JIA 112 (1985) 61-73

SOME EXPERIMENTS WITH SALARY SCALES

BY A. D. WILKIE, M.A., F.F.A., F.I.A.

1. SALARY scales have been widely used in actuarial literature about pension schemes, but they do not seem to have been developed beyond the idea first introduced by Manly (1901) and used in a series of papers following this, including McGowan (1901), Manly (1902, 1903 and 1911), and M'Lauchlan (1908). King (1905), Bacon (1907) and M'Lauchlan (1914) discuss the construction of a salary scale from records of individual employees. King made some valuable observations on how a salary scale may change with time if the observed population is not a stationary one, for example, because the firm is growing or declining, which Bacon also commented on, and M'Lauchlan went into considerable detail about the separation of different grades. Thomas (1913) gave an example of an organization with six ranks, within each of which there was a salary scale, and showed explicit probabilities of promotion in each year of age. His development comes closest to what I shall discuss below. Text books on Life Contingencies, such as Jordan (1952), Hooker & Longley-Cook (1957) and most recently Neill (1977), have followed essentially the definition introduced by Manly, as also have papers and text books on pension funds, such as Porteous (1936), Marples (1948), Heywood & Marples (1950), Crabbe & Poyser (1953) and Lee (1973). Curiously Spurgeon (1922) does not mention salary scales, although his book was written after they had come into use.

2. In principle, the salary scale s_x is such that the ratio s_{x+t}/s_x equals the ratio of the average salaries of employees in the year of age $(x+t,\ x+t+1)$ to their average salaries in the year of age $(x,\ x+1)$, conditional on their surviving in the relevant employment. Such a salary scale allows for the fact that average salary levels typically increase to some extent with age, both because of regular progression up a scale with annual increments, and because of promotion to a higher salary level with increasing age. It is convenient to assume that salaries in this context are measured in terms of a constant general level of salaries, which does not change with time. The effects of general price inflation on the money value of salaries, and the further effects of changes in the real level of salaries, measured in constant price terms, are usually brought in separately. I shall ignore these complications in this note.

3. I shall also ignore variations in the frequency of salary payments and the incidence of salary increases. I assume, for simplicity, that all salary increases take place on birthdays; thus the rate of salary throughout the year of age is constant, and the frequency of payment is irrelevant.

4. The specimen salary scale given by Neill and used also in 'Formulae and Tables for Actuarial Examinations' (1980) is reproduced in the column headed s_x

of Table 1. It runs from 1·00 at age 18 to 5·40 at age 64. It is assumed that everyone retires no later than their 65th birthday. I now define the function r_x, where:

$$1+r_x = s_x/s_{x-1},$$

that is, r_x is the average proportionate increase in salaries on the xth birthday. The values of $100r_x$ are also given in Table 1, and are seen to run from 10% at age 19 to 0·56% at age 64.

5. While the definitions and descriptions so far give us a general idea of what is happening to average salaries among the employees in question it is clear that we would need a more detailed description if we wished to consider the salary progress of any individual. One possible interpretation of the model would be that each individual enters service with a given salary, which may vary according to the individual and his age at entry, but that he then progresses rigidly up the salary scale, receiving increases each year in accordance with r_x, so that at any age his salary depends solely on his starting salary and on the salary scale so far. This seems a fairly unrealistic representation. In reality some employees do receive promotional increases, while others do not.

6. The definition of salary scale makes no reference to the distribution of salary levels at age x, nor to the previous salary history of the relevant employees at age

Table 1. *Salary scale and annual percentage increase*

Age x	s_x	r_x%	Age x	s_x	r_x%
			40	3·58	2·87
			41	3·68	2·79
			42	3·78	2·72
18	1·00	—	43	3·88	2·65
19	1·10	10·00	44	3·98	2·58
20	1·21	10·00	45	4·08	2·51
21	1·33	9·92	46	4·18	2·45
22	1·46	9·77	47	4·28	2·39
23	1·59	8·90	48	4·38	2·34
24	1·73	8·81	49	4·47	2·05
25	1·87	8·09	50	4·56	2·01
26	2·02	8·02	51	4·65	1·97
27	2·16	6·93	52	4·73	1·72
28	2·29	6·02	53	4·81	1·69
29	2·42	5·67	54	4·88	1·46
30	2·55	5·37	55	4·95	1·43
31	2·67	4·71	56	5·01	1·21
32	2·78	4·12	57	5·07	1·20
33	2·88	3·60	58	5·13	1·18
34	2·98	3·47	59	5·19	1·17
35	3·08	3·36	60	5·24	·96
36	3·18	3·25	61	5·29	·95
37	3·28	3·14	62	5·33	·76
38	3·38	3·05	63	5·37	·75
39	3·48	2·96	64	5·40	·56

x. For these factors to be irrelevant implies that the definition applies regardless of which employees at age x are under consideration, and hence applies to each individual employee at age x. We can therefore restate the definition more formally: let the salary of employee i at age x be $Y_i(x)$; then:

$$E(Y_i(x+t)|Y_i(x) = Y \text{ and } i \text{ survives}) = Y \cdot s_{x+t}/s_x,$$

that is, given that the salary of employee i at age x is Y, the expected value of his salary at age $x+t$ is given by Y times the ratio of the salary scale factors, provided he survives. It can be seen that the expected proportionate change in his salary between ages x and $x+t$ depends neither on the level of his salary at age x, nor on his salary history prior to age x. In particular:

$$E(Y_i(x+1)|Y_i(x) = Y \text{ and } i \text{ survives}) = Y \cdot (1+r_{x+1}).$$

7. There are infinitely many distributions of $Y_i(x+1)/Y$ that have a mean of $1+r_{x+1}$. In order to define the distribution of salaries further we need to make further assumptions. The following is only one among many methods, but it is a simple one, and consideration of it may give some insight into possible alternative models.

8. Consider an employee who enters service at the youngest age in the salary scale, 18 in this case, with a salary of 1·0. For convenience I measure all salaries in terms of this base unit. The ladder of possible future salaries in each year is defined by powers of $(1+j)$; I shall choose a value for j later. At each age, x, the employee moves up the ladder either $k(x)$ or $k(x)+1$ steps, with respective probabilities $q(x)$ and $p(x) = 1-q(x)$, such that his expected increase is r_x. That is:

$$\begin{aligned} Y_i(x)/Y_i(x-1) &= (1+j)^{k(x)+1} \quad \text{with probability } p(x) \\ &= (1+j)^{k(x)} \quad \text{with probability } q(x), \end{aligned}$$

and

$$E(Y_i(x)/Y_i(x-1)) = 1+r_x.$$

Since $k(x)$ is integral, it has to be chosen so that:

$$(1+j)^{k(x)} \leqslant 1+r_x \leqslant (1+j)^{k(x)+1},$$

which determines $k(x)$ uniquely unless one of the equalities holds. We can then determine $p(x)$ by:

$$1+p(x)\cdot j = (1+r_x)/(1+j)^{k(x)}.$$

If $(1+r_x)$ exactly equals a power of $(1+j)$, then we can either choose $k(x)$ to equal that power, with $p(x) = 0$, or one less than that power, with $p(x) = 1$. The effect in either case is that, in that year, the employee is certain to move a particular number of steps up the ladder. In other years he may rise for example either 0 or 1 steps, 1 or 2 steps, etc., with the appropriate probabilities.

9. A few examples may make the process clearer. If $j = \cdot 2$, i.e. each step on the

ladder implies a salary 20% higher than the previous step, then at age 19, when $r_{19} = \cdot 10$, the probability of moving up one step on the ladder is ·5 and of staying on the same step is also ·5, giving an average rise of ·10. The same happens at age 20, since r_{20} also equals ·10; at age 21 we have $r_{21} = \cdot 09917$, so the probability of an increment is ·4959, and the probability of staying on the same step is ·5041. If we choose $j = \cdot 10$, then at ages 19 and 20 the employee certainly moves up exactly one step, and at age 21 his probability of an increae is ·9917. If $j = \cdot 05$, then at age 19 the employee may go up one step (ratio 1·05), or two steps (ratio 1·1025) with respective probabilities ·0476 and ·9524, and so on.

10. Now let $f(x,h)$ be the probability that at age x our 18-year-old entrant is on the hth step of the ladder above his starting point, i.e. his salary is $(1+j)^h$. We see that he can reach this position from age $x-1$ either by having been at point $h-k(x)$ and going up $k(x)$ steps, or having been at point $h-k(x)-1$ and going up $k(x)+1$ steps. Thus:

$$f(x,h) = q(x)\cdot f(x-1,\, h-k(x)) + p(x)\cdot f(x-1,\, h-k(x)-1),$$

with initial conditions:

$$\begin{aligned} f(18,h) &= 1 \text{ if } h = 0 \\ &= 0 \text{ otherwise.} \end{aligned}$$

We can thus readily calculate recursively the probability distribution of salaries at age x for our hypothetical 18-year-old entrant.

11. In order to apply the same salary ladder and the same probability distributions to all employees we need to make some further strong assumptions. First, that a new entrant at age x receives a starting salary which is one of those on the ladder, with probability equal to $f(x,h)$. Further, that the probability of dying, retiring either through age or ill health, or leaving service for any other reason, is independent of the step on the salary ladder reached at the time of exit. Then the probabilities $f(x,h)$ give the probability distribution for all members in employment at age x.

12. A model such as this is at least sufficient for one to be able to use a salary scale for pension fund valuation in the usual actuarial way, although the assumptions are not all necessary ones. Thus, many alternative patterns of dispersion probabilities would be equally satisfactory. One does not need to assume that new entrants have the same distribution of salaries as existing staff. But one does need to assume that at least the average salary of those who die, retire, withdraw, etc. is the same as the average salary of those who do not. And one does need to assume that the expected salary progression at any age is independent of the salary at that age and of the previous salary history.

13. If we make a further assumption about the distribution of the population of employees at each age, say, that the proportion of all employees of age x is $g(x)$, then we can calculate a distribution of salaries for the total population of employees. An appropriate distribution will depend on the circumstances of each case, but two impartial ones are:

(a) equal numbers at each age from 18 to 64;
(b) numbers at age x are proportional to ${}_{x-18}p_{18}$, for x = 18 to 64.

The second distribution assumes that we have a stationary population in which everyone was born on the same day of the year, that we count salaries on that date, that everyone enters service at the age of 18, and that the only withdrawal until retirement at age 65 is by death. This would not be very realistic for most individual firms, but it may not be too unreasonable for the total employed population of a country.

14. Now for some numerical results: Table 2 shows the percentage distribution of salaries at ages 24, 34, 44, 54, and 64, for j such that $1+j = \exp(\cdot 10)$, i.e. a ladder with steps of 10·5171%. A dash (–) indicates a zero probability; ·00 indicates a very small, but non-zero probability. It can be seen how at 24 there are only seven possible steps on the ladder, with over half the population being on the top step; the dispersion increases with age, so that by age 64 there are 45 possible steps though the probabilities of reaching the highest steps or having remained on the lowest steps are extremely small, and most of the population is spread across seven or eight steps rather below the middle of the range.

15. Table 3 shows summary statistics at each age for j such that $1+j = \exp(\cdot 05)$, $\exp(\cdot 10)$ and $\exp(\cdot 20)$, i.e. steps of 5·1271%, 10·5171% and 22·1403%. I have chosen these values so that the steps may coincide in later tables. The mean salary at each age is the same for each ladder, and of course is the same as the original salary scale. It is given in the first column. The statistics shown are the standard deviation of salary at each age, and the Gini coefficient at each age, which will be explained later. It can be seen how the dispersion, measured by the standard deviation, increases with age, and increases with increasing j.

16. Table 4 shows certain statistics relating to the total population, assuming that the population is distributed from age 18 to 64 in accordance with assumption (b) above and using A1967–70 ultimate mortality. In fact the figures are not substantially different from those using assumption (a). The figures are shown for the same values of j. The columns for each value of j show:

the cumulative percentage of the population on a particular step of the ladder or below it, and

the percentage of total salaries received by those on or below each step.

At the foot of the table are shown the mean salary, standard deviation of salary and Gini coefficient for each distribution. Formally: we define the distribution for the whole population by:

$$f(h) = \sum_{x=18}^{64} g(x)\cdot f(x,h),$$

the cumulative distribution up to step H by:

$$F(H) = \sum_{h=0}^{H} f(h),$$

Table 2. *Percentage distribution of salaries at selected ages:* $(1+j) = \exp(\cdot 10)$

Step h	Salary $(1+j)^h$	Ages: 24	34	44	54	64
0	1·0000	·00	·00	·00	·00	·00
1	1·1052	·00	·00	·00	·00	·00
2	1·2214	·06	·00	·00	·00	·00
3	1·3499	·97	·00	·00	·00	·00
4	1·4918	8·24	·00	·00	·00	·00
5	1·6487	34·57	·07	·00	·00	·00
6	1·8221	56·16	·44	·03	·00	·00
7	2·0138	—	1·92	·16	·03	·01
8	2·2255	—	5·95	·64	·12	·06
9	2·4596	—	13·19	1·99	·45	·22
10	2·7183	—	20·97	4·83	1·31	·67
11	3·0042	—	23·76	9·29	3·11	1·72
12	3·3201	—	18·93	14·28	6·07	3·67
13	3·6693	—	10·32	17·69	9·86	6·57
14	4·0552	—	3·64	17·79	13·47	10·00
15	4·4817	—	·75	14·59	15·60	13·06
16	4·9530	—	·07	9·78	15·42	14·73
17	5·4739	—	—	5·35	13·08	14·44
18	6·0496	—	—	2·39	9·56	12·38
19	6·6859	—	—	·86	6·04	9·32
20	7·3891	—	—	·25	3·31	6·19
21	8·1662	—	—	·06	1·57	3·64
22	9·0250	—	—	·01	·65	1·89
23	9·9742	—	—	·00	·23	·88
24	11·0232	—	—	·00	·07	·36
25	12·1825	—	—	·00	·02	·13
26	13·4637	—	—	·00	·00	·04
27	14·8797	—	—	—	·00	·01
28	16·4446	—	—	—	·00	·00
29	18·1741	—	—	—	·00	·00
30	20·0855	—	—	—	·00	·00
31	22·1980	—	—	—	·00	·00
32	24·5325	—	—	—	·00	·00
33	27·1127	—	—	—	·00	·00
34	29·9641	—	—	—	·00	·00
35	33·1154	—	—	—	·00	·00
36	36·5982	—	—	—	·00	·00
37	40·4473	—	—	—	·00	·00
38	44·7012	—	—	—	—	·00
39	49·4024	—	—	—	—	·00
40	54·5982	—	—	—	—	·00
41	60·3403	—	—	—	—	·00
42	66·6863	—	—	—	—	·00
43	73·6998	—	—	—	—	·00
44	81·4509	—	—	—	—	·00
45	90·0171	—	—	—	—	·00
46	99·4843	—	—	—	—	·00

Table 3. *Mean, standard deviation and Gini coefficients for salary distributions*

Age(x)	Mean salary	$(1+j) = \exp(\cdot05)$ Std dev.	Gini co. %	$(1+j) = \exp(\cdot10)$ Std dev.	Gini co. %	$(1+j) = \exp(\cdot20)$ Std dev.	Gini co. %
18	1·00	·0	·0	·0	·0	·0	·0
19	1·10	·0159	·43	·0227	·45	·1102	4·98
20	1·21	·0247	·78	·0354	·85	·1718	7·48
21	1·33	·0340	1·12	·0488	1·28	·2319	9·34
22	1·46	·0448	1·47	·0645	1·76	·2947	10·88
23	1·59	·0607	1·98	·0894	2·60	·3594	12·21
24	1·73	·0771	2·39	·1153	3·32	·4290	13·40
25	1·87	·0954	2·78	·1464	4·10	·5012	14·46
26	2·02	·1147	3·12	·1790	4·76	·5793	15·44
27	2·16	·1330	3·40	·2165	5·47	·6559	16·31
28	2·29	·1475	3·55	·2558	6·15	·7301	17·07
29	2·42	·1603	3·66	·2961	6·76	·8058	17·77
30	2·55	·1711	3·70	·3373	7·33	·8829	18·43
31	2·67	·1827	3·78	·3779	7·85	·9562	19·00
32	2·78	·1979	3·94	·4171	8·32	1·0251	19·51
33	2·88	·2152	4·14	·4543	8·75	1·0892	19·96
34	2·98	·2332	4·35	·4917	9·15	1·1540	20·38
35	3·08	·2518	4·55	·5293	9·52	1·2194	20·79
36	3·18	·2710	4·75	·5672	9·87	1·2855	21·18
37	3·28	·2906	4·94	·6054	10·21	1·3521	21·54
38	3·38	·3107	5·13	·6438	10·53	1·4194	21·90
39	3·48	·3313	5·32	·6826	10·83	1·4873	22·24
40	3·58	·3522	5·50	·7216	11·12	1·5557	22·56
41	3·68	·3736	5·67	·7609	11·40	1·6247	22·88
42	3·78	·3952	5·84	·8005	11·67	1·6942	23·18
43	3·88	·4172	6·01	·8404	11·92	1·7643	23·47
44	3·98	·4395	6·17	·8805	12·17	1·8349	23·75
45	4·08	·4621	6·33	·9210	12·41	1·9060	24·02
46	4·18	·4849	6·49	·9617	12·64	1·9776	24·28
47	4·28	·5081	6·64	1·0027	12·86	2·0497	24·53
48	4·38	·5314	6·78	1·0440	13·07	2·1223	24·78
49	4·47	·5536	6·92	1·0818	13·26	2·1884	25·00
50	4·56	·5759	7·06	1·1199	13·45	2.2548	25·21
51	4·65	·5983	7·19	1·1582	13·63	2·3216	25·41
52	4·73	·6191	7·32	1·1928	13·79	2·3816	25·60
53	4·81	·6400	7·43	1·2275	13·95	2·4418	25·77
54	4·88	·6589	7·54	1·2583	14·08	2·4949	25·92
55	4·95	·6779	7·65	1·2893	14·22	2·5483	26·07
56	5·01	·6947	7·74	1·3162	14·33	2·5944	26·20
57	5·07	·7116	7·83	1·3431	14·45	2·6406	26·33
58	5·13	·7284	7·93	1·3701	14·56	2·6870	26·45
59	5·19	·7453	8·01	1·3972	14·67	2·7335	26·57
60	5·24	·7598	8·09	1·4201	14·76	2·7725	26·67
61	5·29	·7743	8·16	1·4430	14·85	2·8116	26·77
62	5·33	·7862	8·23	1·4615	14·92	2·8431	26·85
63	5·37	·7981	8·29	1·4801	14·99	2·8747	26·93
64	5·40	·8072	8·33	1·4941	15·05	2·8985	26·99

Table 4. *Cumulative proportions in total population*

Step h for $1+j = \exp(\cdot 05)$	Salary $(1+j)^h$	$(1+j) = \exp(\cdot 05)$ F%	U%	$(1+j) = \exp(\cdot 10)$ F%	U%	$(1+j) = \exp(\cdot 20)$ F%	U%
0	1·0000	2·222	·636	2·332	·669	5·043	1·446
1	1·0513	2·430	·700	—	—	—	—
2	1·1052	4·454	1·341	4·665	1·408	—	—
3	1·1618	4·840	1·470	—	—	—	—
4	1·2214	6·710	2·125	7·021	2·233	10·689	3·424
5	1·2840	7·263	2·329	—	—	—	—
6	1·3499	8·999	3·001	9·453	3·175	—	—
7	1·4191	9·734	3·300	—	—	—	—
8	1·4918	11·430	4·025	12·101	4·308	17·703	6·425
9	1·5683	12·458	4·488	—	—	—	—
10	1·6487	13·981	5·208	14·897	5·630	—	—
11	1·7333	15·223	5·825	—	—	—	—
12	1·8221	16·758	6·627	18·030	7·267	26·811	11·184
13	1·9155	18·173	7·405	—	—	—	—
14	2·0138	19·767	8·325	21·689	9·380	—	—
15	2·1170	21·458	9·352	—	—	—	—
16	2·2255	23·330	10·547	26·177	12·244	38·212	18·460
17	2·3396	25·352	11·903	—	—	—	—
18	2·4596	27·606	13·493	31·610	16·076	—	—
19	2·5857	30·160	15·387	—	—	—	—
20	2·7183	33·056	17·645	38·076	21·117	51·228	28·607
21	2·8577	36·255	20·266	—	—	—	—
22	3·0042	39·694	23·229	45·529	27·538	—	—
23	3·1582	43·337	26·528	—	—	—	—
24	3·3201	47·184	30·192	53·754	35·369	64·433	41·181
25	3·4903	51·260	34·271	—	—	—	—
26	3·6693	55·598	38·836	62·417	44·485	—	—
27	3·8574	60·236	43·967	—	—	—	—
28	4·0552	65·189	49·727	71·052	54·528	76·247	54·920
29	4·2631	70·419	56·120	—	—	—	—
30	4·4817	75·794	63·030	79·081	64·847	—	—
31	4·7115	81·090	70·186	—	—	—	—
32	4·9530	86·015	77·180	85·947	74·600	85·562	68·151
33	5·2070	90·286	83·558	—	—	—	—
34	5·4739	93·708	88·930	91·292	82·991	—	—
35	5·7546	96·242	93·083	—	—	—	—
36	6·0496	97·913	96·012	95·053	89·516	92·047	79·401
37	6·3598	98·942	97·890	—	—	—	—
38	6·6859	99·511	98·981	97·436	94·084	—	—
39	7·0287	99·795	99·552	—	—	—	—
40	7·3891	99·922	99·822	98·792	96·957	96·040	87·863
41	7·7679	99·973	99·936	—	—	—	—
42	8·1662	99·992	99·980	99·483	98·577	—	—
43	8·5849	99·998	99·994	—	—	—	—
44	9·0250	100·000	99·999	99·800	99·397	98·221	93·507
45	9·4877	100·000	100·000	—	—	—	—
46	9·9742	100·000	100·000	99·930	99·769	—	—
47	10·4856	100·000	100·000	—	—	—	—
48	11·0232	100·000	100·000	99·978	99·926	99·279	96·852
49	11·5883	100·000	100·000	—	—	—	—

Table 4. (*Cont.*)

Step h for $1+j = \exp(\cdot 05)$	Salary $(1+j)^h$	$(1+j) = \exp(\cdot 05)$ $F\%$	$U\%$	$(1+j) = \exp(\cdot 10)$ $F\%$	$U\%$	$(1+j) = \exp(\cdot 20)$ $F\%$	$U\%$
50	12·1825	100·000	100·000	99·994	99·975	—	—
51	12·8071	100·000	100·000	—	—	—	—
52	13·4637	100·000	100·000	99·998	99·993	99·737	98·619
53	14·1540	100·000	100·000	—	—	—	—
54	14·8797	100·000	100·000	100·000	99·998	—	—
55	15·6426	100·000	100·000	—	—	—	—
56	16·4446	100·000	100·000	100·000	100·000	99·913	99·452
57	17·2878	100·000	100·000	—	—	—	—
58	18·1741	100·000	100·000	100·000	100·000	—	—
60	20·0855			100·000	100·000	99·974	99·803
62	22·1980			100·000	100·000	—	—
64	24·5325			100·000	100·000	99·993	99·936
66	27·1126			100·000	100·000	—	—
68	29·9641			100·000	100·000	99·998	99·981
70	33·1155			100·000	100·000	—	—
72	36·5982			100·000	100·000	100·000	99·995
74	40·4473			100·000	100·000	—	—
76	44·7012			100·000	100·000	100·000	99·999
78	49·4025			100·000	100·000	—	—
80	54·5982			100·000	100·000	100·000	100·000
82	60·3403			100·000	100·000	—	—
84	66·6863			100·000	100·000	100·000	100·000
86	73·6998			100·000	100·000	—	—
88	81·4509			100·000	100·000	100·000	100·000
90	90·0171			100·000	100·000	—	—
92	99·4843			100·000	100·000	100·000	100·000
etc. up to						...	...
184	9897·1291					100·000	100·000
Mean		3·487		3·487		3·487	
Standard deviation		1·418		1·605		2·236	
Gini coefficient		23·31%		25·63%		32·18%	

the salary at step h by:

$$Y(h) = (1+j)^h,$$

the total salaries payable up to step H by:

$$T(H) = \sum_{h=0}^{H} Y(h)\cdot f(h),$$

and the proportion of salaries up to step H by:

$$U(H) = T(H)/T(N),$$

where N is the highest step on the salary ladder. The figures shown are then $100F$ and $100U$ for each H.

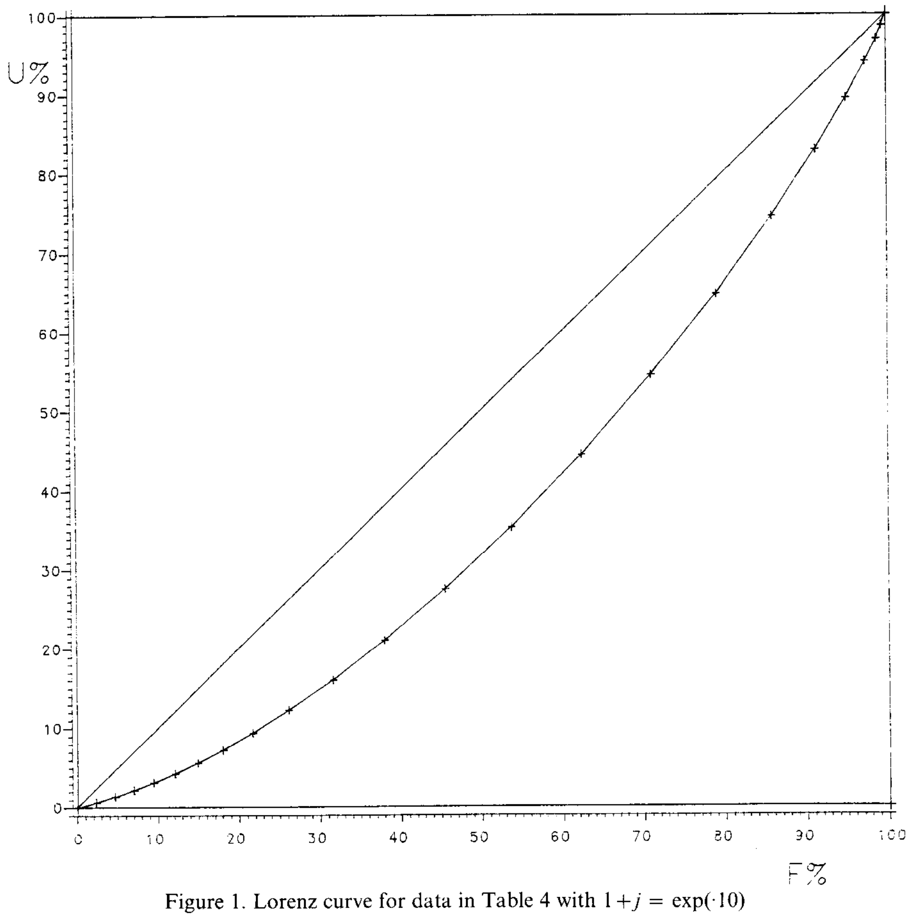

Figure 1. Lorenz curve for data in Table 4 with $1+j = \exp(\cdot 10)$

17. I now need to explain a Gini coefficient. Figure 1 shows an example of a Lorenz curve, which displays the data from Table 4 for $1+j = \exp(\cdot 10)$, with F on the horizontal and U on the vertical axis. The values of $100F(h)$ and $100U(h)$ are plotted as points for each h, and the points are then joined by straight lines. There are enough points for the resulting line to look like a fairly smooth curve, running from (0,0) to (100,100). Such a curve is called a Lorenz curve. It is a suitable way to represent the distribution of any variable that takes only non-negative values, or for comparing the cumulative distributions of two variables.

18. The Gini coefficient is given by the ratio of the area of the segment lying between the curve and the straight line joining (0,0) and (100,100) to the area of the triangle (0,0), (100,0), (100,100). Or, if we scale down so that F and U run from 0 to 1, the Gini coefficient, G, is given by:

$$G = 1-2A,$$

where A is the area under the curve, bounded by the F axis and the line $U = 1$.

19. Another way of calculating the same number is as follows. Consider every pair of values (x_1, x_2) in the distribution of X. The absolute distance between them is $d = |x_1-x_2|$. The expected value of d is called $D = E(d)$. Then, provided X takes only non-negative values:

$$G = D/2E(X).$$

The proof of this is not difficult, and is omitted.

20. The Gini coefficient gives a useful measure of the equality or inequality of a distribution, such as that of incomes or wealth, which is the field in which it is most often used. If all salaries were equal the Lorenz curve would lie along the straight line (0,0) to (100,100) and the Gini coefficient would be zero. As incomes become more unequally distributed the curve moves down to the right, and at the extreme, when everyone has nothing except for one who has everything, the curve follows the two sides of the triangle and the Gini coefficient is 100%. A statement such that x% of the population (small) owns y% of the wealth (large) gives a single point on the Lorenz curve at $(100-x, 100-y)$.

21. It may be of interest to compare the Gini coefficient of some actual distribution of salaries with the theoretical ones derived here. For example, the Gini coefficient of the distribution of *total* taxable incomes (including therefore investment income) in the United Kingdom, 1977–78, given in Inland Revenue Statistics 1980 (the table has been dropped from later editions), allows one to calculate an approximate Gini coefficient of 32·86%. My approximation, which assumes points joined by straight lines, is necessarily an underestimate of the true value, though the error is very small. This value is very close to that of my distribution with $1+j = \exp(\cdot 20)$. Other figures relating to the total population, such as those shown by the New Earnings Survey and discussed and illustrated in Report No. 8 of the Diamond Commission (1979), show that, on average, the incomes of the total population do not rise as steeply as in my specimen salary

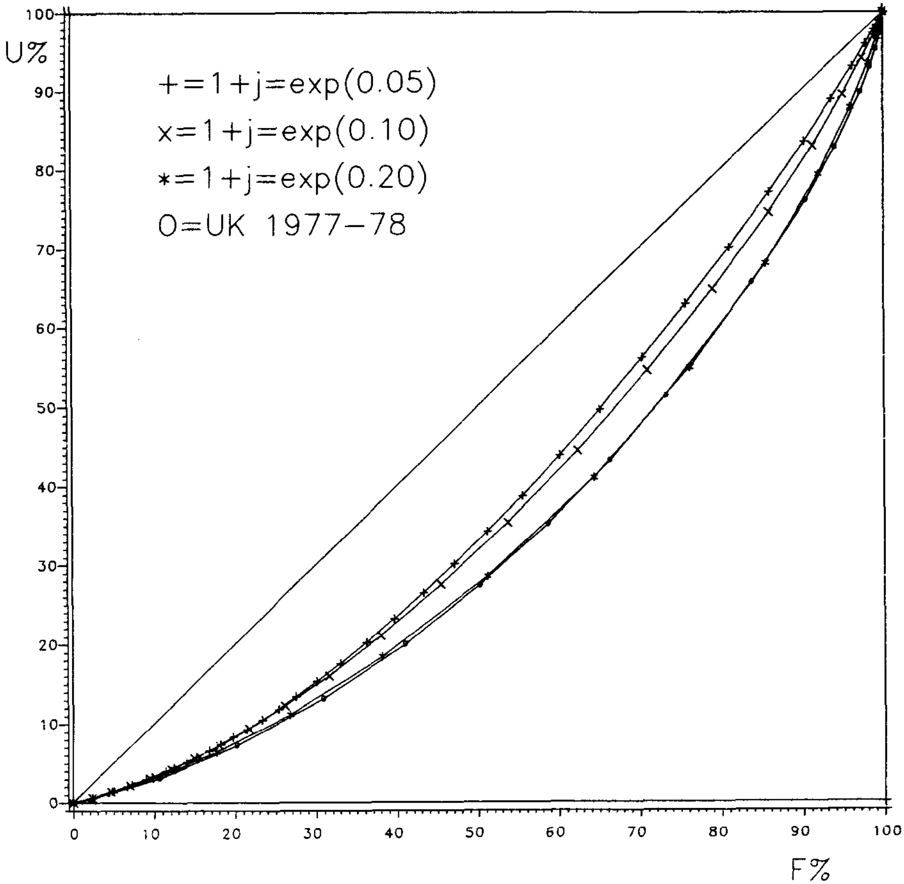

Figure 2. Lorenz curves for all data in Table 4 and also for U.K. incomes 1977–78 (see text).

scale, and indeed generally fall somewhat with age. There are many possible reasons for this, such as the growth of new, higher paid industries employing younger people and the decline of older, lower paid ones employing older ones. One does not conclude from this that there is a falling salary scale within any one firm. However, it is possible that the specimen salary scale I have used continues to rise after about age 40 more than is appropriate for many firms, even though one of my distributions is very similar to that for total incomes for the whole population.

22. Figure 2 shows Lorenz curves for all three of my specimen distributions, along with that referred to above for incomes in the U.K. 1977–78. It can be seen

how close this last distribution is to the specimen distribution with $1+j = \exp(\cdot 20)$.

23. Without using information about the process of salary dispersion in a particular case I can go no further. It would be an interesting exercise to investigate the actual salary increments in some large employer, in order to estimate an actual dispersion pattern.

REFERENCES

BACON, J. (1907) A Pension Fund Problem; with some remarks on the deduction of salary-scales. *J.I.A.* **42,** 36.

CRABBE, R. J. W. & POYSER, C. A. (1953) *Pension and Widows' and Orphans' Funds.* Cambridge.

DIAMOND, LORD, *Chairman* (1979) Royal Commission on the Distribution of Income and Wealth, Report No. 8, Fifth Report on the Standing Reference, Cmnd, 7679. H.M.S.O.

FORMULAE AND TABLES FOR ACTUARIAL EXAMINATIONS (1980). Institute of Actuaries and Faculty of Actuaries.

HEYWOOD, G. & MARPLES, W. F. (1950) Pension Funds: Some Practical Points. *T.F.A.* **19,** 105.

HOOKER, P. F. & LONGLEY-COOK, L. H. (1957) *Life and Other Contingencies*, Vol. II. Cambridge.

INLAND REVENUE STATISTICS 1980, H.M.S.O.

JORDAN, C. W. (1952) *Life Contingencies.* Society of Actuaries.

KING, G. (1905) On Staff Pension Funds. *J.I.A.* **39,** 129.

LEE, E. M. (1973) *An Introduction to Pension Funds.* Institute of Actuaries and Faculty of Actuaries.

MANLY, H. W. (1901) On the Valuation of Staff Pension Funds. *J.I.A.* **36,** 209.

MANLY, H. W. (1902) On the Valuation of Staff Pension Funds, Additions and Alterations. *J.I.A.* **37,** 193.

MANLY, H. W. (1903) On the Valuation of Staff Pension Funds, Part II—Widows' and Children's Pensions. *J.I.A.* **38,** 101.

MANLY, H. W. (1911) On Staff Pension Funds: The Progress of the Accumulation of the Funds; The Identity of a Valuation with the Future Progress of a Fund; The Manner of dealing with Funds which are insolvent; and Sundry Observations. *J.I.A.* **45,** 149.

MARPLES, W. F. (1948) The Effect of Changed Economic Conditions and State Insurance of Private Pension Fund Benefits, Contributions and Valuations. *J.I.A.* **74,** 212.

MCGOWAN, J. (1901) An Enquiry into the Methods and General Principles to be adopted in calculating the Rates of Contribution, or Percentage Deductions from Salary, in the case of Superannuation Funds and Pension Funds generally. *J.I.A.* **37,** 15.

M'LAUCHLAN, J. J. (1908) The Fundamental Principles of Pension Funds. *T.F.A.* **4,** 195.

M'LAUCHLAN, J. J. (1914) On the Construction of Salary Scales for use in the Valuation of Pension Funds, and the Valuation of Prospective Pensions based on Accumulated Contributions. *T.F.A.* **7,** 215.

NEILL, A. (1977) *Life Contingencies.* Heinemann.

PORTEOUS, D.A. (1936) *Pension and Widows' and Orphans' Funds.* Cambridge.

SPURGEON, E. F. (1922) *Life Contingencies.* Cambridge.

THOMAS, E. C. (1913) On the Valuation of Benefits dependent on Promotion to a Higher Status. *J.I.A.* **48,** 217.