STATISTICS IN THE SERVICE OF REBIRTH[1]

Massimo Livi Bacci

1. Introduction

I am very grateful to the President, Salvatore Strozza, and to SIEDS for the invitation to speak at this plenary session. As my friend Salvatore knows, I accepted this invitation with some reluctance not only because the topic is complex, but also because my training and background are traditional, far from the modern algorithms and futuristic analyses made possible by big data.

This is also why I want to begin my talk by reading what Minister Manna wrote in his "Report to the King" [*Relazione al Re*] more than a century and a half ago. On March 11, 1864, he presented this large volume *in quarto* of 501 pages containing the Proceedings of the Census of 31 December 1861 and its results by municipality.

Manna wrote:

"*My Lord, I have the honor to present to your Majesty the results of the General Census of the Population, based on simultaneous surveys carried out obtained on the night of December 31, 1861, in accordance with the Royal Decree of September 8, 1861 and the Law of February 20, 1862. This survey, prepared in the short span of three months and conducted simultaneously and following a standard methodology throughout the kingdom, deserves careful evaluation, my Lord, as it was one of the first and most important administrative acts responding to the new needs of the reconstituted and unified kingdom and also one of the most undeniable manifestations of the strength and spread of the national and united concept*"[2].

---

[1] Talk held during the online plenary session of SIEDS Conference, May 27, 2021.
[2] Original Italian version: "Sire, ho l'onore di presentare a V.M. i risultamenti del Censo generale della popolazione, compilato sulle notificazioni simultanee ottenute nella notte del 31 dicembre 1861, in conformità al decreto reale dell'8 settembre del 1861 e alla Legge 20 febbraio 1862. Codesta operazione che, preparata nel breve giro di tre mesi, poté nondimeno compiersi in uno stesso momento e in tutto il Regno, con norme uguali, merita, Sire, un attento esame, come quella che fu uno de' primi e più importanti atti amministrativi, che rispondessero alle nuove necessità del Regno ricostituito ed unificato e nel tempo stesso una delle più innegabili manifestazioni della forza e della diffusione del concetto nazionale ed unitario"

These words need no further comment.

A century or so later, I was in Israel for a period of study and research visiting Roberto Bachi, the father of modern Israeli statistics. A professor of statistics at Hebrew University, he was placed by Ben Gurion at the head of Israel's Bureau of Statistics shortly after the proclamation of the State of Israel (May 14, 1948) and assigned the task, with the help of a few officials and trucks, of conducting the first census of the new state. Census taking began in November that same year and ended during the Arab-Israeli war, a few months later.

These are two examples—Italy, recently united, and Israel, recently born—for which statistics were placed in the service of the "birth" (rather than "rebirth") of a country. On these occasions, statistics responded quickly and efficiently, producing results that, given the times and such basic means and the rudimentary methods available, were of quite a high quality.

2. "Objective" and "subjective" statistics

Assuming that the pandemic represents a discontinuity with the past, and that there is a "before" and the "after" to think about, I confess that I am offering my remarks in a state of perplexity.

On the one hand, there are "traditional" statistics, let's call them "objective", intent on counting and analyzing empirical quantities: people, births, deaths, and other demographic events, objects, artifacts, goods (produced, purchased, or sold), money, specific activities of legal or economic-social import, and other measurable phenomena (i.e., meteorological, seismic, geological). Interesting synthetic concepts have been defined for economic issues (gross or net domestic product, etc.), but also for demographic and social phenomena (burden of disease, life expectancy in good or poor health). For all these phenomena (and their synthetic concepts), we usually try to quantify empirical indicators.

On the other hand, there are statistics that aim to measure opinions, states of mind, intentions, satisfaction and dissatisfaction, aspirations, and hopes. We could call these "subjective" statistics because they try to measure attitudes, opinions, or inclinations. Scholars have tried to synthesize these too, by using abstract notions such as happiness or unhappiness, well-being or malaise, social utility, and so on.

Since the unwary organizers of this session asked me to express my views on the future perspectives of statistics, I will do so, albeit crudely perhaps, and beg your indulgence. "Objective" statistics can certainly take a large leap forward, and there is ample room for progress in the future. The speed at which data is collected, processed, and disseminated may be improved, thanks to the advances of technology. The ability to connect the ever growing number of databases may

similarly be improved, if only because we live in increasingly complex societies which require new data and information. Again, new technologies and methods are at our disposal. Access to the mass of available data can be made easier by sharing knowledge and research.

In addition to such progress, statistics must, above all, shed light on a number of black holes in our knowledge about the world. Many fields of knowledge are still shrouded in darkness.

For example, the world of "deviant behavior", a term that I use to be understood even though I do not like it, is neither well understood nor accurately described or measured. Furthermore, our "judicial" statistics are stuck in the nineteenth-century, analyzed and made available with great delays. Law enforcement certainly has a huge amount of interesting data for understanding the state of health of society. But these are largely unexplored sources of knowledge. The consumption and use of substances that alter behavior, moreover, are only known with approximation.

There is an enormous amount of data collected on the causes of death, but this is processed with delays and it is difficult to link with the individual characteristics of the deceased persons. The elderly population (aged 80 or more) will double by the middle of this century, along with the number of people with senile dementia, who now number 1.5 million. With regard to this growing population—which represents a huge public and private cost—we need to know much more: the state of health, lifestyles, living environment, etc.

Similar gaps in knowledge exist for environmental statistics, which are poorly developed, despite the obvious need to know much more about the degree of soil, water, and air pollution and their causes (industrial and agricultural activities, housing and transport, individual behavior, etc.).

Italy, a highly seismic area, has a built 30 million housing units, but we know little about the characteristics of these buildings: how many buildings are vulnerable to earthquakes? What interventions should be made to make them safer?

Going further, we know little about mobility and migration and what information we have is generally out of date. Population registers have undoubtedly improved, but have independent checks been made? Even assuming that the information is collected without errors, it represents "delayed" mobility, one might say a "replay" of it. However, cell phones and credit cards are used continuously and are geo-located, tracing our short and long-term movements.

Even population accounting is imperfect; it suffers from a discontinuity of information sources, and does not account for the people who are actually present. According to official tourist statistics, visitors from abroad spend about 220 million days in Italy in a normal year, which represents an average of 600 thousand people every day, perhaps one million, if we include those visitors who are guests of families or have other informal arrangements. Moreover, there are hundreds of

thousands of undocumented immigrants, which only a few worthy private institutions have tried to estimate. These populations are little investigated or known, yet in their own way they represent part of the society in which we live.

These few simple examples confirm the existence of numerous large, deep gaps in our knowledge about our society. Great efforts in highlighting its dimensions and characteristics would be useful for the "rebirth" of the country, making a further cognitive leap forward.

I have not spoken of what I called "subjective" statistics concerning surveys aimed at measuring well-being, satisfaction, or happiness. I believe that official statistics is already faced with a huge task if it wants to take great steps in the "objective" knowledge of the country. Analyses that rest on individual opinions remain very delicate for reasons that we are all familiar with.

I have in front of me the World Happiness Report 2020, which carries the logo of the United Nations and mostly collects the results of Gallup polls on the perceptions and opinions of respondents in different countries. For example, it asks if the respondent was more or less worried and sad on the day before the survey, and comparisons are even made with previous surveys, even though the answers may depend on circumstances (e.g., meteorology, news, victory or defeat of one's favorite team, etc.). According to the synthetic rankings of this Report, Italy would be less happy than Saudi Arabia, which has the death penalty, Sharia law, veiled women, and immigrants in a state of semi-slavery. Greece is close to Libya, and Rwanda (also called the Switzerland of sub-Saharan Africa) is in third-to-last place.

Let's leave the task of procuring "bread and butter" to official statistics and the "champagne and caviar" to other institutions! In other words, let's leave "subjective" statistics to specialized, secular social research institutions—those without preconceived notions about the nature of well-being or happiness—which have flexible programs and are not constrained by the rigid rules of official statistics that must be integrated with those of other EU countries.

3. Concluding remarks

Before concluding, I would like to touch on three final points.

The first concerns the need to "link" our statistics and knowledge with the rest of the world. This is already done, of course, with other European countries, but we need to go further and extend this practice to other countries around the Mediterranean, for example. This would help to extend our knowledge to larger geographical contexts and provide broader information concerning phenomena that are closely related to those in our country.

The second point concerns the need to modernize the system of statistics by opening up to external contexts. Here I would like to quote an excerpt from an article by Julia Lane published in the Royal Statistical Society's magazine. Julia Lane is a scholar whom I would call a libertarian, the author of a successful book in 2020 entitled "Democratizing our data: a Manifesto".

"... *Our economic and social data system should be democratised in terms of measurement, collection and interpretation. The choice of economic and social measures should be decided by people at the state and local level, as well as the federal level. People and institutions at all levels should collect and use the massive amounts of new data that are now available in ways that were not imagined a century ago. Valid interpretation of data is no longer the purview of a small group of highly trained federal statisticians – there are now many trained data scientists and statisticians across the USA. One immediate need is to develop new local, and timely, measures of job loss, business vulnerability, and fiscal impact. But, more fundamentally, we need an infrastructure – what I would call a 'national laboratory for community data' – that is designed to respond to changing needs. This new organisational infrastructure should be based on three core principles: great measures, great people, and great technology. ...*" (Lane 2020: 42).

We need a large open-source laboratory: a place for thinking, a place for generating new ideas that is capable of joining public and private capacities, an external "engine" to push the progress of official statistics.

The third point concerns the limits that we should place on the justified needs of researchers, who are little restrained by the weak institutions defending our privacy. We are photographed, filmed, recorded, shadowed, and profiled; names can be attached to our faces and a name can even be given to our features when we walk with a hood on. Our purchases and earnings are all recorded; credit cards, telephone cells, and various chips all follow and trace us. We are "profiled" for insurance companies, employers, banks, administrators, border guards, and event organizers (think about the "green pass"). Are we entering the dangerous territory of Big Brother? Research can certainly benefit from this cloud of information that is expanding at the speed of light. However, bulimia, even data related, is dangerous and perhaps does not serve the progress of society. It seems to me that it's best to maintain a healthy, even restrained, appetite.

Reference

LANE J. 2020. After Covid-19, the US statistical system needs to change, *significancemagazine.com*, Vol. 17, No. 4, pp. 42-43, https://doi.org/10.1111/1740-9713.01428.

SUMMARY

Statistics in the service of rebirth

Massimo Livi Bacci, during the online plenary session of SIEDS Conference in May 27, 2021, dealt with future perspectives of statistics by distinguishing between "objective" and "subjective" statistics and by describing the main fields of knowledge with poor statistical reporting and/or inadequate information. The reported examples confirmed the existence of numerous, large, and deep gaps in our knowledge about Italian society. Great effort in highlighting its dimensions and characteristics would be very useful for the "rebirth" of Italy, making a further cognitive leap forward.

_____

Massimo Livi BACCI, Università di Firenze, massimo.livibacci@unifi.it