



## The Growth of Demand for New Commodities

A. D. Bain

*Journal of the Royal Statistical Society. Series A (General)*, Vol. 126, No. 2. (1963), pp. 285-299.

Stable URL:

<http://links.jstor.org/sici?sici=0035-9238%281963%29126%3A2%3C285%3ATGODFN%3E2.0.CO%3B2-5>

*Journal of the Royal Statistical Society. Series A (General)* is currently published by Royal Statistical Society.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/rss.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

## The Growth of Demand for New Commodities

By A. D. BAIN

*Corpus Christi College, Cambridge*

### SUMMARY

In this paper we consider the use of the logistic curve in analyses of the demand for new commodities and in studies of innovation. We demonstrate that the theoretical basis for the curve is weak and that the results obtained in empirical analyses have sometimes been unsatisfactory. The choice of a symmetric curve seems unjustified on theoretical grounds, and empirical evidence indicates that a positively skew curve would be preferable. We then present a cross-section analysis of the growth of British television ownership, using a model based upon the cumulative lognormal curve. The results obtained are intuitively plausible and there is no evidence to suggest that a positively skew curve provides an inadequate explanation of the data.

### 1. INTRODUCTION

IN recent years the comparative static has given way to the dynamic approach to demand analysis. The limited applicability of the former, in which the level of demand for a good observed in any period was regarded as an equilibrium level determined by external factors, has been recognized. It seems adequate for well-known perishable commodities which are purchased frequently in the period of observation, but for durable or new commodities the demand observed in even relatively long time periods may not be perfectly adjusted to the prevailing external conditions. In these cases the path of adjustment from one equilibrium level to another must be investigated.

In studies of the demand for new commodities, the adjustment path from the initial to the equilibrium level of demand has usually been summarized by some trend function. Changes in economic factors, such as price and income, either cause deviations from this trend or lead to changes in the trend itself. If this procedure is to give unbiased estimates of the effects of the economic factors the trend function must be correctly specified, since only then will deviations from the trend be unsystematic.

An appropriate trend curve cannot be discovered by simply inspecting the data in any particular analysis. If economic influences have changed during the adjustment process, the observed path will consist of a series of segments of several growth curves whose parameters depend in some way on the economic factors. Thus an apparently skew curve might consist of a series of segments of symmetrical curves.

Though selection on the basis of the observed growth path is illegitimate there are two methods for choosing between curves. In the first place, we can discuss the nature of the growth process and the rationale behind particular curves. This may not lead to any satisfactory generating model: the diverse factors involved may not lend themselves to a simple statistical summary. But we may be able to restrict consideration to a class of curves with some particular characteristics. Secondly, we can examine empirical results and test models incorporating different curves in practice. The first part of this paper is concerned with choosing a class of curves, and is followed by an empirical analysis making use of one member of the chosen class.

## 2. LOGISTIC GROWTH MODELS

In this section we shall consider five empirical studies employing the logistic curve, in which the selection of this curve was justified on grounds other than its simplicity or S-shape. Roos and Von Szeliski (1939) analysed domestic automobile demand in the United States. The change in the car population in any period was the product of the number of potential new owners and the probability that an individual selected at random from this group would purchase a car. To take account of increased knowledge of the product and expansion of sales outlets this probability was assumed to be proportional to the car population. Thus

$$\Delta C = A_1 C(M - C) \quad (1)$$

where  $C$  is the car population and  $M$  the maximum ownership level.  $A_1$  and  $M$  were assumed to depend on economic factors. If economic influences were held constant, the model implies that adjustment to the maximum ownership level would take place along a logistic curve.† The validity of this derivation of the logistic curve depends, therefore, on the assumption that the probability of a potential new owner buying a car in any period is directly proportional to the car population.

In his study of television ownership in the United States, Dernburg (1958) assumed that consumers' tastes, as measured by a threshold index, were normally distributed. The mean value of the threshold index for a group of consumers was a linear combination of objective measurements on a number of variables, including the period for which television had been available. This implied a normal growth path, which Dernburg approximated by the logistic curve. However, his empirical analysis showed that a quadratic function of the period of availability variable would sometimes have been appropriate, which suggests that the logistic curve may not describe the growth path adequately. Since this justification of the curve is only as strong as the assumption of a normal taste distribution, his study adds little to the theoretical basis for the curve and provides some empirical evidence against it.

A second analysis of the growth of television ownership in the United States was carried out by Massy (1961), who argued that changes in the characteristics of the groups buying a new commodity at different stages would affect the growth process. Consequently, in his demand equation he selected the basic logistic process to represent endogenous market growth due to increases in the stock level, multiplied this by terms representing the effects of price and income, and assumed that the elasticities of the rate of growth with respect to these factors varied systematically with the stock level. His estimates implied that the growth curve for constant prices and incomes would be positively skew.

One unacceptable feature of his results was a highly significant positive price elasticity, which may reflect a mis-specification of the growth curve. In the presence of falling prices the use of a symmetric endogenous element, when a positively skew element would have been appropriate, could easily produce a positive estimate of a theoretically negative price parameter.

In a study of the growth of television ownership in the United Kingdom, Bain (1962) developed a logistic model using arguments similar to those of Roos and Von Szeliski. Economic factors and the presence or absence of a second television programme were assumed to affect the saturation level of ownership. However, a

† The differential equation of the logistic curve has the form  $dc/dt = bc(a - c)$  where  $a$  is the ultimate saturation level and  $b$  is the "rate of growth" parameter.

choice of television services appeared in some instances to lower the saturation proportion, an unlikely result which is consistent with a decline in the rate of growth parameter over the range of the curve. A second limitation of the analysis was the exclusion of the early years of television, for which a logistic growth path was clearly unsatisfactory.<sup>†</sup>

In his study of technical change and the rate of imitation Mansfield (1961) developed two models, a deterministic logistic model and a stochastic model. He expressed the probability of a firm adopting the innovation in the next period as a function of a number of variables, including the proportion which had already adopted it. Using the Taylor expansion and assuming that the coefficients of the second and higher order terms in the proportion of adopters were zero, a logistic growth path was obtained. This new rationale depends on the crucial assumption that these coefficients are zero. While Mansfield justifies the assumption for the second-order term by stating that in a test it was rejected (at the 5 per cent. level) in only 3 cases out of 12, such a high rejection rate appears (at the 5 per cent. level) to contradict his hypothesis. However, both the symmetric curve of the deterministic model and the positively skew curve of the stochastic model explained the data adequately, and the assumption may be justified as a simplification even if it is not supported by the data.

The theoretical arguments for the logistic curve used by Dernburg and Mansfield need no more discussion. Studies of the growth of television ownership in both the United Kingdom and the United States indicated that some positively skew curve might be preferable. But before abandoning the logistic curve, the basic formulation due to Roos and Von Szeliski requires further analysis.

### 3. THE ADJUSTMENT PROCESS

The demand for a new good will not usually take on its equilibrium level immediately: the community must first go through a learning process. This is true even of a perishable good sold at a price which enables it to be included regularly in the short-run consumption pattern. Three time lags can be distinguished in the learning process. There is a delay before the consumer is aware of the good's existence; a second delay before he is sufficiently familiar with it to understand fully its characteristics and potential; and a third delay before he decides to fit it into his regular consumption pattern. It is the second which provides the principal argument for the logistic curve. Knowledge of a commodity comes partly from observing its use by other people. The greater the number who use the commodity the more likely is knowledge to be transmitted in this way.

There are additional factors affecting the adjustment process for expensive durable goods. A steady stream of improvements is typical of new commodities in their early years. At the cost of foregoing the use of the product for a time the consumer may avoid teething troubles and obtain a better product later. There may also be an effect on the supply of substitutes for the new commodity. The falling demand for substitutes may induce a fall in their supply, which in turn encourages an expansion of the demand for the new good. An example is the closing of cinemas as a result of the spread of television. Logistic models ignore these additional factors and concentrate on the second lag in the learning process.

<sup>†</sup> The preliminary results obtained in an aggregative time-series analysis of data from 1947 to 1960 suggest that both these defects can be remedied by using a model based upon the cumulative lognormal curve.

Even as a description of growth due to the learning process the logistic model is inadequate, as consideration of the model by Bain (1962) shows. Two assumptions of this model are likely to be invalid. First, the proportion of any group in the population acquiring television sets in any period is determined partly by the proportion of the group “at risk” and partly by the proportion of the *whole population* who already own television sets. But people live and work in social groups, the members of which come in contact more often with each other than with members of other groups. If television is owned by a greater proportion of one group than another it is unlikely that the same proportion of those at risk in both groups will buy sets in the next period. Secondly, for a logistic curve, the influence of every owner on any non-owner must be the same and constant throughout the growth process. But early purchasers are in general more disposed to show off their acquisitions than late, and after the novelty wears off the attraction of demonstrating a new good palls.

These arguments apply equally to logistic models of the growth of demand for other new commodities. How will they affect the shape of the growth curve? The first will be important if, for any given level of ownership within a group, the rate of change of ownership is greater in some groups than in others. Since ownership of highly priced durable goods spreads much more rapidly in high than in low income groups this is likely to be true, and will tend to make the aggregate growth curve positively skew. The decline in the average owner’s influence on non-owners has the same effect. Thus theoretical consideration of the learning process alone suggests that a positively skew curve should be used. However, improvements in the product and changes in the supply of substitutes will increase the rate of growth in the later relative to the earlier years, and will tend to offset the skewness. We cannot say *a priori* which effect will predominate, though there is no reason to believe that the particular case of a symmetric curve will be appropriate. Since our discussion of the results obtained from some of the models based upon the logistic curve suggested that a positively skew curve might be preferable we can select a curve of this type on empirical grounds.

#### 4. TELEVISION OWNERSHIP IN THE UNITED KINGDOM

##### 4.1. *The Cumulative Lognormal Curve*

We shall now consider a study of the growth of television ownership in the United Kingdom, based upon a model incorporating the cumulative lognormal curve. This model has been used to analyse cross-section data relating to television ownership up till 1953. Thus the analysis is concerned with the period which was omitted from the earlier study of the growth of television ownership by Bain (1962).

The properties of the cumulative density function of the lognormal distribution have been fully discussed in Aitchison and Brown (1957) and it is unnecessary to do more than draw attention to a few of them here. The cumulative density function, written as  $\Lambda(t|\mu, \sigma^2)$ , is given by

$$\Lambda(t|\mu, \sigma^2) = \int_0^t \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}} \theta} \exp \left\{ -\frac{1}{2\sigma^2} [\log(\theta) - \mu]^2 \right\} d\theta. \quad (2)$$

The corresponding frequency function has its mode where  $t = e^{\mu - \sigma^2}$  and its median where  $t = e^{\mu}$ . The equation of a lognormal growth curve can therefore be written

$$x_t = \alpha \int_0^t \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}} \theta} \exp \left\{ -\frac{1}{2\sigma^2} [\log(\theta) - \mu]^2 \right\} d\theta \quad (3)$$

where  $x_t$  is the level of ownership at time  $t$  and  $\alpha$  is the ultimate saturation, or equilibrium, level. It has three parameters:  $\alpha$ ,  $\mu$  and  $\sigma$ . The maximum rate of growth occurs when  $t = e^{\mu - \sigma^2}$ , and ownership reaches 50 per cent. of the saturation level when  $t = e^\mu$ .

In order to understand how the parameters  $\mu$  and  $\sigma$  should be interpreted it is useful to define  $y_t$  as the normal equivalent deviate of  $x_t/\alpha$ . Then

$$\log(t) = \mu + \sigma y_t. \quad (4)$$

Consequently,  $\mu$  can be taken as a measure of the speed at which growth takes place: as  $\mu$  increases, for constant  $\sigma$ , the time required to reach any given proportion of the saturation level increases.  $\sigma$  measures the extent to which growth is concentrated in the middle of the curve. For fixed  $\mu$ , the time required to grow from any given

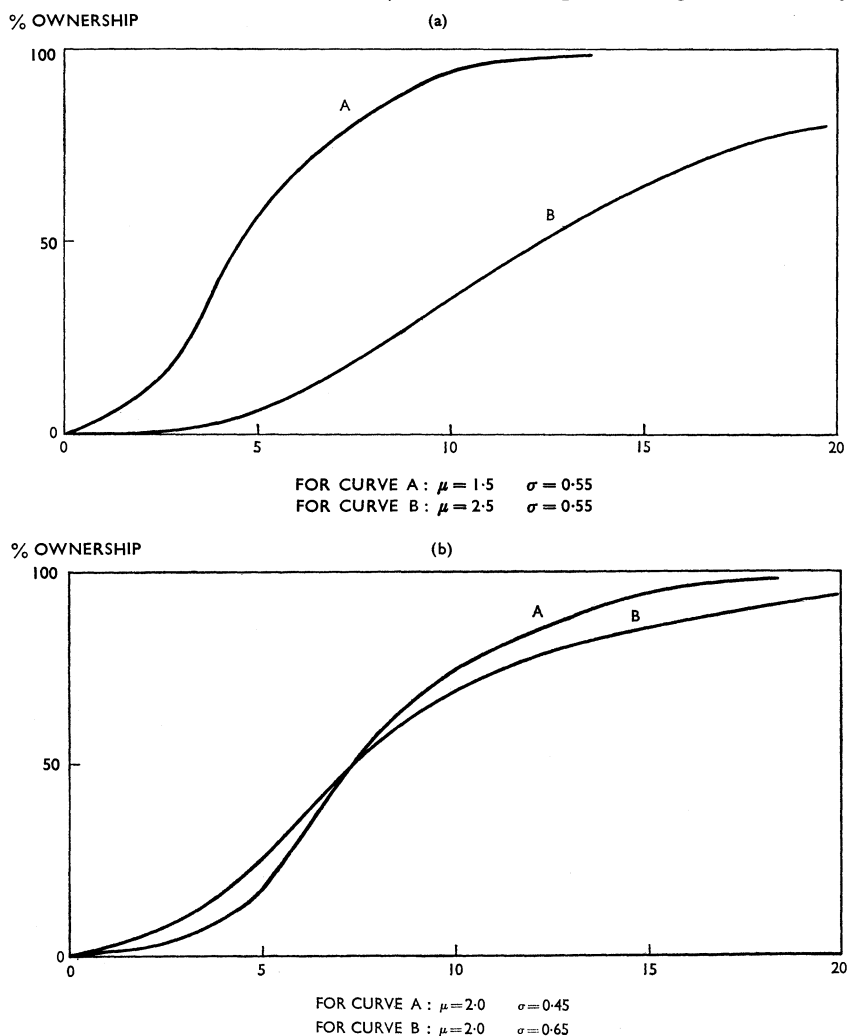


FIG. 1. Cumulative lognormal growth curves.

percentile level below 50 per cent. to an equal amount above increases with  $\sigma$ : the smaller is  $\sigma$  the more precipitous is growth in this part of the curve. Growth curves for two values of  $\mu$ , given  $\sigma$ , and for two values of  $\sigma$ , given  $\mu$ , are shown in Fig. 1.

#### 4.2. *The Model*

Consider any large population of households. Let  $q$  be the proportion which own a television set,  $\alpha$  be the saturation proportion and  $y$  the normal equivalent deviate of  $q/\alpha$ . Then suppose that the rate of growth of ownership is related to the level of ownership through the equation

$$\frac{dq}{dt} = \frac{\alpha}{(2\pi\sigma^2)^{\frac{1}{2}}} e^{-\mu - \sigma y - \frac{1}{2}y^2} \quad (5)$$

where  $\alpha$ ,  $\mu$  and  $\sigma$  are functions of social and economic variables. The growth path defined by this expression will in general consist of a series of segments of cumulative lognormal curves. However, as a simplification for this cross-section analysis, we shall assume that the parameters are constant, and in this special case the growth path reduces to a single cumulative lognormal curve. We shall also assume that  $\alpha = 1$  throughout the population. This may not be strictly true, but since we shall analyse data for a period in which the overall level of ownership remained below 33 per cent. and since we shall exclude the poorest 8 per cent. of the population, it is unlikely to have a serious effect on the results.

Under these assumptions equation (4) holds for the growth path. Thus

$$\log(t) = \mu_i + \sigma_i y_{it} \quad (4a)$$

where  $t$  is the time which elapses from the beginning of the growth process until the level of ownership  $q_{it}$  is reached in the  $i$ th group, and  $y_{it}$  is the normal equivalent deviate of  $q_{it}$ .  $\mu_i$  and  $\sigma_i$  are independent of  $t$ . This equation can also be written as

$$y_{it} = \frac{1}{\sigma_i} [\log(t) - \mu_i]. \quad (4b)$$

The groups in this analysis are defined by a three-way classification, social class (classes A, B, C, D), household size (1, 2, 3, 4, 5 or more persons) and region (London, Midland, North West).

Equation (4b) could be used to investigate the relation between the cross-section variables and the parameters  $\mu$  and  $\sigma$  directly. But an indirect approach has three advantages. First, simultaneous investigations of  $\mu$  and  $\sigma$  can be avoided. Secondly, the weight given to the absolute value of  $t$  in regions other than London, where there is some doubt about the appropriate time origin for the growth curve, can be reduced. Finally, there is no need to use successive values of  $y_{it}$  as the dependent variable. The method by which the data were collected suggests that there would be a high degree of serial correlation between the errors of observation in successive values of  $y_{it}$ .

To investigate the parameter  $\sigma$  we take first differences of equation (4b) and obtain

$$\frac{1}{\sigma_i} = \frac{\Delta y_{it}}{\log(t+1) - \log(t)}. \quad (6)$$

This gives us a series of estimates of  $1/\sigma$  to use as the dependent variable in a regression analysis, as is shown in Section 4.4.

Similarly, if at any time  $t$  we subtract  $y_{jt}$  from  $y_{it}$ , equation (4b) leads to

$$\mu_i - \mu_j = \sigma_j y_{jt} - \sigma_i y_{it}. \quad (7)$$

Using estimates obtained from the analysis of the parameter  $\sigma$  this enables us to investigate variations in the parameter  $\mu$ , as is shown in Sections 4.5 and 4.6.

This analysis has two aims. In the first place we want to try out the lognormal curve in the analysis of data which are known not to conform to the logistic curve. Secondly, we want to see whether the differences between the growth curves for the various groups can be summarized in terms of the parameters of the lognormal curve, and in particular whether they can be summarized by the parameter  $\mu$  alone.

We should certainly expect to find variations in  $\mu$ . Differences in income per head, in household size and in education will be reflected in the rate of growth. But it is hard to find reasons for variations in  $\sigma$  between groups. Although growth may be faster in one group than in another, it is not evident that it should be more or less concentrated. We shall therefore postulate initially that  $\sigma$  is invariant. Our analysis is directed on the one hand to testing the significance of variations in  $\sigma$ , and on the other to estimating the differences between the values of  $\mu$ .

#### 4.3. *The Data*

The data used in the analysis were obtained from the Hulton Readership Survey (1953), which included information on the ownership of television sets. For this survey a quota sample was taken in randomly selected areas. No household appeared more than once in the sample and the information supplied by each respondent can be regarded as information on the television ownership of one household of a given economic status.

The population is divided into five social classes. Although the economic status of each household is not defined in terms of income, there is a high correlation between the social class and total income of any household. Class A, the well-to-do, are roughly  $3\frac{3}{4}$  per cent. of all families; class B, the middle class,  $7\frac{3}{4}$  per cent.; class C, the lower middle class, which includes only a few manual workers,  $17\frac{1}{2}$  per cent.; and class D, the working class, consisting mainly of manual workers, 63 per cent. Class E, the poorest 8 per cent., were excluded from the analysis. Each class is divided into five size groups for this analysis, 1, 2, 3, 4 and 5 or more persons, the average size in the last group being about 6 people.

The information supplied by respondents in three parts of the country was selected for analysis. These regions, consisting of the Registrar General's London and South-Eastern regions, Midland region and North-Western region, correspond to the television services transmitted from the Alexandra Palace, Sutton Coldfield and Holme Moss Stations respectively. Respondents were classified according to whether or not they owned a television set and the number of years for which they had possessed it. The total sample size in these regions was just under 6,000, classes D and E being subsampled by 50 per cent. When the survey took place, television had been available for about  $6\frac{3}{4}$  years in the Alexandra Palace region,  $3\frac{1}{4}$  years in the Sutton Coldfield region and  $1\frac{1}{2}$  years in the Holme Moss region. Television sets were owned by 33 per cent. of households in the London region, 27 per cent. in the Midlands and 17 per cent. in the North-West.

The data provided estimates at yearly intervals of the proportion of the households in each cell which acquired a set during the previous year. Respondents who did not know when their set was obtained are assumed to have had the same pattern of



purchases as those who gave a definite reply. Any group in which “don’t knows” exceeded 10 per cent. of all television owners was excluded.

The cumulative proportions owning sets at any time are the basic data for the analysis. The errors in the different cells are substantially independent but errors in the observations for the same cell at different times are related.

#### 4.4. *The Parameter $\sigma$*

The analysis of the parameter  $\sigma$  is designed to test whether social class, household size and region cause significant variation, and to estimate any significant effects which are found. In equation (6) the relation

$$\frac{1}{\sigma_i} = \frac{\Delta y_{it}}{\log(t+1) - \log(t)}$$

was obtained. Estimates of  $\sigma_i$  from this relation depend on the time origin chosen. But with the exception of the London region, the appropriate time origin is not clearly determined. Knowledge of television in London was very slight before a service was started, so it seems reasonable to regard the beginning of the growth process as coinciding with the first transmissions. But the existence of a service in London implied that people in the Midlands and North-West had some familiarity with television before any service was provided. Hence the effective origin for the growth process should precede the first transmissions, although there is no way to determine *a priori* exactly what date is appropriate. In this analysis the start of transmissions to a region was used as the time origin. If the error introduced is the same for all groups within each region there will be no effect on the estimates of the effects of social class and household size. But variations in  $\sigma_i$  between regions would result from the choice of the wrong time origins, and any test to see whether there are *real* differences in  $\sigma$  between regions is precluded.

Regression analysis and the analysis of variance were used for the estimation and significance testing. The value of  $\sigma_i$  was assumed to depend on sets of dummy variables representing social class, household size and region. Least squares regressions were computed for each set of variables, and the effects of adding the less to the more significant sets were found.

One other set of variables was included to test the hypothesis that variations in  $\sigma$  due to changes in economic variables associated with time could be neglected. This set represents the calendar time intervals of the differences  $\Delta y_{it}$ .

Since  $\log(t+1) - \log(t)$  and  $\Delta y_{it}$  enter the estimate of  $1/\sigma_i$  in a multiplicative rather than an additive way, a multiplicative form of regression model was used. Thus  $\log(\sigma_i)$  was chosen as the dependent variable, while the independent variables are additive sets of dummy variables—a constant, 3 social class, 4 household size, 2 region and 4 time variables. No interaction between the effects of any of these sets of variables was postulated. Hence the regression equation has the following form:

$$\log(\sigma_i) = -\{\log(\Delta y_{it}) - \log[\log(t+1) - \log(t)]\} = a + \sum b_j z_j + \epsilon_{it} \quad (8)$$

where the  $z_j$  are all dummy variables representing social class, household size, region and time, and  $\epsilon_{it}$  is a random error term.† The constant term represents the value of  $\log(\sigma)$  in the North-West region, social class D, household size 5, in the last time period.

† Variable weights, which took some account of the heteroscedasticity of the error distribution, were used in the regression analysis.

To obtain the dependent variable, values of  $y$  corresponding to each  $q$  were taken from tables. The period for which transmissions had been available permitted the use of data for 5 years in the London region, 3 years in the Midland region and 2 years in the North-West region. Since successive observations are not independent, if any observation in a cell had to be rejected because the observed value of  $\sigma$  was infinite, all the observations for that cell were rejected. All the social classes are adequately represented, and of the size groups only the single-person group is inadequately represented in the 109 observations remaining for analysis.

The results of the analysis of variance, which was used to test the significance of the effects of the different factors, are shown in Table 1. Section (a) indicates that the

TABLE 1  
*Variance analysis for the parameter  $\sigma$*

	<i>Source of variation</i>	<i>Sum of squares</i>	<i>d.f.</i>	<i>Mean square</i>	<i>F</i>
(a)	Regions	2.0553	2	1.0276	59.70†
	Residual	1.8246	106	0.0172	
(b)	Regions	2.0553	2		1.96
	Social class	0.0986	3	0.0329	
	Residual	1.7260	103	0.0168	
(c)	Regions	2.0553	2		2.92‡
	Household size	0.1875	4	0.0469	
	Residual	1.6371	102	0.0161	
(d)	Regions	2.0553	2		1.55
	Time	0.1045	4	0.0261	
	Residual	1.7201	102	0.0169	
(e)	Regions and size	2.2428	6		2.00
	Social class	0.0934	3	0.0311	
	Residual	1.5437	99	0.0156	
(f)	Regions and size	2.2428	6		1.66
	Time	0.1039	4	0.0260	
	Residual	1.5332	98	0.0156	

† Significant at 1 per cent. level.

‡ Significant at 5 per cent. level.

two regional dummy variables have an effect which is highly significant at the 1 per cent. level. Sections (b), (c) and (d) show the effects of adding social class, household size and time variables respectively. Only the household size effects give an F ratio which is significant at the 5 per cent. level. In sections (e) and (f) a further test of the social class and time effects is made, when both region and household size variables are included, but again neither is significant at the 5 per cent. level.

Two conclusions can be drawn from this part of the analysis. The estimated value of  $\sigma$  certainly differs between regions, though this may be the result of either a true difference or the use of inappropriate time origins in the later regions. There is no evidence to contradict the hypothesis that  $\sigma$  is independent of social class and does not vary through time, but it seems likely that  $\sigma$  does differ between household size

groups. On account of the lack of independence in the observations this last conclusion should not be put too strongly.

Table 2 shows the regression coefficients and standard errors estimated from the equation including all the variables which contributed a significant amount to the

TABLE 2  
*Regression coefficients for  $\log(\sigma)$*

<i>Mean</i>	<i>Size 1</i>	<i>Size 2</i>	<i>Size 3</i>	<i>Size 4</i>	<i>London</i>	<i>Midland</i>	<i>R<sup>2</sup></i>
+0.7889† (0.1275)	+0.5210† (0.2243)	+0.2746‡ (0.1024)	+0.0834 (0.1009)	+0.1345 (0.1057)	–1.2261† (0.1146)	–0.6385† (0.1444)	0.578†

† Significant at the 1 per cent. level.

‡ Significant at the 5 per cent. level.

explanation. The obelisks indicate that the regression coefficients would differ significantly from zero at the 5 per cent. or 1 per cent. level, if tested individually using a two-tailed “*t*” test. It is clear that all three regions differ significantly from each other, and that the two smallest sizes of households cause the significant household size effect.

The estimated coefficients measure the effect of each variable on the value of  $\log(\sigma)$ . Thus if a coefficient is positive it implies that the variable causes a larger value of  $\sigma$  than the standard (North-West region, 5 or more person households). The comparatively large values of  $\sigma$  in the smaller household size groups imply that the S-shape of the growth curve is less pronounced in these groups than in the larger groups. Similarly, if the estimated regional effects reflected true differences they would imply that growth took place more evenly in the later regions.

The value of  $\log(\sigma)$  for the London region, household size 5, is required for the analysis of the parameter  $\mu$ . It is  $-0.4372$ , with a standard error of  $0.0830$ .

#### 4.5. *The Parameter $\mu$ : (a) Social Class Effects*

To estimate the effect of social class on the parameter  $\mu$  we postulate that it is the same in each region and is independent of the size of household. If it takes class D twice as long as class A to reach 50 per cent. ownership in the London region we assume that it will also take twice as long in the Midland region (although the whole process may take place faster in the Midlands) and whether 1-person or 5-person households are considered does not alter the relative times required. Thus  $(\mu_i - \mu_j)$  is assumed constant for each  $i, j$  (representing social classes) in all regions and in all household size groups.

Equation (7) stated that, for given  $t$ ,

$$\mu_i - \mu_j = \sigma_j y_{jt} - \sigma_i y_{it}.$$

Since the earlier analysis shows that  $\sigma$  can be regarded as invariant between social classes this can be written as

$$\mu_i - \mu_j = \sigma(y_{jt} - y_{it}) \quad \sigma_i = \sigma_j = \sigma \quad (9)$$

$$\text{or} \quad \log(y_{jt} - y_{it}) = \log(\mu_i - \mu_j) - \log \sigma. \quad (10)$$

Since equation (10) is in terms of the difference in  $\mu$  between any two social class groups it is convenient to take one social class as base. Class D was chosen on the

grounds that growth was likely to be slowest, implying that  $(\mu_D - \mu_i)$  would be positive, and because, with the comparatively large sample sizes, the observations were most precise. Let  $(\mu_D - \mu_C) = a$ , and define two dummy variables  $z_k$  representing the additional effects of classes A and B. Then equation (10) becomes

$$\log(y_{it} - y_{Dt}) = a + \sum b_k z_k - \log(\sigma) \quad (i = A, B, C). \quad (11)$$

This equation could be used directly for the investigation of  $\mu$ , the estimated values of  $\sigma$  being incorporated in the dependent variable. Alternatively, we can make a further analysis of  $\sigma$  in conjunction with the analysis of  $\mu$ . In view of the uncertainties connected with the time origin, and the fact that the estimates of  $\sigma$  for each household size are not exact, this approach seems preferable.

In this procedure only one value of the parameter  $\sigma$  need be known in advance. Let this value,  $\sigma^*$ , apply to the London region household size 5. Then if sets of dummy variables,  $z_l$ , represent household size groups and regions, by writing

$$\log(\sigma) = \log(\sigma^*) - \sum c_l z_l \quad (12)$$

we define the coefficients  $-c_l$  as the effects of household size and region on  $\log(\sigma)$ .

Combining (11) and (12) gives

$$\log(y_{it} - y_{Dt}) = [a - \log(\sigma^*)] + \sum b_k z_k + \sum c_l z_l + \epsilon_{it} \quad (13)$$

where  $\epsilon_{it}$  is a random error term. The known value  $\log(\sigma^*)$  must be subtracted from the estimated mean to give  $a$ , which together with both values of  $b_k$  measures the effect of social class on  $\mu$ . The six values of  $-c_l$  measure the effect of household size and region on  $\log(\sigma)$ . These estimates of the regional effects should be preferred to those obtained earlier because they are not subject to any bias through choosing the wrong time origin. Thus, as a by-product, this analysis provides a second opportunity to test the significance of variations in  $\sigma$ .†

The basic data were the values of  $y$  calculated for the analysis of the parameter  $\sigma$ . From these data the 113 values of the dependent variable were found and, with the exception of households of one person, all the dummy variables were adequately represented.

As in the previous section sets of variables were added successively to the regression equation and the significance of each set was tested. The coefficients of the significant sets were then estimated. Table 3 shows the results of the analysis of variance. Section (a) shows that the two social class variables proved highly significant. From sections (b) and (c) it can be seen that the addition of a household size effect on  $\log(\sigma)$  made a significant improvement, but that the regional variables did not contribute significantly to the explanation. Nor were the regional variables significant when fitted in conjunction with social class and household size, as is shown in section (d).

Thus the conclusion of this part of the analysis is that there are significant differences in  $\mu$  between social classes and in  $\log(\sigma)$  between household sizes, but that

† There is a problem of interpretation connected with this by-product. If  $\sigma$  had been assumed constant and interaction had been postulated between the effects of social class and household size on  $\mu$ , the same estimates would be a measure of this interaction. The effect of social class might be less in some size groups than in others. Are these estimates really some mixture of interactions in  $\mu$  and variation in  $\sigma$ ? It is impossible to rule out this theoretical possibility. But it will be shown later that the empirical evidence does not contradict the hypothesis that there is no interaction.

differences in  $\log(\sigma)$  between regions can be neglected. The lack of any regional effect suggests that the errors in the time origin for the Midland and North-West

TABLE 3  
*Variance analysis for the parameters  $\mu$  and  $\sigma$*

	<i>Source of variation</i>	<i>Sum of squares</i>	<i>d.f.</i>	<i>Mean square</i>	<i>F</i>
(a)	Social class	2.3956	2	1.1978	36.21†
	Residual	3.6384	110	0.0331	
(b)	Social class	2.3956	2		3.79†
	Household size	0.4548	4	0.1137	
	Residual	3.1836	106	0.0300	
(c)	Social class	2.3956	2		1.30
	Region	0.0857	2	0.0429	
	Residual	3.0811	108	0.0329	
	Social class and Household size	2.8504	6		
(d)	Region	0.1025	2	0.0513	1.73
	Residual	3.0811	104	0.0296	

† Significant at 1 per cent. level.

regions were responsible for the highly significant effects found earlier.† The evidence certainly does not contradict the hypothesis that there are no regional variations in  $\sigma$ .

The parameter estimates and standard errors which were obtained from the regression including social class and household size variables are shown in Table 4. If the

TABLE 4  
*Regression coefficients for  $\mu$  and  $\sigma$*

<i>Coefficients for <math>\mu</math></i>		<i>Coefficients for <math>\log(\sigma)</math></i>	
Mean	-1.0165† (.1242)	Size 1	+1.385 (.5745)
A-D	+1.0502† (.1136)	Size 2	+5.182† (.1467)
B-D	+ .6215† (.1065)	Size 3	+ .0925 (.1293)
R <sup>2</sup>	.472†	Size 4	+1.808 (.1343)

Notes: (1) Standard errors are shown in brackets.

(2) † Significant at 1 per cent. level.

coefficients are tested individually using a two-tailed “*t*” test the mean and both social class variables show highly significant deviations from zero, but the only significant household size variable is that for 2-person households.

† It is hardly conceivable that real effects of that size could have been offset by interactions in the effects of regions and social classes on  $\mu$ , since the necessary interactions would be larger than some of the direct effects of social classes on  $\mu$ .

Table 5 translates the parameter estimates of the regression equation into the effects of social class on the parameter  $\mu$  of the lognormal growth path. First the component of the mean attributable to  $-\log(\sigma^*)$  is removed, giving estimates of

TABLE 5  
*The effect of social class on  $\mu$*

	$j = A$	$j = B$	$j = C$
$\log(\mu_D - \mu_j)$	-0.404	-0.832	-1.454
"s.e."	(0.151)	(0.144)	(0.149)
$\mu_D - \mu_j$	0.668	0.435	0.234
$\exp(\mu_D - \mu_j)$	1.950	1.545	1.263

$\log(\mu_D - \mu_j)$   $j = A, B, C$ . These are then converted to show the effect of social class on  $\mu$ :  $\mu_A$  is about 0.67 less,  $\mu_B$  about 0.44 less and  $\mu_C$  about 0.23 less than  $\mu_D$ . The estimates of  $e^{(\mu_D - \mu_j)}$  give the ratios of the times required by class D and class  $j$  to reach the 50 per cent. level of ownership. Thus class D takes about twice as long as class A,  $1\frac{1}{2}$  times as long as class B, and  $1\frac{1}{4}$  times as long as class C. This class ordering accords well with theory and none of the relative periods seems unreasonable.

The standard errors in Table 5 apply to the estimates of  $\log(\mu_D - \mu_j)$ , but should be regarded as orders of magnitude rather than precise values. Though probably small, the covariances of the estimate of  $\log(\sigma^*)$  and the parameters estimated above are unknown and have been neglected. As a result of the serial correlation in the observations the standard errors probably exaggerate the precision of the estimates.

The estimated effects of household size on  $\log(\sigma)$  are somewhat different from those obtained earlier, but in no case do the earlier estimates differ significantly from these at the 5 per cent. level.†

TABLE 6  
 *$\sigma$ , by size of household*

Household size	$\sigma$
1 person	1.0337
2 persons	0.9205
3 persons	0.7044
4 persons	0.7521
5 persons	0.6458

Two estimates of each household size effect are now available. If they had been estimated from independent sets of data a weighted average would give a more precise estimate of each effect. In this case the basic data are not independent, but have been processed in very different ways. The distributions of the estimates about the true values are unlikely to be highly correlated. Hence the estimates have been treated as if they were independent and weighted averages have been calculated using weights proportional to the reciprocals of the estimated variances. The resulting values of  $\sigma$  for each household size are shown in Table 6.

† Thus the evidence does not contradict the hypothesis that there is no interaction between the social class and household size effects on  $\mu$ .

4.6. *The Parameter  $\mu$ : (b) Household Size Effects*

For the study of the effect of household size on  $\mu$  we postulate that there is no interaction between the effects of household size and social class, and that the former are invariant between regions. The analysis is again based upon equation (7):

$$\mu_i - \mu_j = \sigma_j y_{jt} - \sigma_i y_{it}.$$

For this part of the analysis the 4-person household size group, which contained the largest sample sizes, has been taken as the base from which deviations were measured and the values of  $\sigma$  for each household size are those shown in Table 6. Then in the regression equation

$$\sigma_4 y_{4t} - \sigma_i y_{it} = \sum a_i z_i + \epsilon_{it} \quad (14)$$

the  $a_i$  ( $i = 1, 2, 3, 5$ ) are estimates of  $\mu_i - \mu_4$  and  $\epsilon_{it}$  is a random error term. A total of 136 observations were available, with all the size groups adequately represented.

The results are shown in Table 7. Only the 1- and 2-person households show

TABLE 7  
*The effects of household size on  $\mu$*

	$i = 1$	$i = 2$	$i = 3$	$i = 5 +$
$\mu_i - \mu_4$	0.8047†	0.3841†	0.0469	-0.0298
"s.e."	0.0635	0.0199	0.0256	0.0252
$\exp(\mu_i - \mu_4)$	2.236	1.468	1.048	0.971

† Significant at 1 per cent. level.

effects which differ significantly from zero, but even allowing for the lack of independence in the observations they are highly significant.† It seems that there is little difference in the values of  $\mu$  for households of three or more persons. It would take single-person households more than twice as long, and 2-person households nearly  $1\frac{1}{2}$  times as long, as larger households to reach the 50 per cent. level of ownership.

The composition of each household size group provides some explanation of these results. The average number of children per household increases with the size of household, but the income per head falls. The results suggest that the specific demand resulting from the presence of children much more than offsets the reduction in income per head. This specific demand emanates partly from the children themselves, and increases as they get older, partly from the fact that where there are children it is less easy for adults to find their entertainment outside the home, and partly because the relative cost of television as a means of entertainment falls with the size of household. For any given household income it is likely that the increase in demand due to an additional child, especially the first, is equivalent to more than the demand due to an additional adult. Since the speed of growth does not increase much beyond 3-person households it appears that these influences are offset by the income effect. The very slow growth among single-person households is also explained by the composition of the group, which consists mainly of comparatively young people who have other activities and whose incomes are possibly below the average for their social class.

† The standard errors have been calculated on the assumption that the values of  $\sigma$  used were correct. Any errors in the values of  $\sigma$  will, of course, be reflected in the estimates of the effect of household size on  $\mu$ .

## 5. CONCLUSION

The empirical analysis has not thrown up any evidence to suggest that the model based upon the cumulative lognormal curve cannot adequately describe the first years of the growth of television ownership in the United Kingdom. It showed that within the population both the parameters  $\mu$  and  $\sigma$  must be regarded as variable. There was no evidence to contradict the hypothesis that  $\sigma$  was invariant with respect to social class and region, but differences between household size groups appeared to be present. The parameter  $\mu$  varied considerably between social classes and household sizes: the lower the social class and smaller the household size the slower the growth. The plausibility of the results which were obtained lends support to our arguments in favour of positively skew curves.

## ACKNOWLEDGEMENTS

I am indebted to J. A. C. Brown, D. G. Champernowne, M. C. Lovell and H. Watts for their comments on earlier drafts of this paper. Any errors are, however, my own. The empirical analysis was carried out while I was a member of the Cowles Foundation for Research in Economics at Yale University, and was financed by the Foundation.

## REFERENCES

- AITCHISON, J. and BROWN, J. A. C. (1957), *The Lognormal Distribution*. University of Cambridge, Department of Applied Economics, Monograph No. 5. Cambridge University Press.
- BAIN, A. D. (1962), "The growth of television ownership in the United Kingdom", *Int. Econ. Rev.*, 3, 145–167.
- DERNBURG, T. F. (1958), "Consumer response to innovation: television", in DERNBURG, Rosett, and Watts, *Studies in Household Economic Behaviour*, Yale Studies in Economics, Vol. 9. Yale University Press.
- HULTON READERSHIP SURVEY (1953). London: The Hulton Press.
- MANSFIELD, E. (1961), "Technical change and the rate of imitation", *Econometrica*, 29, 741–766.
- MASSY, W. F. (1961), *Television and Market Penetration*. Unpublished Ph.D. thesis, Massachusetts Institute of Technology.
- ROOS, C. F. and VON SZELISKI, V. (1939), "Factors governing changes in domestic automobile demand", in *The Dynamics of Automobile Demand*. New York: General Motors Corporation.



## LINKED CITATIONS

- Page 1 of 1 -



*You have printed the following article:*

### **The Growth of Demand for New Commodities**

A. D. Bain

*Journal of the Royal Statistical Society. Series A (General)*, Vol. 126, No. 2. (1963), pp. 285-299.

Stable URL:

<http://links.jstor.org/sici?sici=0035-9238%281963%29126%3A2%3C285%3ATGODFN%3E2.0.CO%3B2-5>

---

*This article references the following linked citations. If you are trying to access articles from an off-campus location, you may be required to first logon via your library web site to access JSTOR. Please visit your library's website or contact a librarian to learn about options for remote access to JSTOR.*

## **References**

### **The Growth of Television Ownership in the United Kingdom**

A. D. Bain

*International Economic Review*, Vol. 3, No. 2. (May, 1962), pp. 145-167.

Stable URL:

<http://links.jstor.org/sici?sici=0020-6598%28196205%293%3A2%3C145%3ATGOTOI%3E2.0.CO%3B2-F>

### **Technical Change and the Rate of Imitation**

Edwin Mansfield

*Econometrica*, Vol. 29, No. 4. (Oct., 1961), pp. 741-766.

Stable URL:

<http://links.jstor.org/sici?sici=0012-9682%28196110%2929%3A4%3C741%3ATCATRO%3E2.0.CO%3B2-W>