Aaron Brown

Statistiks

Why do we use the same word, *statistics*, for numbers put out by the government and the science of applied probability? This question turns out to have a very interesting answer, one that sheds light on attitudes toward quants

Great advances made in statistical accuracy

The English word is taken from the German *statistik*, which was invented by Gottfried Aschenwall and used in his 1748 work *Vorbereitung zur Staatswissenschaft*. Aschenwall's *statistiks* had nothing to do with either numbers or probability. They referred to qualitative reports compiled by agents of a monarch to describe the resources available to him. German political scientists were hostile to numerical data for this purpose because they could not capture essential detail. Adding up total population, for example, lumped together men, women, and children with different skills and social positions, possibly speaking different languages, and following different religions. Without a large bureaucracy to cross-tabulate and analyze data, the prince and his few advisors needed focused qualitative reports more than lists of numbers.

The word worked its way over to France as *statistiques*. In addition to three extra letters for the same sound, the word acquired a different connotation. After the Revolution, France was interested in rationalizing itself, which required measuring everything according to common standards. Part of this effort involved sending questionnaires to the *préfets* of each *département*. These mixed quantitative and qualitative information, but even the qualitative information was categorized. That is, the *préfets* were given systematic questionnaires with boxes in which either numbers or words could be written. The idea was to uncover the ideal of Frenchness, for which the average was usually the best statistic. Individual observations were regarded as error, noise around mean.

This difference came down to a different conception of the State. In pre-Bismarck Germany,

the State had no real legitimacy. There was a prince, and there was some territory he controlled. The prince wanted information about the territory to use for his own benefit. France, on the other hand, was a State of unquestioned legitimacy. Everyone knew there was a France. They might not agree on who ran the place, or even whether the government was a monarchy, republic, or empire, but they didn't dispute that there was a distinct group of people who shared a language and a culture and lived in a patch of Europe with vague borders. The French State wanted information about itself.

England was somewhere in between. For one thing, it was a united (in theory anyway) kingdom of peoples with different languages, religions, and cultures. For another, it lacked the absolutist tradition of France. Britain was an overlapping set of institutions rather than an integrated totality that defined all its subjects. The government lost interest in compiling demographic information after the Domesday Book in 1085 (ordered by a French Norman). Therefore it fell to private amateurs to develop *political arithmetic*. This involved both the compilation and analysis of data. Moreover, in England the entire distribution was studied, not just the mean.

The United States developed in a fourth way. Article 1, Section 2 of the Constitution set the tone by ordering a decennial census used to determine tax assessments and political representation. The two were linked explicitly to give States competing incentives for overcount or undercount. This is the section with the infamous provision to count slaves as three-fifths of a free person (less well known, Native Americans were counted as zero). This was not an evaluation of the human worth of slaves, it was a political compromise between free and slave states to apportion taxes and votes. Government numbers were computed in rigid ways to mediate among differing interest groups. To this day, the census must be an enumeration of individuals. Sampling methods would be far cheaper and possibly more accurate (certainly insisting on enumeration results in some systematic biases) but would open the process up to subjectivity.

Whenever the American Congress has found itself overwhelmed by narrow political interests

and unable to make rational national decisions, it creates an elite group of objective experts to analyze data. Prominent examples are the Army Corps of Engineers that has to pass on all national water projects, and the Congressional Budget Office to provide one set of numbers for fiscal decisions. These groups are never given power, they are only supposed to provide objective analysis according to politicized rules laid down at the beginning. But because they are not political, or more accurately because they have been pre-balanced for political neutrality, their findings carry tremendous weight as natural points of compromise.

An ironic example of the American obsession for objectivity over truth began in 1933 when the newly created Securities and

followed, the CEO and CFO would be required to sign off that statements were accurate. In a doubly ironic move, the government decided to create principles-based accounting with a gigantic new set of rigid rules.

These four national streams fused to create modern statistics in both senses. I organize it by national tradition for ease of exposition. Of course there were contributions from other nations, and citizens of all four nations named contributed to all four traditions. Nevertheless, it is possible to see the differences in national attitudes toward numbers to this day.

The German and French traditions require data collected by someone loyal, to the prince or France respectively, and the results are confidential, to the prince or the bureaucracy respectively.

These four national streams fused to create modern statistics in both senses… Nevertheless, it is possible to see the differences in national attitudes toward numbers to this day

Exchange Commission ordered all public companies to prepare financial statements according to strict rules, and hire independent auditors to certify that the rules were followed. A notorious example of the rules was that all assets were to be valued at historical cost minus pro forma depreciation rather than current value, which had been the usual practice. The SEC wanted objective numbers, not accurate ones. Companies protested that forcing all enterprises into a single format would lead to distortion, but they were overruled. Seventy years later the SEC, and most of the world, decided that public corporations had defrauded the public by issuing misleading 'rules-based' financial statements when every honest person knew that 'principles-based' statements were required; and that it was meaningless to have an independent auditor sign off that the rules were

In both cases the goal is to improve management of the State, for the benefit of the prince or the State respectively. The British and American traditions require the opposite. Statistics must be gathered by disinterested people and are published openly. The British goal is scientific investigation to get at the truth. The American goal is to settle disputes.

The German and British traditions rely on investigators who are not employed by the government. These people need some expertise in their area of investigation, but not rigorous specialized training. The French and American traditions instead use government-employed experts, generally selected by competitive examination from designated educational institutions.

To highlight the differences, consider statistics on unemployment. At any point in time there are people with full-time jobs, people with part-

time jobs, people without jobs who don't want them, people without jobs who do want them, people who are unemployable at any wage they would accept, seasonal workers, people with jobs who expect to lose them soon, people with jobs they hate or that do not provide a decent income, people who have off-the-books jobs, people who derive income from past work (like an author earning royalties) but who are not currently creating new income, people working on projects that do not provide current income (like an as-yet-unpublished author), and many other situations or permutations.

In the American tradition, this led to strong political resistance against computing an unemployment number in the first place. Doing so

The German tradition does not require aggregation. Unemployment would be studied by sending investigators to observe all aspects of the problem. Different people would be unemployed for different reasons, and suffering to varying degrees (some not at all). In this worldview, there is no such thing as unemployment, just people without jobs. Since there is never a shortage of things to do – either for the prince or, in modern times, for the State – unemployment can be cured easily. The jobless can be drafted, or ordered to work on projects, or otherwise employed.

In the English tradition, unemployment can be studied through many statistics to get at its multiple aspects. But there is a strict separa-

All four traditions are alive and well in finance. The modern version of the German attitude is no longer hostile to numbers. Computers and armies of bureaucrats make numbers potentially useful. But numbers are descriptive and illustrative, they are not the reality. A useful report can have charts and figures, but the quant who buries herself in only the charts and figures is blind and dangerous. This is what many people have in mind when they say that quants destroyed the financial system.

The French tradition lives on in central banks and financial regulators. They collect huge volumes of mostly quantitative data, and hire highly-trained experts to interpret them. Most of this is secret, although they do deign to publish some summary values. The experts are universalist, like the metric system. The economic statistics are meaningful in themselves: there are true unemployment rates, inflation rates, and interest rates; these are social facts that affect everyone.

Since both these traditions assume analysts with points of view and are secretive, they tend to invite general distrust. Someone whose argument weaves a lot of special-purpose statistics that cannot be validated independently is not believed. "All figures don't lie, but all liars figure", as the saying goes. A secretive group of mandarins making decisions for everyone's best interest is perfect fodder for conspiracy theories. Therefore, these types of quants need strong prestige to succeed.

The British and American traditions substitute a claim to objectivity for a claim of expertise. You don't need to trust them because they are open about their data and methodology. You can replicate their results yourself. The British tradition is the scientific quant, the American tradition is the quant arbitrator. If you take the British approach, you lack power. All you can do is publish results and hope people are convinced by them. The American quant has power to make decisions, but those decisions are strongly constrained by the need to be transparent and disinterested.

When people say 'statistics' today, they may mean any of the four traditions, or some amalgam of them. Statistics carries a lot of historical baggage, from the nature of reality to the power of the State to the definition of fairness and truth.

Statistics carries a lot of historical baggage, from the nature of reality to the power of the State to the definition of fairness and truth

would necessarily aggregate people in very different circumstances and require discretion on the part of compilers – discretion that could be used to someone's political advantage. When that resistance was overcome during the Depression, rigid definitions were chosen. The point wasn't to get at the true nature of unemployment, but to produce an objective number that could serve as the basis of political bargains. To this day, the United States does not count people who are unemployed but have given up looking for a job, and relies on new claims for unemployment insurance, which can be measured easily but has a complex relation to unemployment. The unemployment rate figures into the Federal Reserve's mandate and there are unwritten political rules for policies that are required at times of high measured unemployment. For example, when unemployment rises above six percent, it is assumed that many of the unemployed cannot find work, so unemployment compensation times are extended.

tion between observation and theory. Everyone should agree on the data, it should be able to be replicated. Two investigators collecting data should get the same result, even if the result has different interpretations in their respective theories. Alone among the four traditions, the work need not lead to policy prescriptions. The goal is to publish the truth, someone else can worry about what to do in response.

In the French tradition, there is a social fact of an unemployment rate. That rate will affect different people differently. If the rate is high, A might lose his job, B might decide to stay in a safe government job rather than trying to start a business, C might stay in school to collect another degree rather than go on the job market, and D's temp services business may thrive. A, B, C, and D are all noisy observations of a social reality. Statistics should collect and refine the data to alert all bureaucrats to the social fact. Then individual bureaucrats can make specific decisions appropriate to the situation.

W