officials. Although it is similar to the ADMIN proposal, it uses seven performance boxes rather than five.

### Statistical Modeling of the Existing System (The AS-IS Scenario)

To keep the confidence of users, simulating the existing system to show what is likely to happen over the next few years if the system remained unchanged was important. The statistical modeling that underpins the simulations is based on the data set of standard officials that we described above. The analysis has three elements—an examination of year-on-year consistency in awarding points and consequent analyses and models for merit points and priority points.

#### Year-on-Year Consistency

The full data set of standard officials consisted of data for the years 2003–2004, 2004–2005, and 2005–2006. For example, 2003 refers to the year in which an official’s performance was assessed and 2004 is the associated promotion year. The aim of the consistency analysis was to check for any shift in how points had been awarded through time. The analysis showed that by 2005–2006, the managers awarding the merit and priority points had learned how to use the system to their advantage. This does not mean that the approach taken to awarding points changed

radically between the first and final year; the managers grew more adept at exploiting awards at the margins. Hence, with the agreement of the officials who commissioned the project, the detailed analysis and modeling of merit points and priority points were based on the latest year—2005–2006. This ensured that the simulation of the existing (AS-IS) system into the future reflected the actual behavior of officials after several years of learning, rather than their early attempts when the system was introduced.

#### Merit Points

Regression analysis of merit points awarded to officials in each grade showed a correlation between the points awarded in successive years to each official. In all years, there were similar relationships ( $R^2$  of approximately 0.6) between merit points in period  $n + 1$  and merit points in period  $n$ ; those relationships depended consistently on whether the official had been promoted in period  $n$ . In both cases, an official who is awarded high merit points in one year is likely to gain high points in the following year; similarly, poor performers tend to receive lower merit points each year. This confirms the view that some officials are likely to be promoted more quickly than others, and this was represented in the simulations. However, with 40 percent of the variability in merit points unexplained, performance in period  $n$  does not prescribe performance in year  $n + 1$ . Hence, the model also reflects the effects that other personal and workplace factors have on year-on-year performance of individual officials.

The models took the form of regression equations in which the merit point awarded to an official in year  $n + 1$  depend on the merit points awarded to the official year  $n$ , whether or not they were promoted in year  $n$ , plus a random error term to represent the unexplained variability. See the appendix for details.

Rather than use these regression equations in a traditional way by sampling for normal errors to represent random variation, the data underpinning the regression models were used to create a transition matrix, each row of which gives the probability distribution of merit points next year, conditional on a particular number of merit points in the current year. Hence, we implemented the merit point model as two

lookup tables, with the choice of lookup table dependent on whether or not the official had just been promoted and the row of the table prescribed by the official's current merit points. The modeling benefit of this approach is that it generates merit points for individual officials consistent with the observed data. The computational benefit is also significant because sampling from lookup tables is very efficient. Onggo et al. (2009) provide details.

This merit point model is not a Markov chain model although it is close to being one. In particular, to determine the probability distribution of merit points awarded in year  $n + 1$ , the state space must represent the merit points awarded in year  $n$  and if the official was promoted in year  $n$ . However, determining if the official should be promoted in year  $n + 1$ , i.e., finalizing the official's state in year  $n + 1$ , also requires information from elsewhere in the overall simulation model. For example, the priority points awarded to the same official and the current grade of the official are needed to determine the relevant promotion threshold. This added dependence means that any attempt to use Markov chain models directly would be very unwieldy, even for this single component of the overall model.

### DG Priority Points

Analysis of the historical data showed that the processes by which DG priority points are awarded is much more complex than those of awarding merit points; we reflect this in how we simulate priority points. The complexity is necessary to reflect the actual processes used in awarding these points, thus ensuring that the simulation of the existing system reflects actual practice.

Our analysis revealed that awarding DG priority points historically reflects three processes. First, officials whose performance is highly rated (in terms of merit points in the current year and DG priority points in the previous year) have tended to be awarded more DG priority points. Second, officials who are very close to a promotion threshold before being awarded DG priority points are usually awarded enough points to see them safely over the threshold. Third, to ensure the distribution of the correct aggregate number of priority points, the number of points awarded are flexed up or down for officials whose chance of promotion this year will not

be affected by the marginal addition or subtraction of points.

Regression analysis showed that priority points in a given year were largely explained by some combination of merit points in the same year, priority points in the previous year, and distance from the threshold after the award of merit points—with the balance of these factors also depending on the official's distance (i.e., number of points still needed for promotion) from the threshold. To use these regression relationships and simultaneously address the problem of nonnormal stochastic variation, we introduced the concept of *promotional strength* to describe the values that the regression equations produced. By rounding promotional strength to the nearest integer, we could use the data underpinning the regression models to create a set of lookup tables in which probabilities reflected the observed distribution of priority points associated with each value of (rounded) promotional strength. Again, the modeling benefit is that this method generates priority points for individual officials consistent with the observed data and the computational benefit of using top-hat sampling is substantial. Onggo et al. (2009) give more details. As with the merit point model, it was again tempting to think in terms of a Markov chain model. However, in this case, DG priority points in year  $n + 1$  depend not only on DG priority points in year  $n$  but also on merit points in year  $n + 1$ , proximity to a promotion threshold, and the aggregate DG priority points available in year  $n + 1$ .

### Simulating the Existing System: The AS-IS Model

Figure 3 shows the operation of the AS-IS model. Within each year, the model completes the following four phases for each standard official.

- (1) Sample the current year's merit points.
- (2) Add this year's merit points to the official's rucksack.
- (3) Sample the DG priority points (DG P pts in Figure 3) and add them to the official's rucksack.
- (4) If the official has sufficient points to pass the promotion threshold, promote the official and recompute that official's year-end rucksack by deducting the threshold value.

Onggo et al. (2009) provide details of the sampling processes used. At the completion of each year, the

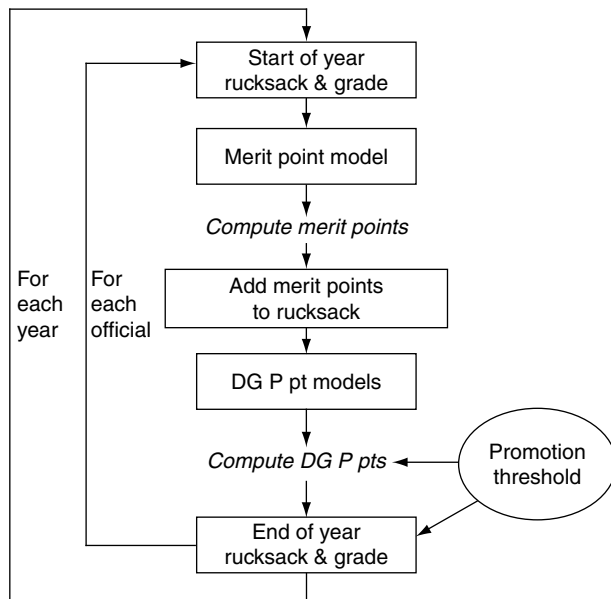


Figure 3: This figure shows the computation sequence in the AS-IS model.

model computes the distributions of merit points and DG priority points to enable the user to check that they follow the required distribution.

Each official’s merit points and DG priority points over the simulation period is available in Excel worksheets at the end of a simulation run. Hence, the simulation results can be subjected to any appropriate ad hoc analyses using Excel or any statistical package.

## Simulating the Options for a Revised System

### A Likely Revised System: The Main Features

At the start of the modeling work, the Commission officials involved had an outline specification for the likely features of a new appraisal and promotion system. However, it is important to realize that as the Lancaster team attempted to model the new system and as early simulations showed how it might operate, the Commission officials developed and adjusted their view of features that would be important in a revised system. The modeling and the models were both part of a learning cycle in which all involved were able to develop their understanding of how a new system might operate.

The result of this learning cycle was an agreement that the revised system would have the following features, operating on an annual cycle.

(1) Each official would have an annual performance appraisal that would include a performance discussion and would be placed in a *performance box*. The boxes would be designed to recognize different performance categories.

(2) The awarding of promotion points would depend on the performance box in which the official is placed and the ranking within the box. The promotion points will be accumulated in a promotion point rucksack (PPR); officials would be promoted when their PPR exceeds the appropriate promotion threshold. Following a promotion, the official’s threshold points would be deducted from the PPR.

As an example, a five-box system might use the performance boxes shown in Table 1.

Hence, the existing system of merit points and priority points would be replaced by one based only on promotion points, which would depend on the performance box at which the official’s performance is assessed. The proportions allocated to each box would guarantee the spread of promotion points, avoiding the narrowing of the range that had occurred using merit marks in the existing system. The distribution of points within each box adds the ability to further discriminate.

### Simulating Possible Replacement Systems

Figure 4 shows the schematic operation of the box model. Because one requirement of our simulation model is simulating different variations of this box system, the user may specify the number of boxes, the percentage of officials in each box, and the promotion point range within each box.

Clearly, continuity between the existing and replacements systems is necessary. A translation between an official’s existing promotion rucksack and the new PPR is required; the simulation model must include a translation from performance measured in merit points and the performance box assessment and promotion points awarded. Hence, in simulating the operation of the box system, *notional merit points* use the same merit point statistical model as the as-is



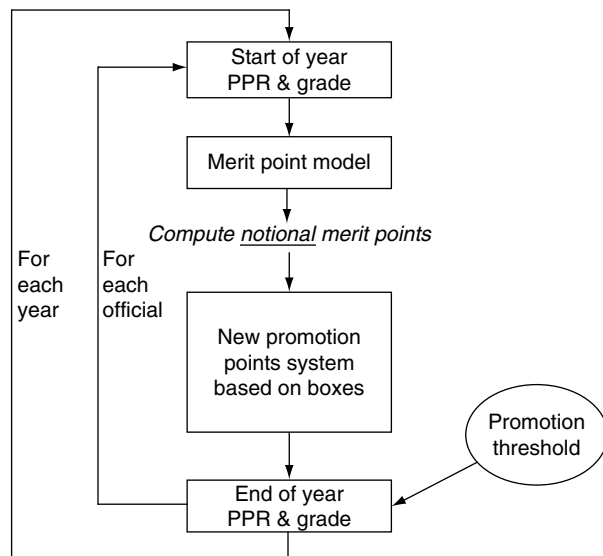


Figure 4: This figure shows the computation sequence of the performance box model.

model. These notional merit points are then translated into performance boxes and hence into promotion points, which allows us to compute an official's PPR prior to any promotion decisions. Using notional merit points in this way allows a proper comparison with the existing system because both are based on the same underlying models of performance.

Conversion from notional merit points to performance boxes and then to promotion points within a performance box should depend on the following three factors:

- (1) The notional merit points awarded to the official, which reflect the official's current performance;
- (2) The seniority of each official within that box in the official's current grade; and
- (3) The gap between the PPR and the announced promotion threshold using the PPR value before the award of this year's promotion points.

Because we assume that any or all of these factors may be brought into play in allocating promotion points, the model allows the user to specify weights for the rankings based on each factor. Once ranked in this way, the top  $x$  percent can be placed in the top box, the next  $y$  percent in the next box, and so on, depending on the percentages that the user specifies for each box. Points within each box are then allo-

cated using the same overall ranking process; however, the user may choose to specify different weights for this second ranking. Thus, the model can simulate the operation of a whole range of possible replacement systems based on performance boxes.

We specified different options for a replacement system by specifying variations of the performance boxes. In analyzing the results of these simulations, the clients considered several criteria. First, the new system must guarantee a mean promotion speed of three years across all standard officials. Second, the range of promotion speeds must be wider than the range in the existing system. As we discussed earlier, Figure 2 shows examples of distributions of promotion speed for the existing system and three proposed replacements that are based on replicated sets of 10-year runs. In this respect, the ADMIN proposal is clearly better than the existing (AS-IS) system and union (ALLIANCE and USF) proposals. Keeping salary costs within budget was also important. Hence, once we established the principle of promotion speeds, we used the model to develop specific proposals to ensure that budgets were unlikely to be breached.

## Reflections on the Project

We learned a great deal from implementing this project. The close relationship between the Lancaster modeling team and Commission officials allowed us to develop the models gradually and allowed the Commission officials to work on successive versions of replacement systems. This was not a project in which a tool was specified, developed, and handed over to users; it was one in which, because of the data analysis and robust discussions, all involved learned what the sensible and useful features in a new appraisal system should be. We modelled these proposed features and then refined them, thus supporting the learning of all involved. The project's sponsors used the final simulation model to negotiate with stakeholders, including trades unions, to gain agreement to implement the performance box system in 2009. Key points in the negotiations included the percentages of officials to be rated in each performance box; the model allowed us to investigate the effects of different percentages and to see clearly the effects on promotion speeds and budgets.

### Modeling HRM Systems

According to the literature, using simulation models to design human resource management (HRM) appraisal systems is unusual; however, experts in manpower planning have long advocated their use (Abdel-Hamid 1989, Bartholomew and Forbes 1979, Blosch and Antony 1999, Ekamper 1997, Weber 1971). There are clear similarities between our work and more traditional manpower planning models in which staff members move between grades according to various “push” or “pull” rules. However, in our work, promotions are driven by the Commission’s appraisal system; therefore, it was important to build a model capable of incorporating appraisal system rules so we could investigate the consequences of modifying these rules on the full workforce.

Many (possibly all) bureaucracies operate using appraisal systems and employment grades that dictate employee promotions. However, the ways that appraisals link to promotions and the requirements for promotion are likely to be many and varied. Hence, it is unlikely that the models we developed for the Commission can be directly moved into another organization. However, the principles underlying the models and the modeling are transferrable. The first, which was crucial although it sounds mundane, was the extremely close cooperation between the Lancaster modeling team and Commission staff and other stakeholders. The second is that the simple Markov structure that is apparent from the points and grade structures must be greatly modified to allow for the real-world decision making and the intelligent game playing that can occur in seemingly rational point-based systems.

### Excel Simulation

Although many technical features of building a simulation model in Excel are routine for someone with good VBA and simulation skills, the memory and speed requirements were important concerns when running with a large number of simulated officials. Hence, we carefully designed the necessary sampling and sorting algorithms to ensure that the models would run quickly on the standard equipment used by Commission officials.

For example, the simulation used memory, rather than a worksheet, to perform calculations because

access to worksheets requires slow disk operations. In an early version of the model, this reduced the run time from approximately an hour to a matter of minutes. As a second example, when the proposed box system required officials with tied ranks to be ordered randomly, the model achieved this by perturbing the ranks using a simple sampling scheme implemented as part of the sorting process, thus avoiding the need for an additional time-consuming shuffling process.

### Regression-Based Sampling Schemes

The merit point and priority point sampling schemes that the simulation used were based on regression models that provided reasonable representations of the data relationships. Although we used regression analysis to understand the relationships that drive the award of merit points and DG priority points, we did not directly use the regression models in the simulations. Instead, we used lookup tables which, in the case of merit marks, linked one year’s awards to those of the previous year and allowed rapid and consistent sampling. In the case of DG priority points, regressions also showed the relationships and led to the concept of promotional strength, which could also be incorporated into lookup tables. They also led to rapidly running simulations that have an acceptable precision, as Onggo et al. (2009) discuss in more detail. The lookup tables have much in common with a Markov chain formulation; however, their inclusion in a simulation model allows the transition probabilities to depend both on merit points (or DG priority points) in the previous year and also on other factors in the model.

### Role of Modeling

As we noted above, our clients clearly understood the difference between scenario exploration and prediction; they agreed that the simulations should provide tools to support Commission officials in considering options for improved performance assessment and promotion. The Lancaster team and the clients met seven times during the project. These meetings included the traditional simulation modeling activities of conceptual model building and model validation and the traditional project management activities of briefings, progress reports, target and milestone refining, and presentation of interim

and final results. These meetings were backed up by e-mails and telephone calls and facilitated an ongoing exchange of ideas between the Lancaster team, the Commission liaison group, and the Commission working group. For example, we gave the working group early data analyses from regression modeling for potential inclusion in the model. These analyses formed the basis of discussions that helped the liaison and working groups to better understand the weaknesses of the existing system and to refine their thinking on the structure and potential benefits of alternative systems. Later, preliminary model outputs that had been produced primarily for model validation purposes again reinforced their understanding of the existing system; hence, it informed their design ideas for possible alternatives.

As the simulation literature (and indeed the model-building literature in general) recognizes widely, validating a model of a system that does not exist is problematic. In this project, the questions necessary for the development of the simulations model prompted the client and the working group to think clearly through the key features of any potential system. Furthermore, preliminary results produced for validation purposes were just as likely to prompt refinements to their proposed alternatives as they were to uncover faults in the model building.

Finally, a key factor behind the project's success was the common understanding of the role of models in this context. Specifying the modeling work in detail at the outset was impossible. A sample of the database and outlines of the existing appraisal system and the likely nature of alternative systems were sufficient to convince all parties that some modeling would be worthwhile. However, they had to agree on the details as the work progressed, considering the emerging issues raised by the client and the total amount of work that the project team had been contracted to perform.

## Appendix. The Merit Point Regression Model

The regression equations took the form

$$M_{i,n+1} = \alpha + \beta \cdot M_{i,n} + \gamma \cdot Prom_{i,n} + \varepsilon_i,$$

where  $M_{i,n+1}$  represents the merit points awarded to official  $i$  in year  $n+1$ ,  $M_{i,n}$  is the merit points awarded

to official  $i$  in year  $n$ ,  $Prom_{i,n}$  is a  $(0, 1)$  variable that indicates whether official  $i$  was promoted in year  $n$ , and  $\varepsilon_i$  is a random error term.

### Acknowledgments

We are grateful for the financial support that European Commission DG Admin A6 provided to conduct this work. We are also grateful to Silke Boger and Philippe Bresson of DG Admin A6 who commissioned the work. They proved to be excellent and understanding clients who were as willing to learn as we were.

### References

- Abdel-Hamid, T. K. 1989. The dynamics of software project staffing: A system dynamics based simulation approach. *IEEE Trans. Software Engrg.* **15**(2) 109–119.
- Bartholomew, D. J., A. F. Forbes. 1979. *Statistical Techniques for Manpower Planning*. John Wiley & Sons, Chichester, UK.
- Blosch, M., J. Antony. 1999. Experimental design and computer-based simulation: A case study with the Royal Navy. *Managing Service Quality* **9**(5) 311–320.
- Ekamper, P. 1997. Future age-conscious manpower planning in The Netherlands: From early retirement to a new perspective on the elderly? *Internat. J. Manpower* **18**(3) 232–247.
- European Commission. 2008. Staff regulations of officials of the European Communities. Retrieved December 17, 2008, [http://ec.europa.eu/civil\\_service/docs/toc100\\_en.pdf](http://ec.europa.eu/civil_service/docs/toc100_en.pdf).
- Onggo, S., M. Pidd, D. Soopramanien, D. J. Worthington. 2009. Modelling career development schemes in the European Commission. Working paper, Lancaster University Management School, Lancaster, UK.
- Pidd, M. 2003. *Tools for Thinking: Modelling in Management Science*, 2nd ed. John Wiley & Sons, Chichester, UK.
- Pidd, M. 2010. Why modelling and model use matter. *J. Oper. Res. Soc.* **61**(1) 14–24.
- Powell, S. G., K. R. Baker. 2004. *Management Science. The Art of Modeling with Spreadsheets*, 2nd ed. John Wiley & Sons, Hoboken, NJ.
- Powell, S. G., R. J. Batt. 2008. *Modeling for Insight: A Master Class for Business Analysts*. John Wiley & Sons, Hoboken, NJ.
- Ragsdale, C. T. 2007. *Spreadsheet Modelling and Decision Analysis: A Practical Introduction to Management Science*, 5th ed. Thomson South-Western, Mason, OH.
- Weber, W. L. 1971. Manpower planning in hierarchical organizations: A computer simulation approach. *Management Sci.* **18**(3) 119–144.
- Winston, W. L. 2004. *Microsoft Excel Data Analysis and Business Modelling*. Microsoft Press, Redmond, WA.

Silke Boger, Directorate General, Personnel and Administration, European Commission, B-1049 Bruxelles, Belgium, writes: "I am writing to support the publication in *Interfaces* of the paper, 'Simulation of Career Development in the European Commission' by Stephan Onggo, Michael Pidd, Didier Soopramanien,

and David Worthington of the Department of Management Science, Lancaster University, UK.

“The paper provides an excellent summary of the work done by the Lancaster team during our interactions in 2006 and 2007. The simulation model they produced was impressive, user friendly while very fast and allowed us to experiment with different options for an appraisal and promotion system for the European Commission which was finally adopted

by the Commission in June 2008 and implemented in 2009 for the first time.

“The modelling work gives us confidence that the new system will not suffer from the problems that beset its predecessor. As you will understand, it is simply not meaningful to estimate the savings from this type of work, which is aimed at achieving equity and transparency in career progression. The Lancaster modelling work was essential in enabling us to do this.”