Int. Stat. Rev., Vol. 43, No. 2, 1975, pp. 211–218/*Longman Group Ltd/Printed in Great Britain*

Engel's Law Revisited

D. Perthel
Institute of Socio-economic Studies of Developing Regions, University of Amsterdam

Summary

The purpose of our study is to investigate how Ernst Engel acquired the data of Table 8, in his original article of 1857, where he gave quantitative material to support his qualitatively stated law elsewhere. In the introduction (section 1) some comments are given on reviews on Engel's article, whereas in section 2 the problem will be outlined. There is a problem because the data of Engel's Table 8, when plotted on a double-log graph paper, form a straight line. In section 3 some possible methods used by Engel will be investigated, with successively increasing plausibility. We were able to obtain strong evidence that Engel used very simple arithmetics to set up his range, which is, by genuine intuition, a very good approximation of his rough material (the 235 budget data collected by Ducpetiaux and Le Play).

1. Introduction

During the time that Ernst Engel was director of the Statistical Office of Saxony he published an article in 1857 entitled: "The relations of production and consumption in the Kingdom of Saxony" [1]. On pages 28 and 29[1] he wrote: "The poorer a family, the greater the part of total expenditures must be spent on food". By this sentence, later referred to as the *Law of Engel*, he obtained immortal reputation among students of economics. Engel derived his empirical law from family budget data collected by Ducpetiaux [2] and Le Play [3]. Ducpetiaux collected 199 family budgets of Belgian workmen and Le Play 36 budgets of workmen all over Europe. As Engel put it, they delivered the pearls but not the string (p. 8).

Later his law has been so often verified with other data, from rich as well as from poor countries, that it could never be our intention to cast doubts on Engel's law. No, the objective of our study is to pay attention to Engel's Table 8 (pp. 30 and 31), where he gave levels of income and the corresponding shares of food expenditure, which must support his law quantitatively. Our problem is: how did Engel arrive at these data, or rather, at the functional relationship? Even after the reprint (38 years later) Engel's findings did not impress economists immediately. H. Higgs, discussing the economic condition of the people in 1899, mentioned Le Play, who, in Higg's words, is an important writer for ascertaining the facts of consumption. Especially Higgs drew attention to the Le Play's family budgets, the receipts and expenses of four English families (out of the 36 families) and concluded that Le Play may fairly be called the father of the scientific family budgets [4].

This little attention to Engel's work was also noticed by J. A. Schumpeter. He wrote that Engel's law was not recognized as an important contribution to economic theory [5]. It seems that real interest in the law began in the first quarter of the twentieth century (although Marshall referred to the work of Engel already in the second edition of his "Principles" in 1891. [6].)

Of course, in the years after Engel's publication in 1857 many articles and comments appeared on his findings and with them many misinterpretations. The critique on Engel's work did not concern *the* law but many other results.

[1] In the following, every page indication, without further note, refers to Engel's original article as reprinted in 1895.

In his excellent survey of Engel's law, C. C. Zimmerman tried to disentangle all these opinions and interpretations, and not without success [7]. The contribution of Engel to economic theory consists of the discovery of a pattern in consumers behaviour, which remains valid for the nineteenth as well as the twentieth century.

2. The Problem

It is customary for reviewers of Engel's law to refer to two tables from his study, from which the relationship between penury and food expenditure is easily seen. C. C. Zimmerman mentioned Table 7 (p. 30), which table was reproduced by Marshall, and G. J. Stigler in his article "The early history of empirical studies of consumer behaviour" even reproduced Table 6 (p. 27) [8]. To our knowledge only J. van der Wijk gave some attention to Engel's Table 8, of which he remarked that it was an econometric regression estimation in tabulated form [9].

We should like to reproduce Engel's Table 8 entirely (see Table I, columns 1 and 2).

Table I. *Annual income and share of food expenditure*

Annual income of a family, in francs	Food expenditure	
	in percentages	in francs
1	2	3
200	72·96	145·92
300	71·48	214·44
400	70·11	280·44
500	68·85	344·25
600	67·70	406·20
700	66·65	466·55
800	65·69	525·52
900	64·81	583·29
1,000	64·00	640·00
1,100	63·25	695·75
1,200	62·55	750·60
1,300	61·90	804·70
1,400	61·30	858·20
1,500	60·75	911·25
1,600	60·25	964·00
1,700	59·79	1,016·43
1,800	59·37	1,068·66
1,900	58·99	1,120·81
2,000	58·65	1,173·00
2,100	58·35	1,225·35
2,200	58·08	1,277·76
2,300	57·84	1,330·32
2,400	57·63	1,383·12
2,500	57·45	1,436·25
2,600	57·30	1,489·80
2,700	57·17	1,543·59
2,800	57·06	1,597·68
2,900	56·97	1,652·13
3,000	56·90	1,707·00

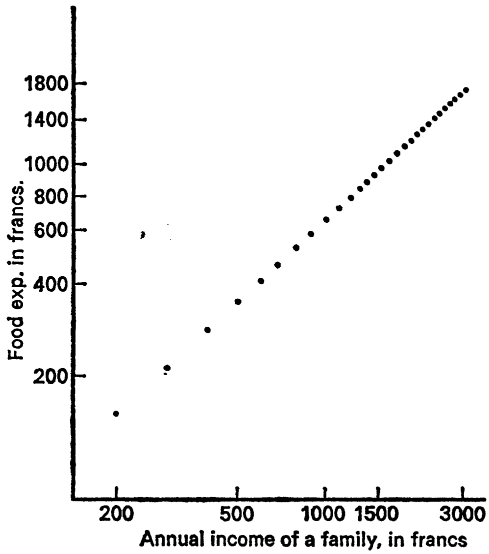
Source: Columns 1 and 2: E. Engel, Table 8, pp. 30, 31.
Column 3 Our computations.

Looking at the columns 1 and 2, the relationship between income and food expenditure as percentage of income is easily recognized and thus Engel tried to obtain some quantitative evidence for his law.

When estimating consumption functions it is common to give both variables in the same kind of units, so we added column 3. Plotting the 29 observations in a double log scaled graph leads to a straight line (see graph 1). We estimated this equation by least squares with the following result:

$$\log C_f = 0{\cdot}1149 + 0{\cdot}896 \log Y \; (R^2 = 0{\cdot}9998). \tag{1}$$

C_f stands for food expenditures and Y for income. This function is highly significant and one must wonder how Engel obtained the 29 observations in his own table. It is difficult to accept the very high correlation coefficient for any consumption function, computed out of budget data, without any further compilation of the rough data. Unfortunately Engel did not mention his method and it seems to us, one may call it idle curiosity, "dogmengeschichtlich" interesting to find the method.



If he could have been acquainted with least-squares estimation it would be simple to compute the regression line of all 235 observations, using subsequently the found coefficients for interpolation at the income levels as given in Table I. Regression over all 235 observations leads to:

$$\log C_f = 0{\cdot}220 + 0{\cdot}862 \log Y \; (R^2 = 0{\cdot}764), \tag{2}$$

which does not differ too much from equation (1).

Grouping the 235 data into seven classes of expenditures (see Table II), we obtained the following double log equation:

$$\log C_f = 0{\cdot}1227 + 0{\cdot}896 \log Y \; (R^2 = 0{\cdot}998), \tag{3}$$

Table II. *Ducpetiaux's and le Play's original data grouped into seven expenditure classes*

Expenditure classes in francs 1	Average expenditure in francs 2	Average expenditure on food, in francs 3
300–500	438·6	297·8
500–700	603·0	418·4
700–900	799·6	540·0
900–1,100	969·7	641·8
1,100–1,300	1,193·1	752·7
1,300–1,500	1,391·7	853·6
1,500 and more	2,039·7	1,209·7

which fits in very well with equation (1). Unfortunately, Engel could not be familiar with the correlation technique. From Pearson's study on the history of correlation it appears that the honour of the invention of this technique must be given to F. Galton, who formulated it some 30 years after Engel published his article [10].

We computed the equations (2) and (3) to show that Engel did a remarkably good job in constructing the table, without using the more sophisticated technique.

In the following section we want to discuss three possible methods out of which we hope to be able to prove which one Engel used.

3. Possible Methods

(*a*) Could it be possible that Engel drew a straight line through two points and determined the other points graphically?

He might have drawn up a diagram. Graphical representation of data was known long before his time. The most commonly used method, however, was to set out an observation against time. One of the first users, in the economic sciences, was A. F. W. Crome, who held a chair in statistics and public finance from 1786 to 1831 in Germany [11]. It is hardly possible that Engel was not familiar with Crome's work. A further step of course was for Engel to transform both axes into a logarithmetic scale. For Engel, as a statistician, this should not have seemed to be too difficult. Two facts weigh heavily against this explanation. First, which observations did he use? In Table 4 (pp. 24 and 25) he gave a summary of Ducpetiaux's material. He broke down the annual incomes into three categories and for each category he computed the average expenditure on food. However, these three observations give a quite different line. The scope in the doublelog equation, equal to the income elasticity of food expenditures, decreased from 0·896 (equation (1)) to 0·667. Taking Le Play's summarized data into account (Table 2 p. 19), the elasticity became 0·927. Second, he remarked (p. 31) that it was not possible for him to construct the exact mathematical formula, which reflects the relation well. However, the formula of a straight line is very exact. Moreover, on the same page he added to this that he was not able to compute the tails of his range.

Out of the rough data of Ducpetiaux and Le Play, he could never arrive at the class averages as given in Table 1. There is no observation in the basic material, which lies below the 400 francs expenditure level and no observation above 2,900 francs. So some observations he certainly must have obtained through extrapolation.

(*b*) Engel's whole study was dealing with the population problem. He started his paper with a section on the law of population density. In the following sections he presented some material to point out that absolute size of the population of a country or region was unimportant if the distribution of labourers among industries is proportional to the distribution of consumer expenditures (p. 50). So with his paper he wanted to set the rather pessimistic view of Malthus in a more unbiased light. In his whole study Engel was considerably influenced by Malthus, not so much in the final outcome, as especially in the chain of reasoning. Let us start with the data as Engel grouped them (pp. 24, 25 and 27).

On page 30 he stated: "The law which is hidden in the material is not an easy one. The level of food expenditure increases with a geometric rate to decreasing income. Because it is permissable to use the small number of categories (see Table III[1]) as the basis for the computations of other categories, the data of the following table (see Table I, columns 1 and 2[1]) reflect the law, although it is not possible to put the law in an exact mathematical expression."

Returning to Table I, it is seen that the level of income increases with a constant amount of 100 francs, whereas the share of food expenditure decreases certainly not at an arithmetic

[1] Author's tables.

rate, but probably at a geometric rate. From Table III we tried to compute Table I, with the geometric progression in our mind. First of all we needed the common ratio of the series. We took the data of the third category. Because income is 1,197·77 francs we multiplied the share of food expenditure (62·42 per cent) with the factor 1,200/1197·77 to get the share for the 1,200 income level. The percentage rose to 62·54 per cent (Table I: 62·55 per cent). To get a second observation we multiplied the average share of all families (65·83 per cent) with 900/913·95, which resulted in 64·82 per cent for the 900 francs income level (table I: 64·81 per cent), both very near Engel's observation, but not exactly equal. The used method is not quite consistent, because income and expenditure are mixed up. To arrive at the 1,200 income level we used the income of 1,197·77, which is correct. To get the second observation, however, the only means was to use the expenditure level of the average family to find the closest approximation of the share of food expenditure, corresponding to any income level out of Table I.

Table III. *Observations of Ducpetiaux, grouped by Engel in three categories*

Categories a	Income in francs 1	Expenditures in francs 2	Percentage of food expenditure of every 100 franc spent 3
1	564·97	648·48	70·89
2	796·71	845·44	67·37
3	1,197·77	1,214·44	62·42
Average of all families	865·56	913·95	65·83

Source: Columns 1 and 2 Engel, Table 4, pp. 24 and 25.
Column 3 Engel, Table 6, p. 27.

Finally, we obtained two observations which Engel certainly used to construct Table I. It must be noted that Engel did not employ Le Play's data. He classified these not according to levels of income or expenditure but according to the nationalities. Le Play's data gave only more qualitative evidence to his final result, and will play no role in our further argument.

Now, assuming that, 62·54 per cent and 64·82 per cent at income levels of respectively 1,200 and 900 francs are two elements of a geometric range of the 29 observations it will be an easy job to compute all elements. In Table I the 1,200 income level is the 19th and the 900 income level is the 22nd observation. Let a be the initial value (starting at the 3,000 income level) and r the common ratio, we got two equations with two unknowns. We got $a = 50{\cdot}4$ per cent and $r = 1{\cdot}012$. The final value, at the 200 income level, became 70·4 per cent. Only between the 900 and 1,200 level of income the percentages became very close to Engel's. Moving away from these points the differences increased.

This method, assuming a geometric series, seemed to us a plausible explanation and it is not quite understandable why Engel did not use it.

Hereafter we shall try to ascertain which method Engel did use.

(*c*) We noted already that it is for us beyond any doubt that Engel did use the two percentages 62·54 and 64·82, respectively. Because the slight difference cannot be attributed to a simple error in arithmetic, with such a serious statistician as Engel was, there must be another reason for it. Looking at Table IV, especially at the columns 2 and 3, we notice an astonishing regularity in these numbers. But this is not the regularity of a geometric series.

It does not surprise us now that Engel talked about a difficult mathematical expression which could not be ascertained. At the 900 and 1,200 income levels we have put our computed percentages between brackets. The absolute difference between them is 2·28, which must be the sum of an increasing range, at a geometric rate. Our sum could be subdivided into $0{\cdot}71+0{\cdot}76+0{\cdot}81 = 2{\cdot}28$ (the summation of the first differences). However, the second differences remain the same, 0·05, whereas Engel assumed increasing second differences, too,

Table IV. *The shares of food expenditure at different level of income in percentages (see columns 1 and 2 of Table I) and the first and second differences of these percentages (own computations)*

Annual income of a family in francs 1	Food expenditure in percentage 2	First difference 3	Second difference 4
200	72·96	1·48	—
300	71·48	1·37	0·11
400	70·11	1·26	0·11
500	68·85	1·15	0·11
600	67·70	1·05	0·10
700	66·65	0·96	0·09
800	65·69	0·88	0·08
900	64·81 (64·82)	0·81	0·07
1,000	64·00	0·75	0·06
1,100	63·25	0·70	0·05
1,200	62·55 (62·54)	0·65	0·05
1,300	61·90	0·60	0·05
1,400	61·30	0·55	0·05
1,500	60·75	0·50	0·05
1,600	60·25	0·46	0·04
1,700	59·79	0·42	0·04
1,800	59·37	0·38	0·04
1,900	58·99	0·34	0·04
2,000	58·65	0·30	0·04
2,100	58·35	0·27	0·03
2,200	58·08	0·24	0·03
2,300	57·84	0·21	0·03
2,400	57·63	0·18	0·03
2,500	57·45	0·15	0·03
2,600	57·30	0·13	0·02
2,700	57·17	0·11	0·02
2,800	57·06	0·09	0·02
2,900	56·97	0·07	0·02
3,000	56·90		—

between these two points. So these are still not the data Engel used. If we start however, from the first difference of 0·70 we get a geometric range. The sum of the first differences,

$$0{\cdot}70+(0{\cdot}70+0{\cdot}50)+(0{\cdot}70+0{\cdot}05+0{\cdot}05+0{\cdot}01) = 2{\cdot}26.$$

Taking the nearest point to 62·54 per cent, 62·55 per cent and add to this 2·26 we get 64·81 per cent. Finally, we obtained four exactly the same percentages out of Engel's Table 8, namely 62·55 per cent, 63·25 per cent, 64 per cent and 64·81 per cent. Engel used the same principle up to the 400 income level, from whereas he added to the first difference a constant factor (0·11).

If Engel would have done the same downwards, the difference in the percentages would soon be a constant, namely 0·50. With such a constant difference, starting at the 1,600 income level, the share of food expenditure would be at the 3,000 income level $60{\cdot}25-(14\cdot 0{\cdot}5) = 53{\cdot}25$. This result is certainly as good as any, but it is quite possible that Engel thought it strange to assume one end of the range growing with a geometric progression and the other end at an arithmetic rate. He found the solution by taking for the second difference a different constant for every observation.

It is now quite understandable that he wrote, on page 31, that the progression at the tails' end is a stronger one. It is possible, he added, that at the 200 income level the share of food expenditure is not 72·96 per cent but 75 per cent (translated in our Table IV this would mean the assumed constancy of the second difference 0·11, must be abandoned) and at the 3,000 income level is the share of food expenditure not 56·9 per cent but only 55 per cent (translated in our Table IV, for example not every five observations a different constant second difference, but every three observations).

Out of Ducpetiaux' material we computed two points which are very close to Engel's, at the 900 and 1,200 francs levels of income. The 1,200 income level has an important meaning. From our Table III it is seen that only the average family in the third category can be called a well-to-do family. The expenditures are about the same height as the income. This point was a distinct key in our reasoning and so for Engel. J. Burnett in his study "A history of the cost of living" noticed that the comfort-line (in England, *ca.* 1850) came at something over £1 earnings a week, depending on the size of the family. £1, about 24 francs earnings a week, gives an annual income of about 1,248 francs, which corresponds fairly well with Engel's observation. The family budget of a skilled worker in Manchester shows that the food bill took only 60 per cent out of the earnings [12].

The expenditures came to £1.7s.11d. a week, which is about 1,750 francs a year. This is remarkable close to Engel's estimation. At the 1,700 income level he found a percentage of 59·79 per cent.

The 1,200 francs of annual income can be regarded as the comfort-line too, from where on upwards the families became savers and downwards dissavers. The expenditure level of a dissaving family can be regarded as a "true" income level, so it was for Engel quite appropriate to use the 900 expenditure level as an estimation of income. Now we can give a more economic explanation to the assumed points of inflection at both ends of Engel's Table 8. He was certainly aware that a "consumption function" for anyone does not start at zero consumption, at a zero income level. Thus, his table would be more realistic if it would start with a higher share on food expenditure than the noted 72·96 per cent. For the higher income groups a lower share might be appropriate. Noting this, the constant elasticity of 0·896 (equation (1)) over the whole range of incomes must be refuted. At both ends the line bends off, at low incomes it will bend upwards, and at the high incomes, downwards. This means that the income elasticity of food consumption is higher at low incomes than on high incomes (compare however equation (1)!).

Engel did not work with elasticities, but he was aware that, like Pareto's law on income distribution, his law, as tabulated, was only applicable for a certain range of incomes, and most likely some points on either side of the mean level of income [13].

We stated on page 212 that Engel's findings were important to economic theory, especially the consumption theory. It might be interesting to note that recently P. A. Samuelson used Engel's laws in a note on real income measures [14]. He applied Engel's laws (food expenditures are less than unity in income elasticity) together with the "Gerschenkron effect" (as a country grows, comparing it with its own past or a poorer country the growth estimation is greater if the poorer state-price weights has been used than the richer state-price weights) to define the real income index number of a country.

Concluding Remarks

To reconstruct Engel's Table 8 exactly out of the 199 data of Ducpetiaux is hardly possible. Every researcher would follow another line of thought. Two things helped us to find the very probable method Engel used. First, trying to answer the question which technique Engel could have known and second, knowing the result we could reason backwards.

It seemed to be an obvious technique in the, let us say, pre-regression time, to work with first, second and more differences, for, as is shown by G. U. Yule, the data to support King's law, has been obtained in a similar fashion [15].

Engel's special statistical feeling helped him to formulate his law in a quantitative form, although he was certainly too serious a statistician to approve of P. Dirac's remark: "It is more important to have beauty in one's equations than to have them fit experiment". [16]

Acknowledgements

The author is grateful to Professor L. J. Zimmerman for the discussions which greatly stimulated this research, and to Professor H. Neudecker and P. Porsius for their critical comments.

Résumé—see page 210.

Notes and References

[1] Engel, E. (1857). Die Productions-und Consumtionsverhältnisse des Königreichs Sachsen, in *Zeitschrift des Statistischen Büreaus des Königlich Sächsischen Ministeriums des Innern*, No. 8 and 9, pp. 1–54. It was reprinted as an appendix to Die Lebenskosten Belgischer Arbeiter Familien früher und jetzt. *Bulletin de l'Institut International de Statistique* (1895), **9**, première livraison.

[2] Ducpetiaux, E. (1855). *Budgets économiques des classes ouvrières en Belgique*. Bruxelles.

[3] Le Play, F. (1855). *Les Ouvriers Européens. Etudes sur les traveaux, la vie domestique et la condition morale des populations ouvrières de l'Europe*. Paris.

[4] Higgs, H. (1899). Some remarks on consumption. *Economic Journal*, **9**, 510. In the same journal, **1**, 759–61, Higgs reviewed "Les Budgets Comparés des cents Monographies de Familles publiées dans 'Les Ouvriers Européens' et 'Les Ouvriers des Deux Mondes', avec une introduction par M. E. Cheysson en collaboration avec M. A. Toqué", Paris, Rome, 1890, which was reprinted from the *Bulletin de "Institut International de Statistique* **5**, première livraison, 1890. So Higgs was familiar with the *Bulletin* and could easily read Engel's article five years later.

[5] Schumpeter, J. A. (1963). *History of Economic Analysis*, p. 961. London: George Allen & Unwin, Ltd.

[6] Marshall, A. (1891). *Principles of Economics* (second edition), pp. 173–174.

[7] Zimmerman, C. C. (1933). Ernst Engel's Law of Expenditures for Food. *Quarterly Journal of Economics*, **47**, 78–101.

[8] Zimmerman, C. C., *op. cit.*, p. 84; Stigler, G. J. (1954). *Journal of Political Economy*, **62**, 98; Marshall, A, *op. cit.*, p. 173.

[9] Wijk, van der, J. (1939). *Inkomens-en Vermogensverdeling* (Distribution of Incomes and Fortunes), p. 194. Haarlem.

[10] Pearson, K. (1970). Note on the history of correlation, in Pearson, E. S., Kendall, M. G., ed., *Studies in the History of Statistique and Probability*', p. 189. London: Griffin. From the study by Harter, H. Leon, The method of least squares and some alternatives – Part 1, *International Statistical Review* (1974), **42**, 147–174, however, one must put the discovery of the regression-technique some fifty years before Engel wrote his article. Especially Laplace, Legendre and Gauss might be called the pioneers in this field. So a more weighted judgment whether Engel was familiar with the technique might be appropriate. Still, Engel himself, in his article, did not give the slightest hint that he tried to estimate any equation. So we are quite convinced that, at any rate, he did not use the method of least squares.

[11] Royston, E., A note on the history of graphical presentation of data, in Pearson and Kendall, *op. cit.*, p. 174.

[12] Burnett, J. (1969). *A History of the Cost of Living*, pp. 263–264. Penguin Books.

[13] Stigler *(op. cit.*, p. 104) quotes Engel (1861) from his study *Die Getreidepreise, die Ernteerträge und die Getreidehandel im preussischen Staate*, in which he dealt with the consequences of changes of the yield on the price of wheat. He wrote: "A decrease in yield of 1 per cent causes a rise in price of $2\frac{1}{2}$ per cent, and an increase in yield of 1 per cent leads to a fall in price of 2 per cent". So, although Engel did not use the term elasticity (or, as in the example, flexibility) he was aware of the concept.

[14] Samuelson, P. A. (1974). Analytical notes on international real income measures, *Economic Journal*, p. 596.

[15] Yule, G. U. (1915). Crop production and price: a note on Gregory King's Law. *Journal of Royal Statistical Society*, p. 296.

[16] Quoted by Koestler, A. (1971). *The Act of Creation*, p. 331. Danube.