

This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: Growth of Industrial Production in the Soviet Union

Volume Author/Editor: G. Warren Nutter assisted by Israel Borenstein and Adam Kaufman

Volume Publisher: Princeton University Press

Volume ISBN: 0-87014-074-4

Volume URL: <http://www.nber.org/books/nutt62-1>

Publication Date: 1962

Chapter Title: The Data: Knowns and Unknowns

Chapter Author: G. Warren Nutter, Israel Borenstein, Adam Kaufman

Chapter URL: <http://www.nber.org/chapters/c1973>

Chapter pages in book: (p. 11 - 51)

## CHAPTER 2

### The Data: Knowns and Unknowns

A STATISTICAL study naturally begins with an appraisal of the underlying data, in this case, official Soviet statistics. Discussion tends to get focused on defects, more easily seen than virtues, and this carries with it the danger that the basic statistics may seem to be worse than they are. Almost every economist in no matter what field of empirical research soon becomes convinced, as he gets familiar with his materials, that no data could be as bad as those he is forced to work with. He has explored the defects more thoroughly than others have. Heeding this lesson, we should weigh the good features with the bad before passing judgment. As we shall see, Soviet statistics, despite their serious shortcomings, do form a basis for studying industrial growth when used with care.

The statistics relevant to a study of industrial growth fall into several categories: output of individual industries, prices and related cost data, labor and capital inputs, and aggregative measures. The discussion here will center on only the first of these, namely, output of individual industries expressed in physical terms. The other types of data will be discussed at appropriate points in other chapters.

The discussion cannot be exhaustive but will concentrate on some of the more significant points. Fortunately, the subject has already been treated very carefully and thoroughly by Professor Gregory Grossman in an earlier report in this series,<sup>1</sup> which should be consulted by those interested in a more detailed analysis. That excellent study is, in fact, the basis of much that will be said here.

#### *Introductory Remarks*

The defects of Soviet statistics on physical output are important and must be understood if the data are not to be misused. There are three major shortcomings, all deriving from the nature of the Soviet political and economic orders. The first is the selectivity of published data, a factor that works in two opposing directions. On the one hand, some areas of poor performance are shielded from view, causing the published data to underrepresent slower-growing sectors of industry. On the other hand, some of the more rapidly expanding economic activities associated with the military sector are also not reported on. It is impossible to determine

<sup>1</sup> Gregory Grossman, *Soviet Statistics of Physical Output of Industrial Commodities: Their Compilation and Quality*, Princeton for National Bureau of Economic Research, 1960.

whether the net effect is to promote an overstatement or an understatement of growth. As we shall see, the degree of selectivity has varied considerably over the years. For a long stretch of time, from 1938 to 1956, almost no data were published on the absolute level of output in any sectors of industry.

The second shortcoming is ambiguity. Primary sources generally do not contain adequate definitions of industries in terms of administrative and territorial coverage, product coverage, and stage of fabrication at which output is being measured. Titles given to industries can be misleading—for example, “silk fabrics” are chiefly rayon—and slight verbal changes may signify a basic change in definition not otherwise described. Things are not always what they seem to be, and the user of Soviet data should beware. In the end, he still will have to use many data whose meaning he does not fully comprehend, and conclusions should be qualified on this account.

The third shortcoming is the general overstatement of absolute levels of output within the Plan period for the sample of industries reported on. The lower the priority of an industry from the Soviet point of view and the less precisely its output can be measured, the greater the overstatement is likely to be, for reasons to be developed later. That much can be said, but no more; we cannot now place an order of magnitude on the overstatement, in the large or in the small. The tendency toward overstatement needs to be taken into account most when levels of output are being compared between the Soviet Union and other countries. It has less bearing on internal measures of growth, since it is doubtful that relative overstatement of output has increased systematically with time, except with respect to prerevolutionary and early Soviet years. Hence growth will, on this count, be overstated relative to, say, 1913 or 1928 but not necessarily relative to later base years. Over later spans of years growth may be overstated, understated, or more or less accurately reflected by the available output data, the effect depending on specific circumstances, some of which cannot now be adequately known.

Offsetting these shortcomings is another feature of the Soviet system: the large volume of economic statistics collected and processed. As Professor Devons has tersely put it, “Without statistics there can be no planning.”<sup>2</sup> It is a curious fact that the United States, lying more or less at the opposite pole from centralized planning, is probably the only other country as figure-minded as the Soviet Union—for quite different

<sup>2</sup> Ely Devons, *Planning in Practice: Essays in Aircraft Planning in Wartime*, Cambridge, Eng., 1950, p. 133.

reasons, of course. The question of quality and reliability aside, the volume of output data flowing out of the Soviet Union during interwar years and since 1956 has been large by normal standards, despite the policy of selective publication. Quantity substitutes to some extent for quality.

When all is said, Soviet data, with their many faults, do provide a basis for assessing Soviet industrial performance and growth, if carefully used and interpreted. This is shown most convincingly by the fact that growth patterns derived from using these data make economic sense. There is a basic internal consistency in the figures; differential rates of growth conform in direction with developments that can be directly observed; certain phenomena appear that are characteristic of economic growth everywhere, such as retardation in growth of individual industries; and changes in industrial structure are shown that are otherwise known to have occurred. These and other lines of evidence on the reliability of the data will be developed more fully at later points in this and other chapters.

But the faults remain to affect the accuracy of measures of growth, and we turn now to discuss them more fully. Since most of the difficulties stem from the nature of the Soviet system and its statistical apparatus, we begin with a review of their salient features as they affect the reliability of statistics.

#### *General Characteristics of Soviet Statistics*

Fault can be found with the economic statistics of every country. They represent, in the first place, a mere sampling of the unbounded volume of data that might be recorded. They have been collected with specific objectives in mind—more varied and far-reaching in some countries than in others—and will therefore be of varying use depending on the purposes they are made to serve. They contain, in the second place, errors introduced at different stages of observation and assemblage. These will depend on the state of statistical literacy among the collectors and suppliers of data, on the effort expended on record-keeping, and on the degree of active competition in gathering and analyzing data. They are, finally, subject to manipulation and distortion by parties with a stake in the figures, checked only to the extent that there are independent fact-seekers and fact-gatherers with competing interests. No government or other statistical agency can be relied upon to resist the temptation to stretch figures to its own account if it feels it can get away with it.

Progress in economic statistics has been driven in the West by two

engines: competition and technical sophistication. An extreme example is perhaps provided by the United States, a country unique in its long tradition of figure-gathering. Thumbnail histories usually mislead, particularly when they treat the causes of some institutional development, but we may perhaps be allowed to speculate very briefly on the evolution of the American statistical system in order to illustrate its basic characteristics and how they differ from those of the Soviet system.

The habit of collecting statistics was formed early, with a constitutional requirement of a decadal census for the purpose of apportioning political representation. Existence of a large market economy led to demands for expanding economic intelligence on the part of legislators who made the laws defining the economy, businessmen who organized it, and scholars who studied it. The government census gradually expanded to cover an increasing area of economic statistics, and special censuses ultimately evolved. At the same time, private agencies arose engaging in a host of specialized activities in economic statistics, each serving the particular interests of its consumers. The long history of statistical activity, together with its competitive nature, provided the experience and pragmatic testing that in turn promoted improvements in technical procedures and competence.

A critical feature of the American statistical system, as it has evolved, is the multiplicity of statistical sources. While the government plays an important role in collecting and disseminating statistics, there is no sustained unity of interest among the governing because of the nature of the political system: federalism, representative government, governmental checks and balances, and the two-party system. No sanctity attaches to the official statistics of the moment; they are subject to challenge and are continually challenged by both ins and outs; they are subject to revision and are frequently revised. There are not only these internal checks, but also the external checks of private statistical organizations and researchers, pursuing their own work as they see it.

Similar conditions prevail in other Western countries in varying degrees. At bottom, representative government, competitive scholarship, and free public discourse are the Western institutions that have counteracted error and misrepresentation in statistics, imperfectly to be sure but at least to an important extent.<sup>3</sup>

<sup>3</sup> The Soviet image of Western statistics is rather different: "It is said that under capitalism comprehensive and truthful economic statistics are not to be expected because of the secretiveness of private firms, the lack of centralized coordination in and authority over the generation and collection of data, the class interests of the governments in power and the mendacity of their statisticians, etc." (Grossman, *Soviet Statistics*, p. 22). Some of

The importance of these institutions is shown by the generally unsatisfactory nature of the statistics gathered and issued during wartime, when public discussion is curbed and large segments of the economy are centrally directed. Devons, in his informative little book on the British experiences of aircraft planning in World War II, concluded his chapter on the role of statistics in planning by saying:<sup>4</sup>

The pseudo-scientific atmosphere which the use of charts and statistics created gave great power to the statisticians. For it was fairly easy by the manipulation of statistics and charts to "prove" a particular case; and the statisticians soon came to realize that many of the officials not used to handling figures were both impressed by this manipulative power and incapable of acquiring it themselves. The department or directorate which had a skilled statistician always had a great initial advantage in any inter-departmental or inter-directorate dispute. And any statistician who was concerned with issues of policy was bound to find himself, sooner or later, selecting and manipulating statistics in such a way as to guide policy along the lines which he had decided, on quite general grounds, were the right ones.

Attempts were made to avoid this danger, by separating the collection and issue of statistics from decisions and discussions of policy. But such attempts invariably failed in M.A.P. [the Ministry of Aircraft Production]. First, because the analysis of data about the past is so intimately concerned with the planning of the future, that any attempt to separate the two functions usually resulted either in the planners paying little attention to the past and so making the most unrealistic plans, or in the planners setting up their own fact-finding staff which by-passed the statistical division and so deprived it of any influence. Secondly, life in a statistics division which was separated from policy was apt to be dull, and there was great difficulty in attracting efficient staff to such a division. In any case, unless the staff of the statistics division were closely concerned with policy decisions, they had no easy means of knowing which were the most significant statistics to collect and analyze; and they had the greatest difficulty in ensuring that some notice was taken of the results of their analyses. The danger that the planners who have a monopoly of the statistics might distort the figures

---

the real shortcomings of American government statistics are discussed in Geoffrey H. Moore, "Accuracy of Government Statistics," *Harvard Business Review*, Spring 1947, pp. 306-317.

<sup>4</sup> Devons, *Planning in Practice*, pp. 163 f.

THE DATA:

to prove their case cannot be avoided. Where planning is necessary, great power must inevitably fall into the hands of the statisticians.

These words might well have been written about the Soviet statistical system. The troubles with Soviet statistics stem, in the first instance, from the system of centralized authoritarian planning—from the nature of what Grossman, following others, has called the “command economy.” Statistics are collected, processed, and issued by only one agency: the state. There are no independent sources to restrain each other or to be used as checks on each other, except to the extent that related figures published by different state agencies might not be fully coordinated before issuance. From the nature of the planning system, everybody seems to have a stake in the figures—those who report them as well as those who process and use them—since performance is judged by them.

One finds in the Soviet responses to difficulties the same dilemmas pictured by Devons under less trying circumstances: statistical and planning agencies are separated, united, and then separated again; internal checks are evolved through a dual reporting system with the administrative and statistical hierarchies supposedly cross-checking each other, later to be abandoned in favor of consolidated reporting through the statistical hierarchy alone; and so on. As Grossman has emphasized:<sup>5</sup>

... one must not exaggerate the specifically Russian or communist elements in these problems. Rather, given the way human beings react in the face of authority and in their quest for material well-being, the problems discussed here arise by and large from the logic of a command economy and a sellers' market. To be sure, many of the details, aspects, and nuances are peculiar to the Soviet scene, and some perhaps even to the Russian “national character,” if there be such a thing. But the broader outlines of these problems can be easily recognized in other authoritarian organizations, especially in other command economies, and in sellers' markets in other countries and at other times.

There is, at the same time, a second set of difficulties with Soviet statistics that originates in circumstances rather specific to communism and the Russian case. The Soviet system embodies an international crusade, and statistics are grist for the propaganda mill. Knowing the ideological views of Soviet leaders, one finds it hard to picture them

<sup>5</sup> Grossman, *Soviet Statistics*, pp. 4 f.

dispensing facts in a passive and detached manner. The official doctrinal concept of statistics as a discipline is considerably at variance with the traditional Western view, statistics being considered "a social science, the theoretical base of which is formed by historical materialism and Marxist-Leninist political economy."<sup>6</sup>

Another set of quite different endemic difficulties, especially in the formative period of Soviet statistics, may be traced to the meager heritage from the Tsarist era of experience and competence in statistical work. The staff conducting statistical work in agencies of the central Tsarist government was notoriously inefficient, and censuses were infrequent and narrow in scope—the first complete population census was taken in 1897. Industrial statistics were largely the by-product of the factory inspection and tax collection systems. Though private trade associations engaged in some statistical activities, they were limited in scope and came into existence late in the nineteenth century, when industrialization first surged forward in Russia. Statistical investigations of high quality were conducted throughout the last four decades of the nineteenth century by professionals working (voluntarily, for the most part) with the *zemstva*, or local and provincial councils; and out of this activity there emerged a nucleus of well-qualified statisticians, particularly in agriculture. But the range of activities and the number of people involved were small. Coupled with this was the crucial fact that educational levels were low in the bulk of the population, around 60 per cent being illiterate in 1914 and most of the rest not far above the threshold.<sup>7</sup> These factors must have had an adverse effect on the quality of statistics at least in the earlier Soviet years, despite the rapidity with which statistical activities grew and illiteracy declined.

Counteracting these detrimental features has been the urgent internal need for reliable statistics to run the economy. In the Soviet economic system, statistics form the basis for making plans, checking on their fulfillment, allocating resources, making technical managerial decisions, assessing performance, and dispensing rewards and punishments—in

<sup>6</sup> A. Yezhov [Ezhov], "Soviet Statistics in the Last Forty Years," *Problems of Economics* (authorized English translation of *Voprosy ekonomiki*), May 1958, p. 34. For further citations, see Grossman, *Soviet Statistics*, p. 23.

<sup>7</sup> The information on statistics is from Bernard Pares, *A History of Russia*, rev. ed., New York, 1944, pp. 402 ff; and A. Yezhov [Ezhov], *Soviet Statistics* (translated from the Russian), Moscow, 1957, pp. 5 ff. As to illiteracy, the census of 1897 listed 79 per cent of the population as illiterate, varying from 20 per cent in the Baltic provinces to 94 per cent in Central Asia. This had apparently fallen to just under 60 per cent by 1914. See M. T. Florinsky, *Russia: A History and an Interpretation*, New York, 1953, Volume II, pp. 1256 f; G. Vernadsky, *A History of Russia*, New Haven, 1951, p. 398; and S. Harcave, *Russia: A History*, 3rd ed., Philadelphia, 1956, pp. 313 ff.



short, for performing virtually every economic function. The pressure for trustworthy statistics comes, so to speak, from the top downward: every agency in the political and administrative hierarchy strives to get truthful reports from subordinate units.

Centralized authoritarian direction of the economy thus generates forces with opposing effects on the reliability of statistics. On the one side, there is a pressure for misreporting moving from the bottom upward: self-interest motivates each subordinate unit to try to mislead its superior, the central government finally being motivated to mislead the outside world. On the other side, there is a pressure for accuracy moving from the top downward, similarly motivated by self-interest. Which force gains the upper hand?

The answer is misreporting, since it does occur—as we shall see—even though it is certainly restricted by the pressure for accuracy. But before moving to the evidence, we may conclude these general remarks by noting the concern of Soviet officials themselves over the question of reliability of statistics. In the words of Grossman once more:<sup>8</sup>

Even a cursory reading of the Soviet literature reveals that the central statistical authorities have been well aware of the imperfect reliability of the data submitted to them. A closer study leaves no doubt that they have been gravely concerned over the problem, and that the question of accuracy of physical output data occupies the very center of this concern. It is also clear that the main source of inaccuracy is believed to be distortion of reported data by interested parties, aided by the negligence, if not abetted by the connivance, of the lower statistical agencies.

The basis of this concern will emerge from the details of the statistical system and the statistics themselves.

*The Statistical System: A Brief Summary*<sup>9</sup>

During early Soviet years the statistical apparatus, called the Central Statistical Administration (*Tsentral'noe statisticheskoe upravlenie*, abbreviated *TsSU*), had an independent status, containing within it a special agency, the Division of Census and Statistics of the Supreme Council of the Economy (*Vysshii soviet narodnogo khoziaistva*, abbreviated *VSNKh*), concerned primarily with large-scale state industry. With the advent of

<sup>8</sup> Grossman, *Soviet Statistics*, p. 49.

<sup>9</sup> This section is a condensation of *ibid.*, Part One.

centralized planning, dissatisfaction arose over the separation of planning and statistical agencies, and in 1930 *TsSU* was made a part of the State Planning Commission (*Gosplan*). The name of the statistical arm was soon changed, in line with its new status, to the lengthy title Central Administration of Economic Record-Keeping Attached to the *Gosplan* of the USSR (*Tsentral'noe upravlenie narodnokhoziaistvennogo ucheta pri Gosplane SSSR*, abbreviated *TsUNKhU*). The merger of the central agencies was strengthened in 1938,<sup>10</sup> after a series of purges associated with the ill-fated population census of 1937, and it was extended to subordinate units late in 1943. The unified structure continued until 1948, when the statistical organization, which had been renamed *TsSU* in 1941, was separated from the *Gosplan* at all levels. It has retained its independent position up to the present.

During its affiliation with the *Gosplan*, the statistical organization was developed into a hierarchical structure on a regional basis. A chain of subordination became established with the central administration at the top, followed by administrations at the level of the republic, territory (*krai*), province (*oblast'*), major city, district (*raion*), and lesser city.<sup>11</sup> This hierarchy has remained in force, apparently being unaffected by the economic and administrative reorganization of 1957, which will be commented on briefly below.

The basic simplicity of this statistical organization belies the complex system of reporting that existed until the reforms of 1957. Data originating in economic enterprises flowed upward through two parallel channels: on the one side, the statistical hierarchy already described and, on the other side, the economic-administrative hierarchy (see Chart 1).<sup>12</sup> The system also provided for cross-reporting and for simultaneous reporting at different levels in the hierarchy. Thus, the enterprise reported in three directions at once: to the local statistical unit, to the next higher statistical unit (at the provincial level if existent, otherwise at the republic level), and to its immediate superior in the economic-administrative hierarchy, typically a chief administration (*glavk*). The chief administration in turn submitted a consolidated report to both its ministry and the central statistical office, and the ministry did the same to the central statistical office. Finally, the statistical offices at every level submitted separate

<sup>10</sup> In this connection, the name was altered to read "of the *Gosplan*" from "attached to the *Gosplan*."

<sup>11</sup> For a short period, these were also subdistrict inspectorates. The "chain of command" given here is simplified. For example, the so-called autonomous republics are subordinate to the union republics to which they are assigned. Major cities are the capital cities of union republics, plus Leningrad.

<sup>12</sup> This chart is also simplified. See the cited source.

THE DATA:

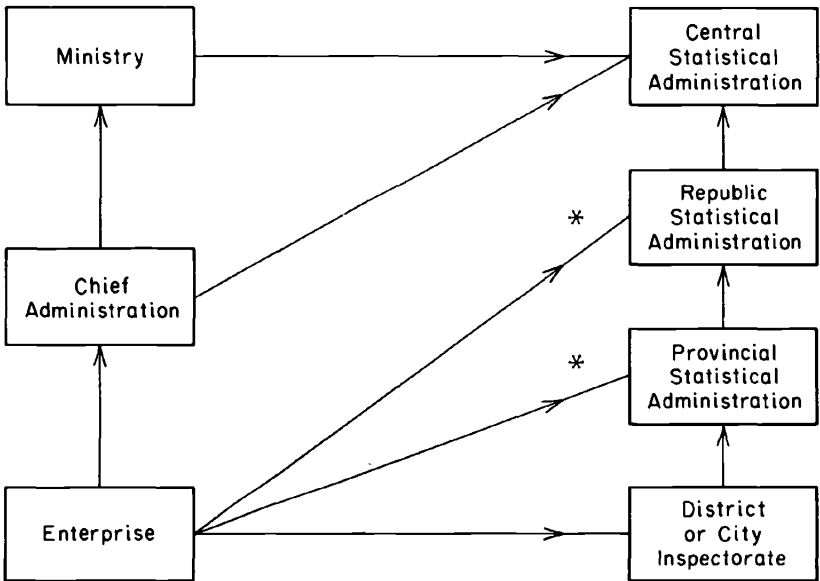
consolidated reports to the corresponding level of the Communist Party, the government, and the planning organization—a flow not shown in our chart.

This complex system could have arisen for a number of reasons, not the least being the desire of every agency to have the most up-to-date figures at its disposal. Whatever its origins, the system abounded in possible

CHART I  
The Soviet Statistical System Until Mid-1957

Economic-Administrative  
Hierarchy

Statistical  
Hierarchy



Arrows indicate direction of reporting.

\* For republics without provinces, reports were made directly to the republic; otherwise, to the province.

Source: Adapted from Grossman, *Soviet Statistics*, p. 37

cross checks. Sums could be checked at almost every level in five different places: the ministerial, statistical, Party, government, and planning organizations. The only figures whose accuracy was not subject to direct checking were the basic data reported by the enterprise itself. Despite the many opportunities for checking figures, it is doubtful that the system did much more than multiply paperwork. The main obstacle to effective auditing is the enormous volume of data that must be rapidly

processed by the Soviet statistical system. Given this fact—discussed more fully below—and the strong incentives to misreport, the interlacing of agencies in the statistical network may have worked in the opposite direction, aiding cooperative misreporting.

The system of parallel reporting was abandoned with the reforms of 1957, occasioned by the administrative reorganization of industry into regional economic councils (*sovnarkhozy*). The details of the new reporting system need not detain us, since it does not generally apply to the period of this study.<sup>13</sup> We may merely note that the upward flow of data to the central government now seems to proceed solely through the statistical apparatus. In line with this change, Starovskii, head of *TsSU*, remarked that “whereas up to now the checking of accounting data has been done by the respective subdivisions of chief administrations and ministries, now this most responsible work will be entirely entrusted to *TsSU* agencies.”<sup>14</sup>

In a more lengthy comment, Starovskii says:<sup>15</sup>

In addition to the state statistical agencies, to which enterprises reported data (on state accounting forms), ministries, departments, and their chief administrations required a tremendous number of different tables, questionnaires, and estimates. Parallel accounts were also sent to financial and banking agencies and to a number of local organizations, and often so-called “wild” accounts (i.e., those not prescribed by law) were compiled. One of the managers of the former Ministry of Heavy Machine Building considered it essential to have, for example, data on the height at which electric light bulbs were hung in factories and other such information without which he thought it was impossible to administer from the center the enterprises under his jurisdiction.

The administrative reorganization of industry and construction enables us to eliminate existing defects. Now the receipt and processing of accounting statistics for these branches of the national economy is centralized in the state statistical agencies. Industrial enterprises, construction works, and economic organizations present their accounts to the appropriate province, territory, or republic statistical agencies. Further processing of statistical data is done by agencies of the USSR Central Statistical Administration. Beginning with the accounts for June 1957, the regional economic councils, the Party and Soviet administrative agencies, and the planning committees will receive the

<sup>13</sup> For the details, see Grossman, *Soviet Statistics*, pp. 38 ff.

<sup>14</sup> V. Starovskii, “Novye zadachi sovetsoi statistiki” [New Tasks for Soviet Statistics], *Kommunist* [The Communist], 1957, No. 14, p. 70.

<sup>15</sup> *Ibid.*, pp. 62 f.

## THE DATA:

statistical material they need directly from the local agencies of the Central Statistical Administration. Within *TsSU* the summary accounts will come not from the ministries but from the local agencies of *TsSU* and *TsSU* will process them and present them to the USSR government, the *Gosplan*, and other central organizations . . . .

The size of the statistical apparatus is indicated by Soviet estimates that nearly two and a half million persons are employed directly in keeping and processing records.<sup>16</sup> Reporting is done on a current basis.<sup>17</sup> All enterprises (except the very smallest producing for local markets) must submit monthly telegraphic reports on physical output, followed by a mailed report sent within three days of the end of the month; they must submit comprehensive monthly and quarterly reports, covering other economic data as well as output, within fifteen days. For products considered particularly important (e.g., fuel, steel, electricity), additional telegraphic reports must be submitted daily or every ten days. Each echelon in the statistical structure must then process within ten to fifteen days the data it receives. With such a flood of data, it is doubtful that much could be done beyond summing and tabulating in this brief period even if the statistical operations were fully mechanized; but only 3 to 4 per cent of those engaged in statistical work had the use of electrical adding and computing machines as late as 1953.<sup>18</sup>

So much for the flow of data into the system. The flow out of it, in the form of published statistics, has been less steady and voluminous. In some respects, the high point of published industrial statistics was reached in the late 1920's. The data, published in many sources, were comprehensive and detailed, and their processing was directed by competent economists and statisticians. Concurrently with the five year plans, the flow of published statistics gradually diminished, the low point being

<sup>16</sup> See Grossman, *Soviet Statistics*, p. 30, n. 21. According to *Pravda*, May 12, 1958, the number engaged in this work was put at "about three million, of whom almost 80 per cent are engaged in so-called primary record-keeping." For the United States, the 1950 census of occupations lists about 376,000 accountants and auditors and 721,000 book-keepers in the employed labor force, a total of 1.1 million. Many of these are engaged in activities not covered by the Soviet concept of record-keeping.

<sup>17</sup> Small-scale enterprises are excepted, their output being estimated through periodic censuses and sample surveys. Current reporting was tried during the period 1949-1954 and then abandoned. Until 1930 all enterprises, large-scale as well as small-scale, were covered by comprehensive periodic censuses. Before the 1957 reforms, each enterprise reported currently on more than a hundred forms, sixty to seventy of them flowing into the centralized reporting system. Even after the reforms, centralized current reporting was in force (in 1957) for more than 10,000 commodities. For details, see Grossman, *Soviet Statistics*, p. 35, n. 15.

<sup>18</sup> Most do have the abacus, a valuable computational aid. For more details on mechanization, see *ibid.*, pp. 55 ff.

reached after 1937 and continuing as late as 1956, when a striking improvement took place. Since then, published statistics have moved toward the coverage characteristic of the late 1920's and early 1930's, but they have not regained that stage yet.

During the First and Second Five Year Plans, published statistics came to be concentrated almost exclusively in a set of annual statistical abstracts, setting the practice for later years. The most important volumes are those bearing the title *Socialist Construction of the USSR* (*Sotsialisticheskoe stroitel'stvo SSSR*), the first being published in 1934 and the last in 1938. These collections of data are roughly comparable in coverage, detail, and amount of explanatory material with summarizing abstracts published in Western countries. They are not comparable with Western primary statistical sources, such as the various census publications of the United States and the United Kingdom.

The most comprehensive of these Soviet abstract-like publications is the volume that appeared in 1936, containing data through 1935. From this peak, the amount of published statistical material fell off sharply. The abstract appearing in 1938, the last of this series, covered only the period of the Second Five Year Plan and a selective group of industries. It was not until 1956, or eighteen years later, that a similar abstract again appeared.

During the years intervening between 1938 and 1956, the only published statistics were those contained in official announcements and directives, political speeches, occasional articles in specialized journals, textbooks, and a handful of books written by Soviet authorities and largely descriptive or polemical in nature. Annual summaries of industrial performance were generally presented in less than a page of the newspaper. Statistics were limited in nature as well as amount: absolute data on output were given out very rarely and for only a very small sample of products; data were usually stated in relative terms, as a percentage of some base figure, itself unknown or obscure. The cryptic information given out during this period caused economic research in the West to take on the characteristics of archaeology.

The appearance in 1956 of a small statistical abstract, *The National Economy of the USSR* (*Narodnoe khoziaistvo SSSR*), markedly improved the statistical picture. Even so, the volume contains only fifty to sixty pages with basic data for industry.<sup>19</sup> More significant additions to industrial

<sup>19</sup> In one English translation, these pages are reduced by more than half without materially affecting readability (*Statistical Handbook of the USSR*, Harry Schwartz, ed., New York, 1957).

statistics were made with the publication of several abstracts in 1957 and 1958, the most important being *Industry of the USSR* (*Promyshlennost' SSSR*).<sup>20</sup> This volume contains 447 pages, with about 168 presenting data on physical output. Virtually all the data on industry in *The National Economy of the USSR* (1956) are repeated in *Industry of the USSR*, while the latter contains many data not in the former.

These recent Soviet abstracts may be compared with statistical sources for the United States. The 1956 edition of the *Statistical Abstract of the United States* contains some 100 pages of industrial data in small type; the *Product Supplement* to the census of manufactures for 1954 contains 259 pages, twice the size of those in the Soviet abstract, of physical output data. If we move to primary sources, the basic volumes of the latter census contain about 3,600 pages, and the *Minerals Yearbook* for any recent year is equally large, though the pages are smaller.

The scope of the most recent Soviet statistical abstracts is perhaps better indicated by the number of industrial products covered. Output data are given for about 90 products in *The National Economy of the USSR* (1956) and 212 in *Industry of the USSR*, in most cases for benchmark years.<sup>21</sup> The product coverage of these recent Soviet abstracts is generally less comprehensive than that of *Socialist Construction* (1936), particularly for chemicals, nonferrous metals, and minerals. In the *Product Supplement* to the U.S. census of manufactures for 1947, physical output data are given for some 6,000 products; the census for 1954 covers about the same number. In the U.S. census of mineral industries for 1954, physical output data are given for more than 750 products. These product coverages are, of course, larger than for earlier years, but in every industrial census of the United States since the turn of the century, the count of products would run at least to many hundred—in most, to several thousand. At the same time, it should be recognized that such counts describe detail more than breadth of coverage. That is to say, the products summarized in the recent Soviet abstracts would be broken down into hundreds of subproducts in U.S. statistics.

<sup>20</sup> For a list of recent statistical handbooks, see *Soviet Studies*, January 1959, pp. 312 ff, and January 1960, pp. 348 ff. The volumes discussed in the text have been followed by steadily improved handbooks, particularly by new editions of *The National Economy of the USSR* appearing in 1959 and 1960.

<sup>21</sup> For the industries covered by *Industry of the USSR*, data are given for every year over 1913–1955, except 1941–1944 inclusive, in the case of 59; only for benchmark years in the case of 76; and only for selected benchmark years in the case of 77. Benchmark years are taken as 1913, 1928, 1932, 1937, 1940, 1945, 1950, and 1955.

In addition to the official Soviet compendiums of statistics, there are a number of secondary sources containing information of one sort or another bearing on industrial output. These range from articles and monographs, such as are found in various professional journals, to general reference books, such as the *Great Soviet Encyclopedia* (*Bol'shaia sovetskaia entsiklopediia*). From the late 1930's to 1956, these sources contained only scattered information on output in relative terms, as mentioned above. Occasionally, an absolute figure might be given. Products were seldom defined and references were not made to related collections of data. Such information is useful only to fill in gaps in other data.

The dearth of statistics in secondary sources reflects the control exercised at the center over release of information internally as well as externally. At the Twentieth Party Congress in February 1956, Mikoyan complained of the absence of large-scale statistical studies in the Soviet Union and remarked that "unfortunately, Comrade Starovsky had these statistical data under lock and key in the Central Statistical Administration. Economists are still deprived of the opportunity of working with them and are condemned to recite and repeat old formulas, old data. This is one reason we do not see creative work from our economists."<sup>22</sup> This statement, it later turned out, was a clue to the forthcoming change in statistical policy. The volume of data has expanded in secondary sources along with official statistical publications, but there still seems to be little available there that is not also in primary sources.

The statistical publications of the Soviet period generally do not reproduce data for the Tsarist period, except for the year 1913, which is used as a basis of comparison for later developments. The only exception to date is the most recent abstract, *Industry of the USSR*, which contains prerevolutionary output series for eight industries. Most data for the Tsarist period must be drawn from the statistical sources of those times. These sources have important shortcomings, primarily traceable to their limited coverage and to the circumstances under which the data were collected (mainly as a by-product of the factory inspection system and as an aid in the administration of taxes). However, there is no indication of widespread distortion or suppression of statistics, either by those providing the primary entries or by those processing the data. The data for the Tsarist period are deficient mainly by virtue of errors, omissions, and poor coverage.

<sup>22</sup> *Current Digest of the Soviet Press*, VIII, 8, p. 10 (original text in *Pravda* and *Izvestia*, February 18, 1956).



*Evidence on Reliability of Data*<sup>23</sup>

## INTRODUCTORY REMARKS

Reliability in a statistical context generally means the accuracy with which quantitative magnitudes measure the things they are purported to measure. Put another way, a statistic is reliable if it is an accurate magnitude for a definite thing. Inaccuracy or ambiguity may be the result of error, distortion, or fabrication, and, needless to say, the distinction between ambiguity and inaccuracy fades at the margins. As an example, consider a magnitude given as the output of coal. Even if the output is measured as accurately as possible, it cannot be adjudged a reliable datum unless we know how the term "coal" is being used.

Ambiguity is a general characteristic of Soviet output data, increasing in degree as the data become more aggregative. At one extreme stands the official Soviet index of industrial production. This index is the result of a set of actual calculations on actual data; but we have only a rough notion of the data and calculations, and hence cannot reproduce the index or fully understand its meaning. Enough is known, however, to be able to say, as almost all Western scholars do, that the index does not represent any of the concepts of aggregate production utilized in Western statistics, though it goes by the same name.

The disparity in these statistical constructs may be shown by tracing

<sup>23</sup> There is an extensive Western literature on this subject, and the discussion here draws much from it. The following is a partial list of specialized monographs and articles: Abram Bergson, "A Problem in Soviet Statistics," *Review of Economic Statistics*, November 1947, 234-242; *idem*, "Reliability and Usability of Soviet Statistics," *The American Statistician*, June-July 1953, 13-16; Colin Clark, *A Critique of Russian Statistics*, London, 1939; Maurice Dobb, "Further Appraisals of Russian Economic Statistics—A Comment on Soviet Statistics," *Review of Economics and Statistics*, February 1948, 34-38; Alexander Gerschenkron, "The Soviet Indices of Industrial Production," *Review of Economic Statistics*, November 1947, 217-226; *idem*, "Comment on Naum Jasny's 'Soviet Statistics,'" *Review of Economics and Statistics*, August 1950, 250-251; *idem*, "Reliability of Soviet Industrial and National Income Statistics," *The American Statistician*, June-July 1953, 18-21; Gregory Grossman, *Soviet Statistics*; Naum Jasny, "Intricacies of Russian National Income Statistics," *Journal of Political Economy*, August 1947, 299-322; *idem*, "Soviet Statistics," *Review of Economics and Statistics*, February 1950, 92-99; *idem*, "International Organizations and Soviet Statistics," *Journal of the American Statistical Association*, March 1950, 48-64; *idem*, *The Soviet 1956 Statistical Handbook: A Commentary*, East Lansing, 1957; Stuart A. Rice, "Statistical Concepts in the Soviet Union Examined from Generally Accepted Scientific Viewpoints," *Review of Economic Statistics*, February 1952, 82-86; *idem*, "Statistics in the Soviet Union," *Bulletin of the Atomic Scientists*, June 1952, 159-162; Harry Schwartz, "On the Use of Soviet Statistics," *Journal of the American Statistical Association*, September 1947, 401-406; *idem*, "The Organization and Operation of the Soviet Statistical Apparatus," *The American Statistician*, April-May 1952, 9-13; V. Tsonev, "Falsification of Soviet Industrial Statistics" (unpublished manuscript), Research Program on the USSR, New York, 1953; and Lynn Turgeon, "On the Reliability of Soviet Statistics," *Review of Economics and Statistics*, February 1952, 75-76.

through the consequences of accepting the Soviet index at face value. According to the official Soviet index, industrial production multiplied 27 times between 1913 and 1955. Over the same period, industrial production in the United States multiplied 4.7 times, according to a standard Western-type index (see Table 61). If these indexes were both taken to measure the same kind of growth, one would conclude that Soviet industrial production had grown almost six times as much as American production. This would imply in turn that Soviet industrial production in 1955 was about 80 per cent of the American level, since in 1913 it was about 14 per cent. In fact, Soviet production in 1955 was, by our calculations, only about 23 per cent of the American level (see Table 63) and, by recent Soviet pronouncements, 36 per cent.<sup>24</sup> One concludes that the Soviet index of industrial production exaggerates growth as that concept is typically measured in the West.

Every industry is a mixture of heterogeneous elements to some degree, and what we call "physical output" is an index number in miniature, even for the more narrowly defined industries. Ambiguity is dispelled only to the extent that product coverage and aggregating methods are described in detail. As we have already noted, Soviet statistical sources are lax in this regard, and the data one must work with are correspondingly ambiguous.

On the other side, the numerical accuracy of many Soviet output data also comes under question, no matter how the industries to which they

<sup>24</sup> Until recently, Soviet statisticians had not fallen into inconsistencies on this score. As late as 1957, a Soviet statistical source (Ia. Ioffe, ed., *Strany sotsializma i kapitalizma v tsifrakh* [Socialist and Capitalist Countries in Figures], Moscow, 1957, p. 8) gave the fractions as 6.8 per cent for 1913 and 47.6 per cent for 1955. Similar fractions for 1913 are given in *Ekonomika sotsialisticheskikh promyshlennykh predpriatii* [Economics of Socialist Industrial Enterprises], Moscow, 1956, p. 7; *Ekonomika promyshlennosti SSSR* [Economics of Industry of the USSR], Moscow, 1956, p. 21; and *Spravochnik komsomol'skogo propagandista i agitatora* [Reference Book for the Young Communist Propagandist and Agitator], Moscow, 1957, p. 126. These figures, though wrong, are at least consistent with comparative growth of the official Soviet index and the American index employed in Soviet sources. For an intriguing sketch of the gyrations followed to preserve such internal consistency, see A. Nove, "1926/7' and All That," *Soviet Studies*, October 1957, pp. 127 ff.

Recently, the picture has changed completely, and the fractions of output claimed are no longer consistent with the official Soviet index. Briefly stated, the latest Soviet position, announced in 1959 by no less an authority than Khrushchev, is that output was 12.5 per cent of the American level in 1913 and 50 per cent in 1958 (*Vestnik statistiki* [Statistical Bulletin], 1959, No. 11, pp. 17 ff). These fractions would imply that Soviet industrial production multiplied 18 times over 1913-1958, not 36 times as shown by the official index. The official claim is thus cut in half at one blow, and it still remains much too large, as we shall see.

Incidentally, the last-mentioned source states that the figure of 7 per cent for 1913 was never given official recognition, merely being an estimate of private Soviet economists. This is contrary to fact, as may be seen by examining the sources cited above.

apply are defined. These inaccuracies result in the main from misreporting—mostly overreporting—generated within the statistical system.

It is difficult for an outsider to appraise the reliability of Soviet statistics, since he must rely almost entirely on reports of émigrés<sup>25</sup> or on internal evidence, in the manner of the historian. With minor exceptions, public discussion of statistics is not allowed within the Soviet Union: they must be accepted without open question. And, since the government has a monopoly of statistics, it is not possible to check independently derived and published figures against each other.

In the discussion that follows, we shall consider, first, elements in the statistical system that promote distortion at various levels; second, examples of published statistical information that must be considered unreliable, by virtue of either distortion or ambiguity; and, third, the inferences about reliability that can be drawn from internal evidence presented by the statistics themselves. The discussion will be only suggestive, for, as Grossman remarks, "It would be futile to attempt to list all the pitfalls in the interpretation of Soviet statistics, even of only the industrial physical output data. In the final analysis each figure must be tested separately and on its own ground for possible descriptive distortion, always bearing in mind what it is that the statistics are 'trying to prove.'"<sup>26</sup>

#### MISREPORTING

Misreporting starts with the enterprise itself. The incentive reaches down to the worker and up to the manager. The worker's incentive derives, in the first place, from the piece-rate system of pay, which applies to almost every job where activity can be measured in physical units. As early as 1928, piece rates applied to more than half the hours worked by all persons engaged (excluding plant managers and superior echelons) in large-scale industry; the percentage rose to 70 by 1935. A comparable statistic is not available for later years, but in 1955 more than three-quarters of persons engaged in *all* industry were paid on a piece-rate basis. Piece rates tend to be progressive: the higher the output, the larger the pay per piece. In addition, special premiums are paid to some workers for economical use of inputs and other savings in unit costs, and non-pecuniary perquisites—such as vacations, better housing, and preferential

<sup>25</sup> The most comprehensive summary and thorough analysis of émigrés' views on this subject is contained in Joseph Berliner, *Factory and Manager in the USSR*, Cambridge, Mass., 1957.

<sup>26</sup> Grossman, *Soviet Statistics*, pp. 117 f.

rations of other types, where rationing is in force—accrue to workers with superior output records. Foremen and other overseers receive similar rewards, based on the performance of those under their supervision.<sup>27</sup> These factors all motivate the worker to exaggerate his output.

Overreporting by workers seems to be widespread.<sup>28</sup> Much of it applies to intermediate activities rather than to the final output of an enterprise, but this may indirectly force management to overreport final output to make it consistent with inflated wage costs. Direct overreporting of final output generally requires the cooperation of management and the independent inspectors. Since there is no conflict of interest between worker and management in this matter, such cooperation may be forthcoming wherever detection is difficult, as in the case of output measured in bulk. Moreover, the pervasive piece-rate system apparently extends in some instances to those who record final output. Thus, one example is known of a clerk whose job was weighing and recording the output of coal and who was paid a piece rate for the amount of coal recorded.<sup>29</sup> This case is perhaps extreme, but one can imagine similar jobs where a worker would be in a position to inflate finished output and where it would be in his interests to do so.

Management's incentive to inflate output derives from the fact that the system of rewards and penalties is geared primarily to its success in meeting or overfulfilling its output quotas. Other goals (such as planned profits) are important, but the manager receives special benefits and privileges to the extent that he accomplishes the output targets set for him—and special penalties to the extent that he fails to do so.<sup>30</sup> His foremost concern is, therefore, with the recorded output, and one would expect the representative manager to be tempted to improve on the actual record by one means or another. This incentive is strengthened by the fact that other indexes by which his performance is judged are improved step by step with the output record. The manager may react by writing up output or skimming on quality, a matter discussed in more detail in the next chapter.

One consequence is a tendency for the product mix to get arranged so

<sup>27</sup> This paragraph is based on *ibid.*, pp. 59 ff.

<sup>28</sup> See, e.g., Berliner, *Factory and Manager*, pp. 172 ff.

<sup>29</sup> Grossman, *Soviet Statistics*, p. 63.

<sup>30</sup> See Berliner, *Factory and Manager*, Chapter III; and A. Nove, "The Problem of Success Indicators in Soviet Industry," *Economica*, February 1958, 1-13. For a description of similar conditions in Hungary, see Bela Balassa, *The Hungarian Experience in Economic Planning*, New Haven, 1959, pp. 132 ff.

that it reflects the highest possible output in terms of the units of measure designated in the planned goal. Alec Nove cites a classic, if apocryphal, example of a nail factory:<sup>31</sup>

... When the plan was established in *numbers*, only small nails were made; so the basis of the plan was changed to weight, and then there were only *large* nails. If the plan is expressed in money, then only those which are cheapest to make will be produced, and probably all of the same size; if each type of nail is to be separately specified in the plan, this would be a glaring case of bureaucratic over-centralization. If the price of nails reflected supply-and-demand conditions, of course, things would be different; but this verges on heresy. Meanwhile, there are repeated appeals to the managers to provide a proper assortment of products.

A number of similar examples are documented by Grossman,<sup>32</sup> and need not be repeated here. The point to be made is that a shift in the unit of measure or an expansion in the coverage of a product category provides an opportunity for the skillful manager to "create" additional output without productive effort, merely by adjusting the product mix. This can be done only over a relatively short period of time, but the fictitious increase in output can occur each time there is a change-over in unit of measure or an expansion in coverage. When output series in different physical units are spliced together—as in the case of flat glass, leather, linen fabrics, and so on—the result may be a substantial exaggeration of the growth in output.

Where such opportunities as these are not present, the manager may resort to simulation. The techniques of simulation are too varied and complex to discuss at length here.<sup>33</sup> It should be noted, however, that devices have been found for "losing" simulated output in inventory and for "passing it on" to customers. While the economic system abounds with seeming built-in checks, these do not prevent widespread misreporting. Officials within an enterprise who are liable for inaccurate records—the

<sup>31</sup> A. Nove, "The Pace of Soviet Economic Development," *Lloyds Bank Review*, April 1956, p. 10.

<sup>32</sup> Grossman, *Soviet Statistics*, pp. 73 ff. For the situation in Hungary, see Balassa, *The Hungarian Experience*, pp. 140 ff.

<sup>33</sup> On them, see Berliner, *Factory and Manager*, Chapters VIII–X; and Grossman, *Soviet Statistics*, pp. 65 ff. Grossman comments (p. 66, n. 23): "Although Berliner's data refer primarily to the thirties, there seems to have been little fundamental change in this regard." This seems to be confirmed by Hungarian experience as related by Balassa (*The Hungarian Experience*, pp. 140 ff).

chief accountant, the head of the planning department, and so on—seem to be dominated by the plant manager and enmeshed in a “web of mutual involvement,” to use Berliner’s expressive phrase.<sup>34</sup> Measuring, counting, and weighing devices tend to be primitive and sparse. Freight is generally not weighed independently by the shipping agent, and sample surveys indicate that it is significantly overreported in weight.<sup>35</sup> Although quality inspection is conducted by an independent organization, its general ineffectiveness is attested to by Soviet authorities. For somewhat different reasons, the transportation system may aid in writing up shipments: its performance is assessed by the volume of traffic it handles. Finally, in the prevailing “sellers’ market” customers refrain from complaining about shortages or defective goods, since they are often happy to get anything at all—in any event, they generally prefer not to incur the disfavor of suppliers.<sup>36</sup>

The widespread practice of overreporting may seem strange for a state as authoritarian as the Soviet Union. Grossman gives the following explanation:<sup>37</sup>

It would seem at first glance that the multiplicity of controlling and auditing agencies . . . , the severity of the punitive measures at their disposal, and the thoroughness of the police system would successfully thwart the commission of such “economic crimes” as the falsification of output data and related illegal acts. Yet even the least acquaintance with Soviet reality leads one to the conclusion that “economic crimes” are extremely prevalent and to the conjecture that for each case that reaches the daylight of publicity there must be many that never do. An important factor is, of course, the inherent advantage that any insider has in concealing irregularities from the outside auditor’s view—what in its more extreme form might be called Pooh-Bah’s

<sup>34</sup> Berliner describes the basis for this involvement as follows (*Factory and Manager*, pp. 324 f): “Awareness of common interests in plan fulfillment often generates within the enterprise a ‘family relationship’ in which Party secretary, chief accountant, and other control officials facilitate or overlook the transgressions of an enterprising and successful director and share in the rewards and prestige that come with plan fulfillment. It is the fact that the control officials perceive their own fates as closely interwoven with the success of the enterprise that explains the endurance of the irregular practices of management.”

<sup>35</sup> See Ernest Williams, *Freight Transportation in the Soviet Union: A Comparison with the United States*, Occasional Paper 65, New York, NBER, 1959, pp. 11–13; and also Grossman, *Soviet Statistics*, pp. 98 f.

<sup>36</sup> For an extensive discussion of these checks and the ways they are thwarted, see *ibid.*, pp. 84 ff. An example of the willingness of low-priority consumers to accept defective sheet metal is given in *Current Digest*, IX, 48, p. 25.

<sup>37</sup> Grossman, *Soviet Statistics*, p. 91.

Law<sup>38</sup>—aided by the complexities of the very paper work that is intended to entrap the culprit, and abetted by the inspector's corruptibility and his reluctance to stir up a possible hornet's nest.

From this discussion one would gather that the possibilities of simulation diminish, the more closely the product in question is related to areas of high priority and the more precisely it can be measured. Thus, it is doubtful that significant distortion of output occurs in enterprises closely related to defense industries. By the same line of reasoning, one may suppose that the worst examples occur in enterprises producing consumer goods, for checks will be weakest here.

This brief survey of statistical misreporting at the enterprise level may be concluded by noting that underreporting also exists, though not as prevalently as overreporting. The most important cause of underreporting is pilferage or other unauthorized use of products.<sup>39</sup> This phenomenon is, however, not unique to the Soviet Union: output is understated in every country to the extent that there is pilferage. Moreover, it is not clear that pilferage will always cause underreporting. If it takes place before output is recorded, then output will be understated on this score. On the other hand, if it takes place after output is recorded—if, for instance, finished goods are taken out of inventory—then underreporting is not only more difficult but also less necessary. The effect here is, from an accounting point of view, the same as would be caused by overreporting of output; that is to say, fewer goods are available for shipment or for storage in inventory than are entered in the production record. If an enterprise can "lose" unproduced goods in its inventory accounts, it can also "lose" produced but stolen ones.

As one moves beyond the enterprise, less and less is known about possible distortions in statistics. Officials in the processing system are more closely related to the top Soviet leadership than are plant managers, and one would suppose that their activities would be less subject to extensive public criticism. It will be recalled that the processing system has had a dual structure. Each ministry in the economic-administrative organization would seem to have an incentive to inflate the output data reported to it, in order to make its performance look better than it actually is. During the period when the statistical organization was subordinate to the *Gosplan*, a similar incentive operated in that side of the structure. Finally, officials at various territorial levels in both the

<sup>38</sup> "... as Paymaster-General, I could so cook the accounts that, as Lord High Auditor, I should never discover the fraud" (footnote in original).

<sup>39</sup> For other causes, see *ibid.*, pp. 78 ff.

economic-administrative and (more significantly) the Communist Party organizations are interested in "improving" statistics for their regions.

Although independent tampering with statistics by only one interested party would be risky,<sup>40</sup> cooperative ventures offer more opportunity for success. One can imagine suitable occasions for such activity, but its prevalence and importance are anybody's guess. The few discussions of this matter to be found in the Soviet literature are essentially exhortations to statisticians to be honest and to resist whatever pressures there might be to get involved in "monkeying" with the figures. Speaking in 1955, Starovskii emphasized that the statistical organization was independent of local political authorities, but went on to say that "independence . . . means only that no local organization may force a worker in a [local] statistical administration or in a district or city inspectorate to change a figure if that figure is correct."<sup>41</sup> In 1956, a newspaper article appeared accusing the Central Statistical Administration of collaborating with political authorities in "adjusting" milk production upward by varying percentages in different provinces.<sup>42</sup> Such accusations are very rare, but they seem to testify that joint distortion is at least feasible.<sup>43</sup>

<sup>40</sup> B. P. Martschenko, an émigré Soviet economist, gives an example from personal experience in which he was able to verify that 1939 population data from Ukrainian provinces (*oblasti*) were faithfully reproduced in a compilation issued for internal use by the Ukrainian Statistical Administration, despite the fact that these data showed large deficits in population as a result of collectivization of agriculture. He goes on to say (as quoted in Grossman, *Soviet Statistics*, p. 114): "It must also be noted that the falsification of census data in the course of their processing in the *oblast'* statistical administrations would have been too unwieldy an operation, which would have inevitably become known to many persons in the statistical administrations, and could not have been concealed." These comments are certainly relevant to the matters at issue, but it must be kept in mind that Martschenko's example is drawn from the field of demography, where the pressures for internal distortion may not be as strong as in the case of industrial output. The pressures are, nonetheless, there, as may be seen from the sweeping purge of statistical personnel after the population census of 1937 produced findings distasteful to the Soviet leadership (see *ibid.*, p. 17).

<sup>41</sup> *Vestnik statistiki*, 1955, No. 1, p. 82, as quoted in Grossman, *Soviet Statistics*, p. 103.

<sup>42</sup> V. Surkov, "Counting on 'Incomplete Accounts,'" *Current Digest*, VIII, 14, pp. 37 f (original text in *Izvestia*, April 6, 1956). During my visit to Moscow in the summer of 1956, I submitted a written inquiry to the Central Statistical Administration about articles on inaccurate reporting of data and received the following reply:

"If you are referring to the article published in *Izvestia* on April 6, 1956 (we do not know of any other articles), the author, obviously not sufficiently informed, expressed the opinion that the Central Statistical Administration determined the milk yield on collective farms incorrectly and made corrections for omissions in collective farm accounting. These omissions lay in not including milk from cows attached to the children's institutions on collective farms or milk used to feed shoats on pig farms.

"In regions with a surplus beef production, the milk consumed on the farm is not included in the records. The milk fed to lambs, the milk consumed by the milkmaids, by the people who transport the milk to dairies, and by the collective farmers in whose quarters the cows are temporarily kept—all this milk is often not recorded.

"The USSR Central Statistical Administration has corrected all the collective farm



When we move to the publishing of statistics, we enter a rather different universe. The motive for misreporting at this level is perhaps more properly viewed as political and propagandist than as personal. The veil of secrecy surrounding the activities of the top Soviet leadership, enforced by a rigorous security apparatus, makes it impossible to know what happens to data between final compilation and publication. In particular, there is no way of knowing conclusively whether Soviet authorities keep two sets of books: one containing statistics for internal use only, the other for dissemination to the outside world. However, most Western specialists have concluded, for a variety of reasons, that dual accounts do not exist, in this narrow sense.<sup>44</sup>

The most direct evidence on the question of dual accounts is provided by a statistical annex to the 1941 Plan<sup>45</sup> that was captured during World War II by the Germans and later recaptured by the Americans. This document is labeled "not for publication," and it therefore presumably represents a compilation of data intended for internal use only. When the planned goals in this document are compared with those publicly announced in 1941, no significant discrepancies are found.<sup>46</sup> Although this conclusion applies directly to planned goals for 1941, it should be noted that they are significantly higher in general than published outputs for 1940 (see Tables 1 and 2).<sup>47</sup>

---

accounting on milk up to 1955. This adjustment amounted to 0.7 per cent of the total milk production in the USSR. Unlike in the U.S.A. and other countries, in the USSR these adjustments are made every year on the basis of a special check.

"At present, the CSA is conducting a routine investigation of the milk yield on collective farms, after which the question of making adjustments on the future data on milk will be discussed.

"The production of grain, meat and other agricultural products is recorded without adjustments."

A similar case involving adjustment of agricultural data was reported by P. Polynsky, "Why are Frauds Shielded in Chernovtsy?" *Current Digest*, IX, 42, pp. 20 f (original text in *Sel'skoe khoziaistvo*, September 12, 1957).

<sup>43</sup> An interesting example of collaboration in statistical misrepresentation, involving officials from the plant level up to Commissar Kaganovich, is recited from personal experience by Victor Kravchenko in *I Chose Freedom*, New York, 1952, pp. 298 ff. Similar cases have been reported in Hungary (see Balassa, *The Hungarian Experience*, pp. 145 and 148).

<sup>44</sup> See Grossman, *Soviet Statistics*, pp. 106 ff.

<sup>45</sup> *Gosudarstvennyi plan razvitiia narodnogo khoziaistva SSSR na 1941 god* [The State Plan for the Development of the USSR National Economy for 1941], Moscow, 1941 (reprinted by the American Council of Learned Societies, 1948).

<sup>46</sup> See Lynn Turgeon, "On the Reliability of Soviet Statistics," *Review of Economics and Statistics*, February 1952, 75-76.

<sup>47</sup> The 1941 Plan seems to have been ambitious, particularly in view of the fact that World War II was in progress elsewhere in Europe, and this supports other evidence that the Soviet Union probably did not expect to get involved in the war (see Chapter 8).

KNOWN AND UNKNOWN

TABLE 1  
OUTPUT FOR 1940 AND PLANNED OUTPUT FOR 1941:  
SOVIET UNION, 119 INDUSTRIES

	Unit	1940 Output <sup>a</sup>	1941 Planned Output	1941 Planned Output as % of 1940 Output
Pig iron	th.m.t.	14,900	18,000	121
Rolled steel	th.m.t.	13,110	15,830 <sup>b</sup>	121
Steel ingots and castings	th.m.t.	18,320	22,450 <sup>b</sup>	123
Quality steel	th.m.t.	3,196	3,914	123
Steel sheets (excl. pickled iron)	th.m.t.	1,786	1,752	98
Steel sheets (incl. pickled iron)	th.m.t.	1,822	1,827	100
Steel wire rods	th.m.t.	512	775	151
Steel beams and channels	th.m.t.	428	765	179
Iron and steel pipes	th.m.t.	966	1,100	114
Copper	th.m.t.	160.9	210	131
Nickel	m.t.	8,660	17,200	199
Electric power	bill.kwh	48.3	54.3 <sup>b</sup>	112
Electric power plants	mill.kw	11.3	12.4	110
Coal	mill.m.t.	165.9	190.8	115
Coke	mill.m.t.	21.1	23.8	113
Crude petroleum	mill.m.t.	31.1	34.6	111
Natural gas	th.m.t.	2,400	3,435	143
Peat	th.m.t.	33,200	39,615	119
Soda ash	th.m.t.	536	673	126
Phosphoric fertilizer	th.m.t.	1,352	1,980 <sup>b</sup>	146
Ground natural phosphate	th.m.t.	381.7	610	160
Synthetic dyes	th.m.t.	33.9	39.5	117
Rosin	th.m.t.	44.1	60.8 <sup>b</sup>	138
Paper	th.m.t.	812.4	969.9 <sup>b</sup>	119
Paperboard	th.m.t.	150.8	208.3 <sup>b</sup>	138
Motor vehicle tires	thousands	3,007	4,000	133
Red bricks	millions	6,723	8,359 <sup>b</sup>	124
Fire-clay bricks	th.m.t.	1,731	1,850	107
Quartzite bricks	th.m.t.	546	670	123
Sand-lime, silica, and slag bricks	millions	732	1,083 <sup>b</sup>	148
Cement	th.m.t.	5,675	7,998	141
Construction gypsum	th.m.t.	892	1,306 <sup>b</sup>	146
Industrial timber hauled	mill.m <sup>3</sup>	117.9	159.0 <sup>b</sup>	135
Lumber	mill.m <sup>3</sup>	34.8	30.3 <sup>b</sup>	87
Roofing iron	th.m.t.	103.4	230.0	222
Asbestos shingles	millions	205.6	253.4 <sup>b</sup>	123
Window glass	mill.m <sup>2</sup>	44.7	62.2 <sup>b</sup>	139
Railroad ties	millions	37.1	46.5	125
Rubberoid roofing	th.rolls	1,700	2,556 <sup>b</sup>	150
Pergamin subroofing	th.rolls	1,190	2,500	210
Tar-paper roofing	th.rolls	3,900	4,495 <sup>b</sup>	115
Railroad rails	th.m.t.	874.8	1,100	126
Sorted asbestos	th.m.t.	147.0	200.0	136
Asphalt	th.m.t.	74.4	150	202
Ginned cotton	th.m.t.	848.6	860.0	101

(continued)

THE DATA:

TABLE 1 (continued)

	Unit	1940 Output <sup>a</sup>	1941 Planned Output	1941 Planned Output as % of 1940 Output
Raw cotton	th.m.t.	2,495	3,010	121
Iron ore	mill.m.t.	29.87	34.03	114
Manganese ore	mill.m.t.	2.6	3.1	121
Automobiles	thousands	5.5	9.0	164
Trucks and buses	thousands	139.9	131.0	94
Diesel and electric locomotives	units	14	16	114
Steam locomotives	units	914	1,300	142
Railroad freight cars	thousands	30.9	60.5 <sup>b</sup>	196
Railroad passenger cars	units	1,051	900	86
Tractors (excl. garden)	thousands	31.6	28.0	88
Plows, tractor-drawn	thousands	38.4	35.4	92
Cultivators, tractor-drawn	thousands	32.3	32.5	101
Drills, tractor-drawn	thousands	21.4	33.5	157
Grain combines	thousands	12.8	13.0	102
Haymowers, tractor-drawn	thousands	3.3	3.0	91
Grain-cleaning machines	thousands	4.3	2.3 <sup>b</sup>	53
Steam boilers	th.m <sup>2</sup>	276.3	272 <sup>b</sup>	98
Water turbines	th.kw	207.7	280.6	135
Diesel engines	th.hp	248.7	368	148
Other internal combustion engines	th.hp	165	165.3	100
Turbogenerators	th.kw	313.5	644.5	206
Hydroelectric generators	th.kw	154.6	379.3	245
Electric motors (a.c.)	th.kw	1,848	2,622	142
Power transformers	th.kva	3,500	5,120 <sup>b</sup>	146
Coal-cutting machines	units	1,256	1,860	148
Machine tools	thousands	58.4	58.1 <sup>b</sup>	99
Bench and engine lathes	thousands	11.5	13.8	120
Spinning machines	units	1,109	2,000	180
Looms	units	1,800	3,150	175
Cotton-carding machines	units	1,312	1,970	150
Typesetting machines, linotype	units	145	120	83
Flat-bed printing presses	units	258	260	101
Industrial sewing machines	thousands	20.3	18.0	89
Excavators	units	274	490	179
Scrapers, tractor-driven	units	2,104	2,000	95
Railroad cranes, steam-operated	units	258	145	56
Automatic switchboards	th.lines	37.5	61.5	164
Metallurgical equipment	th.m.t.	23.7	45.0	190
Equipment for oil industry	th.m.t.	15.5	22.0	142
Macaroni	th.m.t.	324	392.1 <sup>b</sup>	121
Butter	th.m.t.	226	251 <sup>b</sup>	111
Vegetable oil	th.m.t.	798	737	92
Oleomargarine	th.m.t.	121	126.5 <sup>b</sup>	105
Cheese	th.m.t.	38.0	44.5 <sup>b</sup>	117
Meat	th.m.t.	1,183	1,367 <sup>b</sup>	116

(continued)

KNOWN AND UNKNOWN

TABLE 1 (concluded)

	Unit	1940 Output <sup>a</sup>	1941 Planned Output	1941 Planned Output as % of 1940 Output
Sausages	th.m.t.	391.3	395.6 <sup>b</sup>	101
Fish catch	th.m.t.	1,404	1,704 <sup>b</sup>	121
Soap	th.m.t.	700	748 <sup>b</sup>	107
Salt	th.m.t.	4,400	4,780	109
Raw sugar	th.m.t.	2,165	2,745 <sup>b</sup>	127
Yeast	th.m.t.	48	77 <sup>b</sup>	160
Canned food	mill.cans	1,113	1,263 <sup>b</sup>	113
Beer	th.hectoliters	12,130	13,450 <sup>b</sup>	111
Cigarettes	billions	100.4	114.2 <sup>b</sup>	114
Matches	th.crates	10,000	12,270 <sup>b</sup>	123
Vodka	mill.decaliters	92.5	95.7 <sup>b</sup>	103
Confectionery	th.m.t.	790	1,098 <sup>b</sup>	139
Boots and shoes	mill.pairs	211.0	223.6 <sup>b</sup>	106
Rubber footwear	mill.pairs	69.7	82.4 <sup>b</sup>	118
Cotton yarn	th.m.t.	650	716 <sup>b</sup>	110
Cotton fabrics	mill.m	3,954	4,402 <sup>b</sup>	111
Linen fabrics	mill.m	285.2	293.7 <sup>b</sup>	103
Silk and rayon fabrics	mill.m	76.6	80.8 <sup>b</sup>	105
Woolen and worsted fabrics	mill.m	119.7	128.8 <sup>b</sup>	108
Knitted goods	millions	183.0	195.2 <sup>b</sup>	107
Hosiery	mill.pairs	485.4	550.9 <sup>b</sup>	113
Felt footwear	mill.pairs	17.9	18.3	102
Rubber galoshes	mill.pairs	45.0	55.5 <sup>b</sup>	123
Bicycles	thousands	255.0	402.0 <sup>b</sup>	158
Electric light bulbs	millions	139.8	142.0 <sup>b</sup>	102
Phonographs	thousands	313.7	270.0	86
Radios	thousands	160.5	355.0 <sup>b</sup>	221
Clocks and watches	thousands	2,796	3,405	122
Household refrigerators	thousands	3.5	1.5	43

SOURCE: Appendix Table B-2 and *Statistical Abstract of Industrial Output in the Soviet Union, 1913-1955*, New York, NBER, 1956.

<sup>a</sup> On Soviet territory as of end of 1940.

<sup>b</sup> Planned output as given in source adjusted upward to cover acquired Baltic territories. For latter planned output, see *Gosudarstvennyi plan 1941*, pp. 704 ff.

DEFICIENCIES AND DISTORTIONS IN PUBLISHED DATA<sup>48</sup>

Whatever one may conclude about the existence of dual accounts—and the weight of evidence seems to bear against their existence—it is clear that published statistics suffer from lack of reliability because of selectivity, ambiguity, and misrepresentation. For the moment, we shall be concerned primarily with the last two.

<sup>48</sup> This section is based largely on tabular material and notes in *Statistical Abstract of Industrial Output in the Soviet Union, 1913-1955*, New York, NBER, 1956. Examples of defective statistics, in addition to those given here, may be found in Grossman, *Soviet Statistics*, pp. 117 ff.

THE DATA:

TABLE 2  
 FREQUENCY DISTRIBUTION OF PLANNED OUTPUT FOR 1941  
 AS A PERCENTAGE OF ACTUAL OUTPUT IN 1940:  
 SOVIET UNION, 119 INDUSTRIES

1941 Planned Output as % of 1940 Output	Number of Industries
Under 85	4
85 to 95	9
95 to 105	15
105 to 115	22
115 to 125	23
125 to 135	5
135 to 145	13
145 to 155	9
155 to 165	6
165 to 175	0
175 to 185	4
185 to 195	1
195 and over	8
Total	119

SOURCE: Table 1.

An important source of ambiguity is failure to clarify the precise coverage of industries. It is sometimes doubtful whether a published datum refers to the sector of an industry under ministerial jurisdiction or to the whole, to large-scale (or state) industry or to the whole, and so on. In some cases there is doubt about territorial coverage. These shortcomings have been remedied in large measure in the recent Soviet statistical abstracts, but some remain, in particular for that stretch of years in which statistics were most heavily suppressed. It is asserted in the Soviet abstracts for both 1936 and 1957 that all data refer to entire industries except where specifically noted to the contrary. Yet examples can be found where all or a substantial portion of small-scale production is not included in early years (e.g., soap, beer, boots and shoes, silk fabrics, and woolen and worsted fabrics), even though no warning is given.

The treatment of the flour industry gives an example of ambiguity in administrative coverage in earlier years. Until the recent appearance of *Industry of the USSR*, output of flour and groats had been published only for the interwar period. In some years output was given for large-scale industry, in later years for all industry except collective farm mills, and in still later years for all industry producing flour from centralized procurements of grain. This amounted to a temporal expansion in the

coverage of the industry, not pointed out in the statistical sources, and there was an illusion of substantial growth in output, whereas growth was modest, at least according to the recently published data.

Another case of expanded coverage, not yet clarified, is provided by industrial timber. The data apply to haulage out of the State Forest Reserve, accounting for almost all timber now but for only a fraction in the 1920's. The prerevolutionary counterpart used in Soviet statistics—the Crown Forests—accounted for an even smaller fraction. The changing coverage is not described in usual statistical sources, and the published data therefore exaggerate growth in timber haulage from the pre-revolutionary period to the present.<sup>49</sup>

The effects of territorial expansion during World War II are generally not explicitly revealed in output statistics. Data for 1940 and later years cover the expanded territory, while data for earlier years cover the inter-war territory. Recently, output of some industries has been given for 1913 within the expanded territory, but this does not indicate the gains in 1940 through territorial acquisitions.<sup>50</sup>

Product coverage of industries is less well known than administrative coverage. Uncertainty about stage of fabrication and composition of products applies to standard industrial materials as well as to more highly fabricated products. For instance, it is not known whether the recent data for nonferrous metals refer to only primary metal or both primary and secondary metal, nor is it known at what stage of fabrication output is measured. These are matters of some importance: recent output of copper in the United States is more than doubled by moving from a definition covering only blister copper produced from domestic ore to a definition covering all types of refined copper.<sup>51</sup> In the case of more heterogeneous items (such as ball bearings, machine tools, cameras, and so on), vagueness in definitions is even more serious, particularly since output is often reported in units, actual or conventional.<sup>52</sup> Again, the main deficiency of Soviet statistics is inadequate detail, in this case, of product groups. And, again, the situation has improved recently.

Definitions of industries are not only vague but also subject to change

<sup>49</sup> The data on timber haulage used in our study have been adjusted to provide comparable coverage for all years.

<sup>50</sup> It has been possible to estimate those gains for some industries on the basis of output in the acquired territories in 1937 (see Appendix Table B-3).

<sup>51</sup> *Statistical Abstract of the United States, 1956*, Washington, 1956, p. 750.

<sup>52</sup> Ball-bearing units may vary from one used in bicycles requiring fifty seconds to manufacture and weighing a few grams to one used in railway cars requiring twenty-six hours to manufacture and weighing forty kilograms (see *Planovoe khoziaistvo* [Planned Economy], 1956, No. 5, p. 82).

without notice. Changes of this sort are, of course, often unavoidable—even desirable—and are to be condemned only when they are obscured. Usually a change is signaled by a slight alteration in terminology. It may be the dropping or adding of a qualifying phrase. In the course of our study, we did not find a single instance in which attention was directed by statistical sources to a change in definition. The investigator is left to his own devices in finding out whether there has been a change, what it means, and how it affects comparability of data. Frequently, a shift in definition will become known only through curious inconsistencies in fragmentary information uncovered in the course of research. It may be helpful to expand on this matter by giving a few specific examples.

Up to 1949, the “mineral fertilizer” industry covered soluble superphosphates, nitrates, and potassic compounds. The most important product not included was ground natural phosphate, an unprocessed material that is not readily soluble. Coverage was expanded in 1949 to bring in this product, and output was thereby inflated by about an eighth. Aside from some inconsistencies in data that arose, the only sign of a change at the time was the following alteration in title: up to 1949, the industry had been called “mineral fertilizers (superphosphates, nitrates, and potash)”; since 1949, the parenthetical phrase has been dropped. The nature of the change was confirmed when output series for the components appeared in *Industry of the USSR*. No mention is made of the expansion in coverage over series appearing in earlier sources.

The term “canned food” has covered a variety of products, differing in many instances with the sources giving data. Little is known about the composition of products since the middle 1930’s, but a significant relaxation in the meaning of the term took place in the early 1930’s, never described in detail in primary statistical sources. Up to that time, “canned food” had been used to mean food packed in hermetically sealed containers; at some point in the early 1930’s, it came to mean any kind of preserved food, no matter how packed. Thus, processed foods packed in bulk—as pickles in the the barrel and salt pork—apparently came to be taken in under the name “canned food.” In 1934, hermetically sealed products accounted for less than a third of “canned food.” Recent information indicates that “canned food” still includes products not hermetically sealed. At the same time, output for the 1930’s has been revised substantially downward, which suggests that some of the bulk products—we do not know which ones—have been removed from coverage.

A similar shift in coverage of the “confectionery” industry seemingly

took place around the beginning of the Plan period, when cakes and other baked goods were added to the candy already included. During the interwar years, these bakery goods accounted for between 30 and 40 per cent of the output of "confectionery." It seems probable that better grades of bread were also classified as "confectionery" when bread rationing was in effect during the early thirties (there was only one grade of rationed bread). Information for the postwar period shows that bakery goods are still included, but it has never been pointed out that the definition of "confectionery" is considerably broader than for early years.

By tracing through changes in terminology, one notes that the coverage of the "meat" industry has been expanded at least twice. Data for 1930 and later years are given in Soviet sources as applying to "meat and meat products," whereas for earlier years they are given for slaughter weight of meat alone. This expansion in coverage presumably amounted to counting some meat products twice: once at the slaughtering stage and again at the processing stage. A second shift in coverage took place with the publication of *The National Economy of the USSR*; in this source, the industry is called "meat and by-products of Category I," an unexplained expansion in coverage—lard seems to have been added, among other things—that raised output by about a quarter.

Examples could be multiplied, but it is perhaps sufficient to conclude with brief comments on a few other cases. Up to 1928, "soap" included only the common bar soaps used for laundering; after that date, coverage was expanded to include all types of soap. Similarly, the term "leather footwear" originally included only boots and shoes made of leather but later came to include all kinds of footwear—even rebuilt shoes—except those made entirely of felt or rubber. In the case of "vegetable oil," the output for 1928 given in the recent statistical abstracts apparently covers only edible oil, whereas output for later years covers nonedible oil as well. During the pre-Plan period, the "fish catch" included only those fish caught by commercial fishermen; during the Plan period, fish caught in ponds by collective farmers and other local fishermen have also been included, though one may wonder how this is estimated.

In some heterogeneous industries, the output of component products is often aggregated by means of "conventional units." In some cases, Soviet practice differs sharply from Western usage, and the failure of Soviet sources to describe the practice makes it difficult to avoid misinterpretation. For instance, many block-like and brick-like construction materials seem to be counted as "brick," and their output is apparently



## THE DATA:

expressed in some kind of brick equivalents. The output of flat glass is measured in square meters—as in the United States—but only after the different kinds of glass have been converted in an unknown way to conventional units equivalent to window glass with a standard thickness of 2 millimeters. Neither of these procedures is noted or described in primary Soviet sources.

A few specific examples drawn from the technical literature illustrate the complex nature of conventional units. Output of “canned food” is said in statistical sources to be expressed in terms of a conventional can of 400 grams. In fact, the standard unit for hermetically sealed products is a container with a volume of 353.4 cubic centimeters, multiplied by coefficients varying with the product. Thus, beef stew of first and superior grades has a coefficient of 1.13; lamb stew of first grade, 1.2; and lamb stew of superior grade, 1.4. For “canned goods” packed in bulk, the standard unit is a net weight of 400 grams, multiplied by a coefficient varying with density. The rationale for these coefficients is not apparent, unless they are designed to reflect presumed qualitative differences. Similar coefficients are known to be used in the cases of shoes, sausages, lumber, plywood, iron and steel products, producer equipment, agricultural equipment, forest products, and building materials.<sup>53</sup> One Soviet economist, M. A. Tseitlin, states that all but a handful of the output targets listed in the Fourth Five Year Plan were actually expressed in conventional units, involving conversion coefficients of various types, even though they were said to be measured in “physical units.” Among the few exceptions were electricity, petroleum, natural gas, and most processed foods.<sup>54</sup>

The stage of fabrication at which output is measured sometimes does not accord with Western practice, and since it is not revealed in primary sources, one may be misled about productive activity. In most countries, the output of cotton fabrics is recorded at the unfinished or “gray goods” stage. This was also the case with Tsarist statistics. During the 1930's, Soviet statistics began recording output at the finished stage, after dyeing and finishing. This change in practice has taken on significance in the postwar period, since substantial quantities of cotton goods have come to be produced in Poland for export to the Soviet Union. It is quite possible that these Polish exports are gray goods later finished in the Soviet Union and hence counted as Soviet output. The same may also

<sup>53</sup> This paragraph is based on S. A. Gorelik, *Statistika* [Statistics], Moscow, 1956, pp. 29 ff; and Grossman, *Soviet Statistics*, pp. 119 f.

<sup>54</sup> *Ibid.*

be true for railway equipment produced in Poland but "finished" (by, say, painting and labeling) in the Soviet Union.

On the other side, there are cases where it is not made clear that output is being measured at primary stages of fabrication. "Granulated sugar" (*sakhar pesok*), for instance, apparently includes all sugars and syrups (converted into "sugar equivalents") at the crudest processing stage. A part is used directly for household consumption, a part is further processed into "refined sugar," and a part is consumed industrially. Similarly, "vegetable oil" includes that consumed directly and that used in making other products (for instance, margarine).

We have already given incidental illustrations of how the ordinary user of Soviet statistics can be badly mistaken about the meaning of terms, because they diverge from customary usage. Two examples may be added: "silk fabrics" is the title used to identify all fabrics made in whole or in part from artificial and synthetic fibers as well as from silk, and "slate" is the title used to identify asbestos shingles.

Misleading language reached its zenith in the postwar years before 1956, a period in which statistics lost all vestiges of being a science and became instead a linguistic art. The practices then followed are illustrated by the case of machine tools. In the postwar announcements of annual percentage increases in output, data were published under no fewer than four different titles for the machine tool industry, varying from one year to the next. A complete series of percentages was not published under any one of these titles. In the general literature, the product was sometimes times referred to as "machine tools" (*stanki*) and sometimes as "metal-cutting machine tools" (*stanki metallovezhushchie*). It appears from the recent Soviet statistical abstracts that the former include forges and presses while the latter do not. Similarly, output was sometimes referred to as "deliveries" or "sales" (*vypusk*) and sometimes as "production" (*produksiia*). This confused mixture of terms made it impossible to know what was going on in this sector of industry, though careless use of the published figures could lead to an exaggerated picture of performance.

This discussion may be concluded with a few words on the Soviet concept of output itself. According to formal requirements, the product of an enterprise is supposed to be counted as output only when it has passed quality inspection and when it has been delivered to a warehouse or buyer. Goods rejected for failure to meet standards of quality, either by inspectors within a plant or by buyers, are classified as *brak* and are supposed to be excluded from output. But this provision is formally

operative only if the defective goods are discovered and reported within the year in which they are produced, a loophole that would seem to encourage bunching shipments of *brak* around the end of a year.<sup>55</sup> In addition, the standards of quality are low in some industries, and *brak* may mean "most defective."<sup>56</sup> Finally, quality inspection leaves much to be desired, as Soviet authorities complain.<sup>57</sup> As Grossman reports: "A safer and clearly very widespread method of writing up output is the inclusion of *brak* in the reported amount of finished product. Direct references in the Soviet press, eyewitness testimony, and the continual complaints about the substandard quality of industrial products bear such ample and conclusive evidence of the prevalence of this practice in Soviet industry, despite severe criminal and administrative sanctions against it, that it is not necessary to dwell on it further at this point."<sup>58</sup> The concrete effects of this practice will be revealed in more detail in the next chapter.

A more specialized problem of interpretation has to do with the measuring of output in machinery industries. As late as 1938, it was common practice in the power equipment industry to count a complex machine as produced whenever a piece of auxiliary equipment was completed. Thus a steam turbine would be reported as produced when its condenser pump, say, was finished. It was said to be normal for two years to pass between the recording of production of final products (such as turbines) and the actual completion; one case was cited in which five years passed.<sup>59</sup> It is, of course, conceivable that the final product would never be produced. There is no way of knowing whether this

<sup>55</sup> See *ibid.*, pp. 66 and 70 ff. An apparent recent example of such bunching is given in the article on the Altai Tractor Plant, *Current Digest*, X, 3, p. 27.

<sup>56</sup> In the late 1930's, there were three "standard" grades of textiles and at least three "substandard" grades. It appears that only the worst of the latter qualified as *brak*, since the first two "substandard" grades were offered for sale. For a description of the standards of quality for textiles in those years, see P. Fadeev and D. Zamkovskii, "O kachestve standartov tekstil'nykh tovarov" [On the Worth of the Standards for Textiles], *Voprosy sovetskoi torgovli* [Problems of Soviet Trade], 1936, No. 10, pp. 35-42. The following quotation (p. 38), which has to do with varying "standard" grades, is enlightening: "A consumer who buys three meters of drapery fabric that looks moth-eaten and has all the colors of the rainbow—i.e., is completely useless—receives at best a 7 per cent reduction in price if the fabric is third quality. But if this defect is only in those three meters, then the reduction is only 3 per cent because the fabric is second quality, although it makes absolutely no difference to the consumer who buys that piece whether the defect is in all the material or just in his piece. If the defect is only 2.99 meters long, then the fabric is first quality."

<sup>57</sup> Grossman, *Soviet Statistics*, pp. 87 ff.

<sup>58</sup> *Ibid.*, p. 68. For a careful discussion of the problem of *brak*, see Berliner, *Factory and Manager*, Chapter IX.

<sup>59</sup> I. Nelidov, "Somnitel'nye metody planirovaniia" [Doubtful Planning Methods], *Mashinostroeniia* [Machine Building], September 30, 1938.

practice has continued into the postwar years, but it certainly was important in the interwar period.

INTERNAL EVIDENCE ON RELIABILITY

The evidence suggests that data on physical output are generally less accurate in the Soviet Union than in the West. There can be little doubt that Soviet data are generally exaggerated by a significant amount—precisely how much it is impossible to know. Nevertheless, one must not move from this conclusion to a far broader one, namely, that the data are wholly unreliable and useless. They are not a mere collection of numbers taken out of the air. The internal relations among the statistics demonstrate that they are based on reality, even though they diverge from it. In considering this internal evidence, we shall pass from the least conclusive to the most.

The first thing to be mentioned is that there is a basic consistency among data relating to differing administrative coverages: the larger the coverage, the larger the figures. This is in itself not very meaningful, since the first thing that would be attended to in manipulation of statistics would be this kind of elementary consistency. It is more meaningful for the 1920's than for later years, because two agencies (*VSNKh* and *TsSU*), functioning independently in this regard, collected data for different administrative coverages.

The consistency of data for related products is more significant. For example, in the iron and steel complex the series for iron ore, pig iron, coke, steel ingots, and rolled steel move more or less together, and at the same time diverge in accord with known developments. Since 1928, iron ore production has risen more percentagewise than pig iron production because of deterioration in the quality of ore; pig iron has risen less than steel ingots because of increased use of scrap; and steel ingots have risen more than rolled products because of increased use of castings and forgings. Similarly, output of electric power has grown more rapidly than installed capacity, which is consistent with known trends toward a more even consumption of electricity during the day and over the year. In the textile industry, production has grown more rapidly for cotton than for cotton fabrics, while it has grown less rapidly for wool than for woollen fabrics. Both these divergences are consistent with decreased reliance on imports of cotton, with reduced length of staple, and with increased use of cotton in woollen fabrics. Many more examples of this kind could be given, but these suffice to make the point.

The third line of internal evidence turns about the fact that selectivity

and ambiguity are used to conceal whatever it is desired to conceal. Poor performance is habitually masked by silence or evasion. Cases are known of slow-growing and declining industries where no effort has been made to publish data to the contrary; instead, nothing is said at all. In a few cases, like flour milling, data have been ultimately released confirming the worst of Western suspicions. During the postwar years when only annual percentage changes in output were being reported, industries with declines were simply omitted from the list; recently published statistics reveal that some of the declines were substantial (e.g., for many machinery items in 1952). This all merely provides clear evidence that black has not been indiscriminately turned into white in the basic Soviet statistics on physical output.

At the same time the difficulties attributable to the policy of secrecy must not be overlooked. At least until very recently, published Soviet statistics have been carefully selected. To illustrate the selectivity, we may consider frequency distributions of annual relatives of output for three different samples of industries: the first (sample A), as published up to the end of 1955; the second (sample B), as published up to the end of 1956; and the third (sample C), as published up to the end of 1957 (see Table 3). These samples are not strictly comparable in nature. Sample A merely contains all the annual percentage changes in output as announced in reports of plan fulfillment, and the industries covered therefore vary substantially from year to year. Moreover, a number of minor industries and industries with fluctuating product coverage are included. Samples B and C, on the other hand, are composed of industries with essentially continuous output series over the period surveyed. The earlier samples show an upward bias relative to the later ones. The tail of the frequency distributions containing relatives below 100 per cent—i.e., representing industries with annual declines in output—tends to grow increasingly longer as we move from sample A to sample C in each year. In fact, no declines in output are shown in sample A except for 1955. Similarly, the median annual relative—that relative exceeded and fallen short of by half the industries—tends to decline as we move from sample A to sample C. These frequency distributions may be compared with a similar set of distributions for the industries included in the Federal Reserve Board index of U.S. industrial production (see Table 4). The Soviet distributions for sample C accord much more closely in nature with the American distributions than do the Soviet distributions for samples A and B. We note the reduction in bias as more statistics have been revealed. We have no way of knowing whether or how much the bias



THE DATA :

TABLE 3 (concluded)

Annual Relatives <sup>b</sup> (per cent)	Number of Industries					
	Sample A	Sample B	Sample C	Sample A	Sample B	Sample C
	1953			1954		
Under 60			1			2
60 to 70			1			
70 to 80			3		1	3
80 to 90		1	3		1	2
90 to 100		4	14		2	14
100 to 110	22	23	48	27	25	48
110 to 120	37	36	55	37	31	60
120 to 130	13	9	25	11	9	15
130 to 140	5	3	5	5	2	4
140 to 150	5	2	5	3	1	6
150 to 160	2		4	2	1	2
160 to 170				2	2	2
170 to 180	1	1	1	1	1	3
180 to 190				1		
190 to 200				1		2
200 and over	2	2	5	5	1	9
Total	87	81	170	95	77	172
Median (%)	116	114	113	116	113	113
		1955				
Under 60			2			
60 to 70			1			
70 to 80			3			
80 to 90	1	2	11			
90 to 100	2	5	15			
100 to 110	28	23	45			
110 to 120	31	35	55			
120 to 130	18	12	24			
130 to 140	5	2	6			
140 to 150	2	1	6			
150 to 160	1	1	1			
160 to 170	2		2			
170 to 180						
180 to 190	1		1			
190 to 200	3		2			
200 and over	1	1	2			
Total	95	82	176			
Median (%)	115	113	112			

SOURCE: Sample A: *Statistical Abstract of Industrial Output in the Soviet Union*, Supplement, Table 3. Sample B: *ibid.*, Part 1. Sample C: Appendix B.

<sup>a</sup> Sample A refers to output data published up to end of 1955; Sample B, up to end of 1956; and Sample C, up to end of 1957.

<sup>b</sup> Output in specified year as percentage of preceding year.

KNOWN AND UNKNOWN

TABLE 4

FREQUENCY DISTRIBUTIONS OF ANNUAL RELATIVES OF PHYSICAL OUTPUT OF INDUSTRIES IN FEDERAL RESERVE BOARD INDEX OF INDUSTRIAL PRODUCTION: UNITED STATES, 1948-1953

<i>Annual Relatives</i> (per cent)	<i>Number of Industries</i>					
	1948	1949	1950	1951	1952	1953
Under 75	5	17	2	4	7	1
75 to 80	0	9	0	6	0	0
80 to 85	4	18	0	8	6	2
85 to 90	10	21	3	9	18	4
90 to 95	16	38	0	13	17	6
95 to 100	31	28	10	33	46	32
100 to 106	51	32	27	31	53	50
106 to 111	29	17	21	27	22	40
111 to 116	18	4	26	21	9	20
116 to 121	13	3	30	10	2	16
121 to 126	6	0	22	9	3	6
126 and over	10	6	52	22	10	16
Total	193	193	193	193	193	193
Median (%)	103.5	92.4	116.3	103.7	99.5	105.6

SOURCE: Special computation by the Federal Reserve Board.

would be reduced by a full disclosure of data comparable to the practices followed in the United States.

Another aspect of selectivity is suppression of information about industries related to the military effort. Since production in these areas has generally grown faster than the average for all industry, this policy imparts a downward bias to the sample of published output data. In some cases (like nonferrous metals and chemicals), we cannot be sure whether data are suppressed because growth has been fast or because it has been slow. It is even likely that much secrecy is simply due to the traditional Russian love of mystery.

The fourth and most important line of internal evidence on the reliability of Soviet data has to do with the reasonableness of the patterns of growth that emerge from published Soviet data. The sector known as heavy industry is shown to have grown much more rapidly than the sector known as light industry; this certainly accords with general conditions, as every traveler to the Soviet Union can testify. To the person who has studied economic growth in other countries, it is more important to note that there is a general tendency among Soviet industries to grow more slowly percentagewise as they get older and larger, a phenomenon that goes by the name of "retardation in growth"; in this respect, the behavior of the Soviet economy has been quite similar to the



behavior of other economies about which a good deal more is known. Along the same lines, the published Soviet data show that the rates of growth of Soviet industries have been closely related to the stages of development from which they started: in general, those industries that were least "advanced" in the prerevolutionary years relative to other countries have grown most rapidly, while those most "advanced" have grown least rapidly. There is also a general consistency in the stage of development of related groups of Soviet industries, as determined by comparisons with various periods of development in the American economy. Finally, there is a basic consistency between transportation and industrial statistics, similar in important respects to the relation holding for the United States in earlier periods of development. There is also a reasonable relation between industrial employment and output. These matters are discussed more concretely later on and need not be elaborated here. The point to be made at this time is that the available Soviet data on physical output present a picture of growth patterns that makes sense.

*Some Generalizations About Soviet Data*

The evidence bearing on the reliability of Soviet data cannot be summed up in a few words, nor can simple judgments be made. The degree of reliability depends on the purposes for which the data are to be used. In general, absolute magnitudes of physical output are likely to be less accurate than for Western statistics. Similarly, the products to which the data apply are less easily identified. These shortcomings are likely to be less pronounced in industries of high priority, especially if output is subject to rather precise measurement. Thus data on the output of coal are undoubtedly more reliable than those on the output of meat because one has had a higher priority than the other; and data on the output of steel ingots are probably more reliable than those on the output of coal, because one is measured more precisely—and can be checked more precisely—than the other. The shortcomings in absolute magnitudes are most important when levels of output in the Soviet Union are being compared with levels elsewhere. They are somewhat less important when growth trends in Soviet output are being considered, for there is little evidence of a systematic trend in the relative inaccuracy of data, except that data for prerevolutionary and early Soviet years are generally understated in comparison with those for later years. The defects are even less important when percentage movements are being compared among Soviet industries.

We may illustrate with a concrete example. We should allow a wide

## *KNOWN AND UNKNOWN*

margin of error in comparing Soviet and American outputs of cotton fabrics: the products are not the same, the units of measurement are not the same, and the Soviet data have an upward bias in addition. The margin of error is probably less significant if the growth of the Soviet cotton fabrics industry is being considered, particularly if growth is being discussed in terms of annual average rates. There is almost certainly an upward bias in the percentage growth over the Soviet period as a whole (because of relative understatement of earlier data), but probably not over some later stretches of years. Finally, the margin of error is likely to be smaller still when comparisons are made between Soviet growth rates for, say, cotton fabrics and steel. Bearing the necessary qualifications in mind and exercising care along the way, we can use Soviet data on physical output to sketch a picture of Soviet industrial growth.