Lancaster University Management School
Working Paper
2009/007

Modelling Career Development Schemes in the European Commission

Stephan Onggo, Mike Pidd, Didier Soopramanien and Dave Worthington

The Department of Management Science
Lancaster University Management School
Lancaster LA1 4YX
UK

MODELLING CAREER DEVELOPMENT SCHEMES IN THE EUROPEAN COMMISSION

Stephan Onggo, Michael Pidd, Didier Soopramanien, David Worthington
Department of Management Science
Lancaster University
Lancaster LA1 4YX
{s.onggo|m.pidd|d.soopramanien|d.worthington}@lancaster.ac.uk

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. A team from Lancaster University worked closely with Commission officials to develop a simulation model, which is used to demonstrate the performance of the current system and to show how alternative systems might function. As a consequence, the various stakeholders in the Commission's performance appraisal and promotions system have 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. It also uses novel regression-based sampling schemes which prove to be both accurate and efficient. Finally the role of the model in scenario exploration and system redesign is discussed.

KEYWORDS

Simulation, statistical modelling, manpower planning, public sector

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 nations. The Commission is the administrative body employing over 22,000 officials organised into departments known as Directorates General. The EC is a supra-national body that is separate from any national government and is charged with operating across the European Union. The Commission proposes legislation, which may or may not be accepted by 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) and a slightly more detached view can be found in Wikipedia.

Like most large public sector bodies, the Commission is obliged, by law, to treat its employees equally and, as part of this, has an annual performance appraisal system and an annual round of promotions. Such 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 (EC, 2008), which specify average 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 can 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 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 their 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 average ability. Hence, the distribution of merit points across each Directorate General is expected to have a mean value of approximately 14. Merit points above this average 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 following 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 are awarded 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 enough 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 of; 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 to ensure that an average promotion speed can be achieved. Thus, if the average points added to a rucksack were 17, setting the promotion threshold for a grade at 51 would ensure an average 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 system in place. 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, which stemmed from a 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 very difficult to predict promotion speeds because of 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 should 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.

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 EU clients in a variety of ways. These included an initial briefing meeting, analysis of the EU's large workforce 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, hence, the remainder of this paper describes the technical aspects of the work and discusses how the team and client groups interacted with and made use of the modelling work as it developed. The paper concludes with our reflections on the project.

MODEL DEVELOPMENT AND USE

Intended model use

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 figure 1 makes a distinction between models that are intended to replace thinking and those that are intended as tools to support thinking. This reflects the view of Pidd (2003) that models are frequently best considered as tools for thinking rather than ways of automating decisions.

It would, of course, be 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 the two positions shown in figure 1 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.

As a consequence, the Lancaster team and Commission officials agreed that the simulations were intended to provide tools to support Commission officials in their careful consideration of options for improved performance assessment and a revised promotion system.

Data cleansing

The seemingly smooth face of bureaucracy can conceal many surprising variations. The Commission currently employs about 22,000 officials and the appraisal and promotion system is intended to apply to all. However, life implies change and this is true of the Commission. 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 that would underpin the simulations should not be based on this complete set of around 22,000 officials. Thus, the available personnel records had to be reduced to *standard* officials in a multi-stage process.

Alongside the expansion of the Commission, its grading structure was also in flux. Most, but not all officials were now in one of two function groups (assistants and administrators):

- The administrator's function group consists of 10 AD grades, of which two are used for officials who enter employment at the Commission. The AD grades are for higher level officials who are at least graduates and whose posts involve policy drafting.
- The assistant's function group consists of 11 AST grades of which 3 are the grades at which officials enter employment at the Commission. The AST grades are for assistant staff, including clerical grades.

In addition, some assistants were members of other AST grades in force under the old (pre 2003) staff regulations, though some of them were gradually being transferred to current AST grades through internal procedures. The population of officials is unevenly spread across these grades, some of which currently have very few members. Tracking the movement of officials through the grades was an important requirement of the simulation.

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 during the period 2004 to 2006. Thus the following records were specifically excluded:

- Officials who changed function groups or moved to the new assistant's function group via internal procedures known as certification and attestation.
- Officials who joined or left the Commission during the period 2004 to 2006.

- Officials in AST.D grades, since the numbers in these grades were small and they were shortly to disappear.
- Officials not included in standard budgets.

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 final data set includes only 15,000 records and this serves as the basis for the modelling and simulation. This was agreed to be enough to serve the purpose of the simulations.

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 the year in which they complete probation. 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 may specify as part of a scenario to be simulated.

The Excel application

The contract for the work specified that the simulation models must be 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 turn, this suggested that it could be developed as an Excel workbook, with linked worksheets. 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. 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)).

Figure 2 shows the organisation of the main components of the Excel application. On starting the application, the user is presented with a user interface in Excel, written in VBA.

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 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 up to 20,000 officials (15,000 plus joiners) would need to be represented, with 10 or so variables for each of 10 or so years, giving a requirement for approximately 20,000 x 10 x 10 = 2 million cells of information. (All other memory requirements, e.g. look-up tables and performance measures, were negligible in this respect.)

However speed requirements were much more of an unknown as they would depend on the implementation of the appraisal system rules within VBA and on the sampling schemes that would be needed to represent the stochastic behaviour of each of 20,000 or so officials over the 10 year period. In particular the proposed appraisal systems would require annual sorting of officials by grade and perhaps other criteria, and then random shuffling to ensure no bias was caused by any residual ordering of officials from the original order of the database. With the amount of cells of information, frequent access to the worksheets would make the simulation very slow. The simulation would need to avoid this by carefully determining the required cells and loading them into memory at the right time. Also sampling to represent stochastic behaviour would need to be efficient given that there would be a minimum requirement of sampling at least 2 random variables (merit points and priority points) 20,000 x 10 = 200,000 times for each run of the simulation.

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.

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 this agreement was reached 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 the new system, it became clear that some possible features were not needed. Hence the attempt to model the new system 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 replacement system was

developed and used to investigate its likely operation. The results of the simulations were then used in discussions with the stakeholders to gain agreement to make the necessary changes.

Example Results

The basic design of the simulation model in Excel is that each official's future is recorded over the (10 year) simulation period. It is therefore a simple matter to use Excel (or any statistical package) 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 3 summarising repeated simulations of the 10 year period, were one basis for comparison of alternative appraisal systems. For this particular example, Figure 3 shows the resulting different promotion speeds for different appraisals systems for officials in an entry grade. Similarly Figure 4 shows that for one of the senior AD the four systems achieve an average promotion speed of about 4 years (as expected), but with different distributions.

STATISTICAL MODELLING

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. In the consistency analysis, we considered whether there had been any shift in the way that points were awarded through time. By 05/06, the actors awarding the merit and priority points, had learned how to best use the system. This does not mean that the whole approaches taken to awarding points was radically different in the first and final year, 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 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

We examined the merit points awarded to each standard official in each grade across the Commission and observed a correlation between the points awarded in successive years to each official. That is, an official who is awarded high merit points in one year is likely to gain high points in the following year and 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 it is important that this is represented in the simulations. The data analysis also revealed that the promotion of an official in year n affected the points awarded to that official in year $n+1$.

Based on our observation of a correlation between merit points awarded in one year and the next and the impact of recent promotion, we examined in detail the relationship between weighted merit points (if an official is in a grade for only part of the year, she gains only partial points and these need to be weighted as if she had been in grade for a full year) awarded in consecutive years. In doing so, we examined the correlations between years 03/04, 04/05 and 05/06. In all cases, the variance explained by the year on year correlation (R^2) was about 0.6. That is, for all years, there were similar relationships between merit points in

period *n+1* and merit points in period *n* and promotion in period *n* with some differences between AD and AST grades.

DG priority points

Our 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 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, so as to ensure that the simulation of the current system reflects actual practice.

Up to 10 priority points may be awarded to an official, but each Director General has only a limited number of these points that can be awarded (a quota of 2.5 per official). In general, we found that the award of DG priority points is correlated with the merit points awarded to an official, however there are other factors that also affect the award of DG priority points, which include:

- Whether or not the official was promoted the previous year. Typically promotion in the previous year meant less scope for earning DG points in the new grade.
- Whether the rucksack, before the award of DG priority points, is within 10 points of the announced threshold. These are the only officials for whom DG points are crucial for promotion in the current round (i.e. potentially promotable).
- Whether or not the official joined the Commission the previous year.

These observations led us to divide officials, each year, into 6 groups, that seemed to reflect the actual application of the promotion exercise in the Directorates General.

1. Officials who were promoted in the previous year.
2. Officials not promoted the previous year whose rucksack, before the award of DG priority points, is closer than 4 points to the announced threshold. From the data, we observed that these officials are almost always given enough points to ensure they reach the announced threshold and are promoted.
3. Officials not promoted the previous year whose rucksack is between 4 and 10 points from the announced threshold before the award of DG priority points. Such officials will be promoted if they are given enough points. The data suggested that some of these are ‘backed’ and are given the points they need, whilst those not backed are only given a small number of points, if any.

4. Officials not promoted in the previous year whose rucksack takes them to at least the announced promotion threshold before the award of DG priority points. These officials are expected to be promoted whatever the DG priority points awarded.
5. Officials not promoted the previous year whose rucksack is more than 10 points from the announced promotion threshold before the award of DG priority points. These officials cannot be promoted this year.
6. Officials who joined the Commission the previous year.

That is, our analysis of the dataset shows that the historical award of DG priority points reflect three processes. First, officials whose performance is highly rated 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 correct 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.

THE SIMULATION OF THE CURRENT SYSTEM – THE AS-IS MODEL

The as-is simulation model is implemented as an Excel workbook, developed within Excel 2003, though it should run under other versions. The data on which it runs is placed in a worksheet from the cleansed dataset of standard officials. The workbook is rather large, having roughly 20,000 rows to cope with the number of officials whose progress and promotion is simulated. The logic of the model is coded in Excel macros that are presented as VBA modules. The user need not access this code, but runs a simulation by using the input data set and specifying the period for which the model runs. Results are presented as Excel worksheets that are available for analysis in Excel or other suitable software.

The operation of the 'as-is' model

The operation of the as-is model is shown schematically in figure 5. The model does not distinguish between Directorates General, but runs for the Commission as a whole. The model runs on a year-by-year basis, using the merit points actually awarded in 2006 as its starting point. It can be run for any number of years to simulate the operation of the current system. However, it is probably unwise to run the model for more than 10 simulated years due to the parameter shifts that are always present in such models. Since the simulations are based on statistical sampling methods it is wise to replicate each run sufficient times to allow for sampling error. In fact most results turned out to be very stable from one run to another, due to the large number of officials being simulated.

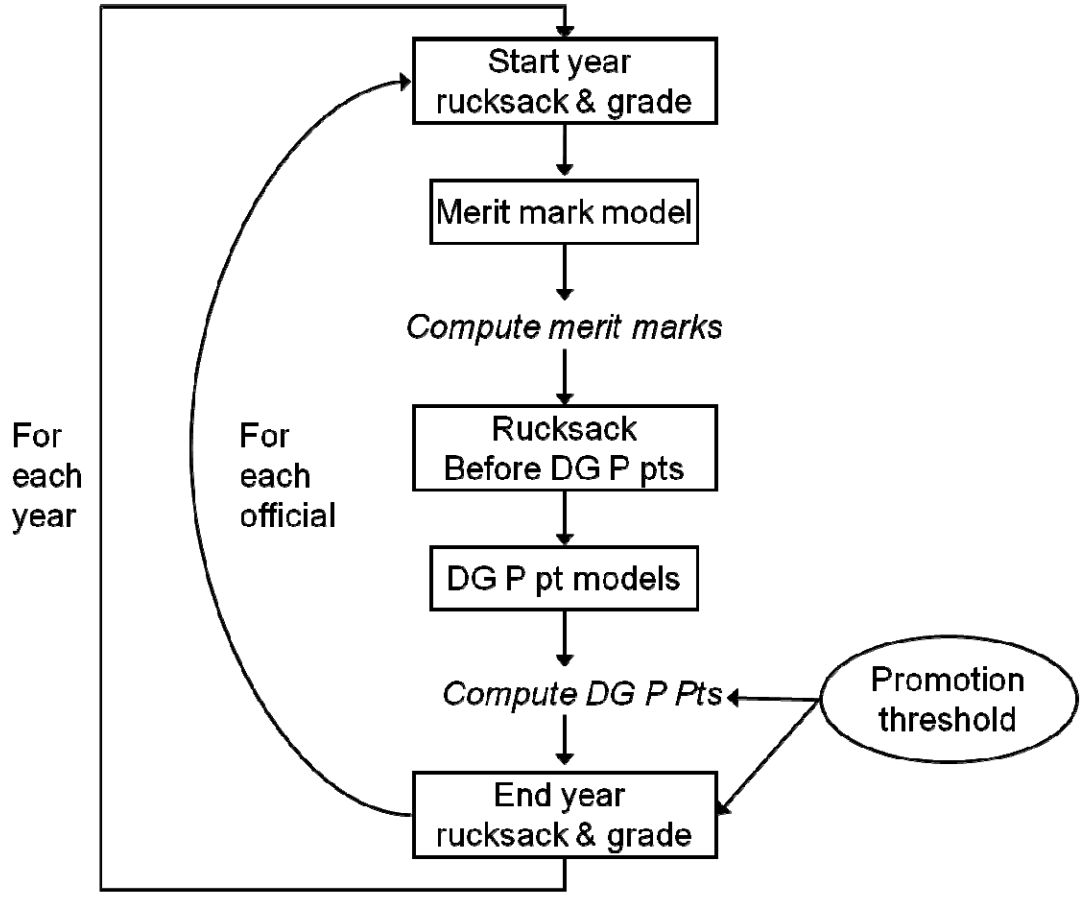

Figure 5: Schematic of 'as-is' model operation

Within each year, the model cycles through the following 4 phases for each standard official.

1. Compute the current year's merit points.
2. Add this year's merit points to the rucksack.
3. Compute 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.

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.

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. Since the simulated track record of each official is stored in Excel worksheets at the end of a run, the simulation results can be subjected to any appropriate ad hoc analyses using Excel or indeed any statistical package.

Simulating the award of merit points

The simulation of merit points uses top-hat sampling rom look-up tables based on the 05/06 dataset. These, most recent, datasets were used so as to ensure that the simulation of merit points reflects the most recent observed behaviour. Figure 6 shows a scatter plot of 2005 weighted merit points versus those in 2006.

Since the data analysis shows that promotion in the previous year also affected the weighted merit points awarded in the current year, a stepwise regression model was developed with two independent variables:

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

Where $M_{i,n+1}$ is the (weighted) merit point awarded to official i in year n+1, $M_{i,n}$ is the (weighted) merit point 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 ε_i is a random error term.

Having fitted regression models to establish the relationships, it was tempting to use the regression equations to predict the expected merit points for the next year and then add Normal error terms to introduce the stochastic variation. However Figure 6 shows that Normal error terms would not be appropriate, as the distribution of error terms is far from symmetric when the predicted points are either high or low. Furthermore merit points

obtained in this way would be continuous whereas actual merits points are discrete. Our solution was to use the data underpinning Figure 6 to create transition matrices, where each matrix contains elements $E_{i,j}$: the probability that an official awarded i merit points last year will be awarded j merit points in the current promotion and evaluation. For example, Table 1 shows an extract from the transition matrix used to define the look-up table for officials not promoted in the previous year. These are then used in top-hat sampling to allow for the random variation represented in the error term. One look-up table was created for those who were promoted in the previous year and another was created for those who were not. In addition, look-up tables for top-hat sampling were developed for joiners injected into the system at specific grades. This top-hat sampling from look-up tables allows very fast execution.

Table 1: Extract from transition matrix for merit points awarded to officials not promoted in the previous year.

		Merit Points in next year										
		…	12	12.5	13	13.5	14	14.5	15	15.5	16	….
Merit Points in previous year	….	…	…	…	…	…	…	…	…	…	…	…
	12	…	0.46	0.19	0.22	0.05	0.03	0.00	0.00	0.00	0.00	…
	12.5	…	0.03	0.37	0.30	0.18	0.08	0.01	0.00	0.00	0.00	…
	13	…	0.01	0.02	0.34	0.28	0.26	0.07	0.02	0.00	0.00	…
	13.5	…	0.01	0.01	0.03	0.29	0.41	0.19	0.05	0.00	0.00	…
	14	…	0.00	0.00	0.01	0.02	0.37	0.36	0.20	0.03	0.01	…
	14.5	…	0.00	0.00	0.00	0.01	0.03	0.39	0.44	0.11	0.02	…
	15	…	0.00	0.00	0.00	0.00	0.01	0.03	0.50	0.31	0.13	…
	15.5	…	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.49	0.37	…
	16	…	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.08	0.63	…
	…											

Simulating the award of DG priority points

Whereas the award of merit points is simulated by exploiting the observed year-on-year correlations in look-up tables, a different approach was need to simulate the award of DG priority points across the Commission. To this end, we devised a three stage process based on the idea of *promotional strength*.

Stage 1: compute promotional strength

Promotional strength is a concept used to represent how likely an official is to be promoted given several factors, including the size of their promotion rucksack before the award of priority points. The simulation stems from five regression equations obtained from the analysis of standard officials, one for each of groups 1 to 5 (see the earlier discussion of modelling these points) to compute the promotional strength of each official. These equations reflect our statistical analysis which showed that, within each group, the DG priority points awarded to each official are related to the merit points awarded in the current year, the DG priority points awarded in the previous year and the gap between the announced threshold and the rucksack before the award of any DG priority points. Joiners (i.e. group 6) are treated differently as they have no promotional strength.

The regression-based equations have the following form, for each of the groups, j=1 to 5.

$$S_{i,n} = \delta(j) + \eta(j).M_{i,n} + \theta(j).P_{i,n-1} + \varphi(j) .Gap_{i,j,n}$$

Where:

– i identifies the official,

– n is the current year

– $S_{i,n}$ represents the 'promotional strength' of official i in year n

– $\delta(j)$, $\eta(j)$, $\theta(j)$ and $\varphi(j)$ are constants that apply to all members of group j,

– $M_{i,n}$ is the merit points awarded to official i in year n,

– $P_{i,n-1}$ is the DG priority points awarded to official i in year $n-1$,

– $Gap_{i,n}$ is the gap between the rucksack and promotion threshold before the award of DG priority points in year n.

Using these equations produces measures of promotional strength for each official that is then used as a basis for sampling to produce DG priority points in stage 3.

Stage 2: adjust promotional strength

However, the fact that each Directorate General has only a limited number of points to distribute is important, but is not captured in the above equations for promotional strength. There is thus a danger that the number of points awarded may be too high or too low in total because of the mix of the current set of officials. Hence it is important to adjust this first-stage promotional strength to give a second-stage promotional strength to reflect the limited availability of promotion points. Hence, we adjust the promotional strength for officials in groups 1, 4 and 5, i.e. those officials for whom priority points are not critical in the current round, to correct the overall mean of points awarded. In essence, this attempts to mimic the way that DG priority point allocations are varied to match the available aggregate number of points.

Promotional strength is a good predictor of average priority points when taken across a number of officials, see for example Figure 7 which shows the S-shaped relationship $f_1(\,)$ between promotional strength and expected priority points for officials in group 1. The S-shaped function means that average promotional strengths by grade can be substituted into $f_1(\,)$ to give an easily calculated indication of average priority points by grade for group 1 officials. In the same way the equivalent functions for the other groups, denoted by $f_2(\,)$, $f_3(\,)$, $f_4(\,)$ and $f_5(\,)$, have similar shapes and are again used to indicate the average priority points by grade.

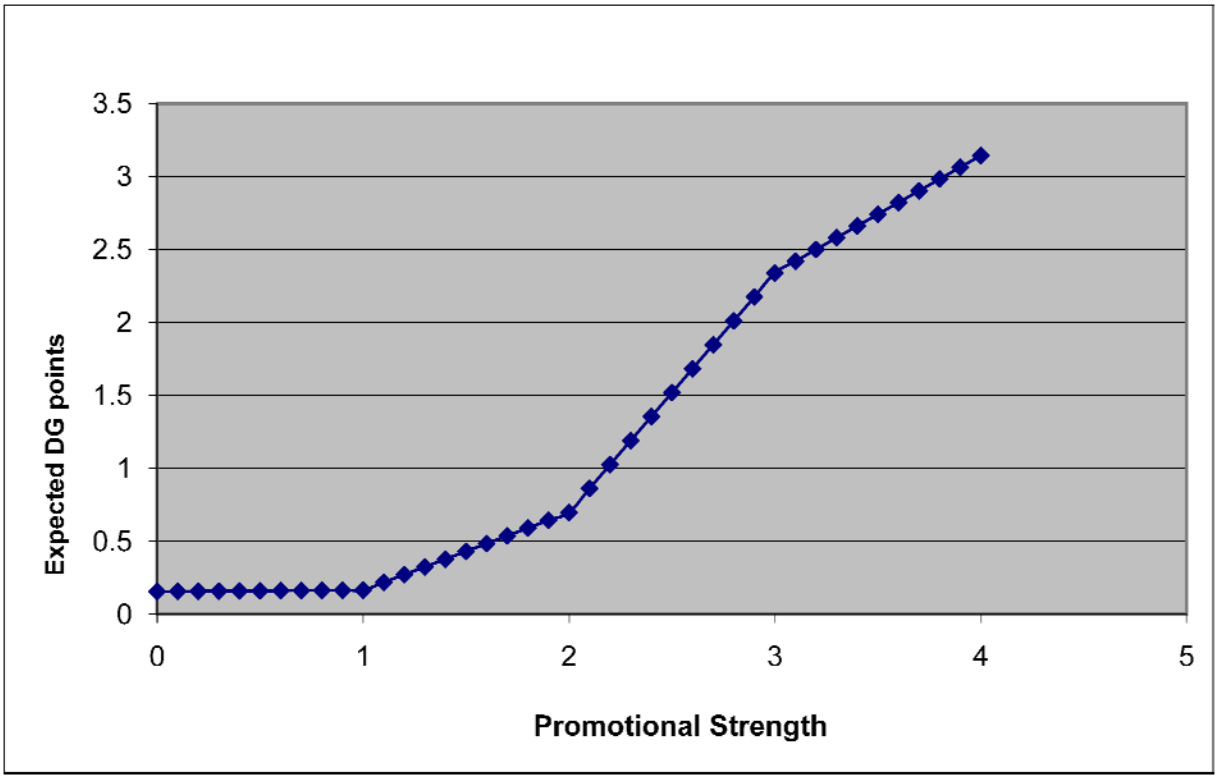

Figure 7: Promotional Strength and DG points

These indications are then used as a basis for increasing or decreasing the promotional strength of officials in groups 1, 4 and 5 in each grade, so that the revised indication of average priority points is set to the EU target average of 2.5 per official. The steps involved are as follows:

1. For each grade *g*, estimate the average priority points, *Average(P_g)*, using:

$$Average(P_g) = \frac{\sum_{j=1}^{5} N_{j,g} \times f_j(Mean(S_{j,g}))}{\sum_{j=1}^{5} N_{j,g}}$$

   where $N_{j,g}$ is the number of officials in grade *g* who are in group *j,* for *j*= 1 to 5; *Mean($S_{j,g}$)* is the mean promotional strength in each of the groups.

2. Suppose this value differs from the target average (2.5) by some factor δ, i.e. δ=. *(Average(P_g)-2.5)*, Calculate the adjustment factor for groups 1, 4 and 5, i.e. δ', using:

$$\delta' = \delta \times \frac{\sum_{j=1}^{5} N_{j,g}}{(N_{1,g} + N_{4,g} + N_{5,g})}$$

3. For each grade calculate constants γ_1, γ_4 and γ_5 by which to modify the promotional strength for officials in groups 1, 4 and 5 in order that the indicated average priority points change by δ'. These constants are estimated using the functions $f_1(\,)$, $f_4(\,)$ and $f_5(\,)$. For example suppose the indicated average priority points in group 1 was X_1 and we wish to increase it by δ', we find γ_1 such that $f_1(X_1 + \gamma_1) = f_1(X_1) + \delta$'.
4. Finally use constants γ_1, γ_4 and γ_5 to modify the previously obtained promotional strengths of officials in classes 1, 4 and 5 of grade g.

Stage 3: compute DG priority points

This stage uses the modified promotional strength scores to produce DG priority points that will be awarded to each official. By rounding promotional strength to the nearest integer the data underpinning the regression-based models could 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 same 6 groups are used as in stage 1 and each group has a look-up table which is used for top-hat sampling to produce the DG priority points awarded to each official. For example, Table 2 shows the matrix of probabilities used to define the look-up table for officials who were promoted in the previous year. These are then used in top-hat sampling to allow for the random variation represented in the error term.

Table 2: Matrix of conditional probabilities (given promotional strength) for priority points awarded to officials promoted in the previous year

	DG priority points										
Promotional Strength	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
<=0	0.96	0.02	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00
1	0.87	0.10	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.60	0.20	0.14	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.00
3	0.29	0.12	0.16	0.11	0.22	0.01	0.05	0.02	0.01	0.01	0.01
>=4	0.21	0.21	0.14	0.07	0.00	0.00	0.29	0.00	0.00	0.00	0.07

.

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 simulation 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 3.

Table 3: one possible set of performance boxes

Performance description	% of staff in a 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

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 operation of the simulation of replacement systems

Figure 8 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.

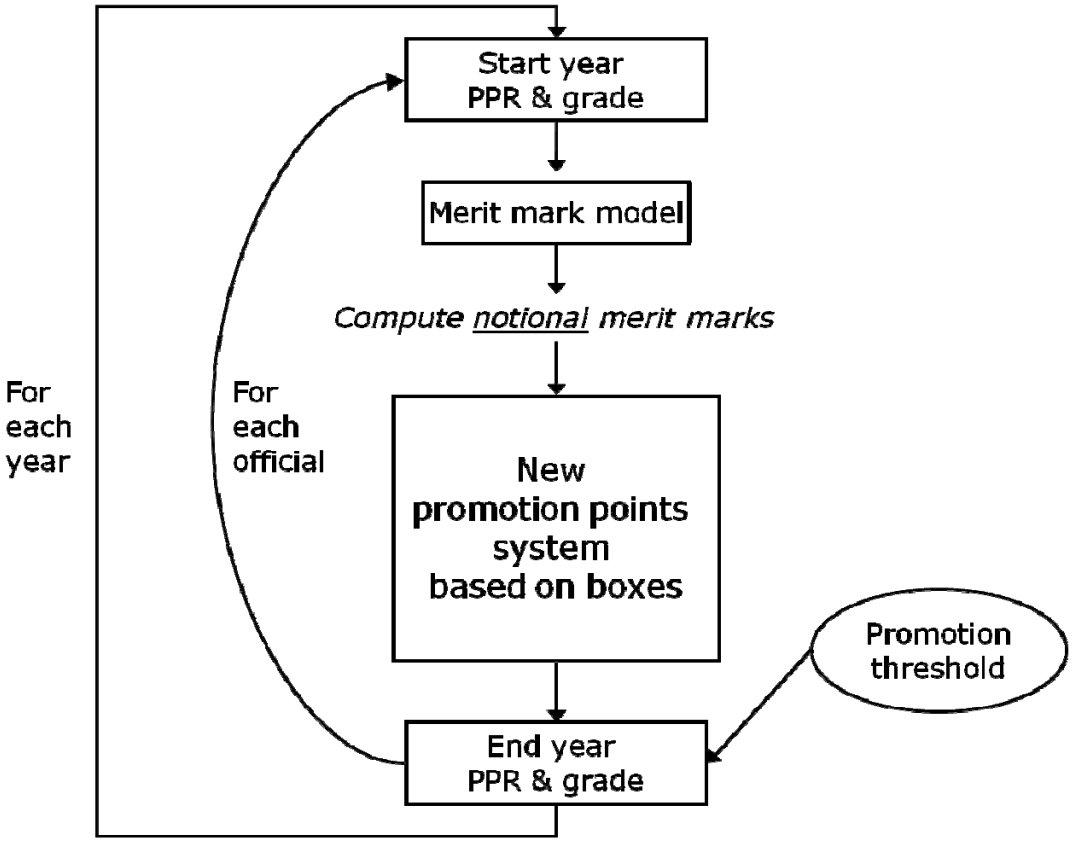

Figure 8: Schematic of performance-box model operation

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. In addition, for the simulation model, there must be a translation between performance measured in merit points, the performance box assessment and promotion points awarded. Hence, in a simulation of the operation of the box system, *notional merit points* are computed in the exactly the same way as merit points in the as-is model. As highlighted in figure 9, 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. It is important to realise that using notional merit points in this way allows a proper comparison with the existing system, since both are based on the same models of performance.

Since each promotion box may cover a range of promotion points, it is necessary to allocate points that lie within the sub-range; that is, to determine how many promotion points are awarded to each official. In the box model, three factors are allowed to affect both the box within which an official may be placed and the number of promotion points awarded to that official.

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 to play in allocating promotion points, the model requires the user to specify weights for each factor. This is done by allocating a percentage to each factor. Thus if all three factors are given a 33.33% weighting, this indicates that they will be equally weighted. By contrast if 100% is allocated to the notional merit points and zero to the seniority and gap, this means that promotion points are wholly dependent on the performance as reflected in the notional merit points.

The three factors are used as follows in the box model.

1. Sort the officials within the grade by their notional merit points this year and award a merit rank (*M*) to each official. Since merit points are awarded as half integers many tied ranks would occur, so a small random amount varying between 0 and 0.4999 is added to

the notional merit points to produce perturbed notional merit points on which the officials are ranked uniquely.

2. Sort the officials by seniority and, on this basis, award a seniority rank (S).

3. Sort the officials by threshold gap and, on this basis, award a gap rank (G). Since threshold gap is an integer, a small random amount varying between 0 and 0.9999 is added to produce a perturbed threshold gap on which the officials are ranked uniquely.

4. Compute a multi-factor rank as ($w_m.M + w_s.S + w_g.G$) for each official where w_m is the percentage weight attached to the perturbed notional merit points, w_s is the percentage weight attached to seniority and w_g is the percentage weight attached to the threshold gap. Use this weighted sum to provide an overall ranking for the officials.

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.

REFLECTIONS ON THE PROJECT

We learned much from this project, and summarise what we believe are our most important reflections under the headings: modelling HRM systems, simulation in Excel, regression-based sampling schemes, and the role of modelling.

Modelling HRM systems

To our knowledge 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), Ekamper (1997), Weber (1971)). There are obvious similarities between our work and more traditional manpower planning models where staff move between grades according to various 'push' or 'pull' rules. However in our case promotions are driven by the EU 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.

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, we took great care in designing the necessary sampling and sorting algorithms to ensure very fast running. For example, we used memory rather than the 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.

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 which had been found to provide reasonable representations of the relationships in the data.

Regression modelling showed a strong year-on-year relationship for merit points. However 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 a set of look-up tables, with the look-up 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, as noted above, is that top-hat sampling is very efficient.

Regression modelling also 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 depending on the official's promotion group. 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. One computational benefit is again the speed of top-hat sampling. An added computational bonus is that the relationship between promotional strength and expected priority points could be used to adjust the promotional strength equations prior to sampling so that average priority points per official was close to the EU target. The alternative would be an iterative sampling process in which the priority points of every official was sampled in order to calculate the overall average, the promotional strengths would then be adjusted before a second round of sampling.

Role of Modelling

Our final reflection is to compare 'actual model use' with 'intended model use'. As noted earlier the EU clients well 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 we met with the EU clients 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 ourselves, the EU liaison group and the EU working group. For example, early data analyses from our model building process were fed back to the clients for potential inclusion in the model. These formed the basis of a discussions that helped the EU 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 well-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 we asked in order to develop the simulation model prompted the client and the working group to think clearly through the key features of any potential system. Furthermore preliminary results that we produced for validation purposes were just as likely to prompt refinements to their proposed alternatives as they were to uncover ‘faults’ in our 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 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 EC 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 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 on software engineering, 15, 2, 109-119

Bartholomew D.J. and Forbes A.F. (1979) Statistical techniques for manpower planning. John Wiley, Chichester.

Blosch M and Antony J. (1999) Experimental design and computer-based simulation: a case study with the Royal Navy. Managing service quality, 9, 5, 311-320.

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)

Ekamper (1997) Future age-conscious manpower planning in the Netherlands. Intl Jnl Manpower, 18, 3, 232-247.

Pidd M (2003) Tools for thinking: modelling in management science (2nd Ed). John Wiley, Chichester.

Powell S.G. and Baker K.R. (2004) The art of modeling with spreadsheets: management science, spreadsheet engineering and modeling craft. John Wiley, New York.

Powell S.G. and Batt R.J. (2008) Modeling for insight: a master class for business analysts. John Wiley, New York.

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

Weber W.L. (1971) Manpower planning in hierarchical organizations: a computer simulation approach. Mgt Sci, 18, 3, 119-144.

Winston W.L. (2004) Microsoft Excel data analysis and business modelling. Microsoft Press, Redmond, Washington.