PART II

Reports on Selected Bureau Programs

SOME IMPLICATIONS OF THE GROWING IMPORTANCE OF THE SERVICE INDUSTRIES

"The economics of tertiary industry," wrote Colin Clark in 1940, "remains to be written. Many, as yet, feel uncomfortable about even admitting their existence." In the last twenty-five years, pioneering studies at the National Bureau by Friedman and Kuznets, Fabricant, Barger, Stigler, Kendrick, and others have done much to change that picture. Our present research, begun in the summer of 1963 with the aid of a three-year grant from the Ford Foundation, is an effort to build on this earlier work and to extend our knowledge of the service sector—defined for this study to include trade, finance, insurance and real estate, personal, professional, business, and repair services, and general government. In particular, we are attempting to measure and analyze output, input, and productivity in these industries.

Our work has been organized along three major lines. First, we have undertaken studies of individual industries, such as retail trade, state and local government, personal services, and health. Second, we have planned studies designed to shed light on problems that cut across all or most of the service industries—problems such as changes in skill mix, changes in hours, and the role of the consumer in the production process. Finally, we are attempting to view the sector as a whole and to contrast it with the rest of the economy.

This last is not to posit some bland homogeneity for all the service industries, or to deny the existence of conflicting trends within the sector. David Schwartzman's study of retailing, for instance, indicates that productivity gains have varied widely from one store type to another, and Jean Wilburn has shown that even two such apparently similar industries as barber shops and beauty parlors can behave

very differently with respect to productivity. We are aware of many *intrasector* differences and intend to discuss them in various parts of the project.

In addition to such differences, however, we find that the service sector differs significantly from the rest of the economy in a number of important ways. Some of these *intersector* differences are discussed in *Productivity Trends in the Goods and Service Sectors, 1929-61: A Preliminary Survey,* Occasional Paper No. 89, published last October. Others will be developed in a second paper, "The Growing Importance of the Service Sector," which is now in preparation. My purpose here is to report a few tentative conclusions concerning the growth of services and to indicate some implications of this growth for the economy and for economic analysis.

1. THE GROWTH OF SERVICES

I begin with the observation that all of the net growth of employment in the United States in the postwar period has occurred in the service sector. As a result, this sector now accounts for more than half of total employment and more than half of gross national product. The U.S. is now a "service economy" —i.e., we are the first nation in the history of the world in which more than half of the employed population is not involved in the production of food, clothing, houses, automobiles, and other tangible goods.[1]

Although the shift of employment to the service industries has been particularly dramatic in the postwar period, it was also in evidence prior to the war, as may be seen in Table II-1 and Chart II-1.

In addition to sector comparisons, Table II-1 presents data by industry group in explicit recognition of the partly arbitrary character of the sector definitions. These definitions arise in part from our interest in a group of industries that have not received much attention in the past from economists concerned with productivity analysis. The boundary between service and goods production is very difficult to draw, and probably no division based on industrial classifications would be completely satisfactory. In practice, however, the basic point concerning the growing relative importance of services would be unaffected by any reasonable changes in definition.

Not only is there a significant difference between sector aggregates, but most individual service industries grew rapidly while most goods industries grew slowly. One-third of the goods industries experienced an absolute decline in employment between 1929 and 1963. If a service and a goods industry are chosen at random, the odds are greater than two to one that employment in the service industry will have grown faster. If we look at "service type" occupations compared with "goods type" occupations, the differential rate of growth is even greater than when the basis of classification is the industry, because service-type occupations have grown relatively, even within goods industries.

This shift of employment to services did not start in 1929. For as long as we have records on the industrial distribution of the employed population, we find a secular tendency for the percentage accounted for by the service sector to rise. Between 1870 and 1930 the differential in rates of growth averaged 1.4 per cent per annum. Since 1929 the average differential has been 1.7 per cent per annum. Between 1870 and 1920, the shift could be explained entirely by the relative decline of agriculture. Since 1920, however, the nonagricultural goods sector has not grown as rapidly as services, and in the last decade has not had any employment growth at all.

This pervasive and persistent trend has generally been attributed to sector differences in income elasticity of demand and productivity. Neither explanation can be tested precisely because accurate measures of real output are not available for many service (and some goods) industries. By making reasonable alternative assumptions about sector differentials in rates of growth of real output, however, we

[1]One dramatic example of this shift is that the *increase* in employment in education between 1950 and 1960 was greater than the *total* employment in primary metal industries in either year.

TABLE II-1

PERSONS ENGAGED BY SECTOR AND MAJOR INDUSTRY GROUP, SELECTED YEARS, 1929-63
(thousands)

	1929	1937	1948	1953	1957	1963
Goods	27,561	25,989	31,764	33,286	32,767	31,445
Service	18,655	21,167	26,812	31,779	33,807	37,962
Agriculture, forestry, and fishing	9,205	8,864	7,012	5,885	5,470	4,725
Mining	1,017	993	1,021	896	858	654
Construction	2,306	1,738	3,262	3,801	4,161	4,305
Manufacturing	10,556	10,686	15,468	17,462	17,054	16,767
Transportation	3,034	2,333	3,000	2,997	2,846	2,546
Communications and public utilities	1,034	901	1,281	1,403	1,514	1,461
Government enterprise	409	474	720	842	864	987
Wholesale trade	1,744	1,857	2,712	2,971	3,205	3,391
Retail trade	6,077	6,305	8,597	9,311	9,775	10,537
Finance and insurance	1,207	1,065	1,349	1,705	2,040	2,437
Real estate	368	455	574	615	681	763
Households and institutions	3,249	3,060	3,051	3,246	3,749	4,316
Professional, personal, business, and repair services	3,235	3,369	4,449	4,780	5,303	6,182
General government (including armed forces)	2,775	5,056	6,080	9,151	9,054	10,336

SOURCE: Office of Business Economics, *Survey of Current Business*, July 1964; *U.S. Income and Output*, 1958; *National Income, 1954 Edition*.

NOTE

Goods = Agriculture, mining, construction, manufacturing, transportation, communications and public utilities, and government enterprise.

Service = Wholesale and retail trade; finance, insurance, and real estate; services; and general government.

have formed some judgments concerning the relative importance of different explanatory factors for the period 1929-63.

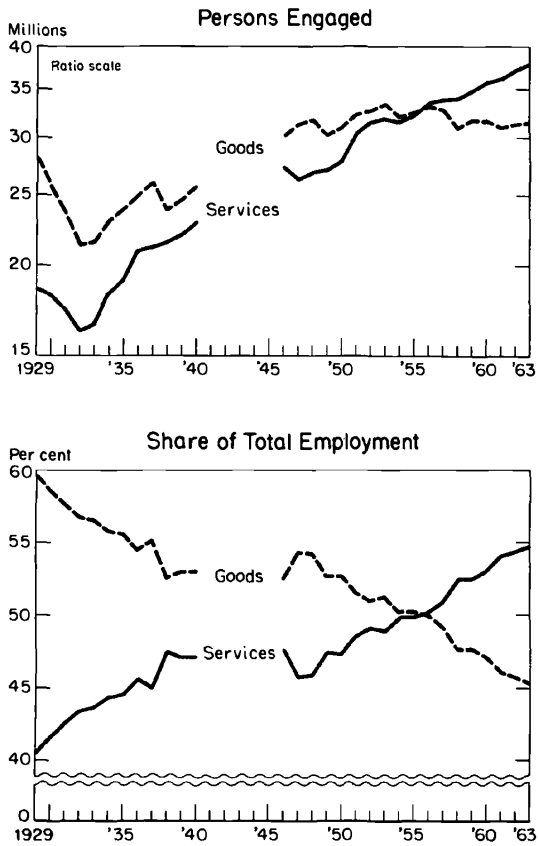
1. The income elasticity of demand may have been slightly higher for services than for goods (but not higher than for goods excluding agriculture), but it is very unlikely that this was a major cause of the shift in employment. Sector differences in the rate of growth of real output were probably very small; differences in the rate of growth of real output per man were probably very large.

2. The differential in the rate of growth of real output per man probably reflects a moderate differential change in productivity in the sense of efficiency in the use of resources, but this is not the only or major explanation. It also reflects a more rapid decline in hours per man in services, a more rapid rise in the quality of labor in goods, and a more rapid rise in capital per worker in goods.

The slow growth of quality of labor and physical capital per worker can be observed in most individual service industries. An interesting instance is state and local government, where Ernest Kurnow finds that there has been no increase in capital per worker since 1929, and little change in labor skill mix.

CHART II-1

Employment, Goods and Service Sectors, 1929-40, 1946-63



SOURCE: U.S. Department of Commerce, Office of Business Economics.

NOTE: See Table II-1 for sector definitions.

The observed differential change in capital intensity and quality of labor are in the same direction as the differential in technological change and may be complementary to it. Alternatively, they may be a response to changes in relative factor prices. Since 1929 the price of labor has tended to rise relative to the price of capital, and the price of unskilled labor has risen relative to the price of skilled labor. Such changes would have tended to alter sector employment shares because the distribution of factors was not the same in the two sectors in the initial period. Moreover, these trends were probably stronger in the goods sector than in the service sector because of the growth of unions in goods but not in services, and because many of the service industries were exempt from minimum-wage legislation.

We plan to undertake work this summer on the elasticity of substitution between skilled labor and unskilled labor and on income elasticity of demand in order to gain some insight into these questions.

2. SOME IMPLICATIONS FOR THE ECONOMY

The shift from primary to secondary production has had profound consequences for every industrial nation; the shift of employment to services may also have significant implications for the economy.

To be sure, such an attempt to look into the future is subject to important qualifications. A shift in the relative importance of different industries is only one of many changes that are occurring simultaneously in the economy, and these other changes may tend to offset the effects of interindustry shifts. Also, these shifts themselves may set in motion changes with implications different from those discussed here. Nevertheless, given the rapid growth of the service industries, it is useful to consider some of the differences between them and the rest of the economy.

LABOR

Several important sector differences in labor force characteristics are summarized in Table II-2. Probably the most significant difference is that many occupations in the service sector do not make special demands for physical strength. This means that women can compete on more nearly equal terms with men, and we find females holding down almost one-half of all service jobs compared with only one-fifth of those in the goods sector. We also find proportionately more older workers in services despite the fact that the more

TABLE II-2

LABOR FORCE CHARACTERISTICS, GOODS AND SERVICE SECTORS, 1960

Row		*Percentage of U.S. Total in*		*As Percentage of Sector Employment*	
		Goods	Service	Goods	Service
1.	All employed[a]	50	50	100	100
2.	Females	29	71	19	46
3.	Over 65	41	59	4	5
4.	Part-timers	41	59	19	27
5.	Self-employed	50	50	13	13
6.	Union members	85	15	48	7
7.	More than 12 years of school	32	68	13	30
8.	Fewer than 9 years of school	63	37	38	22

SOURCE: Rows 1-5, *U.S. Census of Population, 1960;* Row 6, H. G. Lewis, *Unionism and Relative Wages in the United States,* Chicago, 1963, p. 251; Rows 7, 8, *U.S. Census of Population, 1960, 1/1000 sample.*

NOTE: For sector definitions see note to Table II-1.

[a] Data in this table for civilian employment only.

rapidly growing sector would tend to have a younger work force.

An additional reason women and older workers are attracted to the service sector is that it provides greater opportunities for part-time employment. Trade and services, in particular, have employed large numbers of part-timers, and the number has grown appreciably in the postwar period. If data were available on those working part-time *voluntarily,* the difference between the sectors would probably be even greater than that shown in Table II-2.

The situation with respect to self-employment is complex. According to the 1960 Census of Population, the two sectors have approximately equal numbers of self-employed. Agriculture accounts for the lion's share (63 per cent) of the goods sector, while self-employment opportunities in services are widespread throughout the sector, with the exception of government and nonprofit institutions. The Census of Population undoubtedly understates the number of self-employed in services relative to goods, because corporate employees are classified as wage and salary workers, regardless of the size of the corporation. The officers of small, owner-managed corporations are, for analytical purposes, similar to partners or individual proprietors. About three-quarters of such corporations are in the service industries.

It is widely believed that opportunities for self-employment are diminishing in the United States, but if one excludes the decline of agriculture, this is no longer true. In recent years, due largely to the growth of services, the self-employed have grown absolutely and have been a constant fraction of total nonagricultural employment.[2]

The role of self-employment in the future will be determined by several conflicting

[2] See Special Labor Force Report No. 27, "Self-Employment in the United States, 1948-62," John E. Bregger, U.S. Department of Labor, Bureau of Labor Statistics, *Monthly Labor Review,* January 1963, Reprint No. 2410.

trends. A continued shift to service industry employment will tend to favor self-employment, but this may be offset by the influx of young workers and women into the labor force, since these groups are predominantly wage and salary workers. There may also be some tendency toward larger firms within each individual industry, but there is little reason to think that the door to self-employment will be closed as long as services continue to grow.

Given the importance of females, part-time employment, and self-employment in the service sector, it is not surprising to find a vast difference in the importance of unions in the two sectors. The continued growth of services may mean a decline in union influence in the United States.[3]

The last two rows of Table II-2 reveal interesting sector differences in education. The service industries make much greater use of workers with higher education, and make relatively less use of those with only limited schooling. This is not true for all service industries, of course, but it is true for the sector on average.[4]

There is another implication concerning labor which is not readily apparent in the statistics but which is potentially of considerable importance. For many decades we have been hearing that industrialization has alienated the worker from his work, that the individual has no contact with the final fruit of his labor, and that the transfer from a craft society to one of mass production has resulted in depersonalization and the loss of ancient skills and virtues.

Whatever validity such statements may have had in the past, a question arises whether they now accord with reality. The advent of a service economy may imply a reversal of these trends. Employees in many service industries are closely related to their work and often render a highly personalized service offering ample scope for the development and exercise of personal skills.

This is true of some goods-producing occupations as well, but there is little doubt that direct confrontation of consumer and worker occurs more frequently in services. Within each service industry there is some tendency for work to become less personalized (e.g., teaching machines in education, self-service counters in retailing, and laboratory tests in medicine), but with more and more people becoming engaged in service occupations the net effect for the labor force as a whole may be in the direction of the *personalization* of work.

Industrial Organization

The shift of employment to the service sector carries with it important implications for industrial organization in the United States because the size of the "firm" and the nature of ownership and control are typically different in the two sectors.

In goods, with some notable exceptions, such as agriculture and construction, most of the output is accounted for by large profit-seeking corporations. Ownership is frequently separate from management, and significant market power held by a few firms in each industry is not uncommon. In the service sector, on the other hand, and again with some exceptions, firms are typically small, usually owner-managed, and often noncorporate. Furthermore, nonprofit operations both public and private account for one-third of the sector's employment.

In wholesale trade, retail trade, and selected services, accounting for more than 50 per cent of the service sector, half of the employment is in companies with fewer than twenty workers. In finance, insurance, and real estate, 40 per cent is in very small firms. Another large fraction of service sector employment is accounted for by self-employed professionals and domestic servants, representing the ex-

[3]Cf. Leo Troy's NBER paper, *Trade Union Membership, 1897-1962,* in press.

[4]The higher *level* of education of service industry employees should not be confused with the fact that *changes* in the level of education have been greater in the goods sector.

treme in small size of employer. Even government includes many small employers. One-half of local government employment (comprising more than one-quarter of all civilian government employment) is in governmental units with fewer than 500 employees.

One statistic that epitomizes some of the trends discussed is the percentage of the national income originating in business corporations. Ever since the development of the private corporation, its role in the economy has tended to grow, but its relative importance apparently reached a peak in 1955, when corporations accounted for 55.8 per cent of total national income. Since then there has been a tendency for this fraction to decline, and in 1963 it was 53.8 per cent, approximately the same as in 1948.

Other things being equal, the shift to services tends to increase the relative importance of small firms in the economy. There are, however, forces within many industries that tend to increase the size of the average "firm." The pressure for consolidation of school districts and other local government units is a notable example. Bank mergers is another. The net effect of these countertendencies is difficult to predict.

Industries in which small firms account for the bulk of the output typically do not present industrial control problems of the "trust-busting" variety. On the other hand, there may be more need to guard against possible restrictive practices of trade associations and professional organizations. Small firms pose another problem for the economy because they may not allocate sufficient resources to research and other activities with large external benefits.

The growing importance of the nonprofit sector will probably raise some disturbing questions about how to promote efficiency and equity in such organizations. (Cf. the problems with Blue Cross.) When nonprofit operations represent only a minor exception to an essentially private-enterprise economy, the problem is not very serious; but if we ever reach the stage where nonprofit operations tend to dominate the economy, we probably will be faced with the need for radically new instruments of regulation and control.

Demand for Capital Goods

There are some portions of the service sector that use large quantities of physical capital. Real estate and the services provided by government roads and highways are notable examples. By and large, however, goods industries tend to be more capital intensive than services. In recent years (1960 through 1963) expenditures for new plant and equipment in goods industries were approximately three times as great as in profit-seeking service industries; the comparable ratio of output levels in the two groups of industries was only 1.25 to 1.00. Corporate plus noncorporate depreciation charges as a percentage of industry gross product reveal a two-to-one ratio in favor of the goods sector, and balance-sheet data from the *Statistics of Income* also suggest that capital intensity in the goods sector is roughly double that of the service sector.

In pointing out the relatively lower capital intensity of most service industries, I am not attempting to revive a "stagnation" theory in any form. The maintenance of high levels of employment and a rapid rate of growth are logically consistent with a decline in the relative importance of physical capital in the economy. The important point is to recognize that if such a decline occurs because of interindustry shifts, it may be a proper and useful adjustment to new circumstances, with important implications for relative profit levels. While the national rate of savings may be just as high as before, other forms of investment, such as education, that are not customarily included in savings-investment estimates may take on increased importance.[5]

Cyclical Fluctuations

The greater stability of service industries compared with goods has been observed by Daniel Creamer in his study of *Personal Income*

[5]Cf. F. Thomas Juster, "Trends in Consumer Investment and Consumer Credit, 1897-1962," NBER proposed Occasional Paper, in preparation.

TABLE II-3

UNEMPLOYMENT RATES OF WAGE AND SALARY WORKERS BY SECTOR AND INDUSTRY GROUP, 1948-63

		Average 1948-63	1948	1949	1950	1951	1952
(1)	Goods	5.8	3.9	7.3	6.1	3.4	3.0
(2)	Service	3.8	3.4	4.6	4.6	2.9	2.4
(3)	Goods excluding agriculture	5.7	3.8	7.4	5.9	3.4	3.0
(4)	Service excluding public administration	4.3	3.7	5.1	5.1	3.2	2.8
(5)	Agriculture	7.0	4.7	6.5	8.2	3.9	3.9
(6)	Mining, forestry, and fisheries	7.6	2.9	8.5	6.6	3.8	3.4
(7)	Construction	10.1	7.6	11.9	10.7	6.0	5.5
(8)	Manufacturing	5.3	3.5	7.2	5.6	3.3	2.8
(9)	Durable goods	5.3	3.4	7.4	5.2	2.6	2.4
(10)	Nondurable goods	5.3	3.6	6.9	6.0	4.0	3.3
(11)	Transportation and public utilities	3.7	3.0	5.2	4.1	1.9	1.9
(12)	Wholesale and retail trade	5.1	4.3	5.8	5.8	3.7	3.1
(13)	Finance, insurance, and real estate	2.1	1.6	1.8	2.0	1.3	1.5
(14)	Service industries	3.9	3.5	5.1	5.0	3.1	2.6
(15)	Public administration	2.1	2.0	2.9	2.8	1.6	1.1

during Business Cycles, (Princeton University Press for NBER, 1956), and by Geoffrey Moore in an unpublished study of fluctuations in employment during four postwar (1945-61) business cycles. Evidence of the greater stability of services can also be found in unemployment rates of wage and salary workers. Table II-3 presents the average rate by sector and industry group annually for 1948-63, and the average rate for the sixteen years.

We see that unemployment in goods has been consistently higher than in the service sector; the average rates over the period were 6.0 and 3.8 per cent respectively. Much of this differential can probably be explained by a greater amount of *seasonal* unemployment in goods-producing industries. The Bureau of Labor Statistics estimated that in 1957 the unemployment rates for seasonal reasons alone were as follows: agriculture 2.7 per cent, construction 4.2 per cent, manufacturing 1.5 per cent, and transportation .8 per cent, whereas in trade the rate was .6 per cent and in other private services only .3 per cent.

Of greater interest in the present context than the difference in level is the fact that unemployment in goods is more sensitive to business conditions, as may be seen in Chart II-2. The rate for each sector in each year has been plotted as an index number with the sector's average rate 1948-63 equal to 100. We note that the index for goods fluctuates more sharply than for services over the business cycle.

The larger cyclical amplitude of unemployment in goods presumably reflects larger swings in output. One of the reasons for the stability of output in services is the fact that

TABLE II-3 (concluded)

1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	
3.0	6.7	4.9	4.6	5.4	9.2	6.7	6.8	8.3	6.4	6.4	(1)
2.3	3.8	3.4	3.0	3.4	4.9	4.3	4.3	5.1	4.4	4.5	(2)
2.9	6.6	4.8	4.5	5.3	9.2	6.6	6.7	8.2	6.3	6.2	(3)
2.6	4.4	3.9	3.5	3.8	5.5	4.9	4.8	5.9	5.2	5.1	(4)
4.7	8.0	6.4	6.5	6.7	9.9	8.7	8.0	9.3	7.3	8.9	(5)
4.9	12.3	8.2	6.4	6.3	10.6	9.7	9.5	11.6	8.6	7.5	(6)
6.1	10.5	9.2	8.3	9.8	13.7	12.0	12.2	14.1	12.0	11.9	(7)
2.5	6.1	4.2	4.2	5.0	9.2	6.0	6.2	7.7	5.8	5.7	(8)
2.0	6.5	4.0	4.0	4.9	10.5	6.1	6.3	8.4	5.7	5.4	(9)
3.1	5.7	4.4	4.4	5.3	7.6	5.9	6.0	6.7	5.9	6.0	(10)
1.8	4.8	3.5	2.4	3.1	5.6	4.2	4.3	5.1	3.9	3.9	(11)
3.0	5.2	4.3	4.1	4.5	6.7	5.8	5.9	7.2	6.3	6.2	(12)
1.6	2.0	2.1	1.4	1.8	2.9	2.6	2.4	3.3	3.1	2.7	(13)
2.4	4.0	3.8	3.2	3.4	4.6	4.3	4.1	4.9	4.3	4.4	(14)
1.2	2.0	1.8	1.6	2.0	3.0	2.3	2.6	2.7	2.2	2.5	(15)

SOURCE: *Manpower Report of the President,* Washington, 1964, Table A-11.

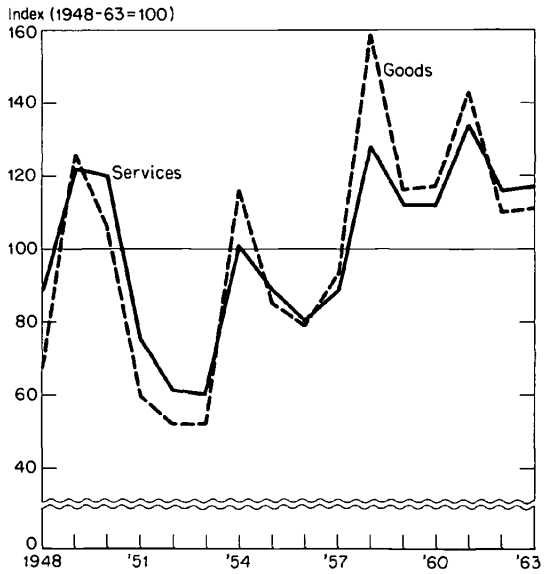
the output cannot be stored. This sector, therefore, is spared the effects of swings in inventory investment, swings which make a major contribution to the cyclical fluctuations of the economy. Similarly, some service industries do not experience cyclical changes in demand comparable to the fluctuations in consumer or producer demand for durable goods. We find that cyclical swings of unemployment in nondurable goods manufacturing industries are less marked than in durable goods manufacturing; they are not as stable, however, as in services even with government excluded.

It is difficult to obtain accurate data on cyclical swings in service industry output as distinct from employment, but in my judgment the amplitude of fluctuation in output is almost certainly greater than that of employment. Thus, inferences about stability, based on employment data, should be tempered in discussing output swings.

Reasons for the discrepancy stem from the nature of the labor force in services. First, there are large numbers of self-employed; their employment is almost completely insensitive to cyclical fluctuation in output. Second, the role of salaried employees, as opposed to hourly workers, is much larger in services than it is in goods. Also, the educational level is higher and the costs of hiring are probably greater. This means that dismissals or layoffs during recessions that are expected to be short-lived will be less frequent. Finally, it should be noted that there is a substantial number of service industry employees classified as "wage and salary workers" who are actually compensated on a "piecework" basis. Their wages, in whole or in part, are determined by their output, and take the form of commissions, tips, or a share of "profits." Employers have little reason to fire such employees when business falls off. This group

CHART II-2

Annual Indexes of Unemployment Rates, Goods and Service Sectors, Relative to Their Average Rates, 1948-63



SOURCE: *Manpower Report of the President*, 1964, Table A-11.

NOTE: See Table II-1 for sector definitions.

includes real estate, insurance and security brokers, waiters and waitresses, barbers and beauticians, and most salesmen of durable goods. Because their earnings are more sensitive to cyclical fluctuations in spending than are their hours of work, we can think of these workers as having "flexible wages."[6]

3. IMPLICATIONS FOR ECONOMIC ANALYSIS

One lesson that our study of productivity in the service industries keeps forcing upon us is the importance of the consumer as a cooperating agent in the production process. To the best of my knowledge, this point is neglected in the analysis of productivity in goods-producing industries, as well it might be. After all, productivity in the automobile industry is not affected by whether the ultimate drivers are bright or stupid, or whether they drive carefully or carelessly.

In services, however, the consumer frequently plays an important role in production. Sometimes, as in the barber's chair, the role is essentially passive. In such cases the only conceptual adjustment called for is to recognize that the time of the consumer is also a scarce resource. But in the supermarket the consumer actually works, and in the doctor's office the quality of the medical history the patient gives may influence significantly the productivity of the doctor. Productivity in banking is affected by whether the clerk or the customer makes out the deposit slip—and whether it is made out correctly or not. This, in turn, is likely to be a function of the education of the customer, among other factors. Productivity in education, as every teacher knows, is determined largely by what the student contributes, and, to take an extreme case, the performance of a string quartet can be affected by the audience's response. Thus, we see that productivity in many service industries is dependent in part on the knowledge, experience, and motivation of the consumer. Consider, for instance, what would happen to service industry productivity in the United States if technology and capital and labor inputs remained as they are, but the consumers were exchanged for 190 million consumers chosen at random from India.

In a similar vein, productivity can be and often is affected by the level of honesty of the consumer. If consumers can be trusted to refrain from stealing merchandise, to report prices and costs properly at check-out counters, to honor verbal commitments for purchases and other contracts, and so on, there can be tremendous savings in personnel on the part of producers of services. These savings are probably important when comparisons are made with productivity in other countries or with the same country at different points in

[6] I am grateful to Jacob Mincer for this formulation.

time. It may be that qualities such as honesty are themselves functions of the general level of productivity and income. A full analysis of productivity, therefore, requires consideration of these interrelations. We expect to begin work soon on the role of the consumer in production.

Another area where the growth of services may require some refinement of concepts is in the analysis of the relation between changes in demand and changes in productivity. In many service industries it is not enough to know by *how much* demand has changed in order to predict the effect on productivity. At least two other dimensions of demand in addition to quantity must be specified.

One source of variation arises because output is frequently uneven, with peaks coming at particular hours of the day, particular days in the week, and even particular weeks in the month. During nonpeak times there is usually idle capacity. An increase in demand, if it occurs at these times, may result in very substantial gains in productivity. On the other hand, an increase in demand, if it occurs at times of peak demand, will probably not result in much or any increase in productivity.

A second source of variation is the "size of transaction," which has received some attention with respect to goods production (especially in articles by Armen Alchian and Jack Hirshleifer), but which seems to be of special importance in many service industries, such as retailing, banking, and insurance. We hope to explore this question in greater detail next year.

My final example of how the growth of services may affect economic analysis concerns the gross national product in constant dollars. This statistic is the keystone of many studies of productivity and economic growth. Unfortunately, it probably is becoming increasingly less useful for such purposes. The reason is very simple. Measures of real output in the service sector have always been unsatisfactory; as this sector becomes more important, the aggregate measure must become less satisfactory in the absence of significant improvements in the measures for individual industries.

Another trend working in the same direction is the decrease in market labor as a fraction of all time spent in productive activity. A small increase in the fraction of the adult population in the labor force has been more than offset by decreases in average hours per week and increases in vacations and holidays. Some of the increased free time may be spent in pure leisure, but probably the bulk of it is spent in the nonmarket production of goods and services and in consumer participation in the market production of services. As I have already suggested, how well or poorly these activities are carried out will surely influence economic well-being. Furthermore, both the output and inputs involved should be included in any comprehensive measure of productivity.

Economists have long been aware that the value of real GNP as a measure of economic well-being differs depending upon the level of economic development. There has been a presumption that the measure becomes more useful the more highly developed the economy is.[7]

Up to a point it is probably true that the higher the real GNP is, the more reliable it is as a measure of economic welfare. But the trend may now be in the other direction because at high levels of GNP per capita a large fraction of productivity effort will be devoted to services (where output is very difficult to measure) and to other activities that are presently not measured at all.

One aspect of the problem emerges from my study of mortality (undertaken because of our interest in the productivity of the health industry). At low and moderate levels of economic development, there is usually a negative correlation between real GNP per capita

---

[7]Cf. Simon Kuznets, "the importance of domestic activities relative to those that are part of the business system declines in the long run." *National Income and Its Composition, 1919-1938,* New York, NBER, 1941, p. 432.

and death rates. Now we have a situation where U.S. GNP per capita is 50 per cent above the Swedish level, but life expectancy is considerably lower in the U.S., and the death rate for males 50-54 is double the Swedish rate. The reasons for this huge difference are not known, but are probably related to the pace of work, diet, exercise, as well as to the productivity of the health industry.

I conclude that even as we increase our efforts to measure real output in the service sector, we must recognize that these efforts are likely to leave considerable margins of uncertainty. Future studies of growth and productivity will probably find it necessary to develop auxiliary measures of economic welfare and "output" to be used in conjunction with the gross national product.

VICTOR R. FUCHS

ON THE ACCURACY AND PROPERTIES OF SHORT-TERM ECONOMIC FORECASTS

Measure what is measurable and make measurable what is not measurable.—GALILEO GALILEI
Discontent is the first step in progress.—OSCAR WILDE

Forecasts of economic conditions aim at magnitudes that are exceedingly difficult to measure or even to define, such as the nation's aggregate output. They also often involve factors of presumed importance which are very elusive, such as the state of "business confidence." When measurement is difficult and estimates are subject to substantial errors, prediction can be particularly hazardous. Impressed also by the complexity of economic processes and the pervasiveness of chance, some would indeed question the very "predictability" of the course of economic events. This, however, misses the real point, which is that predictability is a matter of degree, and that the problem of the forecaster is to seek ways to exploit most effectively all the pertinent information available. Since this precept is easy to recognize but hard to implement, forecasters can ill afford to be perfectionists; however, in general they ought to be, and one suspects usually are, discontented with their products. If combined with a belief that it is possible to improve forecasting methods and results, the discontent is meaningful and potentially productive. It calls for systematic appraisals of the forecasts, which alone may lead to lasting improvements through a process of learning from past errors.

The need for a comprehensive and objective review of the record of their performance appears, in fact, to be generally recognized by economic forecasters of all persuasions. When the National Bureau, with the aid of a grant from several industrial companies, initiated research designed to result in such a review, it received excellent cooperation from sources of the forecasts that were to be compiled and analyzed. Work began in July 1963.

1. BACKGROUND AND SCOPE OF THE PROJECT

The primary purpose of the National Bureau study is to assess the accuracy of short-term forecasts of aggregate economic activity in the United States. This is a broad but separable category of forecasts, in which there is much general interest and which had a large share in the recent expansion of the volume and variety of economic forecasting. The restriction to "short term" means that the forecasts analyzed refer to the next year or two, or a sequence of a few shorter periods, e.g., four

or six quarters. "Aggregate economic activity" refers to several different variables, all comprehensive in scope, such as the gross national product (GNP) and its major components, or the index of industrial production. Forecasts of events of special significance to the nation's economic fortunes, such as the major turning points in general business conditions, are also included. The final restriction is to authentic forecasts from reputable sources, published or unpublished, but in any case recorded before the event. The sources include individuals and small or large groups from each of three spheres that are important in generating forecasts: business, academic institutions, and government. Except for the turning-point predictions, the analysis has so far covered only forecasts stated in specific numerical terms.

In what follows, I shall report briefly on the organization of the project as a whole and at some length on that phase of it for which I am primarily responsible. But, before going into particulars, it is well to recall the general questions that were formulated at the very start of the research venture. Assuming we could assemble a sufficiently large and varied collection of authentic forecasts, and develop an adequate methodology of forecast evaluation, the following are the main substantive problems we chose to explore: How good have the forecasts been? What are the comparative merits and shortcomings of the different assumptions, techniques, and models used? To what extent and in what sense are the forecast errors systematic rather than random? What improvements are feasible to make for better and more dependable forecasts?

2. WORK IN PROGRESS

Two parts of the project are concerned mainly with the accuracy analysis of numerical forecasts. One, under my direction, deals with forecasts made without a formal specification of the underlying model or method. The other, carried on by Jon Cunnyngham, covers forecasts which have an explicit basis in econometric models, i.e., in sets of mathematical equations designed to represent the interrelations among economic variables operating over time.

The accuracy analysis results in a quantitative description of forecasting errors—their magnitudes, types, and structure. The problem is how to measure and evaluate the errors in ways that lead to meaningful inferences about the dependability and usefulness of the forecasts. Jacob Mincer's recent work on methods of forecast evaluation furthered the progress of this phase of the study.

One important factor in the forecasting process is the quality of the data inputs. Rosanne Cole is engaged in a study of the characteristics and effects of data revisions (with particular reference to GNP accounts), in addition to supervising most of the data processing for the project.

Rendigs Fels has been studying forecasters' ability to recognize turning points in the business cycle. His materials include largely qualitative forecasts, which are often difficult to interpret and evaluate because of vagueness or hedging.

Geoffrey H. Moore and Julius Shiskin have conducted research in another phase of the forecasting study, dealing with improvements in the selection and classification of business cycle indicators. They have also investigated the properties and uses of diffusion indexes with different spans.

The rest of this report will concentrate on the aggregative forecasts I have been studying. The reports of my colleagues (see Part IV) provide information on progress in the other areas of the project.

3. A SCORECARD FOR FORECAST EVALUATION

Most economic forecasts are annual and have a seasonal pattern, with a peak at year-end when predictions are made for the following calendar year. These periodic flurries of speculation about the future generate considerable excitement both within the forecasting indus-

TABLE II-4

EIGHT SETS OF ANNUAL FORECASTS OF GROSS NATIONAL PRODUCT, INDIVIDUAL AND AVERAGE ERRORS AND COMPARISONS WITH SIMPLE EXTRAPOLATIONS, 1953-63

	Actual Values of GNP and Errors of Forecasts in Each Year, in Billions of Dollars											Average Error, All Years	
	1953 (1)	1954 (2)	1955 (3)	1956 (4)	1957 (5)	1958 (6)	1959 (7)	1960 (8)	1961 (9)	1962 (10)	1963 (11)	With Regard to Sign (12)	Without Regard to Sign (13)
1. Actual GNP[a]	367.2	357.2	387.2	412.4	434.4	437.7	479.5	503.2	521.3	553.9	585.0		
2. Forecast errors,[b] set A		−10.2	−25.5	−11.4	−6.4	+1.8	−8.6	+7.7	−11.0	+9.3	−17.2	−7.2	10.9
3. B	−7.2	+8.8	−17.2	−10.4	−4.4	+11.3	−4.5	+5.8	−7.3	+12.1	−18.0	−2.8	9.7
4. C						−2.3	−9.1	+8.5	−16.3	+12.4	−12.3	−3.2	10.2
5. D				−8.7	−7.0	+1.0	−17.0	+8.6	−13.5	+0.6	−19.2	−6.9	9.4
6. E	−19.0	+2.0	−25.8	−20.1				+5.3	−10.3	+4.6	−20.0	−10.4	13.4
7. F	−6.2	−5.2	−16.1	−6.4	−6.9	−4.4	−4.5	+6.8	−12.5	+11.1	−9.0	−4.8	8.1
8. G[c]	−12.0	+5.2	−9.6	−6.5	+1.0	+4.4	+3.3	+10.7	+1.8	+12.9	+8.4	+1.8	6.9
9. H		−8.7	−22.7	−12.2	−5.4	+1.1	−16.9	+5.6	−9.0	+6.3	−15.6	−7.8	10.4
Average errors, all sets:													
10. With regard to sign	−11.1	−1.4	−19.5	−10.8	−4.8	+1.8	−8.2	+7.4	−9.8	+8.7	−12.9	−5.0	
11. Without regard to sign	11.1	6.7	19.5	10.8	5.2	3.8	9.1	7.4	10.2	8.7	15.0		9.6
12. Error in extrapolating last year's level	−19.2	+7.7	−26.7	−21.5	−19.7	+2.6	−37.8	−21.1	−16.9	−35.2	−30.3	−19.8	21.7
13. Error in extrapolating past average change	−3.6	+34.2	−14.7	−3.5	−3.8	+19.3	−23.0	−4.3	−0.1	−10.1	−11.4	−1.9	11.6

[a] First estimates by the Department of Commerce for the preceding year, which appear in February. These figures are not comparable from year to year because revisions of prior years are not entered.

[b] The annual figures for sets A, C, D, and E were in some cases obtained by averaging forecasts for the first and second half or for the four quarters of the coming year. Sets D and H are means and sets A and E are medians of a group of forecasts.

[c] These forecasts are typically made in terms of base period prices. For the purpose of these comparisons, they have been converted to current dollars, using actual prices for the forecast year. Since this eliminates the possibility of error in forecasting prices, it probably reduces the error in this set relative to the others.

try and without, among its customers and observers, but the interest in the products of all this activity typically fades away soon. Yet forecasts, to be more useful, should be treated as a crop planted at the opening of the predictive "season" (here a year) and harvested after its close, when the actual magnitudes in question can be estimated and compared with the forecasts, thus permitting appraisal of them.

In Table II-4, this treatment is applied to eight sets of forecasts of GNP for the period 1953-63. There is a column for each year covered, which starts at the top with a figure representing the "actual" level of GNP (line 1, cols. 1-11).[1] Underneath that value are listed the deviations from it of the predictions of GNP for the given year, that is, the forecast level minus actual level (lines 2-9). Below these, averages of the individual forecast errors are shown, with and without regard to sign (lines 10-11). This arrangement enables us to compare the forecasters' performance in and between any of the years. This is instructive because the state and the prospects of the economy vary from one year to another under the impact of both economic and other changes and disturbances, so that forecasters face different problems and difficulties at different times. Furthermore, forecasters communicate with one another, and the result of this communication may show up in the similarity of the errors they make.

The figures at the right end of the table summarize the record of each forecast set over the entire period covered; again, averages of the forecast errors are given here with and without regard to sign (cols. 12 and 13). Thus, reading the table line by line, one can see how a forecaster has performed in any year and compare his individual errors with those of others and with the corresponding averages. Reading down the columns, one can compare the accuracy of different forecast sources in any year and on the average over time.[2]

In addition, the table presents the errors of two simple types of extrapolation which provide common standards for screening the forecasts. The first (line 12) consists in projecting forward the last known (or estimated) value of GNP: the level of the series next year is assumed to be equal to that of this year. The second, which is far more effective (line 13), is an extrapolation of the past average change as it could have been computed from the postwar record of GNP available up to the time the forecast was made.

Table II-4 illustrates a simple and effective way of organizing basic data on forecasts. Makers and users of forecasts can readily adopt such a form to keep a running score of their own or other people's predictive successes and failures. It can suit a variety of needs by application to forecasts of different variables for different time units and periods, along with other types of summary measures of error and benchmark extrapolations. We shall use this "tableau" in the following sections to discuss, first, the over-all record of the forecasters as reflected in their average accuracy in the periods covered, and, second, the short-run variations in their performance.

The results shown in the tableau are based on eight sets of forecasts which will be used throughout this report. The forecasts represent a variety of sources and types. Four are company forecasts which are typically products of small professional teams (coded B, C, F, and G). One comes from a small group of forecasters from various industries, government, and academic institutions (set A). Another set (D) is from a group that presently

[1] These are early estimates which have probably more in common with the data inputs used by the forecaster than the subsequently revised figures. When the latter are used, which amounts to making the forecaster responsible for estimating the future revisions of the data, somewhat larger errors tend to be obtained. The figures in line 1 are not intended and should not be treated as a continuous annual series, because they do not give effect to the historical revisions that render them comparable.

[2] Note, however, that not all of the averages in cols. 12 and 13 relate to the same periods, which impairs their comparability.

includes about fifty business economists, for which both summary measures and forecasts by each individual participant were examined. There is a median forecast (set E) based on a poll of a large number of economists (presently about three hundred). Finally, there is a set of averages compiled from an annual summary of forecasts available since 1954 (set H). Altogether, the table records the efforts of some three or four hundred economic forecasters.[3]

4. SUMMARY MEASURES OF ERROR

Average Errors of Annual Forecasts and Extrapolations

The errors in our collection of annual forecasts of GNP averaged nearly $10 billion disregarding sign (Table II-4, line 11, col. 13). They appear small—about 2 per cent—when compared with the average level of GNP, but they are big enough to make the difference between a good and a bad business year. The average year-to-year change in GNP over the period 1953-63 amounted to about $22 billion (line 12). Thus the errors were, in terms of absolute averages, not quite one-half the size of those errors that would be produced by assuming that next year's GNP will be the same as last year's.

Furthermore, the errors of these forecasts were, on the whole, not much smaller than those which would have resulted had the forecasters simply assumed that GNP would advance next year by the average amount it had advanced in preceding years. The average error on such an assumption would have been nearly $12 billion (line 13). This exceeds the corresponding figures for all but one forecast set, but often by rather slight margins (column 13).[4]

Errors of Levels, Changes, and Base Values

The simplest measure of error is obtained by comparing the predicted and the actual levels, but it is perhaps more important to compare predicted and actual *changes*. The error in the change will be the same as that in the level, when the actual level at the time of the forecast is known. As a rule, however, this is not so and the two will differ by the amount of error in the estimated current position, which is the base of the forecast.

The errors in the estimated current position are, of course, typically smaller than the errors in forecasts proper, since the present is generally better known than the future. However, they are by no means negligible. They average about one-fourth or one-third of the corresponding errors of the level forecasts for GNP, and from 12 to 50 per cent for industrial production.

For GNP, the errors of change forecasts tend to be smaller than those of level forecasts. The latter, as shown in Table II-4, average about $10 billion when taken without regard to sign; the former average about $9 billion, or 10 per cent less. For industrial production, there is less regularity in the relation between the level and the change forecasts because errors of base and of level often differ in sign.

Chart II-3 compares the actual year-to-year changes in GNP and industrial production with the changes forecast by one of the better, but not unrepresentative, forecast records. Clearly the pattern of predicted changes is broadly similar to that of the actual changes. This is true of each of the eight forecast sets examined here.[5]

---

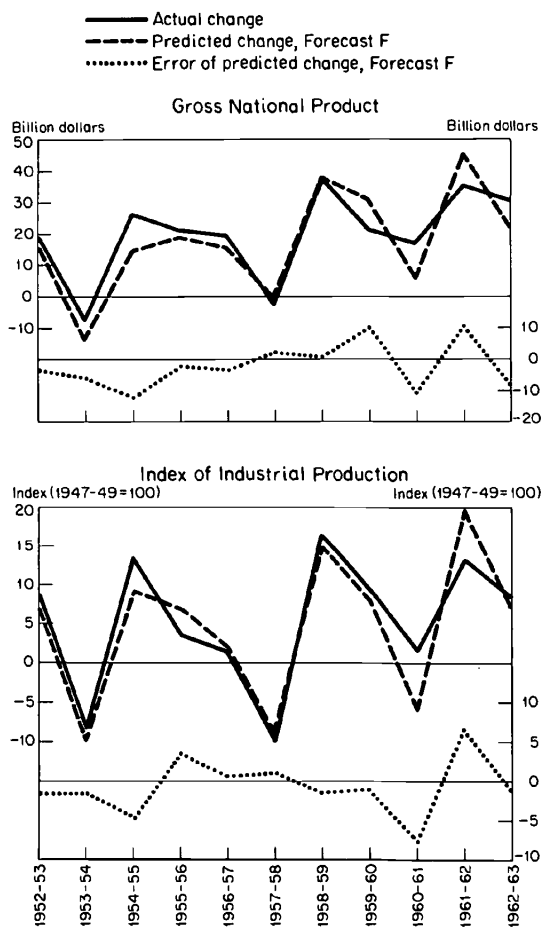
[3]Some forecasters may be in more than one group.

[4]An extrapolation of the average historical change produces in effect good trend estimates for series that tend to grow smoothly, such as postwar GNP. Such extrapolations prove then to be a rather severe, if simple, yardstick. They would be less adequate for economic magnitudes that fluctuate more widely, such as unemployment.

[5]The significance of this relation will be better appreciated when compared with the definitely inferior results obtained for extrapolations. Thus the use of the average historical change, as described above, produces relatively good projections for GNP, in terms of the over-all average errors, but predicted change here is nearly constant from one year to the next. Graphically, it would approximate a straight line cutting through the "actual" GNP changes in Chart II-3 at the $20 billion level.

CHART II-3

Actual and Predicted Changes and Errors, Gross National Product and Industrial Production, 1953-63



NOTE: Actual changes are based on first estimates published by the source after the end of the forecast year.

Underneath the paired curves for changes, Chart II-3 shows the discrepancies between them, i.e., the errors of the change forecasts. These are on the whole much smaller than the corresponding changes themselves: as already noted, the forecasts are typically better than simple extrapolations of last year's level, which assume no change and therefore produce errors equal to the actual changes.

It should not be inferred from this chart that changes in GNP and industrial production are effectively forecast a full year ahead. As noted below, our records of quarterly forecasts suggest that forecasters typically achieve a considerable measure of success in forecasting the next one or two quarters, but limited success beyond that. The annual record can likewise be interpreted as implying an effective forecasting range of no more than two quarters. Not only are the forecasts typically made late in the preceding calendar year, but a record of accuracy very similar to that shown here could be attained merely by accurately forecasting the first two quarters of the ensuing year and assuming no change beyond that.

THE UNDERESTIMATION BIAS

The tendency of forecasters to underestimate changes, which has been observed elsewhere,[6] is evident in most of our data. Thus the mean changes in actual values exceed those in predicted values for all but one of the forecast sets examined; the actual changes in GNP average about $22 billion and the predicted changes $19 billion.

When underestimation applies to increases in the predicted series, then the future levels of the series will also as a rule be underestimated. (This follows necessarily if the current position is known or itself understated.) On the other hand, underestimation of decreases would tend to result in overestimation of levels. In series with upward trends, such as GNP or industrial production, increases are more frequent than decreases, so that the levels are understated most of the time. (Note that all but one of the average errors in Table II-4, col. 12, are negative.)

However, in the GNP forecasts it is primarily the increases that were underestimated, not the decreases. Recent recessions have been sufficiently mild and short to cause only very small declines, or even merely retardations of growth, in the *annual* values of GNP. According to early estimates, declines occurred in

[6] H. Theil, *Economic Forecasts and Policy*, Amsterdam, 1958, Chaps. III-V.

1953-54 and 1957-58.[7] On the former occasion, most forecasts showed larger declines than actually occurred; and on the latter, most of them missed the downturn. In the years when GNP increased, on the other hand, underestimates were almost twice as frequent as overestimates. Industrial production declined much more in the recessions than did GNP, and its contractions were in fact underestimated by most forecasters (though a few greatly overestimated the drop in output in 1954).

If underestimation of changes was merely the result of failure to forecast irregular or random variations in the actual values, it would not constitute an error or "bias" that forecasters could or should guard against. Forecasting errors that can be traced directly to short random movements must be regarded as unavoidable: a sudden outbreak of war or a strike started without warning can hardly be events that an economic forecaster should be expected to predict (though his job certainly does include an evaluation of the effects of such events, once known, on the course of the economy). To put it differently, in principle the requirement of a good forecast is that it predict well the systematic movements of the given variable—not that it predict perfectly the actual values of the latter. It could not do that, except by accident, for economic series where random elements are virtually always present.

However, the observed underestimation is, in fact, partly systematic. This is indicated by the strong prevalence of negative signs among the errors of predicted changes: the means of the change forecasts are, with few exceptions, smaller than the means of the actual changes.[8] Also, there is direct evidence that not only the erratic changes but also the systematic cyclical movements in comprehensive economic aggregates are typically underestimated (see Table II-5, below).

Larger Errors of Longer-Span Forecasts

Five of our forecast sets provide chains of predictions made at a given date for two or more successive periods, e.g., for the first and second half, or the four quarters, of the coming year. These data have been used to analyze the performance of forecasts over different intervals between the current base and the future target (which we call the *span* of the forecast).

The evidence shows clearly that the accuracy of short-term forecasts diminishes as the span increases. Both levels and changes are predicted better over the next three months than over the next six, and better over six months than over nine or twelve. Only a few examples of this highly regular and pervasive relation can be given here, but they are representative and telling. In a semiannual forecast of GNP for 1955-63 by a rather large group of business economists, the mean absolute errors of change are, for a six-month span, $6.7 billion; for a twelve-month span, $12.3 billion. In a quarterly forecast by the staff of a large company for the same years, the average errors, in billions of dollars, are for one quarter 5.0, for two quarters 8.4, for three, four, five, and six quarters 11.9, 13.7, 15.5, and 17.1, respectively. Again, in a semiannual forecast of the Federal Reserve Index of Industrial Production for 1947-63, the mean absolute errors are 3.6 index points (1947-49=100) for six, and 7.1 points for twelve months. Similar progressions in the size of errors appear also in forecasts of other variables, such as the components of GNP.

In general, extrapolations of various types also tend to worsen with the lengthening of the predictive span. Since many forecasters use such extrapolative techniques, this is presumably a part of the explanation for the observed association between longer forecasts and larger errors. Other ingredients of forecasts—vari-

[7] Subsequent revisions reduced the decline of GNP in 1953-54 and converted the small decline in 1957-58 into a very small increase.

[8] Smoothing out the irregular movements would simply tend to make the *variance* of predicted changes less than that of actual changes; averages of predicted and of actual changes, taken with regard to sign, should not be systematically different.

ables that typically move in advance of the predicted series, anticipations data, and the forecaster's own judgment—also are probably decreasingly reliable over longer spans, hence contribution to the deterioration of the resulting forecast.

While the annual and shorter forecasts are generally superior to extrapolations, those with longer spans (say of twelve to eighteen months) are often worse than the more effective types of extrapolation, including projections of the average historical change as used in Table II-4.

5. SHORT-RUN VARIATIONS IN FORECASTERS' PERFORMANCE

Year-to-Year Comparisons of Forecasting Errors

It is often said that forecasters have their "off years"—periods in which they generally go wrong because of some difficulties inherent in the type of economic change that occurred at that time. In other periods, it is believed, no such special difficulties exist, and forecasters tend to be "right."

A year-by-year survey of forecasters' performance will help appraise this view and its implications. Our records show that the increases in both GNP and industrial production were underestimated most in the boom year 1955. Consequently, the level of economic activity was also underestimated most in that year (Table II-4, line 10). Errors of the same kind were made also in 1961, following the latest recession, and in 1963, when the economy showed unexpected vigor after a retardation. On the other hand, in the recession years 1958 and 1960, the predicted levels of GNP and industrial production were too high, either because the changes were understated or because no declines were expected. The retardation of 1962 was widely missed, with the result that the forecast levels were again too high. Finally, the 1954 recession, as already noted, turned out to be milder than many forecasters apparently anticipated, particularly in terms of GNP.

It should be noted that the timing of recent business cycle downturns was early enough to make the presence of the recession rather generally known by the end of the peak year (1953, 1957, and 1960). This, plus the presumption that the contractions would continue to be short, made the task of predicting annual changes relatively easy.

Another fact, illustrated in Chart II-3, is the correspondence between the year-to-year changes in the mean forecast errors of GNP and in those of industrial production. The patterns of changes have actually been quite similar in the two variables, and the forecasts we have reviewed reflect this similarity.

Forecasting and Business Cycles

The year-to-year comparisons of the preceding section suggest that predictive errors are affected by the cyclical characteristics of the forecast period. Table II-5, which utilizes forecasts for quarters and half-years, demonstrates this still more clearly. The levels of GNP and industrial production are underestimated most in the first year of expansion, when the increases in these series are very large. Later in the expansion, when the increases are usually smaller, the levels are underestimated much less; and they may even be overestimated, as happened in the "retardation" of 1962. In contractions, overestimation of levels is the rule, sometimes because the downturn is missed and sometimes because the decline turns out to be larger than predicted.

To understand how such results might come about, it may be helpful to distinguish two predictive patterns, both of which would underestimate current cyclical changes, though in very different ways. First, imagine a series of forecasts which traces the fluctuations in the actual series, but with a short lag. The observed cyclical amplitudes need not be underestimated, but the current changes at certain stages of the cycle will be. Second, imagine a series of forecasts which faithfully reproduces the trend, but not the cyclical movements, of

TABLE II-5

AVERAGE ERRORS IN FORECASTS OF GNP AND INDUSTRIAL PRODUCTION DURING THREE STAGES OF BUSINESS CYCLES, 1947-63

CODE OF FORECAST AND PERIOD COVERED	Variable Predicted and Unit	MEAN ERROR IN FORECAST LEVEL DURING		
		First Year of Expansion	Rest of Expansion	Contraction
A—1948, 1955-63	GNP ($ billion)[a]	−19.2	−4.4	+1.0
C—1958-62	GNP ($ billion)[b]	−14.0	+2.9	+4.6
D—1947-62	Indus. prod. (1947-49=100)[c]	− 7.1	−4.4	+3.2

[a] Forecasts for the first and second half of the coming year.

[b] Forecasts made quarterly for sequences of up to five quarters ahead.

[c] Forecasts made twice in the year for two or three semiannual periods ahead. The forecasts for the eighteen-month spans cover the period from 1948 through the first half of 1956.

the actual values. The predictions simply cut across the fluctuations, so that here both the cyclical amplitudes and the current short-period changes are underestimated.

Some forecasts look more like the lagging cyclical model, others the trend-projection model. In Chart II-4, the former is illustrated by forecast C, and the latter by forecast G, for GNP during 1958-61. Both these sets consist of forecasts made for several quarters ahead, which in the chart are linked together into chains that fan out to the right from points representing the forecasters' estimates of the current position. The forecasts are made twice or four times a year, and hence the chains overlap. The C chains have "kinks" in 1958 and 1960-61, which lag the turning points in GNP; the G chains show no kinks at all.

Forecasts of the trend-projection type, such as G, come out rather well in terms of measures of average errors, largely because they are relatively optimistic and avoid the large underestimation bias of the "cyclical" forecasts. It is clear, however, that forecasts such as these must be regarded as failures with respect to the recognition of turning points. The chain forecasts with cyclical patterns can be more useful in this role, even when they are tardy.

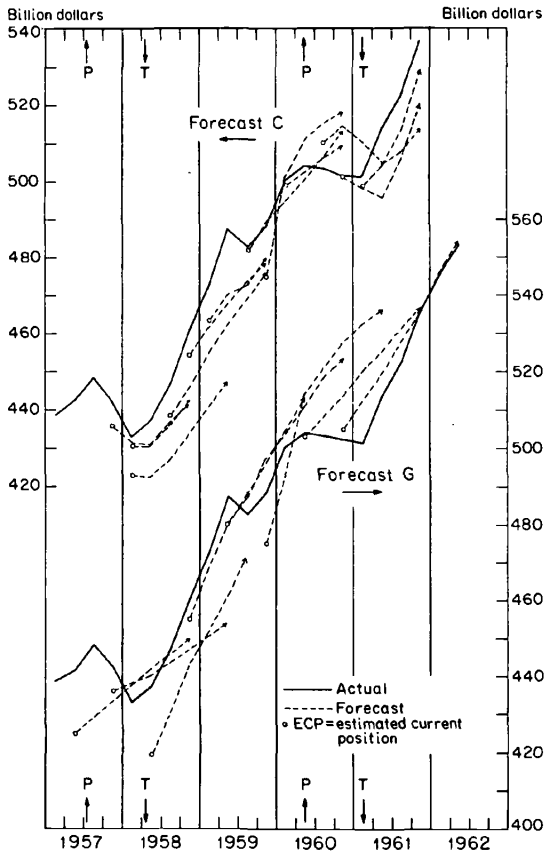
6. USES OF FORECAST EVALUATIONS

The services that forecasts are expected to render depend on the needs of the user and are not easily defined by an outside observer. Surely, however, the usefulness of a forecast is in the first instance a function of its accuracy. With the growth of public interest in the expanding activity of economic forecasting, there is increasing need for objective and comprehensive evaluations of forecasting accuracy—a need as yet largely unsatisfied. Our study will help to fill it by working out and applying systematic methods of testing short-term economic predictions.

Makers of forecasts, as well as users, have a vital interest in a review of their record of performance. Let me give two examples of

CHART II-4

Two Sets of Chain Forecasts of GNP for 1958-61 and the Corresponding Actual Values



NOTE: P and T represent business cycle peaks and troughs, respectively.

simple improvements in forecasting methods and results that have so far been suggested by our study. First, we have found a great deficiency in forecasters' records of the current position at the time of forecast. These records are essential to a proper appraisal. Second, better utilization of the historical content of the series is also needed. It seems desirable that, at some stage, trend projections should be incorporated in the forecasting process, since our study shows that many forecasts looking ahead four quarters or more are inferior to simple trend extrapolations. As our work proceeds, we expect that advantages and disadvantages of different forecasting techniques will come increasingly to the surface.

The general economist, finally, can benefit from a study of the properties of forecasts, since they throw light upon the formation and revisions of economic expectations. With this in mind, we are giving considerable attention to forecasts of chains of short periods that partially overlap in time, since such materials are particularly informative here. Also, our study reveals a large degree of interdependence, though often a wide dispersion, among forecasts made at a particular time (cf. Table II-4). This has important implications for theories of economic change.

VICTOR ZARNOWITZ

CONSUMER CREDIT MARKETS: A PROGRESS REPORT

The current consumer credit study reflects the National Bureau's continued interest in this segment of the credit market. Until the first Studies in Consumer Instalment Financing were published by the Bureau in the 1930's, little was known about the nature and operations of institutions serving the consumer's credit needs, only scattered statistical estimates measuring the volume of credit existed, and little had been done to ascertain the risks and costs inherent in consumer credit operations. Since the 1930's, consumer credit has grown enormously in absolute and in relative terms. The present study was designed to reexamine some of the results of the earlier study in the light of the additional experience and statistical data available; to extend the research into subjects, such as rates of charge, which

had been treated sparingly if at all; and to assess the significance of consumer credit developments in the changed economic circumstances of today.[1]

The work on the individual studies has begun to produce tentative answers to various questions.[2] How rapidly have consumer credit markets grown in importance to the national economy? How do these markets function? What rate and cost relations exist among the institutions serving consumers? What is the nature of the demand for consumer instalment credit? Do consumers know the levels of rates of charge and can they discern differences in rates of charge? What kinds of credit information are given consumers and what procedures are necessary to make them comparable? Where do the funds lent to consumers originate and are these sources affected by changing monetary conditions? Finally, does access to consumer credit play a stabilizing or destabilizing role when unemployment occurs?

CONSUMER CREDIT AND THE NATIONAL ECONOMY

F. Thomas Juster's preliminary findings on trends in consumer investment and consumer credit between 1897 and 1962 indicate that instalment credit has become increasingly important to the national economy through its contribution to the rapid growth of the stock of capital owned by consumers—automobiles, kitchen appliances, and so forth. Chart II-5 shows the rapid rate of growth of two forms of household capital (dwellings and major durable goods), and also indicates that both long-term mortgage debt and short-term instalment debt grew even more rapidly. The implication is clear that the household sector has increasingly relied upon credit to finance a growing stock of consumer capital.

Juster points out that gross additions to consumer capital have been a rising fraction of gross national product, while gross additions to capital owned by business enterprise have been a declining fraction (Chart II-6). These differential trends necessarily imply a shift in the composition of gross fixed capital formation. In the decade from 1899 to 1908, his figures show, business enterprises accounted for roughly two-thirds of gross fixed capital formation, households for about one-quarter, and government for about one-tenth. By 1949-58, the share of the enterprise sector had declined to about 40 per cent of the total, the household sector accounted for nearly one-half, and the government share had risen to 15 per cent (excluding military capital expenditures). Gross capital formation is now larger in the household than in the enterprise sector.[3]

The relative contribution of the household and enterprise sectors to cyclical fluctuations in capital goods output has changed as well. From the turn of the century to the beginning of the First World War, short-run variations in the output of capital goods purchased by business enterprises were four to five times as great (in dollar magnitude) as those in consumer capital formation. In the post-World War II period, the variations in output of capital goods purchased by households were 50 per cent larger than the corresponding fluctuations for the enterprise sector.

Marked changes in credit markets and financial institutions have accompanied these structural changes in the nature and composition of capital formation. There has been an almost uninterrupted liberalization of the credit terms made available to consumer borrowers as the market came to embrace larger and larger segments of the population. Although the rates

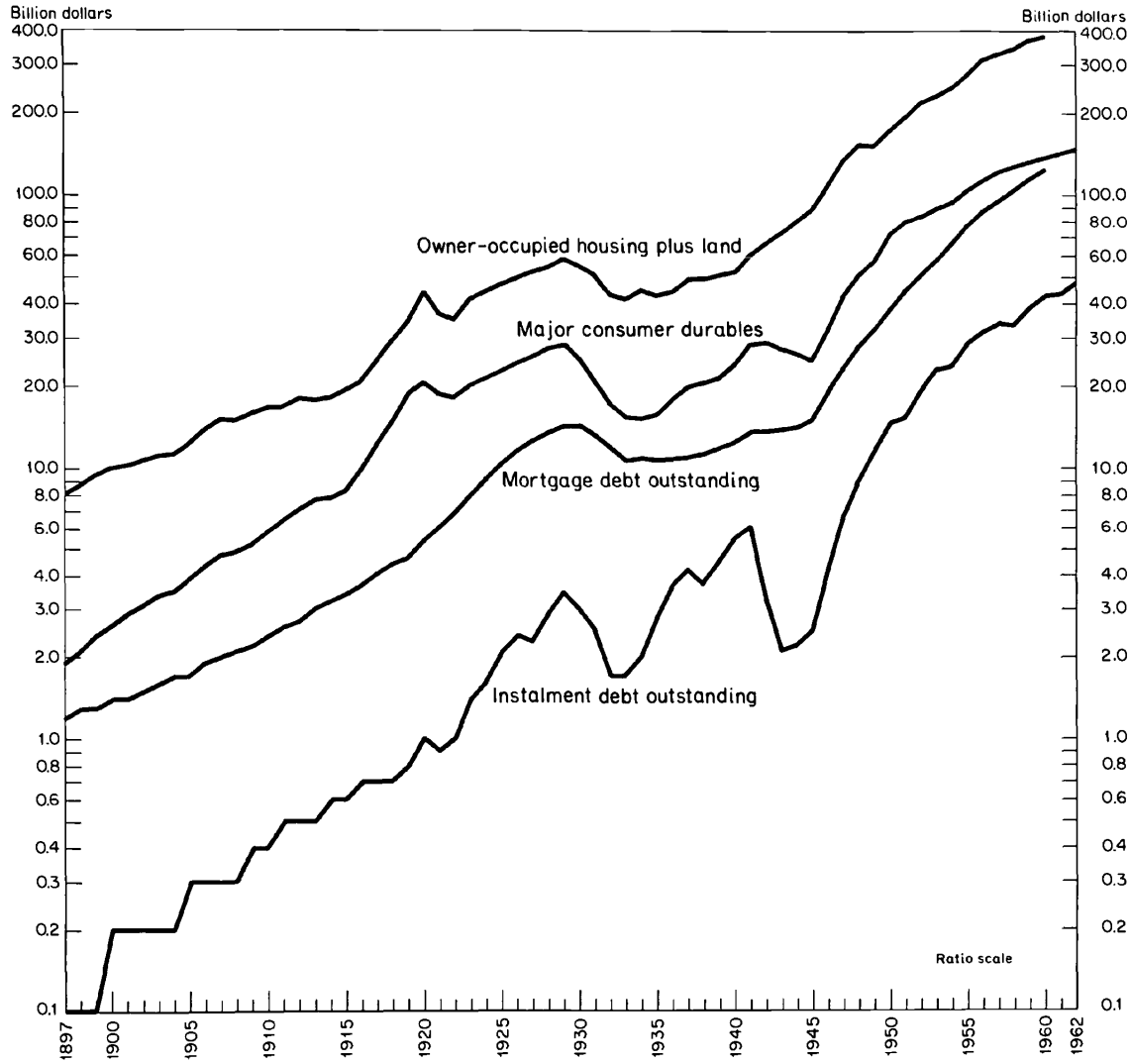
[1]The current study was made possible by grants from four large sales finance companies. We have benefited from the assistance of an advisory committee under the chairmanship of Paul W. McCracken, University of Michigan.

[2]At the present time one book and four Occasional Papers have been published, a second book and fifth Occasional Paper are in press, and four other reports are in preparation. For a list of these reports and further detail on the studies in progress, see Part IV, section 4.

[3]It does not necessarily follow that net capital formation is also larger in the household sector.

CHART II-5

Major Types of Household Tangible Assets and Related Debt, 1897-1962



SOURCE: F. Thomas Juster, "Trends in Consumer Investment and Consumer Credit, 1897-1962," NBER, in preparation.

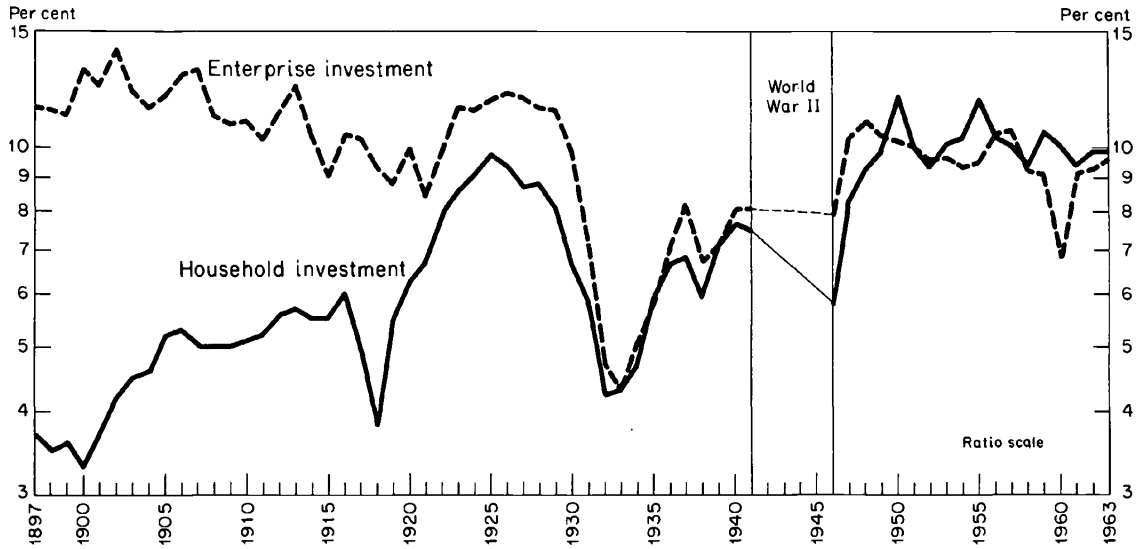
of charge on consumer debt have apparently declined over the decades (see below), our studies suggest that the change in credit terms, rather than lower finance rates, was the more important element in the rise in borrowing.

THE FUNCTIONING OF THE AUTOMOBILE CREDIT MARKET

A large portion of our investigation of how consumer credit markets function has been devoted to the automobile market, partly because

CHART II-6

Gross Investment by Households and Business Enterprises, 1897-1963



SOURCE: Same as Chart II-5; 1963 estimated.

of its importance and partly because data concerning the price of credit, which we defined as the finance rates, were more readily available here. A first step was to construct estimates of finance rates on automobile credit contracts. Three large samples of sales finance company contracts with reasonable geographic coverage were used as benchmarks, covering 1935-38, 1954-55, and 1958-59. My portion of the work was to construct time series estimates of finance rates and to analyze factors affecting rate levels. Wallace P. Mors, using the same data, has studied the effects of state legal rate ceilings upon finance rates.

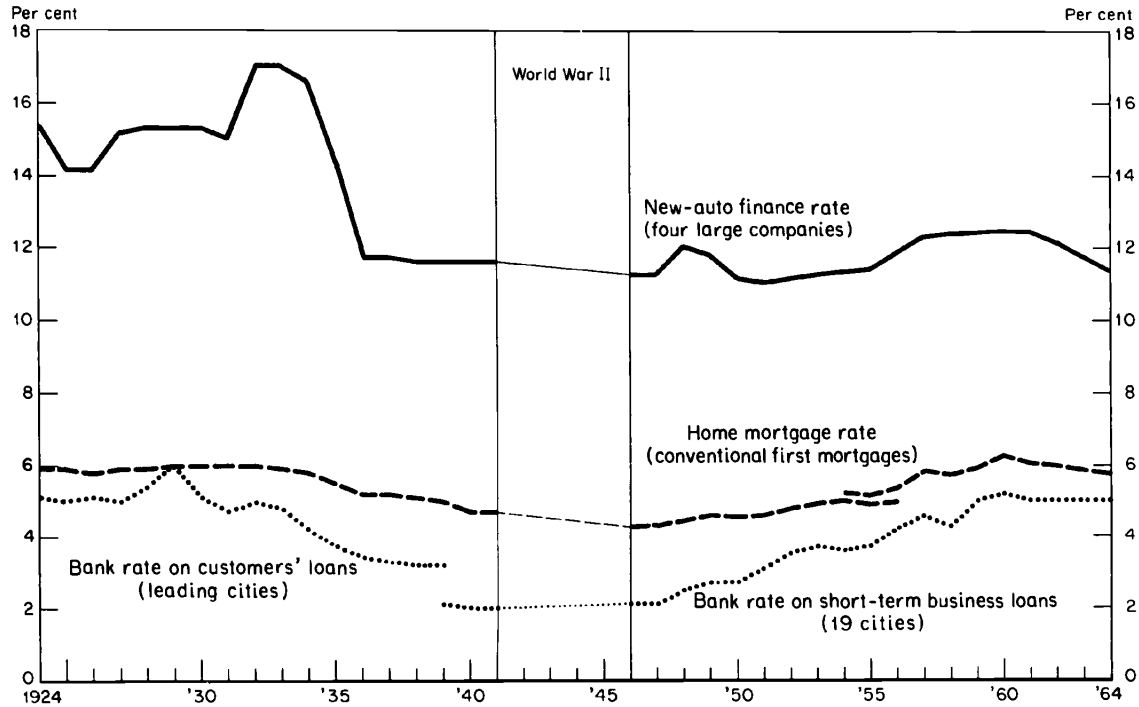
We were surprised by the variation in new-auto finance rates over time. Rates apparently declined from levels approximating 15 per cent in the 1920's to about 11 per cent in the early 1950's, then rose to about 12.5 per cent in 1960.[4] Since then, sales finance company new-auto rates have declined more than a full percentage point, a greater decline that that observable in residential mortgage rates (Chart II-7). The recent decline in finance rates and mortgage rates in relation to other borrowing rates may reflect the increased flows of funds into savings accounts in commercial banks and thrift institutions. This, along with other factors, appears to have increased rate competition among banks and sales finance companies for automobile credit and lowered finance rates.

The automobile credit market is patronized by borrowers who either finance their obligations through dealers or obtain loans directly from lending agencies. Commercial banks make loans directly and purchase paper through dealers. As a result, three major types of credit sources can be distinguished: sales finance companies (which purchase dealer paper), indirect financing banks, and direct lending banks. The 1954-55 sample of credit contracts provided us with an opportunity to compare factors affecting rates among these sources. The factors that appear to be most significantly related to the level of finance rates

[4]The finance rate series is presented in Part IV, section 4.

CHART II-7

New-Auto Finance Rates, Home First-Mortgage Rates, and Bank Rates on Business Loans, 1924-64



SOURCE: Finance rate series from Robert P. Shay, *New-Automobile Finance Rates, 1924-62,* New York, NBER, 1963, Table 8, with minor revisions; mortgage rate series: 1924-56, Saul B. Klaman, *The Postwar Residential Mortgage Market,* Princeton for NBER, 1961, Table A-6; 1954-62, FHA; 1963-64, new homes only, HLBB in cooperation with FDIC, *Federal Reserve Bulletin,* February, 1965; and business loan rates for 1924-28 from Winfield W. Riefler, *Money Rates and Money Markets in the United States,* New York, 1930, and *Federal Reserve Bulletins* for the years thereafter.

NOTE: With the greater prevalence of second and third mortgages in the earlier years, home mortgage borrowing costs probably declined over the period. All of the three series shown above reflect nominal rates and are not adjusted for fees, product prices, and compensating balances. The 1939 data for business loan rates are overlapped; the higher rate is a two-month average and the lower rate is a ten-month average from the revised series.

are type of credit source, size of city, legal rate ceiling, dealer's markup over factory price of auto, purchase of credit life insurance, and borrower's occupation. For the market as a whole, it appears that elements of the credit contract itself, legal restraints, and institutional differences among credit sources are of primary importance to rate levels.

Wallace Mors's study of the effects of legal ceilings upon finance rates and the dealer's share of the rate provides important additional insights into the automobile credit market. Mors finds, in his investigation of the 1958-59 sample of sales finance contracts for both new and used automobiles, that consumer rates are generally well below ceiling, indicating that competitive forces are operative. He also finds that legal rate ceilings exert some influence upon the level of finance rates charged borrowers.

Although these matters are still under study, the results so far suggest the following observa-

tions about the functioning of the automobile credit market. First, active price competition among sellers of credit has resulted in variations in finance rates over time and kept rates well below legal ceilings. Second, the pairing of buyers and sellers according to the price of credit does not appear to be governed directly by the creditworthiness of the individual borrower, but rather by differentiating elements in credit contracts, legal restraints, and general institutional policies. On balance these factors seem to result in a reasonable sorting of credit customers according to risk, that is, in a manner consistent with rate of charge.

INSTITUTIONAL COSTS OF PROVIDING CREDIT

While the analysis of the automobile credit market suggests that competition among sellers of credit exercises a large influence upon finance rates, the relation between income and expenses among the institutions providing credit also is indicative of the extent to which consumers receive protection through the operation of competitive forces under existing legal constraints. Paul F. Smith's book, *Consumer Credit Costs, 1949-59,* revealed that the costs of providing consumer credit are generally high in comparison with other kinds of credit, and that net profit rates are broadly comparable with those in manufacturing industries.

Operating expenses among four types of institutions studied (consumer finance companies, sales finance companies, commercial banks, and credit unions) were extremely wide, ranging from $14.25 to $3.30 per hundred dollars of outstanding credit in 1959. Many factors contribute to the level and to variations in costs. The method of acquiring business, differences in character of risks assumed, variations in size of credit contracts, legal controls upon operating practices, and some implicit costs inherent in credit union operations all contributed to the wide range in recorded costs. During 1949-59 the spread in average rates among the four types of institutions declined, a development that seems attributable both to competitive factors and to a narrowing of the differences among the kinds of consumer loan services provided to borrowers.

THE DEMAND FOR CONSUMER INSTALMENT CREDIT

Even if rate competition exists and if profits are at competitive levels, it does not necessarily follow that credit markets function perfectly. In *Consumer Sensitivity to Finance Rates,* F. Thomas Juster and I undertook an experimental questionnaire sampling of Consumers Union subscribers to test the hypothesis that the market for consumer instalment credit is imperfect—in the sense that the bulk of consumer borrowers are "rationed" with respect to the amounts of credit they can obtain at going rates of charge. Because of institutional limits upon rates, down payments, maturities, and loan sizes, we believe that most consumers accumulate more equity than they desire when purchasing durable goods instalment credit. As a result, they are insensitive to finance rates because they value equity funds more than the marginal costs of borrowing in the market. Borrowers in these circumstances are strongly impelled to obtain the largest amount of credit available—usually associated with the longest maturity and lowest monthly payment.

Our data are consistent with the hypothesis that there are substantial differences in rate elasticity between groups of borrowers classed as "rationed" and "unrationed." The former group included younger married families and those with small liquid asset holdings and a relatively favorable attitude toward credit use. The rate elasticity of this group was virtually nil. On the other hand, in the unrationed group —older married families or single persons with higher incomes and liquid assets, and those with an unfavorable attitude toward credit use —rate elasticity was considerably higher. Since the majority of consumer borrowers probably fall in the rationed category, our results suggest that little rate sensitivity would be observable in the population as a whole.

In this same work, Juster and I attempted to ascertain the extent of consumer knowledge of finance rates actually paid on recently financed durable goods purchases. The majority of borrowers, some 82 per cent, were unable to give rates estimates of reasonable accuracy. Yet respondents did seem to know that certain types of credit were more expensive than others. They knew, for example, that automobile loans were less costly than furniture loans and that small loans were more expensive than large loans. We termed this "institutional" knowledge because other research has suggested that borrowers may draw such broad distinctions by associating high and low rates with particular types of institutions.

RATE INFORMATION AND QUOTATION

Wallace P. Mors's book, *Consumer Credit Finance Charges: Rate Information and Quotation,* now in press, traces the origins and development of the various forms in which consumers receive finance charge information, reviews the procedures they must follow to make comparisons among credit alternatives, and evaluates the extent to which the variety of information given can be used by consumers to make effective comparisons of credit cost. Mors's major conclusion is that no single method of quoting charges, whether in dollars or rates of charge, is a reliable guide to credit cost, taken by itself. Among the several forms of information—monthly payment size, effective monthly or annual rates, add-on or discount rates on the amount borrowed—each may serve as a useful guide to credit cost in certain situations but have little value in others. Mors notes problems confronting consumers in using the varied information effectively and concludes that the diversity of state legislation contributes to the lack of uniformity encountered among the systems used to compute and quote charges.

In this connection it is of interest to note that the National Conference of Commissioners on Uniform State Laws has established a project for the purpose of drafting comprehensive uniform or model state legislation to govern consumer credit. Members of their staff have shown considerable interest in our reports and have stated that they have found them useful for their purposes.

SOURCES OF FUNDS FOR CONSUMER CREDIT

Richard T. Selden, in Occasional Paper 85, *Trends and Cycles in the Commercial Paper Market,* explored the rapid resurgence of this market in the postwar period. He found that the demand for commercial paper funds was determined by the amount of receivables which finance companies could acquire profitably and by the relative cost of commerical paper and bank loans. The supply of funds appeared to be quite elastic with respect to the differential between paper yields and Treasury-bill yields. Both demand and supply were sensitive to cyclical conditions. The divergent cyclical behavior of the volume of directly placed paper compared with that of dealer paper—the former rising and falling with the business cycle and the latter showing countercyclical movements—is a major finding which Selden explains in terms of differential demand and supply hypotheses. Selden's remaining work on sources of funds to the finance industry is still in preparation.

FINANCIAL ADJUSTMENTS TO UNEMPLOYMENT

In an Occasional Paper now in press, Philip A. Klein analyzes the financial adjustments to unemployment made by families receiving unemployment compensation, in order to determine how and to what extent consumer credit affects these adjustments. Working with questionnaires originated by the Department of Labor in six field surveys between 1954 and 1958, the study ascertains the manner in which these families were able to adjust to the loss of income when unemployment occurred. Unemployment compensation was the major source of funds to offset the decline in income, reducing the gross loss in family earnings by

32 per cent. To offset the drop in family income net of unemployment compensation, families reduced their net worth to the extent of 34 per cent of the income decline and reduced consumption expenditures the remaining 66 per cent. Reductions in liquid asset holdings accounted for two-thirds of the net worth reduction, and debt adjustments accounted for the remaining third. Debt adjustments took the form of borrowing from friends, relatives, or financial institutions, running up unpaid bills, or allowing delinquencies and repossessions to occur on instalment obligations. While impairment of net worth has long-run consequences for those who suffer unemployment, Klein points out that in the short run it does permit a less drastic cut in consumption expenditures and cushions the aggregate effects of unemployment upon the rest of the economy. Klein also found that the longer unemployment lasted, the greater was the reduction in consumption expenditures, since resources get used up and other adjustments in net worth become more difficult to make. He concludes that unemployment compensation, the liquidation of assets, and adjustments in debt position all acted as stabilizing influences, although unemployment compensation has far greater quantitative importance and has the advantage of not impairing the net worth of unemployed families.

SUMMARY REPORT

A summary report on the study's major findings is being prepared. It will review trends and cycles in credit; compare the relative importance of the several institutions supplying credit; discuss the effects of changes in rates, terms, and sources of funds; and, finally, assess the contribution of consumer instalment credit to the nation's economic performance.

ROBERT P. SHAY