RETURNS TO SCALE IN RETAIL TRADE*

BY GUR OFER

Department of Economics, The Hebrew University of Jerusalem

This paper attempts to estimate *genuine* scale effects in retail trade from a cross section of retail stores in Israel. This is done by estimating a simple production function for several retail branches and employing the faithful old direct Cobb-Douglas structure with value added as output and labor and capital inputs. And indeed despite the well-known peculiarities of the retail industry, a cross section estimation produces "normal" production-function estimates with reasonable input elasticities. The estimates also identify marked increasing returns-to-scale parameters, higher in food and lower in branches less affected by consumer participation and geographical dispersion. These increasing returns may explain a good part of the increase in sales per unit of inputs observed in time series.

I. INTRODUCTION

Like other distributive services, retail trade enjoys increasing returns to scale when output increases within a given district but suffers from increasing costs as the lines of distribution become longer and thinner, i.e., as they serve fewer consumers. Also like other services, retail trade is subject to an increasing-returns illusion when increasing consumer participation in performing the service reduces the amount of service actually performed by the firm.

This paper attempts to estimate genuine scale effects in retail trade from a cross section of retail stores in Israel. This is done by estimating a simple production function for several retail branches and employing the faithful old direct Cobb–Douglas structure with value added as output and labor and capital as inputs. As far as I know, this is the first time that such an approach has been used to investigate the production characteristics of retailing.[1] The theoretical and statistical task of the paper is thus not to construct an elaborate estimating procedure but to demonstrate the applicability of the approach to the case of retail trade with its well-known peculiarities. The first of these, the undefinable nature of the product, lies at the root of the efficiency problem: is an increase in "efficiency" genuine or does it reflect a reduction in the amount of service? Next is the related phenomenon of varying consumer participation in the production process. This is the industry's market structure: an imperfectly competitive market imposed by its distributive function. And fourth is the impact of the predominance of small proprietors.

*This paper is part of a project on the retail industry undertaken and financed by the Maurice Falk Institute for Economic Research in Israel and in part supported by the Ministry of Commerce and Industry. I would like to thank Michael Bruno for his comments on an early draft and Zvi Griliches for the many helpful suggestions he made in the course of long discussion on the final draft. Unfortunately, his book (written with Vidor Ringstad) on *Economies of Scale and the Form of the Production Function* reached me only after the paper was almost complete. While there is much similarity of approach, I have benefited greatly from the thorough and precise treatment of many problems, and some of the ideas suggested have been incorporated in this paper. I would also like to thank my research assistants Itzhak Tal and Yosef Tawil, and the editor of the Falk Institute Susanne Freund.

[1]Some studies on retail trade using other approaches are [2], [6], [9], [14], [20], [21], [22].

Despite all these a cross-section estimation produces "normal" production-function estimates with reasonable input elasticities. The estimates also identify marked increasing returns-to-scale (RTS) parameters, higher in food and lower in branches less affected by consumer participation and geographical dispersion.

Section II takes up the principal theoretical and statistical considerations mentioned above while Section III describes the data and defines the variables. The problems of comparing proprietors' with employees' labor and estimating the input of fixed capital are taken up. Section IV presents results based on inputs measured in physical units. The RTS findings are discussed, RTS being distinguished from the efficiency advantage of modern selling techniques associated with large stores and the efficiency disadvantage resulting from the backwardness of small proprietors. Section V contains some remarks on the relation of RTS to the increase in sales per input unit.

II. The Production and Market Characteristics of Retail Trade

In this section we try to establish on the basis of theoretical considerations (a) that there are genuine increasing RTS in the retail industry; (b) that increasing RTS are likely to show up in a cross section of individual stores; (c) that when measured as value added in constant prices, retail output discriminates between real and spurious inter-store efficiency differentials: the first are estimated and the second excluded; (d) that existing price differentials between stores tend to cause an underestimation of the RTS parameter; and (e) that, on grounds additional to those in (c), value added is a more appropriate measure of output than gross output or sales.

Retail stores enjoy increasing returns of three types. First, many retail activities contain a fixed-cost element, and an increase in scale distributes this element over an increasing number of service units (whatever they may be). Second, retailing is confronted with uncertainty about the stream of customers and their specific demands. By the rule of large numbers, the cost of uncertainty per unit of service is reduced as scale increases. It should be emphasized that such increasing RTS do not represent a reduction in the amount of service per unit of goods sold.[2] A third source of increasing RTS is the association between size of store and size of transaction;[3] since it is an observed fact that there are increasing returns on transaction size to stores and consumers alike.[4] This factor depends partly on the demand for the services of various stores and only partly on the nature of "production" of these services, like the bringing of many kinds of goods into an individual store (supermarket). It could be claimed that only the latter part of the increased transaction size should be included in the estimation of RTS. Aside from the practical problem of separating the two effects, it should be pointed out that the demand side effect is similar to scale effects enjoyed by many industries due to larger "run" sizes which to a large extent depend upon the nature of the demand.

[2]Henceforth referred to as "service level."

[3]See Sharir [21].

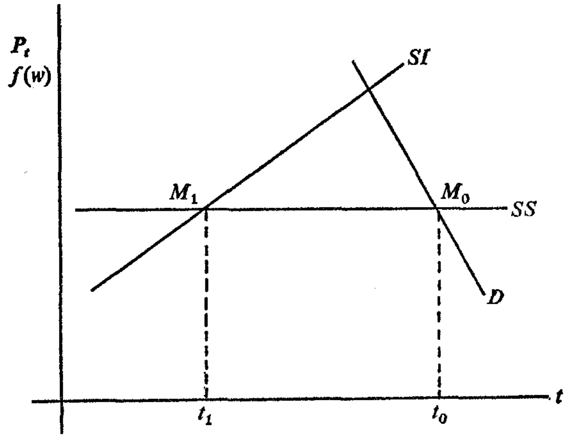
[4]See, among others, *ibid.;* Schwartzman [21]; and Hall, Knapp, and Winston [9].

If there is such a scale advantage why do stores not take advantage of it and expand? Why are there so many small stores? The main answer to this is that, with given technology, population density, and purchasing power, any increase in store size involves a reduction in the service *level*, to below what consumers—as utility maximizers—seek to buy. To see this let us look on the production of retail services as a production process carried out by the household and the industry jointly. The retail industry may be looked upon as producing "consumer time".[5] Accordingly a unit of retail service is here defined as the amount of services that will save one hour for the representative consumer operating at a given technology. The consumer has a demand function for retail services (from either source) and a self-supply function. His demand for commercial retail services is derived as the difference at each price between his total demand and his self-supply. The equilibrium amount of retail services consumed is determined where demand equals total supply. The amount of retail services *bought* is determined by the point of equality of the marginal cost per unit of services from each of the two sources. This last equilibrium exists because, as the service level of the industry increases—as the industry get "closer" to the consumer—the cost advantage over self-supply declines. This follows from the existence of increasing RTS together with the fact that getting "closer" to the consumer involves a reduction in scale and thus increases the cost per service unit. "Getting closer to the consumer" is in the first place meant literally, in the sense that when the store is near the consumer's home there is a saving in travelling time; but it also means home delivery of goods, reducing customer waiting time, and increasing the choice and therefore the stock of goods so as to better serve individual consumer tastes (thereby making it unnecessary to visit more than one store). This can be presented diagrammatically. Service units (hours of consumer time) are measured along the horizontal and costs along the vertical axis, the volume of goods being held constant. D is a consumer's total demand for retail services, SS is his self-supply schedule [horizontal since both technology and alternative time costs—a function of wages, $f(\overline{w})$—are assumed constant], and SI is the retail industry's increasing-costs supply schedule. Equilibrium is at points M_0 and M_1, where t_0 is total services consumed, t_1 is purchased services, and $t_0 - t_1$ is, according to our definition, the distance between store and consumer. Note that the level of SI is determined also by the size of the market; a denser population with greater purchasing power would shift SI downwards.

It is the increasing cost of increasing the service level for a given volume of goods sold in a district of given population density, purchasing power, and cost of self-supplied services that determines store size.[6] Consequently optimum or equilibrium store size varies according to location and the character of the clientele. Specifically, the denser the population, the higher its purchasing power, and the lower the cost of self supply—the bigger will be the optimum store. It thus follows that production-function estimation of a cross section of stores of

[5]Following Becker [1], this approach is used by myself [19] and, in a more systematic way, by Sharir [22]. In what follows I rely on both and confine myself to the essentials relevant to the present analysis.

[6]The increasing costs of electrical companies resulting from transmission losses are somewhat similar. See Nerlove [17, pp. 103–104].



different sizes should find increasing returns to scale, provided output is measured properly (see below). There is also a dynamic reason why returns to scale should show up in a cross section. Over time, incomes go up, population density increases, and shopping techniques (especially home storage facilities and travel) improve; all these tend to shift consumers to bigger and more distant stores or shopping centers and shop-owners must follow. If there is a lag in their response—and there is evidence that there is, especially in the case of small independents—bigger stores will be more efficient. This argument, it is true, assumes that there are short-run elements in the cross section, although a cross section is usually thought of as representing the long run.

If one could assume perfect competition in the entire retail industry a store's value added would be the exact measure of its service output. Of course, equal service prices entail different goods prices in accordance with the service level. (For this reason "sales at constant prices" is a bad measure of retail output since it will overstate the output of goods with a smaller service element. If service level is positively correlated with store size it will also inflate the RTS parameter.) Under competitive service prices, the price per unit of time saved increases as the service approaches the consumer, a trend that corresponds to the decline (in the same direction) in the store's advantage over the consumer. In a sense, a time unit saved near the consumer can be thought of as a better quality unit than a distant one. Competition thus means equal prices for service units at equal "distances".

Service prices (at equal distances) need not be equal thoughout a cross section which covers several districts. Since output must be measured by value added at actual prices it is necessary to estimate any biases that arise. Two possibilities come to mind. First, there may be more than one price if the cross section extends over more than one market, each with its own (competitive) set of prices; in this case one would expect that in communities or markets where conditions enable large stores to exist, prices will be lower at each distance, and this will tend to understate the cross-section scale effect. Second, prices may vary within a market or community, as a result of increasing monopolistic powers as one moves from the commercial center to the more isolated neighborhoods; this erosion of competition tends to favor small stores and again to bias the scale effect downward.[7] However, the observed fact that goods prices are higher in small groceries than in shopping centers, supermarkets, and the like is, at least partly, due to higher service level and it is difficult to know if there is any surplus ascribable to monopolistic effects.

Finally there is the issue of gross output versus value added.[8] In retail trade, more than in a "normal" industry, value added is appropriate and sales unacceptable even without regard to the price issue raised before.[9] Although sales may be looked upon as produced from two inputs—goods purchased and services produced (by the factor inputs)—it is the consumer rather than the store owner who decides on the ratio between them; that is, the ratio is determined by demand considerations and not by input prices and technology. The appropriate production function is thus the sum $s = f(n, k) + q$ [where s is sales, $f(n, k)$ is value added and q is goods purchased] or $y = s - q = f(\mathrm{n}, k)$, which is the general form of value-added estimation.[10] Thus, retail trade does have a measure of output almost as appropriate as that for other aggregative sectors of the economy.

The estimation centers on a Cobb–Douglas production function using value added (y) as output and labor (n), fixed capital (k_1) and working capital (K_2) as inputs:[11]

$$y = An^{\alpha}k_1^{\beta 1}K_2^{\beta 2} \cdot \Pi\, e^{\lambda_i D_i}\, e^{u_i}. \tag{1}$$

This form and the above theoretical considerations bring up problems of identification and bias or consistency of the estimated coefficients. Identification is guaranteed by exogenous factors that ensure variation in scale (to identify the expansion path) and input ratios (to identify the isoquants). In our case the expansion path is identified by the fact that output is exogenously determined. The isoquant is identified by, first, actual or expected inter-firm differences in input prices. Small stores, for example, usually pay less for labor and may be paying more for capital. Following Zellner *et al.*, expected price differences are a sufficient condition for identification.[12] Second, it is identified by the short-run restriction on the size of stores mentioned above. Stores willing to expand but unable to extend their space will try to increase output by using greater amounts of other inputs. As shown later, this may be relevant for both capital and labor.

For lack of data, we cannot use reduced-form methods which involve input and output prices, and are confined to direct estimation of the production function.

[7]This holds even though some small stores are concentrated in markets and traditional commercial centers.

[8]For an extensive discussion of value added versus gross output, see Griliches and Ringstad [8, pp. 108–123].

[9]A difficulty that could be overcome by defining appropriate prices for sales.

[10]Griliches and Ringstad [8, p. 109].

[11]Throughout this paper lower-case letters denote variables in physical terms and capitals denote value terms. Thus $Y = P_y \cdot y$; n and k_1 are measured in physical units; and K_2 is in value terms.

[12]Zellner, Kmenta and Drèze, [23, pp. 323–335].

The only option remaining is therefore to try to assess the possible direction of the biases or inconsistencies created by the omission of other structural equations and for other reasons. First, if output is exogenously determined, and thus determines inputs, a direct least-squares estimate of output or inputs will produce downward-biased estimates of the input coefficients and also of their sum, the RTS parameter.[13] Second, if we accept the assumptions of the traditional model in which inputs are correlated with the disturbance element in the production function, biases or inconsistencies in both directions may be created in the input coefficients, and of course in RTS.[14] If, for example, part of the disturbance reflects a missing input such as management, or a disregard to labor or capital quality differentials, then it has been shown that it is very likely that RTS will be underestimated and the input coefficients will be biased in specific directions.[15] Simultaneity of errors in variables, both of which make for correlation between an independent variable and the disturbance, could create other biases in both inputs and RTS. What little there is to say about such biases will be said later; here I add only that in the Griliches-Ringstad investigation of Norwegian manufactures [8] it was found that they did not damage the RTS findings.

On the other hand, if, under simultaneity, Zellner's "new model" (Zellner *et al.*, *op. cit.*) is accepted, inputs are independent of the disturbance term in the production function and the direct estimators are at least consistent and may be unbiased. This is true even if input and output prices actually differ between firms.[16] The intervention of short-run factors would further increase the independence of the inputs from other structural equations and help to reduce biases.

The last bias considered here is that created when retail service prices (which are included in $Y = P_y y$) vary between stores and are correlated with input levels. The preceding discussion has indicated that store size and P_y are negatively correlated and this suggests that P_y may be negatively correlated with inputs. The result would be to underestimate the input and RTS parameters.[17] Although much is still obscure, it seems reasonable to assume that the sum of the input parameters (i.e. RTS) is at least not overestimated.

III. The Data and the Variables

In its major characteristics Israel's retail industry does not seem to be very different from the retail industries of other countries at a similar level of development. It may be somewhat backward in the predominance of the small single proprietorship with few paid employees and in being slow to introduce modern

[13]This conclusion is an extension of the reciprocal relation between parameters in the two-variable case.

[14]See Zellner, *et al.* [23, pp. 324–329]; Griliches [7, pp. 8–20], and Griliches and Ringstad [8, pp. 92–103, 194–198].

[15]Griliches [7, pp. 8–20].

[16]This is shown in an appendix to an earlier version of this paper (obtainable from the author).

[17]The same type of bias may be created if the price of working capital varies between stores. At this point we assume that it does not.

selling techniques.[18] Nevertheless most of the new methods are being used and are spreading rapidly. Without going into details, we can say that Israel's retail industry is "normal" enough to serve as a case study from which general conclusions can be drawn.

The production function estimations are based on observations for individual stores in 1967/68 and 1968/69. The main source of the data is the trade surveys of the Central Bureau of Statistics (CBS) which have been carried out annually since 1966/67 on a sample of wholesale and retail establishments. Unfortunately, only stores employing hired labor were included until 1967/68, thus excluding not only the large majority of stores and the bulk of sales but also the possibility of comparing the bigger, more modern stores with the small proprietorships. This can be done for the first time with the 1968/69 survey which is for this reason our main source.[19]

The following are the major variables defined from the survey data:

S	annual sales	
Q	annual purchases[20]	
R_t	end-of-year goods inventories	
n_1	employees	annual average
n_2	working owners	annual average
n_3	unpaid family members	annual average
N_1	compensation of employees (wage bill).	

In the 1967/68 survey and for some firms in the 1968/69 survey, we also have:

R_0	beginning-of-year goods inventories
X	end-of-year consumer debt.

We then define:

Value added

$$\text{(a)}\ Y = S - Q \quad \text{or} \quad \text{(a}'\text{)}\ Y' = S - Q + (R_t - R_0);$$

working capital[21]

$$\text{(b)}\ K_2 = R_t \quad \text{or} \quad \text{(b}'\text{)}\ K_2' = \tfrac{1}{2}(R_t + R_0) + X;$$

[18]Only 28 per cent of the industry's total labor force are employees, while 50 per cent are proprietors and the rest are unpaid family workers. In 1960 over 50 per cent of the retail labor force were employees in almost all Western European countries, and in the 1950's over 40 per cent in all countries except Belgium (1955), Italy (1955), and Greece (1951). Jefferys and Knee [14, p. 17] and OECD [18, p. 155]. Less than 1 per cent of all retail stores are self-service stores in Israel, compared with almost 3 per cent in Europe in 1970 or more than 1 per cent as far back as 1955. Israel, Central Bureau of Statistics (CBS), [12], and Jefferys and Knee [14].

[19]CBS [12]. The data of this survey are based on two samples (differing in sampling procedure, coverage, and reliability), one for large stores (with 4+ employees) and one for small stores.

[20]It is not clear from the Trade Surveys whether purchased materials other than the stock-in-trade are included. To the extent that they are not, our measure of Y departs from value added in the strict sense.

[21]Biases may originate from the use of value terms for the inputs in the estimation (as here with K_2); I do not think that the biases involved in K_2 are important.

variants (a′) and (b′) are of course preferable to (a) and (b), but lack of data forced us to use the latter. Regressions using the primed variants were run on part of the observations and proved not to differ materially from those using (a) and (b) whose results are presented in this paper.

Lastly, labor input

(c) $n = a_1 n_1 + a_2 n_2 + a_3 n_3$,

where a_i are weights, explained below. Dummy variables for sub-branch and location (D_i) are also derived from the surveys.

The surveys provide no data on fixed capital. Of the few possibilities available we chose store space (k_1) as proxy.[22] However, store-space data were obtainable only for sample firms in the three main cities (Jerusalem, Haifa, and Tel Aviv).[23] The price of this procedure is the loss of all observations for smaller localities and thus of the possibility of investigating their characteristics; furthermore, the procedure makes the sample less representative. Not that k_1 is an ideal measure of fixed capital. It does not include equipment or any indication of the quality of the space. We shall come back to possible distortions due to this factor.

The next problem is that of weight for the different categories of labor input. The obvious question is why worry about weights instead of running each type of labor as a separate input. This was tried, but was not fruitful. Specifically, in a regression of the type:

$$\log Y = \log A + \alpha_1 \log n_1 + \alpha_2 \log n_2 + \alpha_3 \log n_3 + \beta_1 \log k_1 + \beta_2 \log K_2 + u \tag{2}$$

very low (or negative) output elasticities were found for n_2—much lower than expected. Assume for example that a working owner is only as productive as an employee. One would then expect the output elasticities of n_2 and n_1 to be in the same ratio as the volume of n_2 and n_1.[24] A comparison shows that this is definitely not the case. Some results for the 1968/69 sample of food, clothing, and furniture stores are presented in Table 1. While owners are 27 per cent of total employed persons in food stores, 11 per cent in clothing, and 14 per cent in furniture, their output elasticities are practically zero or even slightly negative. When only stores with employees are considered, as in the 1967/68 sample, the results are even more paradoxical.

Unpaid family workers (n_3) are of course much less productive than employees and this should be reflected in the results. What we find is a very low

[22]One possibility is to assume that fixed capital remains constant for two successive years and to apply the technique used by Mundlak [16] to estimate "firm effect" in order to estimate "capital and firm effect." Another is to use the CES production function together with the assumption of labor-market equilibrium.

[23]The number of observations left in the sample (see Table 3 below) after all those with insufficient information have been eliminated allows us to estimate production functions for food (61 in the CBS classification), clothing (62), and furniture (63; includes household goods and appliances) from the 1968/69 sample and for food and clothing from the 1967/68 sample. There are not enough observations left for estimates of the other retail branches.

[24]The average number of self-employed per store is as follows (with the corresponding standard deviation in parentheses): Food 1.00 (0.52); clothing 0.87 (0.61); furniture 0.66 (0.64). The variation is large enough to preclude the possibility that self-employed are a fixed factor.

TABLE 1

SELECTED CHARACTERISTICS OF PERSONS EMPLOYED IN RETAIL TRADE: 1968–1969[a]

	Food	Clothing	Furniture
Employees (n_1)			
α_1	0.36	0.33	0.17
	(0.04)	(0.04)	(0.07)
$n_1/\Sigma n_i$	0.62	0.84	0.80
$r_{n_1,Y}$	0.98	0.83	0.78
Proprietors (n_2)			
α_2	−0.11	−0.01	−0.27
	(0.07)	(0.05)	(0.09)
$n_2/\Sigma n_i$	0.27	0.11	0.14
$r_{n_2,Y}$	−0.42	−0.40	−0.45
Unpaid family members (n_3)			
α_3	0.08	0.04	0.11
	(0.35)	(0.04)	(0.06)
$n_3/\Sigma n_i$	0.11	0.05	0.06
$r_{n_3,Y}$	−0.16	−0.16	−0.16

[a]Symbols are as in equation (2), p. 370. Figures in parentheses are the standard errors.

output elasticity that does not differ significantly from zero in two out of three cases. The picture is here perhaps not distorted, but it is certainly more obscure. The explanation for all this is not that the marginal productivity of working owners is close to zero or negative but that owners—and unpaid family members—are for obvious reasons concentrated in small stores; thus n_2 (and to a lesser extent n_3) is negatively correlated with size (measured by output or total labor input); this "inferior input" behavior creates, in the regression process, identification of small-store inefficiencies with inputs concentrated in small stores. In other words, the lower efficiency of small stores reduces the n_2 (and n_3?) elasticities, giving the false impression that these inputs are less efficient. The simple correlations, r_{n_iY}, of n_1, n_2, and n_3 with size (Y) are shown in Table 1, where the difference in signs is fully expressed.[25]

Since we do want to separate the quality of the various types of labor from the general efficiency of the store they work in we have to find a way to overcome this difficulty. This has been done by assigning quality weights to each type of labor (with the weight of employees defined as $a_1 = 100$) and aggregating the weighted inputs.[26] Thus n becomes labor input in adjusted employee equivalent units. The quality weights for owners (a_2) and unpaid workers (a_3), with a_1 (employees) = 1, are as follows:

	a_2	a_3
Food	102.2	60.7
Clothing	135.0	79.4
Furniture	124.2	60.7

[25]Note also that the returns-to-scale (RTS) parameters in these regressions are close to unity.

[26]For the reasons discussed in the text, Hodgins' approach of introducing n_1/n_2 and n_3/n_1 as variables in equation (1) is also self-defeating. See Hodgins [11, p. 25].

They are based on information on hours worked, education, sex, and age of persons in the three labor categories and on quality weights for these characteristics taken from both of Denison's works, as well as on studies of wage differentials in Israel.[27] The weights are at best crude estimates and the results based on them should be judged accordingly. Also, the use of average weights for the individual stores distorts inter-store efficiency differences. Lastly, if the RTS parameter is biased by the weighting of n_2 and n_3, I believe that the bias is downwards. For one thing, the weights assigned to owners underestimate their real quality or productivity since no account is taken of the managerial and ownership responsibilities which must make them better qualified and more willing to put effort into their work. And any understatement of a_2 biases the RTS parameter in the same direction.

All the equations contain dummy variables for each of the three main cities: Tel Aviv, the country's biggest city and its commercial center (385,000 inhabitants in Tel Aviv proper and about 750,000 in Greater Tel Aviv), and Haifa, the main port and industrial center (212,000 inhabitants), are compared with Jerusalem, the administrative capital (275,000 inhabitants). The three cities may be considered as separate markets for most purposes, so that each may have a different set of prevailing retail-service prices. *A priori*, one would expect the Tel Aviv retail market to be the most competitive, especially in comparison with the more provincial Jerusalem. One can also expect Tel Aviv to be more forward in introducing new sales techniques. If the same prices prevailed in all three cities, this would show up as greater efficiency in Tel Aviv.

Dummy variables for sub-branches were introduced into each branch equation in order to take account of differences in selling conditions and the products sold. In the 1968/69 sample we also distinguish between large and small stores, the dividing criterion being whether the store employs hired labor (large store, *LS*) or only self-employed and unpaid family labor (small stores, *SS*). As the terms suggest, there is almost complete correspondence between size and the employment of hired labor.[28] As mentioned, the whole of the 1967/68 sample falls into the *LS* group.

The general features of the samples are shown by Table 2. The strong association between size and hired labor makes it difficult to distinguish between size and type of ownership, but several other observations may be made:

1. In all branches the size range is very wide, the difference between *SS* and *LS* being particularly marked. Measured by sales, the average large store is from ten to twenty times the size of the average small store. The difference between extreme observations is of course much greater, especially within the *LS* group.

2. Input ratios vary quite widely within branch and size group. If genuine, this variation is certainly sufficient to identify the isoquants, but it does raise doubts regarding the input measurements, whose weaknesses were pointed out

[27] Denison [3] and [4]; Hanoch [10, pp. 35–130]; Klinov–Malul [15]. The weights to be inferred from the last three sources are in general very close to Denison's.

[28] Classification into $n < 3$ and $n \geqslant 3$ would involve shifting 42 stores (out of the total sample of 504) from *LS* to *SS* and 5 from *SS* to *LS*.

TABLE 2

AVERAGE OUTPUT, INPUT, AND INPUT RATIOS, BY BRANCH AND SIZE GROUP: 1968–1969[a]

	Average Output and Input[b]					Output per Worker[c] Y/n	Average Input Ratios		
	Sales	Value Added Y	Employed Persons n	Space k_1	Stocks k_2		n/k_1	n/K_2	k_1/K_2
Food									
T	264	51	3.7	45.3	10.4	7.1 (4.6)	0.112 (0.071)	0.347 (0.634)	3.30 (6.35)
SS	42	9	1.4	17.2	2.3	5.9 (3.4)	0.108 (0.067)	0.322 (0.460)	2.94 (4.02)
LS	753	145	8.8	107.1	28.2	9.8 (5.7)	0.120 (0.078)	0.402 (0.809)	4.08 (9.69)
LS (1967–1968)	497	97	6.8	124.6	19.8	9.7 (5.9)	0.107 (0.110)	0.262 (0.583)	2.09 (4.55)
Clothing									
T	325	83	7.6	60.6	63.7	9.1 (9.3)	0.155 (0.141)	0.093 (0.222)	1.30 (4.70)
SS	37	9	1.8	19.7	8.6	4.8 (2.6)	0.136 (0.118)	0.166 (0.296)	2.45 (6.49)
LS	607	156	13.4	100.7	117.7	11.3 (9.7)	0.174 (0.158)	0.022 (0.041)	0.18 (0.32)
LS (1967–1968)	414	115	10.2	104.3	92.1	9.4 (5.5)	0.141 (0.110)	0.019 (0.035)	0.18 (0.24)
Furniture									
T	454	104	4.9	70.0	70.2	16.3 (20.0)	0.115 (0.086)	0.166 (0.182)	2.47 (4.97)
SS	48	12	1.7	25.9	13.0	7.4 (3.7)	0.093 (0.061)	0.217 (0.220)	2.78 (2.56)
LS	792	180	7.5	106.8	117.9	23.4 (24.5)	0.132 (0.100)	0.121 (0.126)	2.22 (6.34)

[a] T is total sample; SS, small stores; LS, large stores. The number of observations in each group is shown in Table 3. Figures in parentheses are the standard deviations.

[b] Sales, value added, and stocks in *IL* thousand; space in m^2.

[c] Average of Y/n ratios of individual observations.

above.[29] At least part of the variation is due to price differences and short-run rigidities. As regards the latter, the standard deviations in Table 2 show that the input ratios involving the more variable inputs (especially K_2) vary more than those which are less flexible in the short run.[30]

3. Large stores as a rule use more stocks per unit of labor and space than small stores. This can be seen from the size-group averages in Table 2 and is also supported by regressions of input ratios on size (measured by both n and Y). In both cases the variation of the ratio with size is slight and can be explained both by input-price variations with size and by short-run considerations. As regards

[29] Griliches and Ringstad [8, pp. 97–100, 196–197]. Since we have more inputs it is very difficult to estimate the probable resulting biases as they did. Nor can we assume, as they do, that there are no errors in the measurement of N and that there is equilibrium in the labor market.

[30] The high variability of ratios involving K_2 can also be explained by the fact that the K_2 as measured include variation in prices.

the first, such data as we have suggest that wages go up with size faster than does the price of space, and one would also expect this.[31] Short-run considerations lead stores to make short-run changes in the volume of sales by changing inputs in order of flexibility. It can be shown that such behavior tends to create some correlation with size. One exception (not visible in the table) is that *within SS*, labor input declines relative to space as size increases. The explanation may be that small stores have difficulty in varying their labor input or that there are self-imposed restrictions on doing so: on the one hand, it is difficult to reduce labor input to less than one person;[32] on the other, there may be inhibitions about hiring employees as the business grows. As if to compensate for this labor rigidity, stocks increase with size more strongly within *SS* than in the whole sample. In connection with the issue of errors in variables raised in point 2 above, one may add that the variation of input ratios with size explains only a very small part of their total variation and therefore does almost nothing to alleviate the burden of explaining it.[33]

4. Labor productivity increases strongly from *SS* to *LS* in all branches. This result is obtained despite the *decline* (from *SS* to *LS*) in fixed capital per unit of labor in all branches and of stocks per unit of labor in food. The decline in the n/K_2 ratio in clothing and furniture is the only result that helps to explain the productivity trend. Otherwise, the entire explanation should lie in the increasing returns to scale that we expect to find.

5. Interbranch differences in input ratios are fully consistent with one's general knowledge of the various branches. Clothing stores use more workers per unit of space on the average and, in *LS*, more stocks per unit of space and per worker: variety is much greater in clothing than in food and much more service is needed in the process of making a choice. On the average, furniture stores use more stocks per unit of space and per worker than food stores but less than clothing stores, while they use more space per worker than clothing stores. All this seems plausible.

IV. The Results of the Estimation with Inputs in Physical Units

The results of the basic equations (equations I) are presented in Table 3. Almost all the equations exhibit three features. First, the respectable values of $\bar{R}_2$, which are especially surprising considering that the observations are for a cross section and that the estimation is for such an odd product; the exception is *SS*, where $\bar{R}_2$ ranges from 0.4 to 0.6, not low either.[34] The second general feature is the "normality" or reasonableness, in the most general sense, of the estimated equations and parameters: retail trade is no *enfant terrible* and should be regarded like any other product—a conclusion that may be the most important contribution of this inquiry. The third major result is the existence—in both groups and all branches—of increasing returns to scale, with RTS between 1.2

[31]Part of the increase in wages may, however, reflect an unmeasured increase in labor quality. Griliches and Ringstad [8].

[32]Even if it is, this is usually not caught by our data.

[33]The comparison is between the variances of the input ratios and the variance of the residuals of the regressions referred to in point 2. See Griliches and Ringstad [8, p. 54].

[34]See however Griliches and Ringstad [8, p. 64].

and 1.3 in half the equations and even higher in some of them, which supports the initial hypothesis. As mentioned, the RTS parameters are probably somewhat understated.

Let us turn to a more detailed analysis of the results.

Input shares: In order to move from elasticities to relative shares in the case where RTS is considerably above unity, one has to know which factors "lose" relatively more from this. The usual assumption is that labor gets more or less its marginal productivity and its share is equal to the estimated elasticity; capital, being a fixed factor, gets the remainder, which is then much smaller than its elasticity.[35] In our case, however, there are good reasons for assuming that the RTS "burden" is more equally divided between labor and capital. This is so because owner's labor can be considered a residual "claimant" just as fixed capital is; except for the dead portion, inventories are variable and probably get their marginal product. We cannot assume that there is any dead inventory in

TABLE 3

PRODUCTION FUNCTIONS WITH PHYSICAL INPUTS (TYPE I)—REGRESSION RESULTS BY BRANCH AND SIZE GROUP: 1968–1969[ab]

	T		*SS*	*LS*		*LS* (1967–1968)	
Food	(a)	(b)	(b)	(a)	(b)	(a)	(b)
Number of observations	237	237	163	74	74	77	77
Coefficient of							
log A	7.3	7.0	7.2	7.4	7.2	7.9	6.7
	(0.19)	(0.30)	(0.25)	(0.25)	(0.43)	(0.30)	(0.52)
log n	1.07	0.89	0.88	1.04	0.80	0.98	0.78
	(0.07)	(0.07)	(0.17)	(0.10)	(0.11)	(0.11)	(0.12)
log k_1	0.28	0.31	0.30	0.35	0.34	0.32	0.20
	(0.06)	(0.06)	(0.10)	(0.08)	(0.07)	(0.10)	(0.10)
log K_2	0.07	0.11	0.10	0.03	0.06	−0.01	0.17
	(0.03)	(0.04)	(0.04)	(0.03)	(0.05)	(0.05)	(0.07)
D_{611}		0.40	—		0.70		0.48
		(0.20)			(0.14)		(0.23)
D_{614}		0.35	0.14		0.47		0.81
		(0.11)	(0.13)		(0.18)		(0.21)
D_{613}		0.17	0.23		−0.02		0.53
		(0.14)	(0.17)		(0.29)		(0.28)
D_{Res}		0.28	0.22		0.27		0.54
		(0.10)	(0.12)		(0.15)		(0.16)
$D_{\mathrm{Tel\ Aviv}}$		−0.25	−0.38		0.04		−0.17
		(0.08)	(0.10)		(0.10)		(0.12)
D_{Haifa}		−0.23	−0.33		−0.11		0.21
		(0.09)	(0.13)		(0.11)		(0.15)
RTS	1.42	1.31	1.28	1.42	1.20	1.29	1.15
	(7.60)	(4.40)	(1.68)	(7.92)	(2.90)	(3.90)	(1.43)
$\bar{R}^2$	0.83	0.85	0.38	0.92	0.94	0.84	0.88

[35] $S_k = 1 - \eta_n$, where S is the share and η is the estimated elasticity.

TABLE 3—*continued*

	Clothing				Furniture		
	T	*SS*	*LS*	*LS* (1967/68)	*T*	*SS*	*LS*
Number of observations	190	94	96	97	77	35	42
Coefficients of							
log A	6.2	6.8	5.6	5.7	3.6	6.0	4.1
	(0.29)	(0.30)	(0.74)	(0.71)	(0.83)	(0.92)	(1.56)
log n	0.70	0.77	0.50	0.84	0.58	0.58	0.36
	(0.07)	(0.17)	(0.10)	(0.12)	(0.15)	(0.28)	(0.20)
log k_1	0.36	0.40	0.27	0.16	0.15	0.17	0.24
	(0.07)	(0.08)	(0.11)	(0.08)	(0.11)	(0.12)	(0.14)
log K_2	0.20	0.09	0.33	0.27	0.56	0.27	0.49
	(0.03)	(0.03)	(0.07)	(0.08)	(0.10)	(0.12)	(0.18)
$D_{621;\ 630}$	0.12	—	−0.02	0.38	−0.25	0.17	—
	(0.20)		(0.22)	(0.23)	(0.24)	(0.22)	
$D_{622;\ 631}$	−0.12	0.15	−0.28	−0.23	−0.03	0.34	0.44
	(0.13)	(0.20)	(0.19)	(0.16)	(0.21)	(0.21)	(0.34)
$D_{623;\ 632}$	−0.32	−0.17	−0.51	−0.56	−0.32	0.45	−0.34
	(0.12)	(0.12)	(0.30)	(0.22)	(0.19)	(0.27)	(0.22)
D_{628}	−0.05	−0.07	−0.14	−0.21			
	(0.11)	(0.14)	(0.17)	(0.15)			
D_{Res}	−0.28	−0.09	−0.31	−0.25			
	(0.13)	(0.14)	(0.21)	(0.17)			
$D_{Tel\ Aviv}$	0.03	−0.17	0.17	0.18	0.14	−0.24	0.66
	(0.11)	(0.13)	(0.15)	(0.13)	(0.18)	(0.20)	(0.24)
D_{Haifa}	−0.01	0.13	−0.00	−0.08	−0.01	−0.78	0.26
	(0.13)	(0.15)	(0.18)	(0.15)	(0.23)	(0.31)	(0.28)
RTS	1.26	1.26	1.10	1.27	1.29	1.02	1.09
	(4.73)	(1.43)	(1.41)	(3.50)	(2.90)	(0.72)	(5.45)
$\bar{R}_2$	0.89	0.56	0.82	0.84	0.83	0.46	0.81

[a]*T* is total sample; *SS*, small stores; *LS*, large stores. The subscripts for the sub-branch dummy variables (*D*) are as follows:

Food-supermarkets, 611; greengrocers, 613; butchers, 614; other specialized food stores (confectionary, beverages, etc.), Res; these are compared with groceries, 612.

Clothing (first subscript)—men's clothing, 621; women's clothing, 622; children's clothing and haberdashery, 623; footwear, 628; other specialized clothing stores, Res; these are compared with general clothing stores, 620. Department stores, which sell mostly clothing, belong to a different retail branch (65), and are therefore not included.

Furniture (second subscript)—household utensils, 630; electrical and gas appliances, 631; art and gift shops, 632; these are compared with furniture, 633–636.

[b]Figures in parentheses are the standard errors of the coefficients, except in the case of RTS, where they are the *t*-values of the difference from unity.

food, but there may be in clothing and furniture. In Table 4 we present several variants of shares based on these considerations: It seems reasonable to assume that the labor and inventory shares estimated by the regressions are somewhere between variants (i) and (ii), and the shares of fixed capital somewhere between (ii) and (iv). The total capital share is estimated at between (ii) and (v).[36] These

[36]See note *a* to Table 4 for the definition of the share variants.

TABLE 4

VARIANTS OF INPUT-SHARES ESTIMATE

Variant[a]	Food				Clothing				Furniture		
	T	*SS*	*LS*	*LS* (1967/68)	*T*	*SS*	*LS*	*LS* (1967/68)	*T*	*SS*	*LS*
Labor (n)											
(i)	0.89	0.88	0.80	0.78	0.70	0.77	0.50	0.84	0.58	0.58	0.36
(ii)	0.68	0.69	0.67	0.68	0.56	0.61	0.45	0.67	0.45	0.57	0.33
(iii) (5)	0.85	0.96[b]	0.60	..	0.80	1.02[b]	0.57	..	0.67	0.86[b]	0.50
Capital											
(i)	0.41	0.40	0.40	0.37	0.56	0.49	0.60	0.43	0.71	0.44	0.73
(ii)	0.31	0.31	0.33	0.32	0.44	0.39	0.55	0.34	0.55	0.43	0.67
(v)	0.11	0.12	0.20	0.22	0.30	0.23	0.50	0.16	0.42	0.42	0.64
Fixed capital (k_1)											
(i)	0.31	0.30	0.34	0.20	0.36	0.40	0.27	0.16	0.15	0.17	0.24
(ii)	0.24	0.23	0.28	0.17	0.29	0.32	0.25	0.13	0.12	0.17	0.22
(iv)	0.00	0.02	0.14	0.05	0.10	0.14	0.17	−0.11	−0.14	0.15	0.15
Inventories (K_2)											
(i)	0.11	0.10	0.06	0.17	0.20	0.09	0.33	0.27	0.56	0.27	0.49
(ii)	0.08	0.08	0.05	0.15	0.16	0.07	0.30	0.21	0.43	0.26	0.45

[a]The variants are defined as follows:

(i) Elasticity as estimated in Table 3.

(ii) Shares based on equal allocation of the excess of RTS over unity: (i)/RTS.

(iii) S is the proportion of total wages (including imputed) in value added. Following Griliches and Ringstead (*op. cit.*, p. 73), we use the formula

$$S_n = \exp\left(\bar{S} + \tfrac{1}{2}\frac{N-1}{N}G_S^2\right)$$

where $S_t = \log(wn/Y)_t$ and $\bar{S}$ and G_S^2 are the sample mean and variance of S_t and N is the number of observations. This estimate is based on the assumption that the factor shares are distributed log-normally (see also text).

(iv) Shares based on the assumption that k_1 is the only residual claimant, that is (iv) $= 1 - \eta_n - \eta_k$ (η is elasticity).

(v) The total capital share as the sum of k_1 share (iv) and K_2 share (i).

[b]These estimates are heavily based on imputed wages and they are therefore not very reliable.

labor and capital shares seem reasonable and so are the interbranch and intra-branch differences between them. First, labor is in general most important in food (probably with shares of between 0.7 and 0.8) and least important in furniture (probably around 0.4–0.5), with clothing somewhere in between. Accordingly, the share of total capital is highest in furniture and lowest in food, a result reflecting mainly the high share of working capital in furniture (probably between 0.3 and 0.45) compared with clothing (not above 0.2) and food (not above 0.1). For fixed capital, the highest shares are in clothing and food. These results all correspond with what is generally known about the relative importance of the inputs in the different sub-branches. Comparing size groups, labor shares are on the whole higher and capital shares are lower in *SS* than in *LS* (though in food the difference is negligible). This may reflect, among other things, the rigidity of *SS* labor input referred to above. The *LS* "advantage" in capital shares is due entirely to the high share of working capital in clothing and furniture; the shares of fixed capital show no clear pattern of interbranch differences.

There is little that we can add here on whether these estimates are seriously biased and in which direction. Their reasonableness may indicate absence of a net bias. One way of checking is to compare the labor shares estimated from the regressions with those estimated from wage data on the assumption that wages represent value of marginal productivity. Admittedly this is hardly the case here —it was necessary to impute the wages of owners and family workers in order to arrive at the total wage bill.[37] The labor shares computed in this way are presented as variant (iii). In most cases they are above our estimated range but, except for *SS*, not very far from it.[38] This may indicate the existence of a small downward bias in the labor shares estimated by the equations; such a bias may result, among others, from simultaneity (labor inputs depend on wages), and from the neglect of labor quality differentials in the estimations. The capital coefficients, and thus shares, may be somewhat overestimated as a result of the exclusion of management (and possibly also capital equipment), as well as labor quality, from the estimated equations.[39]

Sub-branch differences in efficiency: In food, both supermarkets and specialized stores appear to be more efficient than groceries.[40] Supermarkets are more efficient than groceries by between 50 (in the total sample, *T*) and 100 per cent (in *LS*), butchers (*T* and *LS*) by between 40 and 60 per cent, and other specialized stores (Res) by at least one third; greengrocers appear to be more efficient than groceries only in the 1967/68 sample, which gives higher efficiency advantages to all specialized stores. Since supermarkets are much bigger than groceries and other food stores—there are no food stores as big as the smallest supermarket—in a situation where increasing returns to scale prevail, the efficiency advantage of the supermarket may well be a manifestation of its size. Therefore this point

[37]Lack of space precludes our going into the imputation procedures. See note *a* to Table 4 for labor share (*S*) formula.

[38]See note *b* to Table 4.

[39]Griliches [7, pp. 10–16] and Griliches and Ringstad [8, pp. 92–103, 194–198]. Note that while the biases in the estimated labor and capital coefficients may be independent of each other, those in the share calculated from the coefficients necessarily offset each other.

[40]The relative efficiency advantage is the term $e^{\lambda i}$ in equation (1), where i stands for the dummy variable considered.

is discussed below together with returns to scale. Specialized food stores on the other hand are no larger than ordinary groceries; indeed their dummy variables correlate negatively with size in all groups and their efficiency advantage must be sought elsewhere.

In clothing there are few statistically significant efficiency differences between sub-branches. It is only in haberdashery and children's clothing, which is from 60 to 80 per cent as efficient as general clothing, that we consistently find any relative inefficiency, due apparently to the fact that this is a decaying branch. It was not possible to compare ordinary clothing stores with department stores; had the latter been included in the regressions we might have obtained a higher RTS. Lastly, there are no significant efficiency differences between sub-branches in furniture.

Differences between cities: The only inter-city efficiency difference worth discussing emerges in food, where Jerusalem *SS* (but not *LS*) are more efficient than those in Tel Aviv and Haifa: small stores are only about 70 per cent as efficient in Tel Aviv and Haifa as in Jerusalem; for *T* the rate goes up to around 80 per cent. Size as such cannot explain this difference, nor to my mind can it be explained by the use of modern techniques, since Jerusalem is if anything backward in this respect. It seems that here the explanation must be sought in market structure and prices. Jerusalem stores have suffered least from supermarket competition, especially in food: only 5.2 per cent of Jerusalem food purchases were from supermarkets in 1968 (when there were only three in Jerusalem), compared with 17.7 per cent in Tel Aviv and 13.6 per cent in Haifa. The argument offered is that less efficient stores were able to keep up their prices because there was less competititive pressure than in Haifa and Tel Aviv.

Returns to scale: On the basis of *a priori* assumptions we have concluded (p. 368 above) that the RTS estimates are most likely not biased upward and that they may even be downward biased. Indeed all the inputs missing in the estimated equations—management, indexes for quality of labor and for capital equipment (which may be considered a capital quality indicator)—seem to vary less than proportionately with the included inputs, and thus contribute to an underestimation of RTS.[41]

The estimation of possible biases that originate from the dependence of the labor input on wages or from other errors of measurement (of capital), biases of the kind discussed by Griliches and Ringstad (8, pp. 92–100, 194–198), involves mathematical developments which are beyond the competence of this author. As can be seen in their study, the two input case demands quite complicated procedures and even so depends in its conclusions on a number of strong assumptions. Although we have made several very crude estimates of possible biases and have satisfied ourselves that the orders of magnitude are very small, we shall have to ask the reader to rely mainly on the abovementioned theoretical considerations on this point.

As mentioned, the estimated RTS in food stores is closely tied up with the estimated efficiency advantage of supermarkets. At the other end of the size range it is also tied up with the problem of the relative efficiency of stores owned

[41]See further discussion on biases resulting from differentials in quality of labor in p. 382 below.

by small proprietors; there is almost complete identification in the sample between small stores and small proprietorships, a feature common to all subbranches. The question is to what extent the returns-to-scale parameter represents genuine scale effects and to what extent it reflects differences in ownership, selling methods and so on.

To start with food and the supermarket problem: for the 1968/69 sample, RTS_T is estimated at 1.42 when there is no supermarket dummy variable [Table 3, equations(a)] and 1.31 when there is [equation (b)]. The corresponding figures for LS are 1.42 and 1.20.[42] A simple interpretation of these results is that supermarkets are more efficient than ordinary food (or grocery) stores on two counts: first they are bigger and thus benefit from increasing returns, as estimated by the RTS of equations (b); and second they use modern selling methods and thus are more efficient at a given size—as estimated by λ_{611} of the same equation. According to this interpretation the RTS parameter of equations (a) is biased upwards by the second factor. Unfortunately, however, the identification of "large size" with "being a supermarket" creates a situation in which the estimated RTS and λ_{611} do not necessarily represent the right allocation between size and "pure" supermarket factor. More light may be thrown on this problem when one considers the rate at which RTS changes with size. If one assumes that RTS is constant with respect to size then the comparison of RTS_{SS} with RTS_{LS} [both in equations (a) and (b)] suggests that while there probably is a "pure" supermarket advantage it is smaller than estimated by $\lambda_{611} = 0.70$.[43] This value of λ_{611} is consistent with an RTS which declines with size; the "pure" supermarket effect will vanish only if RTS is assumed to increase with size, which is unlikely since both theory and empirical results support a constant, if not a declining, RTS.[44] Regressions run on the LS group *without* supermarkets and on supermarkets separately provide additional support for the existence of a "pure" supermarket efficiency advantage (results not shown).[45] While in many respects inconclusive, these regressions do suggest that each group has increasing RTS of a magnitude similar to that of the two together. Although the statistical weakness of the data do not allow us to make any firm claim, the efficiency advantage of supermarkets is, if anything, greater than the one estimated by the regressions.

There is a more fundamental—and as yet unsolved—problem than the statistical ones discussed above. This is the question of whether it is possible to increase the size of groceries without changing the selling-mode, that is, without going over to the self-service-open-shelf system and to having most everyday household requisites supplied in the same store. As the obvious answer to this is no, the supermarket must get the credit for increasing store size beyond what is otherwise possible and in this way exploiting economies of scale. The conclusion to be drawn is that while increasing returns do exist in food even

[42]Step-by-step examination of the regressions shows that it is the introduction of the supermarket dummy rather than the others that is responsible for the decline.

[43]Compare $RTS_{SS} = 1.28$ with RTS_{LS}: in equations (a), $1.42 > 1.28$ and in equations (b), $1.20 < 1.28$.

[44]See for example Nerlove [17, pp. 119–125].

[45]At present, the data are not good enough to run separate production functions for supermarkets. This will be attempted later.

without supermarkets, supermarkets may well be responsible for keeping RTS high at large size.

The other branches have lower RTS—much lower than food equations (a) and somewhat lower than food equations (b). The difference is particularly marked in *LS* (RTS is 1.20 for food and 1.10 and 1.09 for clothing and furniture respectively), and for furniture also in *SS*.[46] To the extent that this is a firm result it can be explained by the fact that self-supplied services are much more important to the consumer in food and their proportion in total services can vary much more than in the other branches. Consumers buy food every day and spend more time on buying it than on all other goods together, and they are therefore more sensitive to the difference in economic distance or distribution costs which are the main source of increasing returns.

Within clothing the 1968/69 sample indicates that RTS declines with size, with $RTS_{SS} = 1.26$ and $RTS_{LS} = 1.10$. Though the difference is not statistically significant, the result nevertheless is a reasonable one: personal service is much more important in clothing than in food and the net gain from an increase in size or a move towards self-service is much more limited and tends to diminish as size increases, since the possibility of keeping service on a personal level diminishes and since opportunities for drastic changes in selling methods are limited. The two techniques used to overcome this difficulty—the bargain basement (which is a combination of self-service and discount pricing) and the lease of small areas in big stores to concessionaires—are almost non-existent in Israel.

Although the same arguments should hold for furniture we do not find there the same trend of decline in increasing returns to scale. In this branch returns to scale show up mainly as an efficiency difference between *LS* and *SS* and not within groups: compare $RTS_T = 1.29$ with $RTS_{LS} = 1.09$ and $RTS_{SS} = 1.02$.[47] In the case of furniture this may be due to the heavy concentration of household utensils in *SS* and furniture in *LS*. More generally, it may indicate that *SS* are less efficient because of type of ownership as well as small size.

While this possibility cannot be tested by direct efficiency comparison of small proprietorships and single-employee stores, the sample includes hardly any of the latter,[48] the results for the other branches also support it: In both food and clothing we observe the same "jump" in RTS between *SS* and *LS* indicated by RTS_T being larger than expected on the basis of the scale parameters of the individual groups: in food [equation (b)] RTS_T (1.31) exceeds both RTS_{SS} (1.28) and RTS_{LS} (1.20); true these differences are not statistically significant. And in clothing RTS_T is at 1.26, equal to RTS_{SS}, greater than RTS_{LS} (1.10), and greater than the RTS one would expect when both groups are pooled and there is no

[46]*LS* 1967/1968 is again an exception, with RTS higher in clothing than in food equation (b).

[47]The differences are quite significant statistically. The conclusion is reinforced when it is observed that the intercept of the equation is lower for *LS* than for *SS*. Taking this into account, one would expect to get a much lower "average" RTS for *T*.

[48]Most of the single-employee stores in the country belong to the consumer cooperative organizations; only a few of them are in the three main cities and even fewer are in our sample. Any comparison between them and small proprietorships must take account of their advantage in belonging to a chain.

independent "group effect."[49] The observations that in clothing only RTS_T is significantly different from unity and that in food RTS_{SS} is on the border of significance reinforce the same conclusion. Some of this small proprietorship's inefficiency may be offset by the nonpecuniary benefits that the store owner derives from his independent status.

It should be emphasized, however, that small proprietorship as such is not the only main reason for the increasing returns result; as we have seen, increasing returns occur in all *LS* and—statistical significance excused—in most *SS* groups.

In conclusion it can be said that the RTS_T estimated here may take account of the pure inefficiency of small proprietorships at one end of the size range and, in food equations (a) of the pure efficiency advantage of supermarkets at the other end.

Up till now the scale effect was related to the output of the retail store as a whole. Total output, however, can be looked upon as a product of the volume of goods sold (at pre-trade prices) multiplied by the trade margin $[Y \equiv Q \cdot (Y/Q)]$, and it is the former, the store's turnover, that is responsible for the entire scale effect found. There are no increasing returns to scale of margin. That it should be so can be deduced from the model of consumer behavior sketched in section II above; it also emerges from empirical tests where the scale effect is estimated separately for each of the two components.[50]

In the original study a good deal of effort was invested in trying to correct the estimations for non-measured variations in the quality of inputs, and especially for such variations as might affect the RTS results.

While theory and empirical results suggest that different factors work in different directions (that is, they may contribute to a positive or negative relation between input quality and store size), the balance of the evidence, in general and in this study, favors a positive correlation.[51] Owing to the lack of appropriate data the effort was hardly justified by the results because of the large number of new difficulties raised. I shall therefore sketch only briefly the main problems and results.

In the absence of direct information on input quality we had to resort to input prices and try to distinguish between price differences representing quality variation and those representing (pure) price variation. On the one hand, the use of input values (with output in any case measured in value terms) brings the estimating equation close to identity. On the other hand, wages had to be imputed for self-employed and family labor; clearly any mistake here immediately affects the RTS results. Also, correct prices had to be found for (fixed) capital; here one must be especially careful not to include any locational rents or other payments reflecting size advantage. Municipal tax rates, far from ideal for the purpose, were used as capital prices.

[49]Here again, as in furniture, the intercept either reinforces the conclusions or does not destroy them.

[50]This has been accomplished either by adding, one at a time, Q and Y/Q to the original equations—thus keeping one of them constant—or by estimating two separate production functions each with one component of y as its dependent variable.

[51]See for example Fuchs [5, chapter 6]; Schwartzman [21, chapters 4 and 6]; Cynog [2, pp. 130–131], and Hall [9, pp. 53–57].

Two additional sets of equations were estimated. The first (equations type II), is the same as equations (I), except that all inputs are measured in value terms; the second (equations type III) has input quantities and prices entered separately. In the type III equations, a price elasticity lower than the physical input elasticities is interpreted as indicating that price variations reflect at least some pure price as well as (probably) quality variations, which in turn leads to an underestimate of the RTS in the value (type II) equations, and vice versa. The main results are, first, that even the value equations show increasing returns to scale, though as a rule at a lower rate than the physical equations. Second, the wage elasticities are mostly lower, in the majority of cases significantly so, than the physical labor elasticities (type III equations), at least with respect to labor, indicating that the (type II) RTS coefficients are underestimated.[52]

V. A Concluding Remark

One of the most difficult problems of retail trade is to determine to what extent an increase (usually over time) in sales per worker or, as it should be stated, per unit of combined input, is due to less service or to increased efficiency or to both.[53] Although we do not have a full answer to this question, especially not on time trends, some insight may be gained into changes with size if we compare two measures of sales per input unit while keeping the markup under control. On the basis of the physical equations (type I) define S/x (where $x = n^{\alpha}k_1^{\beta_1}K_2^{\beta_2}$ is a measure of combined input) and S/x^* [where $x^* = x(\alpha+\beta_1+\beta_2)^{-1}$] and α, β_1, and β_2 are taken from equation (b) of the corresponding regression (Table 3 above). x is a measure of combined input that includes the scale factor (RTS is the sum of input coefficients) while x^* is a measure of combined input that does not take RTS into account (the sum of input coefficients is unity). Two linear equations are then estimated for physical inputs:

$$\frac{S}{x} = A + b_1 Y + b_2 \frac{Y}{Q} + u \tag{3}$$

$$\frac{S}{x^*} = A + b_1^* Y + b_2^* \frac{Y}{Q} + u \tag{4}$$

and two corresponding equations for value inputs.

While all b_1^* coefficients are positive and significant—that is, sales per unit of factor inputs increase with size—the b_1 coefficients are either virtually zero (in most cases) or—if positive and significant—represent much smaller elasticities than those represented by the b_1^* coefficients. Increasing returns thus explain most, if not all, of the increase in sales per input unit as store size increases, even after changes in markup are accounted for. The b_2 and b_2^* coefficients are of course all negative and significant. To the extent that developments over time approximate the movement from small to large size and from old to new methods, organization, and ownership, increasing returns as measured here should explain

[52]A full presentation of work done in this direction may be obtained directly from the author.

[53]This is the topic of Schwartzman's study [20, pp. 201–209]. Also see Fuchs [5, pp. 99–107].

much if not all of the commonly observed time trend of increasing sales per input unit.

REFERENCES

[1] Becker, Gary S. "A Theory of the Allocation of Time," *Economic Journal*, LXXV (September 1965), 493–517.

[2] Cynog Jones, T. W. "Wholesale and Retail Trade," in *Manpower Problems in the Service Sector: Supplement to the Report* (Paris: OECD, 1966), pp. 117–34.

[3] Denison, Edward F. *The Sources of Economic Growth in the United States and the Alternative Before Us.* (Supplementary Paper No. 13.) New York: Committee for Economic Development, 1962.

[4] ———. *Why Growth Rates Differ: Postwar Experience in Nine Western Countries.* Washington D.C.: The Brookings Institution, 1967.

[5] Fuchs, Victor R. *The Service Economy.* (National Bureau of Economic Research: General Series No. 87.) New York: Columbia University Press, 1968.

[6] George, K. D. *Productivity in Distribution.* (Department of Applied Economics Occasional Paper 8.) Cambridge: Cambridge University Press, 1966.

[7] Griliches, Zvi. "Specification Bias in Estimates of Production Functions," *Journal of Farm Economics*, XXXIX (February 1957), 8–20.

[8] ——— and Ringstad, V. *Economies of Scale and the Form of the Production Function.* Amsterdam: North-Holland, 1971.

[9] Hall, M., Knapp, J., and Winsten, C. H. *Distribution in Great Britain and North America.* Oxford: Oxford University Press, 1961.

[10] Hanoch, Giora. "Income Differentials in Israel," *Fifth Report 1959 and 1960.* Jerusalem: Falk Project, 1961, pp. 35–130.

[11] Hodgins, Cyril D. "On Estimating the Economies of Large-Scale Production: Some Tests on Data for the Canadian Manufacturing Sector." Unpublished Ph.D. thesis, University of Chicago, 1968.

[12] Israel. Central Bureau of Statistics (CBS). *Trade Survey 1967/68.* (Special Series No. 299.) Jerusalem: 1969.

[13] ———. *Trade Survey 1968/69.* (Special Series No. 341.) Jerusalem: 1971.

[14] Jefferys, James B. and Knee, D. *Retailing in Europe*, London: Macmillan, 1962.

[15] Klinov-Malul, Ruth. *The Profitability of Investment in Education in Israel.* Jerusalem: Falk Institute, 1966.

[16] Mundlak, Yair. "Empirical Production Function Free of Management Bias," in *Readings in Economic Statistics and Econometrics.* Ed. Arnold Zellner. Boston: Little, Brown & Co., 1968, pp. 336–48.

[17] Nerlove, Marc. *Estimation and Identification of Cobb-Douglas Production Functions.* Chicago and Amsterdam: Rand McNally and North-Holland, 1965.

[18] OECD. Manpower and Social Affairs Directorate. Social Affairs Division. *Manpower Problems in the Service Sector: Supplement to the Report.* (International Seminars 1966–2.) Paris: 1966.

[19] Ofer, Gur. *The Service Sector in Soviet Economic Growth: A Comparative Study.* Cambridge, Mass.: Harvard University Press. 1972.

[20] Schwartzman, D. "The Growth of Sales Per Man-Hour in Retail Trade, 1929–1963," in *Production and Productivity in the Service Industries.* Ed. Victor R. Fuchs. (National Bureau of Economic Research: Studies in Income and Wealth Vol. 34.) New York: Columbia University Press, 1969, pp. 201–29.

[21] ———. *The Decline of Service in Retail Trade: An Analysis of Growth of Sales per Manhour, 1929–1963.* (Bureau of Economic and Business Research: Study No. 48.) Pullman, Wash.: Washington State University, 1971.

[22] Sharir, Shmuel D. "Consumer's Buying Behavior and the Structure of Retail Industries." Unpublished Ph.D. thesis, Columbia University, 1970.

[23] Zellner, A., Kmenta, J., and Drèze, J. "Specification and Estimation of Cobb–Douglas Production Function Models," in *Readings in Economic Statistics and Econometrics.* Ed. Arnold Zellner. Boston: Little, Brown, & Co., 1968, pp. 323–35.