SIMULATION OF CAREER DEVELOPMENT IN THE EUROPEAN COMMISSION

ABSTRACT
The European Commission employs over 22,000 officials who see to the administration of the European Union. In 2003 the Commission introduced a performance appraisal and promotion system based on points earned each year. After about five years of operation it became clear that this system was not satisfactory and needed to be revised. Working closely with Commission officials, a team from Lancaster University developed a simulation model that allows examination of the future performance of the current system and shows how alternative systems might perform. As a consequence, the various stakeholders in the Commission’s performance appraisal and promotions system agreed to implement an improved system in 2009. The simulation model is unusual in the field of manpower planning because of the requirement to model the consequences of appraisal system rules. Novel sampling schemes based on regression models of the underlying relationships in the data are used, and prove to be both accurate and efficient. The model was a crucial part of the renegotiation of the appraisal and promotion scheme and the implementation of a new scheme in 2009.

KEYWORDS
Simulation application, statistical data analysis, manpower planning, government
INTRODUCTION

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 is the administrative body employing over 22,000 officials organised into departments known as Directorates General (DGs). The Commission is a supra-national body, separate from any national government and charged with operating across the European Union. The Commission proposes legislation, which may or may not be accepted by Council of the European Union and the European Parliament, and is responsible for ensuring its implementation by member governments or through pan-European bodies. Most functions of the Commission are based in Brussels, Belgium. More details of the Commission and its operations can be found on its web-site (http://ec.europa.eu).

The Commission is a large public sector body and is legally obliged to treat its employees equally. To ensure this, it has an annual performance appraisal system and an annual round of promotions. The performance and promotion systems are intended to encourage excellent performance by officials by offering more rapid promotion to officials who perform particularly well. A promoted official moves from one grade to another and gains a salary increase. The rates at which officials are to be promoted are enshrined in the staff regulations (European Commission, 2008), which specify mean promotion speeds (time to transit a grade) for most grades.

When an official is promoted, she is not obliged to change jobs or functions as a consequence. Unlike a position system where promotion is linked to the successful application to a higher post, a change of jobs within the Commission is independent of promotion and has no impact on the grade of an official. However, some posts may only be held by officials in certain grades – mainly high level or management posts.

The appraisal and promotion system introduced in 2003

Under the scheme introduced in 2003, each official is awarded points each year that are cumulated in a promotion rucksack. Two main sets of points are awarded: merit points and priority points. There are other additional points that can be awarded, but these only apply in a minority of special cases.

Each official has an annual performance appraisal conducted by her line manager. As a consequence, the official is awarded merit points, which can range from 0 to 20. On this
scale, an award of 10 points is intended to indicate performance that is just acceptable. The scoring system also assumes that the Commission employs, as a norm, officials of above mean ability compared to the general population. Hence, the distribution of merit points across each Directorate General is expected to have a mean value of approximately 14. Merit points above this mean indicate very good and outstanding performance whereas a score of between 10 and 13 points indicates satisfactory performance. As a result of the appraisal, the merit points of the official are cumulated in her promotion rucksack.

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

The official will be promoted if she has sufficient points in her rucksack. This is defined by a promotion threshold that states how many points are needed for an official to be promoted to the next grade. The threshold values vary by grade and depend, each year, on the available budget. Thus, in each year, between 0 to 30 points are added to the promotion rucksack; some of which are based on performance appraisal and some of which are based on the promotion criteria in the relevant Directorate General. Were the system to be 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 the promotion threshold for a grade at 51 would allow a mean promotion speed of 3 years.

**The reason for the study**

In 2006, the Vice-President Siim Kallas, who is responsible for staff policy in the Commission, decided to review the current system. After four rounds of the appraisal and promotion system several shortcomings were identified. First, there was appraisal drift, an inflation of the target average of merit points, following an earlier decision to increase the merit point average from 14 to 14.5. Furthermore, the distribution around the average had narrowed over time, so that most officials are awarded between 14 and 16 merit points, which allowed only small differentiation between officials. A further issue was that the criteria for the award of priority points were not always perceived as clear and transparent. Finally, it was
proving to be very difficult to predict promotion speeds because the promotion thresholds were not constant across years.

In order to support the review process, the Commission invited tenders for a project to develop simulation models that could be used to project the performance of the current system into the future and could also show how alternative systems would be likely to behave. This was to be done to a very tight time-scale and was intended to result in simulation models that could be used by appropriately trained officials of the Commission of the unit responsible for the appraisal and promotion system.

Any replacement appraisal and promotion system had to meet several criteria of which the most important were:

1. The mean promotion speeds (time for an official to transit a grade) should be three years in all grades.
2. The points awarded and accumulated to provide the basis for the promotion system should make full use of the range rather than cluster around a few values, so as to properly discriminate between different performance levels.
3. The costs of promotions should remain within limits set by the staffing budget of the Commission.

The model developed on this project became a crucial part of the renegotiation of the appraisal and promotion scheme and the implementation of a new, promotion box, scheme in 2009. It allowed different replacement schemes to be tested and convinced stakeholders that the new ‘performance box’ system would be an improvement over the existing scheme.

Outline of remainder of paper
The work was conducted over a five month period, during which the Lancaster project team worked closely with and for the Commission clients in several ways. These included an initial briefing meeting, analysis of the Commission’s large appraisal database, conceptual model building, progress meetings to discuss progress and ideas, statistical modelling, simulation model building and validation, scenario analyses and presentation and delivery of a final model. The project itself was an interwoven mixture of these technical and non-technical aspects, and the remainder of the paper is structured around the technical aspects of the work: namely overall model development and use, statistical modelling of the current system, and simulating options for a revised system. The interactions between the team and client groups were triggered by, and made use of, the modelling work as it developed, and are described alongside the technical aspects. The paper concludes with reflections on the project.
**Model development and use**

*Intended model use: to support thinking*

In any study, it is important to be clear about the intended use for the simulations and it is critical that the analysts and clients come to a single mind on this. The Commission officials who liaised with the Lancaster team were themselves technically highly trained and well understood the difference between scenario exploration and prediction. There are many ways of discussing model use and one approach 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. This reflects the view of Pidd (2003) that models of the latter type are best considered as tools for thinking.

It is, of course, a parody of reality to argue that there are only two ways in which models can be used to support decision making. Hence, it is best to treat these two approaches as archetypes, but the discussion with officials about those archetypes was extremely valuable. Both modes of model use assume that models will be successively refined until fit for purpose, but there are major differences. As a result of this discussion, there was wholehearted agreement that the simulations were not intended to produce predictions but to demonstrate the likely consequences of particular scenarios. Within this, they would allow performance comparison between the current system and possible options for change. That is, though the simulations may be subject to bias due to various modelling assumptions, this same bias would affect all scenarios. Hence, all involved agreed that the simulations were intended to support Commission officials in their careful consideration of options for improved performance assessment and a revised promotion system.

*Data cleansing*

The Commission currently employs about 22,000 officials and the appraisal and promotion system is intended to apply to all. However, the seemingly smooth face of bureaucracy can conceal many surprising variations. Some officials are seconded to other institutions or national government departments, on sabbatical leave, or have only recently joined and have no history of appraisal and promotion. Hence, it quickly became clear that any statistical modelling should not be based on this complete set of around 22,000 officials. Thus, the available personnel records (which were anonymised) had to be reduced to *standard* officials in a multi-stage process.

After a series of meetings and analyses, the *standard* data set was defined as consisting of full-time officials who have neither joined nor left the Commission, nor been
seconded elsewhere, during the period 2004 to 2006. In addition, like most databases, the Commission’s personnel system contained some records with anomalous values and these were removed from the dataset. Thus, though the Commission has over 22,000 officials, the standard data set included only 15,000 officials and this was agreed as the basis for the modelling and simulation.

**Joiners and leavers**

Though the records of officials who joined the Commission during 2004-2006 and those who left in that period were not present in the standard data set, such joiners and leavers must be represented in the simulation models. Hence, the simulation models allow users to inject new officials (joiners) into the system at specified grades, if this is appropriate. Once injected into the model, the progress of such joiners is based on the progress of typical standard officials, but with different behaviour 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, numbers of joiners and rates of leaving are parameters that the user specifies as part of a scenario to be simulated and are subject to careful scrutiny to ensure that they are realistic.

In the scenarios discussed later, the numbers of joiners was specified by the Commission officials. Their estimates of joiners were based on projected recruitment plans and the knowledge that, in year 1 of the simulation (2007), Romania and Bulgaria were due to join the European Union. Hence, in years 1 and 2 of the simulation, the number of joiners was specified at about 1000/year and much lower in subsequent years. Future leaving rates were assumed to follow recent patterns, in which turnover was very low (around 3% for higher grades and close to 0% in lower grades).

**The Excel application**

It was important that the simulation models were in a form that could be used by trained Commission officials. Since the appraisal and promotion systems are based on annual rounds, it was clear that a time-sliced simulation based on annual updates was appropriate. In essence, the models would take 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 it could be developed as an Excel workbook, 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 but within the user interface provided by
Excel. There is nothing particularly novel about using Excel in this way, since it is widely used for business modelling (Powell and Baker (2004), Powell and Batt (2008), Ragsdale (2000), Winston (2004)) and has the advantage of using software that is familiar to users.

Figure 1 shows the organisation of the main components of the Excel application. On starting the application, the user is presented with a user interface. This allows the user to examine the data set and to specify which promotion rules and statistical models will be used in a simulation and also to specify other parameters such as joiners, leaving rates, the run length and number of replications. Once a run is complete, summary statistics and graphs are available in a report generator and, should the user be so inclined, she can access the full data set to conduct any detailed investigations that are deemed necessary.

Whilst the problem itself fitted rather naturally into Excel, there were nevertheless some concerns about the memory and speed requirements when running the simulation with a large number of simulated officials. The memory requirements could be estimated in advance, as it was known that about 15,000 officials would need to be represented, with 10 or so variables for each of 10 or so years, giving a requirement for approximately 15,000 x 10 x 10 = 1.5 million cells of information. (All other memory requirements, e.g. look-up tables and performance measures, were negligible in this respect.) Speed requirements emerged as the project developed and are discussed later.

Actual model use
As in many modelling projects, a series of models, firstly very simple, were developed. Each was discussed with the clients and modified or discarded as appropriate. It is important to realise, though, that during this process both the Lancaster team and Commission officials were learning about the operation of the current system and about options for the replacement system. Hence, it would be a mistake to see this work as being the development of a model by a technical team that then presented the model to users to do with it whatever they chose. The model building was a co-operative effort in which all involved learned about what was feasible and sensible and the models progressed on this basis.

At an early meeting with Commission officials, one was heard to say “Well, I think what you’ve taken on is impossible.” There were stages in the modelling work when the Lancaster team agreed with this! However, as the statistical analysis became more refined and as officials realised what might be possible, a model of the as-is system (i.e. the appraisal and promotion system in place from 2003 to 2008) gradually emerged and was refined until all
agreed that it was fit for purpose – that is, it showed what was likely to happen to standard officials if the current system remained in place for a further 10 years.

Only when all agreed that the current, as-is, situation was satisfactorily modelled was attention turned to modelling possible replacement systems and, when this occurred, Commission officials realised that their initially loose specification needed to be tightened up. In this tightening and the consequent attempts to model alternative new systems, it became clear that some possible features were not needed. Hence the attempt to model new systems led to a helpfully parsimonious view of the main features of such a system. Based on this, the final version of a model of the possible replacement systems was developed, tested and handed over to the Commission officials for their detailed investigations of likely performance. The simulation results were then used in negotiations with stakeholders in the Commission, including trades unions, to gain agreement to make the necessary changes.

**Example model output**

In the Excel model, each official’s future is recorded over the (10 year) simulation period. It is therefore a simple matter to extract statistics such as number of officials, number promoted, number leaving, number joining and time to promotion, and to summarise these by grade, by year or over a number of years.

For example, one of the main requirements for any new appraisal system was that it should be capable of achieving a greater range of promotion speeds than the current system. It was accepted that year-on-year variations might lead to some deviations from the target and that the promotion speeds must vary by grade (in line with the legal basis set out in the staff regulations). Hence results such as those in Figure 2 summarising repeated simulations of the 10 year period, were one important basis for comparison of alternative appraisal systems. For this particular example, Figure 2 shows the resulting different promotion speeds for different appraisal systems for officials in an entry grade. The 4 appraisal systems represented are:

- **AS-IS**: the current system, with merit marks awarded on a range of 0-20 and an intended mean of 14.5 in each grade, together with DG priority points. It is clear from figure 2 that this results in a narrow range of promotion speeds for entry grade officials.

- **ADMIN**: an appraisal system proposed by the Commission clients, which results in a much wider range of promotion speeds, though retains a mean speed of 3 years. It employs 5 ‘performance boxes to award promotion points (see Table 1).
- **ALLIANCE**: an appraisal system proposed by one of two groups of trades unions that represent officials. This 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 systems, has no performance boxes and awarded 1, 2 or 3 points to each official each year, most officials would be awarded 2 points each year.

- **USF**: an appraisal system proposed by the other group of trades unions, which results in a range of promotion speeds for entry grade officials which lie between the AS-IS and ADMIN systems. It is similar to the ADMIN proposal, but has seven performance boxes rather than five.

*Insert figure 2 about here*

**STATISTICAL MODELLING OF THE CURRENT SYSTEM (AS-IS)**

In order to keep the confidence of users, it was important to develop a simulation of the existing system that could be used to show what is likely to happen over the next few years if the system remains unchanged. The statistical modelling that underpins the simulations is based on the data set of standard officials described earlier. The analysis itself has three elements: an examination of year-on-year consistency in the award of 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 03/04, 04/05 and 05/06, where xx/ refers to the year for which an official’s performance was assessed, and /yy is the promotion year. The aim of the consistency analysis was to check whether there had been any shift in the way that points were awarded through time. The analysis showed that, by 05/06, the actors awarding the merit and priority points had learned how to use the system to their advantage. This does not mean that the whole approach taken to awarding points was radically different in the first and final year, but rather that the actors had grown more adept at exploiting awards at the margins. Hence, with the agreement of the officials who commissioned the project, the detailed analysis and modelling of merit points and priority points was based on the latest year – 05/06. This ensures that the simulation of the current (AS-IS) system into the future reflected the latest expertise in its practical application.

*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 $n$, and those relationships depended consistently on whether or not 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 and similarly poor performers tend to receive lower merit points each year. This is important, because it 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% 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 fact that year-on-year performance of individual officials will be affected by other personal and workplace factors.

The regression equations took the form:

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

(1)

where $M_{i,n+1}$ is 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. 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 was 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 the merit point model was implemented as two look-up tables, with the choice of look-up table dependent on whether or not the official has just been promoted, and the row of the table prescribed by the official’s current merit points. The modelling benefit of this approach is that it generates merits points for individual officials consistent with the observed data. The computational benefit is also significant as top-hat sampling is very efficient. (See Onggo et al (2009) for details).

This merit point model is not a Markov chain model although it is close to being one. In particular, the state-space must represent the merit points awarded in year $n$ and whether or not the official was promoted in year $n$ in order to determine the probability distribution of merit points awarded in year $n+1$. However to determine whether or not the official is promoted in year $n+1$, i.e. to finalise their 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 the award of merit points and this is, in turn, reflected in the way that their award is simulated. The complexity is needed so as to reflect the actual processes used in awarding these points, to ensure that the simulation of the current system reflects actual practice.

This revealed that the historical award of DG priority points 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 and that this process has continued through time. Secondly, officials who are very close to a promotion threshold before the award of DG priority points are usually awarded enough points to see them safely over the threshold. Finally, 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 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 from the threshold. In order to make use of these regression relationships and at the same time deal with the problem of non-Normal stochastic variation the concept of ‘promotional strength’ was introduced to describe the values produced by the regression equations. By rounding promotional strength to the nearest integer, the data underpinning the regression models could again be used to create a set of look-up tables in which probabilities reflected the observed distribution of priority points associated with each value of (rounded) promotional strength. The modelling benefit is again that this method generates priority points for individual officials consistent with the observed data, and the computational benefit of using top-hat sampling is again substantial. More details are given in Onggo et al (2009). 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 promotion threshold and the aggregate DG priority points available in year $n+1$. 
The simulation of the current system – the as-is model

The operation of the as-is model is shown schematically in figure 3. Within each year, the model completes the following 4 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 and add these to the rucksack.
4. If the official has sufficient points to pass the promotion threshold, promote the official and re-compute their year-end rucksack by deducting the threshold value.

Details of the sampling processes used can be found in Onggo et al (2009). At the completion of each year, the model computes the distributions of merit points and DG priority points to enable the user to check that these follow the required distribution.

Insert figure 3 about here

The complete simulated behaviour of each official, in terms of 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 indeed any statistical package.

Simulating the options for a revised system

The main features of a likely revised system

At the start of the modelling work, the Commission officials involved had an outline specification for the likely features of a new appraisal and promotion system. It is, though, important to realise that, as the Lancaster team attempted to model this and as early simulations showed how this might operate, the Commission officials developed and adjusted their view of important features of a revised system. That is, the modelling 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. Annual performance appraisal which, as well as allowing discussion about performance, will result in each official being placed in a ‘performance box’. These boxes will be designed to recognise different categories of performance.
2. The award of promotion points that depend on the box at which the official's performance is assessed, and their ranking within the box. The promotion points will then be accumulated into a Promotion Point Rucksack (the PPR) and an official will
be promoted when her PPR exceeds the appropriate promotion threshold. Following promotion, the threshold points will be deducted from the PPR.

As an example, a 5-box system might use the performance boxes shown in table 1.

*Insert table 1 about here*

Hence, the current 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 with merit marks in the current 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. Since one requirement from the simulation models is the ability to simulate different variations on this box system, the user may specify the number of boxes, the percentage of officials in each box and the promotion points range within each box.

*Insert figure 4 about here*

There must clearly be some continuity between the existing system and anything that replaces it. In real-life there will be a translation between an official’s existing rucksack and the new PPR and the simulation model must include a translation from performance measured in merit points and the performance box assessment and promotion points awarded. Hence, in a simulation of the operation of the box system, *notional merit points* use the same merit points statistical model as the as-is model. These notional merit points are then translated into performance boxes and hence into promotion points, which allows us to compute their PPR before any promotion decisions. Using notional merit points in this way allows a proper comparison with the existing system, since 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 is intended to depend on three factors:

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

Since it is assumed 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 $x\%$ can be placed in the top box, the next $y\%$ in the next box and so on. Points within each box are then allocated using the same overall ranking process, although the weights selected can be different from those used for allocating between boxes. Thus the model can simulate the operation of a whole range of possible replacement systems based on performance boxes.

Different options for a replacement system were simulated by specifying variations of the performance boxes. In analysing the results of these simulations, the clients were considering several criteria. Firstly, the new system must guarantee a mean promotion speed of 3 years across all standard officials. Secondly, the range of promotion speeds must be wider than the current system. As discussed earlier, Figure 2 shows examples of distributions of promotion speed for the current system and 3 proposed replacements, based on replicated sets of 10-year runs. In this respect the ADMIN proposal is clearly better than the current (AS-IS) system and union (ALLIANCE and USF) proposals. It was also important that salary costs were kept within budgets. Hence, once the principle of promotion speed was established, the model was used to develop specific proposals to ensure that budgets were unlikely to be breached.

**Reflections on the project**

Much was learned from this project, particularly because of the close relationship between the Lancaster modelling team and Commission officials. This allowed the models to be developed gradually and also allowed the Commission officials to work on successive versions of replacement systems. That is, this was not a project in which a specific tool was specified, developed and handed over to users, but one in which all involved learned, due to the data analysis and robust discussion, what might be sensible features in a new appraisal system. These proposed features were modelled and then refined, supporting the learning of all involved. The final simulation model was used by the project’s sponsors 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 and the model allowed the effects of different percentages to be investigated and the effects on promotion speeds and budgets to be clearly seen.
Modelling HRM systems

Judging by the literature, it is unusual for simulation models to be used in the design of HRM appraisal systems, though they have long been advocated in manpower planning (Abdel-Hamid (1989), Bartholomew and Forbes (1979), Blosch and Antony (1999), Ekmper (1997), Weber (1971)). There are obvious similarities between this work and more traditional manpower planning models in which staff move between grades according to various ‘push’ or ‘pull’ rules. However in this work, promotions are driven by the Commission’s appraisal system, making it important to build a model capable of incorporating appraisal system rules so that the consequences of modifying them on the workforce as a whole can be investigated.

Many, possibly all, bureaucracies operate with appraisal schemes and employment grades through which officials may be promoted. However, the ways that appraisal links to promotion and the requirements for promotion are likely to be many and varied. Hence, it is unlikely that the models developed for the Commission can be directly transplanted into another organisation. However, the principles underlying the models and the modelling are transferrable. The first, which sounds thoroughly mundane but was crucial, was the extremely close cooperation between the Lancaster modelling team and Commission staff and other stakeholders. The second, is that the simple Markov structure that is apparent from the points and grade structures needs to be greatly modified to allow for the real-world decision-making and the intelligent game-playing that can occur in seemingly rational points-based systems.

Simulation in Excel

Whilst many of the 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 great care was taken to design the necessary sampling and sorting algorithms to ensure very fast running on the standard equipment used by Commission officials.

For example, memory was used, rather than a worksheet, to perform the necessary calculations because access to worksheets requires slow disk operations. For an early version of the model this reduced the runtime from an hour or so to a matter of minutes. As a second example, when the proposed box system required that officials with tied ranks were ordered randomly, this was achieved by perturbing the ranks using a simple sampling scheme implemented as part of the sorting process, so avoiding the need for an additional time-consuming shuffling process.
Regression-based Sampling Schemes
The sampling schemes for merit points and priority points used in the simulation were based on regression models that had been found to provide reasonable representations of the relationships in the data. Though regression analysis was used to understand the relationships that drive the award of merit points and DG priority points, the regression models were not used directly in the simulations. Instead look-up tables which, in the case of merit marks, linked one year’s awards to the previous year’s, allowed very fast and consistent sampling. In the case of DG priority points, regressions also revealed the relationships and led to the concept of promotional strength which could also be incorporated in look-up tables. This is discussed in more detail in Onggo et al (2009) and leads to simulations with an acceptable precision that run extremely fast. The look-up tables have much in common with a Markov chain formulation, but their inclusion in a simulation model allows the transition probabilities to depend not only on merit points (or DG priority points) in the previous year, but also on other factors contained elsewhere in the model.

Role of Modelling
As noted earlier the clients clearly understood the difference between scenario exploration and prediction, and agreed that the simulations were intended to provide tools to support Commission officials in their careful consideration of options for improved performance assessment and promotion. In total the Lancaster team and the clients met seven times during the project. These meetings included the traditional simulation modelling activities of conceptual model building and model validation together with the traditional project management activities of briefings, progress reports, refining targets and milestones, and presentation of interim and final results. However alongside these activities these meetings (backed up by emails and telephone calls) also facilitated an ongoing exchange of ideas between the Lancaster team, the Commission liaison group and the Commission working group. For example, early data analyses from regression modelling were fed back to the clients for potential inclusion in the model. These formed the basis of discussions that helped the clients and the working group to better understand the weaknesses of the current system, and hence to refine their thinking on the structure and potential benefits of alternative systems. Later, preliminary model outputs produced primarily for model validation purposes again reinforced their understanding of the current system and hence informed their design ideas for possible alternatives.
As is widely recognised in the simulation literature (and indeed in the model building literature in general), validation of 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 one of the key factors behind the success of this project has been the common understanding of the role of models in this type of context. It was not possible to specify the modelling work in detail at the outset. A sample of the database and outlines of the current appraisal system and the likely nature of alternative systems were sufficient to convince both parties that some modelling would be worthwhile. However the details had to be agreed as the work progressed, with due regard to the emerging issues raised by the client and the total amount of work that the project team had been contracted to do.

ACKNOWLEDGEMENTS
We are grateful for the financial support of European Commission DG Admin A6 when conducting this work. We are also grateful to Dr Silke Boger and Mr Philippe Bresson of DG Admin A6, who commissioned the work and who proved to be the excellent and understanding clients who were as willing to learn as we were.

REFERENCES


EC. 2008 Staff regulations of officials and conditions of employment of other servants of the European Community, articles 43 and 45.
[http://ec.europa.eu/civil_service/docs/toc100_en.pdf](http://ec.europa.eu/civil_service/docs/toc100_en.pdf)


Ragsdale C.T. 2000 *Spreadsheet modelling and decision analysis* (4th Ed). South-Western Thomson Learning,


<table>
<thead>
<tr>
<th>Performance description</th>
<th>% of staff in a grade who are placed in this box</th>
<th>Promotion point range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exceptional performance</td>
<td>10</td>
<td>10 – 12</td>
</tr>
<tr>
<td>Excellent performance:</td>
<td>15</td>
<td>7 – 9</td>
</tr>
<tr>
<td>Very good performance</td>
<td>65</td>
<td>4 – 6</td>
</tr>
<tr>
<td>Adequate performance</td>
<td>8</td>
<td>1 – 3</td>
</tr>
<tr>
<td>Performance needing improvement</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 1: an example of performance boxes
Figure 1: the organisation of the Excel application
Figure 2: Comparison of promotions speeds of entry grade officials for 4 appraisal systems
Figure 3: Schematic of ‘as-is’ model operation
For each year

Start year PPR & grade

Merit point model

Compute *notional merit points*

New promotion points system based on boxes

For each official

Promotion threshold

End year PPR & grade

Figure 4: Schematic of performance-box model operation