European Economic Review 11 (1978) 387-393. © North-Holland Publishing Company

THE INDIVIDUAL WELFARE FUNCTION OF INCOME

A Lognormal Distribution Function?

Berend WIERENGA*

Agricultural University of Wageningen, Wageningen, The Netherlands

Received October 1977, final version received April 1978

The procedure used by Van Praag and Kapteyn to test the theory that the individual welfare function of income is a lognormal distribution function is critically evaluated. It is shown that random data give the same test results as the income levels actually collected from real consumers. Thus it is doubtful whether Van Praag and Kapteyn's results, even although they are based on data collected from as many as 12000 consumers, provide any support for the lognormal model. It is concluded that more powerful methods are required to test the lognormal model thoroughly.

1. Introduction

Van Praag (1968) has developed a very ingenious theory with the major implication that the individual welfare function is a lognormal distribution function (in this article this is further to be referred to as 'the lognormal model'). Using this model a number of interesting questions can be answered, such as:

- If person A's welfare function of income is a certain amount lower than that of person B, how much higher should A's income be if A and B are to share the same welfare level?
- If a family is to be kept at the same welfare level, how much income compensation should it receive as the family size increases?

Van Praag and Kapteyn (1973) discuss these and other applications of the model.

After the theory had been developed, Van Praag and Kapteyn collected income evaluation data from 6 different samples of consumers in Belgium and the Netherlands (more than 12000 persons in total) for an empirical test of the lognormal model. Two reports on this research appeared in this Journal: Van Praag (1971) and Van Praag and Kapteyn (1973); a comprehensive presentation of results is given in Kapteyn's recent dissertation (1977,

*The author wishes to thank Arie Kapteyn for his comments on an earlier version of this article, and Jaap Bijkerk for writing the computer program for the simulation study.

ch. III). On the basis of the very good fit of the model in this test, for all cases, the authors concluded that the shape of the individual welfare function is approximately lognormal.

Unfortunately however, from a simulation study it appears that income data randomly drawn by a computer, without any income evaluation notion, produce a fit to the lognormal model in Van Praag and Kapteyn's test that is as good as the fit obtained for the income evaluation data collected from real consumers. The implication is that the empirical data used in Van Praag and Kapteyn's research, although perhaps interesting in themselves, constitute little if any support for the lognormal model.

This article describes the simulation study and its results, discusses the major weakness in Van Praag and Kapteyn's testing procedure, and also suggests some more powerful ways of testing the lognormal model.

2. Van Praag and Kapteyn's testing procedure

The methods of data collection and analysis are fully described by Kapteyn (1977, pp. 59-72). They are summarized briefly here. Respondents in a sample of consumers were asked to provide actual values (in Bfrs. or Dfl., respectively) for eight income levels, qualitatively described in a questionnaire. The first level was the income level above which the consumer would call his income 'excellent'; the second level was the income level above which the income would be called 'good', and so on. The lowest level was that below which the respondent would call his income 'very bad'. The complete list of qualifications in the questionnaire was: 'excellent', 'good', 'amply sufficient', 'sufficient', 'barely sufficient', 'insufficient', 'very insufficient', 'bad' and 'very bad'. Van Praag and Kapteyn assumed that a consumer would supply values for these eight income levels in such a way that the corresponding income intervals would correspond to equal quantities of his welfare function. Given this assumption and the lognormal distribution hypothesis, for the eight income levels of a particular consumer,

$$\ln y_i = \mu + \sigma w_i, \qquad i = 1, \ldots, 8. \tag{1}$$

Here y_1 to y_8 are (from low to high) the income levels given by the respondent and w_1 to w_8 are the points of the normal distribution with p-levels: 88.9%, 77.8%, ..., 11.1%, respectively. The parameters of the normal distribution, μ and σ, are to be estimated. By applying least squares to eq. (1), μ and σ can be found and the fit of the lognormal model can be estimated for every individual consumer. The universal good fit [the average multiple correlation coefficient R (unsquared) ranged from 0.97 to 0.98] led Van Praag and Kapteyn to conclude that these results supported the lognormal model.

Now it is clear that with this data collection procedure one will get from every respondent a series of eight increasing income levels (looked upon from 'very bad' upwards: the respondent provides them in reverse order). The presumption is then that whether or not they reflect the income evaluation of the respondent in the way expected by Van Praag and Kapteyn, almost all of such series of income levels will correlate strongly with the series w_1 to w_8 (which also increases monotonically). This was examined in the simulation study described in the next section.

3. The simulation study

In this study the computer generated pseudo-respondents each with eight different income levels. These income levels were drawn randomly from the range Dfl. 15000-50000. After rearranging them in increasing order, these income data were analysed in the same way as Van Praag and Kapteyn analysed their income evaluation data from real consumers.

The income levels generated were rounded off to multiples of Dfl. 1000. We stipulated that successive income levels should be at least Dfl. 1000 apart; thus if two income levels happened to coincide, one of them was increased by Dfl. 1000.

The limits for the income interval were determined as follows. We wanted to compare the simulation results with the results of Van Praag and Kapteyn's most recent studies for the Netherlands which refer to 1974 and 1975. The lower limit of the income range (Dfl. 15000) approximates to the 1974/75 net income level (after tax and social security deductions) of a family with two children, earning the government-imposed minimum wage. The upper limit (Dfl. 50000) approximates to the 1974/75 net income level of a Dutch government official (also two children) being paid according to the highest rank in the officially published government salary scales. Thus in 1974/75 there were few families earning less than Dfl. 15000 or more than Dfl. 50000.[1]

In Van Praag and Kapteyn's data not every respondent provided values for all eight income levels requested. Accordingly, for every pseudo-consumer we simulated the number of income levels to be used in the analysis: 8 ($p=0.70$), 7 ($p=0.06$), 6 ($p=0.06$), 5 ($p=0.06$), 4 ($p=0.06$) or 3 ($p=0.06$). The probability distribution used here is an approximation of the last column in Kapteyn's Table 1 (1977, p. 74). For pseudo-consumers with income levels missing we randomly determined which levels were omitted from the analysis. For the simulation a sample size of 1000 pseudo-consumers was used.[2] The following quantities were computed for every pseudo-consumer:

[1]However, in a simulation run where the income interval was Dfl. 5000-50000 this different range practically did not affect the fit of eq. (1) to the simulated data.

[2]It transpired that with this sample size a different starting number in the random number generating process had virtually no effect on the numerical results.

Table 1

Comparison of the fit of the lognormal distribution function for random and real consumer income data.

	(1) Income levels randomly drawn	(2) Income levels randomly drawn, difference between successive levels ≤ Dfl. 7500	(3) Income levels provided by real Dutch consumers (1974)[a]	(4) Income levels provided by real Dutch consumers (1975)[a]
Average correl. coefficient: R	0.96(0.03)[b]	0.97(0.02)	0.97(0.04)	0.98(0.02)
Average value of $\hat{\mu}$	10.34(0.12)	10.33(0.12)	9.78(0.36)	9.75(0.34)
Average standard error of $\hat{\mu}$	0.04(0.02)	0.03(0.01)	0.03(0.03)	0.02(0.02)
Average value of $\hat{\sigma}$	0.39(0.10)	0.39(0.09)	0.43(0.03)	0.40(0.17)
Average standard error of $\hat{\sigma}$	0.05(0.03)	0.04(0.02)	0.04(0.03)	0.03(0.02)
$\hat{\rho}_1$ (autocorrelation coeff.)	−0.05(0.36)	0.02(0.38)	0.04(0.38)	0.03(0.37)
$\hat{\rho}_2$ (autocorrelation coeff.)	−0.04(0.37)	0.03(0.39)	0.03(0.39)	0.02(0.39)
Residual variance: s^2	0.0	0.0079	0.0068	0.0030
n	1000	1000	919	1748

[a]Kapteyn (1977, p. 75).
[b]Sample standard deviations of the statistics are given in parentheses.

the multiple correlation coefficient R (unsquared), $\hat{\mu}$, $\hat{\sigma}$, standard error of $\hat{\mu}$, standard error of $\hat{\sigma}$, $\hat{\rho}_1$ and $\hat{\rho}_2$. They correspond with the quantities given by Kapteyn in his Table 2 (1977, p. 75). Here $\hat{\rho}_1$ and $\hat{\rho}_2$, Kapteyn (1977, pp. 71-72), are estimated autocorrelation coefficients used to test the linearity of the model. For all pseudo-consumers in the sample, averages for these quantities were computed. These are presented in column 1 of table 1.

In the procedure for drawing incomes randomly, described above, differences between successive income levels can become unrealistically large. For example, one of the pseudo-respondents had Dfl. 15000 as his lowest income level (which would correspond to 'very bad' in Van Praag and Kapteyn's questionnaire), while his next lowest level (corresponding to 'bad') was Dfl. 27000. Another pseudo-consumer has as his highest two levels: Dfl. 36000 ('good') and Dfl. 49000 ('excellent'). Since such differences are very unlikely in practice, as well as carrying out a simulation in which there was no upper limit for the intervals between successive income levels, we did a simulation in which the maximum permissible limit for these intervals was Dfl. 7500. The results for this simulation are given in column 2 of table 1.

4. Discussion of the results

Table 1 presents the results from the random income data next to the corresponding statistics obtained by Van Praag and Kapteyn in their two most recent studies for the Netherlands [Kapteyn (1977, table 2, p. 75)]. The samples used in the latter studies are described by Kapteyn (p. 84).[3] Considering $\bar{R}$, it appears that the fit of the lognormal model is practically as good for the simulated data as for the real data. This is also reflected in the very small standard deviations of $\hat{\mu}$ and $\hat{\sigma}$ in the simulated data which are in the same order of magnitude as in the real data. The estimated autocorrelation coefficients are also of the same order of magnitude (absolute values). The values of $\hat{\mu}$ and $\hat{\sigma}$ do not measure the fit of the lognormal model, but indicate how well the income interval chosen reflects reality. By comparing their values in the simulation with those from the real study this appears to be satisfactory.

The somewhat larger residual variance in the simulated data compared with the real data is due to the assumed homogeneity of consumers in the simulation, i.e., with respect to the interval from which the incomes are drawn.[4] In the simulation study every pseudo-consumer was assumed to have the whole range from Dfl. 15000–50000 as his reference income range, but in reality some consumers would have a reference range of, for example, Dfl. 15000–35000 (such consumers consider Dfl. 35000 to be an excellent income): other consumers would have a reference range from Dfl. 30000–50000, and so on. For such individuals the variance of the dependent variable in the regression equation (1), and generally s^2 too, will be smaller than when the reference income range is Dfl. 15000–50000. In a simulation run with a mixture of 500 consumers drawing their income levels from the interval Dfl. 15000–35000, and 500 consumers drawing their income levels from the interval Dfl. 30000–50000, s^2 turned out to be as low as 0.0033.

The uniform drawing interval also results in a relatively small sample standard deviation of $\hat{\mu}$ in the simulation study: 0.12 compared with 0.36 and 0.34 in the real data. In the 'mixed' simulation just mentioned this sample standard deviation increased to 0.25.

In spite of the very unrealistic jumps in income levels that occur, the fit of the lognormal model in column 1 of table 1 is very good. Column 2 of the table shows that the fit improves even more when the most extreme differences are removed. An additional simulation run showed that when the differences between successive income levels are further restricted, i.e., to Dfl. 5000, $\bar{R}$ becomes as high as 0.98. Since this type of restriction has the effect

[3] For reasons of conciseness only these two studies by Van Praag and Kapteyn are referred to here. However, the results are typical for their complete set of six studies.

[4] Since the left-hand values are different, s^2 is not the appropriate measure to compare the fit for the various cases of table 1 though; see Theil (1971, p. 544).

of equalizing the income intervals, it appears that any tendency for the consumer to provide equal income intervals - which seems to be very natural in a task like this - increases $\hat{R}$ in Van Praag and Kapteyn's test. It should be noted, however, that in the lognormal model equal income intervals generally do not correspond with equal utility intervals, that are the basis of Van Praag and Kapteyn's test. So a high $\hat{R}$ can obtained under conditions that are very divergent from those assumed by Van Praag and Kapteyn. Since the expected income intervals are equal, also in the simulation of the unrestricted model the assumption of equal utility intervals is violated, though.

In Van Praag (1971) and Van Praag and Kapteyn (1973) the fit of the lognormal model is judged on the basis of the quantities R, $\hat{\rho}_1$ and $\hat{\rho}_2$ given in table 1. Moreover, Kapteyn (1977) has developed a procedure to estimate how much of the residual variance s^2 is caused by rounding-off errors (indicated as: $\frac{1}{3}c^2$) and how much is accounted for by other types of error, including specification errors. It is favorable for the lognormal model when the latter fraction is low. This appeared to be the case in the two Dutch studies of 1974 and 1975, where Kapteyn estimated this fraction to be 0.06 and 0.07, respectively.

We did not have at our disposal the numerical procedure for estimating $\frac{1}{3}c^2$, but thanks to Kapteyn's kind cooperation we were able to process the income levels of 100 of our pseudo-consumers using his method. It appeared that these random data also easily passed this specification test: the part of s^2 due to specification and other errors was estimated to be as low as 0.04.

5. Conclusion

It appears that in the procedure used by Van Praag and Kapteyn to test the lognormality hypothesis random data produce the same fit as the income data actually collected from real consumers. If the consumers that were interviewed would have responded by dividing their reference income range randomly into nine intervals, the resulting $\hat{R}$ would have been of the same order of magnitude as the values actually found by Van Praag and Kapteyn. Any tendency for the consumer to divide the income range into intervals of approximately equal length would further increase $\hat{R}$. Such random data clearly violate the assumption of equal utility intervals which is the basis of Van Praag and Kapteyn's test.

As a consequence the results, reported with respect to the fit of the lognormal model obtained by this test, even although they are based on the data from more than 12000 consumers provide no substantial support for this model.

On the other hand, the simulation results reported in this paper do not imply that Van Praag and Kapteyn's results are at variance with the lognormal model. A test in which random data produce a multiple correlation coefficient of 0.97 simply cannot be expected to provide conclusive evidence either to support or to refute a model. Furthermore it is possible that the assumption of equal utility intervals in the income data provided by consumers is wrong. This would remove the rationale behind Van Praag and Kapteyn's test and make their results irrelevant with respect to the question whether or not the lognormal model is correct. Van Praag and Kapteyn did not test this assumption. They merely show that equal utility intervals will occur when the respondent minimizes average inaccuracy (as defined by the authors). The assumption implies that the income levels mentioned by the respondent are independent from the specific verbal statements (such as 'excellent', 'good', etc.) in the questionnaire. This is not very plausible: it seems unlikely that a respondent would state the same income levels if a different set of nine verbal qualifications would be used.

If it is at all possible to test the lognormal model using statements made by respondents in a survey, a more elaborate research procedure is required. The effort should be directed towards obtaining more information from each respondent, instead of to increase the number of respondents in the sample to thousands of individuals. For example, it should be ascertained whether a person is consistent with the model when evaluating 'new' income levels (i.e., levels not used to estimate his welfare function parameters) and when evaluating and comparing income differences. The sensitivity of the estimates to different verbal statements in the questionnaire needs to be checked and it should be examined if a person produces the same welfare functions when interviewed on different occasions. Of course, this requires a more complicated interviewing: it would be difficult to collect all this information by means of a postal survey. A model that makes such specific assumptions about individual welfare evaluation requires specific testing procedures to test its validity.

References

Kapteyn, A., 1977, A theory of preference formation (Pasmans, The Hague).

Theil, H., 1971, Principles of econometrics (Wiley, New York).

Van Praag, B.M.S., 1968, Individual welfare functions and consumer behavior (North-Holland, Amsterdam).

Van Praag, B.M.S., 1971, The welfare function of income in Belgium: An empirical investigation, European Economic Review, Spring, 337–369.

Van Praag, B.M.S. and A. Kapteyn, 1973, Further evidence on the individual welfare function of income: An empirical investigation in the Netherlands, European Economic Review 4, 33–62.