Volume Title: Orders, Production, and Investment: A Cyclical and Structural Analysis

Volume Author/Editor: Victor Zarnowitz

Volume Publisher: NBER

Volume ISBN: 0-870-14215-1

Volume URL: http://www.nber.org/books/zarn73-1

Publication Date: 1973

Chapter Title: Appendix G: Regressions of Shipments on New Orders: First Results Based on Data for 1948–58

Chapter Author: Victor Zarnowitz

Chapter URL: http://www.nber.org/chapters/c3567

Chapter pages in book: (p. 718 - 726)

APPENDIX G

REGRESSIONS OF SHIPMENTS ON NEW ORDERS: FIRST RESULTS BASED ON DATA FOR 1948–58

THESE REGRESSIONS use the OBE series compiled before the 1963 revision. The results are inferior to those of the more advanced and complete analysis based on the revised Census data and discussed in Chapter 5. They must be treated with caution but are nevertheless of interest as supplementary evidence.

Relations with Variable Discrete Lags and with Several Lagged Terms

Table G-1 draws on Hyman Steinberg's calculations for the National Industrial Conference Board, which cover the period from 1952 to mid-1957.[1] Columns 1–7 in the first part show the simple correlations between new orders and shipments when the former series are assumed to lead the latter by intervals varying from 0 to 6 months. The highest coefficients, and the leads that yield them, are identified. These coefficients vary from .769 for primary metals to .834 for machinery. In terms of the proportion of the variance of shipments accounted for by new orders (r^2), the corresponding range is .591 to .696 (columns 8–9). As leads longer or shorter than the "optimal" are taken, the correlations decline, but slowly. The declines are continuous, with few exceptions. In contrast to these relatively high correlations, the single

[1] Hyman Steinberg, "Influence of New Orders on Sales," "More on Relating Durables Orders to Sales," and "Relationship Between Ordering and Sales, Part III—Transportation Equipment Industry," *Conference Board Business Record*, September 1957, October 1957, and January 1958, respectively.

r coefficient reported for transportation equipment is as low as .49.[2] This is disturbing even for this highly heterogeneous industry, where particularly large aggregation errors may mar estimates of the relationship between total S and N.[3]

Apart from transportation equipment, the maximum-correlation lags of S in Table G-1 agree fairly well with their counterparts in Table 5-1. According to both tables, two-month lags yield the best results for the total durable goods sector, and zero or one-month lags for the metalworking industries. Three-month lags work best for nonelectrical machinery in Table 5-1 and for total machinery in Table G-1.

When new orders of several past months are used jointly as independent variables (Table G-1, lower panel), the highest partial regression coefficients turn out to be associated with the same lags of shipments as those that yielded the highest simple correlations.[4] This might suggest the existence of certain well-behaved linear relationships between S and N, involving unimodal lag distributions.[5] However, the information provided at this point is certainly insufficient to support such inferences, and perhaps all that can be said here is that the intercorrelations among the independent variables (i.e., the autocorrelations of new orders) are apparently not such as to disturb the correspondence between the correlation and regression coefficients observed for these samples.

The equations in Table G-1 include from two to four N_{t-i} terms

[2] The five-month lag of S behind N is said to maximize simple correlation, and Steinberg does not show the corresponding r coefficients for other lags (*Conference Board Business Record,* January 1958, pp. 23–24). However, working with earlier data, Steinberg obtained the highest r (.42) for a two-month lead of N and observed generally higher correlations for the short leads (of 0 to 3 months) than for the longer leads (of 4 to 6 months), as shown by a graph in the September 1957 issue of the *Business Record* (p. 426). Furthermore, in this book, the highest r for transportation equipment in 1953–65 was obtained for simultaneous timing of N and S; the lowest, for the six-month lead of N (Table 5-1). These coefficients show a relatively wide range, from .772 to .865, but they are all much higher than Steinberg's coefficient for 1952–57.

[3] It will be recalled that in the automotive and nonautomotive parts of this industry the relative importance of production to order is sharply different (see Chart 3-4 and text, Chapter 4, "Timing Differences Among the Major Industries"). The distribution of lags in shipments for total transportation equipment could therefore be bimodal, with short lags prevailing for automotive and long lags for nonautomotive orders. If so, the over-all lag may be quite unrepresentative and unstable. It would strongly depend upon changes in the product mix of the industry group as a whole, and these may at times be large, particularly due to shifts in the weight of the important defense-goods component of transportation equipment output.

[4] See the items on lines 1–5 of Table G-1 that are included in notes b and c and the items on lines 6–10 that are included in note e.

[5] See Chapter 5, "Regression Estimates and Turning-Point Estimates," text and note 40, for a statement of conditions under which the maximum-correlation timing would correspond to the mode of the lag distribution.

Table G-1

Correlations Between New Orders (N) and Shipments (S), Assuming Various Lags, Durable Goods Industries, 1952–mid-1957

Line	Industry [a]	Simple Correlation Coefficients (r) for Assumed Lead in Months of New Orders Over Shipments: 0 (1)	1 (2)	2 (3)	3 (4)	4 (5)	5 (6)	6 (7)	Timing That Maximizes Simple Correlation: Lead of New Orders (mos.) (8)	(r^2) (9)
	A. SIMPLE CORRELATIONS									
1	Durable goods, total	.763	.792	.819 [b]	.812	.798	.778	.741	−2	.671
2	Primary metals	.769 [b]	.754	.693	.709	.696	.587	.583	0	.591
3	Fabricated metal products	.805	.814 [b]	.784	.760	.747	.735	.750	−1	.663
4	Machinery, total	.778	.806	.830	.834 [b]	.831	.823	.804	−3	.696
5	Transportation equipment, total						.49 [c]		−5 [c]	.24

B. MULTIPLE REGRESSIONS OF S_t ON SELECTED VALUES OF N_{t-i} $(0 \leq i \leq 5)$

Line	Industry	Constant Term (1)	Regression Coefficients of N_t (2)	N_{t-1} (3)	N_{t-2} (4)	N_{t-3} (5)	N_{t-4} (6)	N_{t-5} (7)	R (8)	R^2 (9)
6	Durable goods, total [d]	6,103	−.0435	.0117	.3019 [e]	.2700			.929	.863
7	Primary metals	614	.4282 [e]	.2759					.813	.661
8	Fabricated metal products	467	.2116	.2341 [e]	.2133				.856	.733
9	Machinery, total	1,450			.0905	.2062 [e]	.1407	.1678	.886	.785
10	Transportation equipment, total	1,675				.1183	.1318	.1743 [e]	.61	.37

Notes to Table G-1

Source: *Conference Board Business Record*, October 1957 and January 1958 (articles by Hyman Steinberg).

[a] The estimates for total durable goods are based on data for January 1953–July 1957 (see note d). The estimates for the other industries are based on data for January 1952–June or July 1957. Monthly seasonally adjusted series have been used in all these calculations.

[b] Figures denote the highest correlation coefficients and thus also the leads that maximize the simple correlations between S_t and N_{t-i} ($i = 0, 1, \ldots, 6$ months). These leads are listed in column 8, and the squares of the highest correlation coefficients are listed in column 9, lines 1–5.

[c] The results for other leads are not available, but they are said to be worse in terms of correlation between S and N. See note 2 in the text.

[d] This regression equation is based on data for January 1953–July 1957.

[e] The highest regression coefficient in each of the fitted equations.

with selected lags i. These are presumably significant, but the reliability of the regression coefficients cannot be appraised in the absence of calculated standard error statistics. The proportion of the variance of S_t that is statistically explained does increase substantially when two or more N_{t-i} terms are used instead of one (compare the r^2 and the R^2 coefficients in column 9). However, for transportation equipment R^2 is still only .37. And the sums of the regression coefficients vary from 0.424 to 0.704 for the five equations in Table G-1, lines 6–10, thus falling short of unity by large margins.[6]

Applications of a Modified Koyck Model of Lag Distribution

My own first attempts to analyze the distributed-lag relationships between new orders and shipments centered on regression equations of the form

$$S_t = k + aN_{t-j} + bS_{t-1} + u_t. \qquad \text{(G-1)}$$

This is analogous to the well-known model by Koyck (see equations 2–4 and their explanation in Chapter 5), except that N_{t-j} is used instead of N_t, and a constant term k is admitted. The lead j is

[6] Adding up the entries in columns 2–7 in each of the lines 6–10 of Table G-1 gives the following figures: durable goods, 0.5401; primary metals, 0.7041; fabricated metal products, 0.6590; machinery, 0.6052; and transportation equipment, 0.4244.

an estimate of the timing that maximizes the simple correlation between S and N. The (G-1) model would be a logical one in pure production to order, with j representing the minimum period needed for production and delivery and also the "normal" or most frequent delivery lag.[7]

Ideally, if these hypothetical conditions were fully satisfied and the S-N relations were linear and stable on the aggregation levels used, the sum $(a + b)$ and the intercept k would equal, or at least closely approximate, the values 1 and zero, respectively. Actually, Table G-2 shows positive and in some cases large values of k and sums $\Sigma = a/(1 - b)$ that are considerably smaller than 1, except for the paper industry (columns 3 and 7). In principle, $\Sigma(= a + ab + ab^2 + \cdot \cdot \cdot)$ should show the complete cumulative response of S to a unit change in N maintained "forever." Given that N represents net new orders and that both N and S are expressed in the same units (millions of dollars), this "total effect" should equal unity.[8] In fact, the estimates of Σ in Table G-2 vary from 0.373 to 0.982; three exceed 0.8, six exceed 0.6, and two are less than 0.5.

It is clear that these results leave a great deal to be desired. In particular, they are definitely inferior to estimates from Koyck distributed-lag regressions that have the same form as equation (G-1) except that j is taken to equal zero. Thus most of the Σ estimates in Table 5-5 exceed 0.9, and several are not significantly different from 1.0. Also, the constant terms in these regressions are small, in most cases probably not different from zero (see also text in the sections on geometric-lag models in Chapter 5). It seems unlikely that this contrast between the estimates in this appendix and those in Chapter 5 is due to the differences in vintage and coverage between the data used in the two analyses. Rather, the principal reason for the inferiority of the results shown in Table G-2 lies probably in the difference between the models, that is, in the omission of the terms N_{t-i}, where $0 \leqslant i \leqslant j$, from equation (G-1).

Errors from this source should be particularly large where j is unduly high. The eleven-month lag of shipments of nonautomotive transportation equipment presents a drastic case, although such long delivery

[7] See Chapter 5, note 19 and accompanying text. Equation (G-1) is the estimated form of the equation shown in the note.

[8] See "Estimates of Geometric Lag Distributions" in Chapter 5.

periods are undoubtedly quite prevalent in this industry. The four-month lag for primary metals raises doubt in view of the different results obtained elsewhere.[9]

The new-order variables with zero or short leads, N_{t-i}, which are not included in the (G-1) model, may well be correlated with S_{t-1}. Such correlations would cause the estimated coefficients of S_{t-1} in Table G-2, column 5, to be overstated. Indeed, these b coefficients are generally higher than the b' coefficients of S_{t-1} in Table 5-5, column 3. Furthermore, the a estimates of the coefficients of N_{t-j} in Table G-2, column 4, are generally lower than the a' estimates of the coefficients of N_t in Table 5-5, column 2. This presumably reflects the importance of production to stock and for relatively short delivery periods, which is allowed much greater and more direct expression in Table 5-5 than in Table G-2.

A larger value of b indicates that more time is required to account for any given proportion q of the total effect Σ, according to the formula $q = 1 - b^n$, where n is the time interval required. Columns 8–10 of Table G-2 list the values of $n = \log(1 - q)/\log b$, for $q = 0.5$, 0.7, and 0.9. These lags are on the whole much larger than their counterparts in columns 6–8 of Table 5-5. The latter are underestimates, according to our "best" average-lag measures for the corresponding industries, which include the results of second-order distributed-lag functions and the instrumental-variables approach (Table 5-8, column 4). Still, there can be little doubt that the figures in Table G-2 definitely overstate the lags of shipments for the same reason that they overstate the b coefficients. Moreover, according to the lag structure assumed in Table G-2, the count of the n intervals starts from the month $t - j$, the time index of the new-order variable in these regressions. If this were allowed for, the bias of overestimation of the lags would appear still larger.

Despite these deficiencies, some aspects of the results reported in

[9] According to Table G-1, line 2, $j = 0$ for the corresponding series in 1952–57. According to Table 5-1, $j = 1$ for the new Census data on primary metals orders and shipments in 1953–63. These figures do not necessarily imply either inconsistency or calculating errors, but they do indicate that the correlations between S_t and N_{t-i} can deviate little for different lags i in any given period and that the maximum-correlation lags j may depend sensitively on the choice of the period covered. For example, the second highest r coefficient for primary metals in 1948–58 equals .794 for $i = 3$, as compared with $r = .798$ as shown in Table G-2. The maximum-correlation lags in 1953–65 could well be shorter than their counterparts in the earlier postwar years. This would be consistent with some evidence of the turning-point comparisons in Chapter 4 (Table 4-6 and text) and of the U/S ratios in Chapter 6 (Charts 6-4 and 6-5).

Table G-2
Distributed-Lag Regressions of Shipments on New Orders, Eight Major Manufacturing Industries, 1948–58

		Simple Correlations[a] for Selected Lags		Distributed-Lag Equations[b]							
				Constant Term	Regression Coefficients[c]			Sum of Implicit Coef.[d]	Lags[d] Necessary to Account for		
									50	70	90
		Lag j (mos.)	r	k	a	b	R^2		Per Cent of Col. 7		
Line	Industry	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	Primary metals	4	.798	213.30	.2202 (.0681)	.6705 (.0767)	.789	.668	1.7	3.0	5.8
2	Fabricated metal products	2	.853	33.06	.0895 (.0172)	.8887 (.0200)	.952	.801	5.9	10.2	19.5
3	Machinery exc. electrical	4	.854	35.51	.0419 (.0099)	.9438 (.0118)	.983	.745	12.0	20.8	39.8
4	Electrical machinery	2	.815	17.37	.0608 (.0122)	.9298 (.0137)	.976	.866	9.5	16.5	31.6
5	Nonautomotive transport. equipment	11	.572	37.48	.0327 (.0137)	.9323 (.0210)	.952	.484	9.9	17.2	32.9
6	Other durable goods[e]	1	.942	65.55	.0735 (.0147)	.9018 (.0156)	.968	.749	6.7	11.7	22.3
7	Textile-mill products	3	.640	111.82	.0609 (.0204)	.8367 (.0368)	.829	.373	3.9	6.8	12.9
8	Paper and allied products	2	.982	8.41	.3069 (.0128)	.6876 (.0128)	.979	.982	1.9	3.2	6.1

Notes to Table G-2

Note: The data come from the OBE Industry Survey and are monthly and seasonally adjusted; the unit is $1 million for both new orders (N) and shipments (S). They cover the period 1948–58 for all industries, except nonautomotive transportation equipment, where the period covered is 1949–58. The numbers of correlated observations per industry vary from 108 to 131. Adjustments for numbers of observations and constants lower the R^2 coefficients in column 6 only slightly, to figures ranging from .777 to .967.

[a] For zero-order correlation between S_t and N_{t-j}.

[b] Measures in columns 3–10 are based on least-square regressions of S_t on N_{t-j} and S_{t-1}: $S_t = k + aN_{t-j} + bS_{t-1} + u_b$.

[c] Figures in parentheses are calculated standard errors.

[d] In column 7, the sum equals $a/(1 - b)$. For explanation of the measures in columns 7–10, see text. The lags are in months.

[e] Includes professional and scientific instruments; lumber; furniture; stone, clay, and glass; and miscellaneous industries.

Table G-2 are acceptable and instructive. First, new orders taken with leads j retain substantial effects upon current shipments S_t in face of the strong autoregressive terms S_{t-1}. The coefficients a, while small in comparison to the b's, are all significant according to conventional statistical criteria.[10]

Second, the table displays pronounced interindustry differences that are consistent with other evidence. At one extreme, there is nonautomotive transportation equipment (line 5), an industry in which production is predominantly to order, delivery periods are typically long, and highly irregular inflows of orders are translated into relatively smooth outflows of shipments. The correlations between S_t and N_{t-j}, while moderate, are much better for long lags (j of 9 to 11 months) than for shorter lags. Also, the transition from simple to distributed lags improves the association greatly in this case (where $r^2 = .327$, $R^2 = .952$). At the other extreme, the correlations between S_t and N_{t-j} for the paper industry (line 8) already are very high when discrete lags of one or two months are assumed. They are not much increased by the addition of the S_{t-1} term (from $r^2 = .966$ to $R^2 = .979$). The regression coefficients and related lag-distribution measures indicate that the delivery periods for paper products are typically short.

[10] The ratios of these estimates to their standard errors exceed 4 for five industries and 2.38 for all (Table G-2, column 4). By the one-tailed t test, this means that all these coefficients (expected to be positive) differ significantly from zero, at least at the 1 per cent level. However, no conclusive tests can be offered here, because the appropriateness of both the model and its estimation by simple least squares can be questioned (see also the section on "Estimates of Geometric Lag Distributions" in Chapter 5).

Timing comparisons at turning points and ratios of unfilled orders to shipments lead to the same general conclusions about these industries.[11]

Nonelectrical machinery has the second longest delivery periods according to the estimates in Table G-2. Electrical machinery and the metalworking industries have generally shorter lags. For the group of "other durable goods," as for paper, the correlations yielded by simple lags are very high and the gain from using the distributed-lag formula is small. This, it will be recalled, is a group of industries working predominantly to stock. The textile industry shows the second lowest correlation coefficients (both simple and multiple) in the set, but the improvement due to the application of the distributed lag is here large. This is consistent with our earlier inference regarding the highly heterogeneous product mix of this industry, but it also suggests that a substantial proportion of textile output is produced to order with varied but generally not very long delivery periods.

Using the OBE series for 1948–58, the eight industries included in this analysis were ranked according to the average lead of new orders at cyclical turns in shipments. These ranks show a positive correlation with the ranks based on lags j in Table G-2, column 1. (The Spearman coefficient, adjusted for tied ranks, is .720.) They also show a correlation of .762 with ranks assigned according to the sums of j and the corresponding entries in column 8 of the table.[12] The correlation of the latter ranks with ranks based on the U/S ratios for the corresponding industries (see "Backlog-Shipment Ratios" in Chapter 6) is as high as .905. According to all of these various measures, paper, textiles, the other durables group, and primary metals have relatively short delivery lags and low U/S ratios, while fabricated metal products, the two machinery industries, and nonautomotive transportation equipment show increasingly long delivery periods and high U/S ratios.

[11] On the relative timing of cyclical turns in these and other OBE series for N and S, see "Major Industry Aggregates and Their Components" in Chapter 4, with Tables 4-6–4-8. The backlog-shipments ratios are discussed in Chapter 6 with the aid of Tables 6-5 and 6-6. On the particular characteristics of the paper industry, see also Chapter 2, note 15, and the accompanying text.

[12] It may be noted that this result does not depend on the (arbitrary) choice of column 8 ($q = 0.5$): one might just as well have used the figures from column 9 or 10, for example. While the lags n increase with q, the industry ranks according to n are the same for any q.