

THE PROBLEM OF ACCURACY OF ECONOMIC DATA

PHILIPP BAGUS*

In his classic book *On the Accuracy of Economic Observation* Oskar Morgenstern deals with a common, yet widely neglected problem with which economic historians are faced, namely the quality of economic data. For the economic historian in the Austrian tradition, the quality of economic data is of utmost importance, since false data or belief in inaccurate data can lead the economic historian to faulty interpretations of the past.

The quality of economic data is at least as important for economists who adhere to positivism in economics, since they use economic data to confirm or falsify their models.

Likewise, Morgenstern's insights are relevant for mathematical economists, as it makes sense to perform computations and solve a system of mathematical equations only if one has reliable data. Morgenstern illustrates this in the following example.

The equations

$$\begin{aligned}x - y &= 1 \\x - 1.00001y &= 0\end{aligned}$$

have the solution $x = 100001$, $y = 100000$, while the almost identical equations

(*) University Rey Juan Carlos in Madrid.

$$\begin{aligned}x - y &= 1 \\x - 0.9999999y &= 0\end{aligned}$$

have the solution $x = -99999$ $y = -100000$.

The coefficients in the two sets of equations differ by at most two units in the *fifth* decimal place, yet the solutions differ by 200,000.¹

Morgenstern's sample equations show the significance of a small error in the observation. Yet, in more complex equations with extensive mathematical operations the extent of error due to unreliable data may increase (or, depending on the equation, the errors may cancel out). It is indeed surprising to note how much the problem of accuracy in economic data has been neglected.

This is not so in the physical sciences. There the error of observation is always explicitly mentioned. Yet in economics there is simply no error estimate. This means that we do not know the accuracy of the economic data presented to us. This is even more troubling when we consider that in social or economic data there are more possible sources of error than in the physical sciences. We therefore face the question of why the problem of accuracy of economic data is rarely mentioned or passed over in silence in economics, while in the physical sciences this problem is widely acknowledged.

SOURCES OF ERRORS IN ECONOMIC STATISTICS

Oskar Morgenstern names several sources of error that influence the accuracy of economic observation. One is a lack of designed experiments. The observations are not produced by the user of an experiment, as in the natural sciences, but rather, statistics

¹ See Morgenstern (1963), p. 109.

are simply a byproduct of business and government activities. There is a complete lack of incentive to provide accurate information for government statistics and economic researchers on the part of companies, because to do so would require a costly and burdensome process.

In addition to the lack of accurately designed collections of data, there exists a related problem, also absent in the physical sciences – namely, the possibility of hiding of information or outright lying.

Companies have strong incentives to hide information or lie in order to mislead their competitors about their competitive strategy or strength. Companies also have an incentive to lie to the tax authorities and to the government in general in order to seek subsidies or avoid taxation. Sometimes companies manipulate profits in order to pay out fewer dividends.

Likewise, governments themselves have an incentive to falsify statistics, thereby improving their economic record. Doing so improves the ruling party's chances of staying in power. Falsification of economic statistics can also improve the likelihood of receiving some kind of foreign aid or foreign recognition. A recent example involved the Greek government, whose officials falsified the Greek budget deficit in order to gain entrance into the European monetary union.

Another potential source of error consists in the inadequate training of those who observe economic data. Whereas in the physical sciences the observers are the scientists conducting the experiment, the observers of economic data are often not trained at all. A lack of training can lead to error in data collection. For instance, errors may stem from questionnaires. The conductor of the research, does not normally conduct all interviews. Instead, the interviews are likely conducted by different persons. As a result, the delivering of the questions, the setting up, the interpretation and the recording of the answers are additional sources of error. The errors in mass observation do

not necessarily cancel each other out. Frequently, such errors are cumulative.

An additional potential source for errors is the lack of clear definitions or classifications. These problems apply, for instance, in the classification of goods, types of employment, or classification of companies within industries. Companies like General Electric operate in various industries, making it difficult to assign its revenues or profits to distinct industries.

PRICE STATISTICS

One of Morgenstern's examples of the questionable accuracy in which economic observations are presented is that of price statistics. Almost all possible sources of error mentioned above apply to price statistics: the desire to hide or lie about the true price, problems of classification or definition, and quality changes.

Moreover, in reality a certain good has multiple prices. The price changes when the goods are sold in different units, at different times and different qualities. Which price should be chosen? There are also non-monetary components to prices, for instance the quality of service before, during, and after the sale, which might vary. These, however, are not taken into account by merely measuring the monetary price.

When observed prices enter the calculation of index numbers, further problems are created. For one thing, the method of calculation itself is arbitrary, since many methods of calculating averages or price indexes exist. They all lead to different results. Furthermore, the components and their (changing) weight in the index is arbitrary.

Keeping all of those problems in mind, it is surprising that no error estimate of price level statistics is provided. Even more surprising is that economists take changes in price indexes up

to 1/10 of one percent at face value, without questioning their validity. However, those changes in price indexes are totally irrelevant for practical life. As Ludwig von Mises points out:

A judicious housewife knows much more about price changes as far as they affect her own household than the statistical averages can tell. She has little use for computations disregarding changes both in quality and in the amount of goods which she is able or permitted to buy at the prices entering into the computation. If she «measures» the changes for her personal appreciation by taking the prices of only two or three commodities as a yardstick, she is no less «scientific» and no more arbitrary than the sophisticated mathematicians in choosing their methods for the manipulation of the data of the market.²

NATIONAL INCOME STATISTICS

Another of Morgenstern's examples is that of national income statistics. National income statistics are widely considered to be relevant. They supposedly reflect the success of the government and are used in econometric models. These statistics are also of international importance. Morgenstern notes that, shortly after World War II, Japan and the United States «negotiated» the national income of Japan, because the national income influenced the size of economic help by the United States.

Morgenstern mentions several conceptual problems with national income statistics. The first involves the difficulty of the imputation of value. The problem lies in assigning a monetary value to goods and services produced. As Morgenstern states:

A classical illustration is that of persons living in houses they own themselves. If these same houses were owned by others,

² Mises (1998), p. 224.

rent would have to be paid (in money, goods, or services), thereby swelling the national product. To avoid this, a value has to be imputed to owner-occupancy. This is, obviously, a tricky affair, with less certain results than finding out about rent payments made in money. These estimates are uncertain and many arbitrary decisions have to be made.³

A similar problem arises when domestic help, which involves money payments, is substituted by housewives' labor, which does not involve money payments. Money payments are also reduced when the amount of barter in an economy increases.

A second problem in calculating national income statistics arises from the treatment of government services. They are not sold on the market. How should we account for them in the national income? The common practice is to account for them with factor costs. However, this seems arbitrary. The monetary cost of a service is not important as a measure of wealth production. Important, rather, is what people are willing to pay for a service on the free market. One could even make the case that government expenditures should instead be subtracted from national income, because the government withdraws resources from the productive private sector and uses them for its purposes.⁴ As an example of the absurdity of adding government services positively into national income statistics, consider the case of a government that builds a bomber and a bomb and destroys a newly built house in its own country. In today's national income statistics, the costs of building the bomber and the bomb are added into the national income, as is the house.

A third problem arises from depreciation allowances. Estimates of depreciation are made by corporations themselves and are guided by tax considerations and sometimes misleading

³ Morgenstern (1963), p. 246.

⁴ See Rothbard (2000), pp. 253-5.

ideas about the inflation process. Companies, therefore, fail to give a realistic accounting of the depreciation of capital in an economy.

Besides these conceptual problems, there are, as Morgenstern notes, three principal types of errors in constructing the statistics of national income. First, there are errors in the basic data that occur because they are a mere byproduct of other activities, because of classifications difficulties, lying, hiding of information, transmitting errors, etc. A second type of error results from the adjustment of the basic data to a conceptual framework, as the collected data is not directly suitable for use in national income statistics. A third type of error arises when gaps must be filled where basic data is not available, for example for a range of years or for industries where estimates are not known.

With all these difficulties in mind, would it not be very important, not to mention more honest, to provide an error estimate for national income statistics? However, nothing is said about the degree of accuracy in the publications of the national income statistics. We have to rely on our own estimates about their accuracy or about the expertise of those who make these judgments.

Simon Kuznets, an expert on national income statistics, argues that an average margin of error for national income estimates of about *10 percent* is reasonable.⁵ Considering this, it makes no sense to state changes in GDP with an accuracy of 1/10 of one percent! That is like having a yardstick and stating that a certain distance would be 4,312 yards. It aspires to an accuracy that is impossible. However, many economists take national income statistics at face value and use them, for instance, to confirm or falsify econometric models of the business cycle. In the light of Morgenstern's analysis this is completely futile.

⁵ See Morgenstern (1963), p. 255.

International comparisons of national income statistics are even more difficult to conduct due to different classifications, definitions, different hidden non-monetary incomes, interventions of the government into their respective price systems, and different measurements of inflation and deflation in the respective countries.

From the difficulties of national income statistics, it also follows that growth rates too should not be taken at face value. Obviously, the choice of the basic year introduces ambiguity and the base year estimate will contain error. The margin of error in the base year (again Kuznets suggests an average error of 10 percent) has a huge influence on the growth rate. For international comparisons the problem increases again. Morgenstern concludes that one can only make qualitative judgments about growth over longer periods of time.

CONCLUSION

In contrast to physics, there is still no estimate of statistical error within economics. The various sources of error that come into play in the social sciences suggest that the error in economic observations is substantial. This is a widely neglected problem and should be taken into account by the economic historian. Economic statistics cannot be accepted at face value.

Moreover, Morgenstern's *On the Accuracy of Economic Observation* has an important implication for modern economics. It shows that the solution of a system of economic mathematical equations or econometric models is, due to the quality of the data, completely devoid of meaning.

REFERENCES

- MISES, L. von (1998), *Human Action, The Scholar's Edition*. Auburn, Ala.: Ludwig von Mises Institute.
- MORGENSTERN, O. von (1963), *On the Accuracy of Economic Observations*. 2nd ed. Princeton, NJ: Princeton University Press.
- ROTHBARD, Murray N. (2000), *America's Great Depression*. 5th ed. Auburn, Ala.: Ludwig von Mises Institute.