From Cournot to Public Policy Evaluation: Paradoxes and Controversies involving Quantification

Alain Desrosières[1]

Prisme N°7
April 2006

[1] Alain Desrosières is an administrator at the Institut National de la Statistique et des Études Économiques (INSEE) and a member of the International Statistical Institute. He teaches the history of statistics at the École Nationale de la Statistique et de l'Administration Économique (ENSAE) and at the Alexandre Koyré Centre .

Summary

The French mathematician, economist and thinker Augustin Cournot inaugurated the philosophical treatment of the new probabilistic and quantitative modes of reasoning that emerged in the first half of the 19th century. The text reviews the legacy and implementation of Cournot's intuitions concerning the distinction between so-called objective and subjective probabilities, and the interpretation of the categories constructed by statisticians according to "equivalence conventions". Suggestive clues emerge for the empirical study of current statistical practices, in particular, those transactions that take place in the "contact zone", where quantified assertions recorded in more or less formal models replace unquantified assertions formulated in natural language. Examples of these exchanges are illustrated in the cases of risk management, macroeconomic policy and public service performance evaluation.

In conclusion, the paper highlights how the ambivalence of Cournot's thought is echoed in the controversies raised in some recent sociology of science, polarized between diverse forms of "realism" and "constructivism". Questions suggested by Cournot are the starting point for an exploration of the sense in which quantification can be said to create objectivity.

Introduction

Augustin Cournot (1801–1877) is frequently presented as one of the founding fathers of mathematical modelling in economics. In contemporary terms, mathematical modelling of economic phenomena involves either purely theoretical or hypothetico-deductive constructions, or, more commonly, the testing of theoretical hypotheses against statistical data, using econometric tools. Yet, the combination of theory and empirical data in this way is a recent development. Modern econometrics in its unified form took off only in the 1930s (Morgan 1990; Armatte 1995). Cournot's work in economic modelling, by contrast, divided into two separate parts, each represented by a book. His first book, from 1838, concerned the "theory of wealth", whereas the second book treated a different topic, the "theory of chance and probability", and there was no connection between the two. The same dichotomy appeared in the case of other authors, such as Edgeworth and Keynes. Claude Ménard (1977) has examined the "resistance to statistics" of three 19th-century economists: Say, Cournot and Walras. In Cournot's treatise of 1843, which broke new ground in the theory of knowledge, Ménard found "an epistemological representation in which the method of investigation and the object of knowledge are perceived as independent". Cournot, however, did not see in statistics a satisfactory instrument for the support of theoretical hypotheses. Statistics presupposed conventions and comparisons that could not be replicated fully. Paraphrasing Cournot, Ménard remarks:

> *If the exact same experiment can never be repeated, how can the specificity of social facts be compared? These insoluble problems have been the cause of an "exuberant proliferation" of statistical information […]. The importance of space and time, that is to say, history, in social phenomena, only serves to accentuate the difficulties. […] How can one compare data collected in human contexts and environments that are so completely heterogeneous?*

> *[…] How can the observer's contribution be circumscribed? On the basis of which tools?*
> (Ménard 1977, p. 422)

Drawing support from the notion of an "equivalence convention", I shall describe how researchers and more generally social policy makers since Cournot have handled, if not resolved, these difficulties. This notion will serve to show that since that period, the objections of 19th-century economists and their resistance to statistics have been overcome, not logically, in the epistemological universe that Cournot inhabited, but socially, in a world where people *agree* to compare incomparables, treating heterogeneous situations as equivalent *for practical ends*.

Comparing the incomparable[2]

In the first half of the 19th century, the pioneers of the so-called moral sciences (the future social sciences) took a lively interest in the probabilistic and quantitative modes of reasoning that had originated in the traditions of the 18th-century philosophers and astronomers. Two figures symbolize this historical moment: the Belgian statistician Adolphe Quetelet (1796–1874), and the French philosopher and mathematician Augustin Cournot (1801–1877). Unlike Quetelet, however, the advocate of offices of national statistics, Cournot devoted himself less to the effective implementation of quantitative methods in science and society, and more to the study of the philosophical implications, in terms of the theory of knowledge, of recourse to these new

[2] The expression is borrowed from the title of the book by the historian Marcel Detienne (2000), which is a critical analysis of the fact that history, often written from a nationalist perspective, considers the historian's nation as radically incomparable to other nations, blocking all historical comparativism. Statistics is precisely a conventional means, among other methods, of comparing the incomparable. Several controversies surrounding statistics bear on this exact question of comparability.

ways of reasoning and argumentation. His work of 1843, ***Exposition de la théorie des chances et des probabilités,*** treats these questions in detail.

The modern reader cannot fail to observe that Cournot distances his position subtly from the scientistic objectivist interpretation adopted by frequentist statisticians, led by Quetelet[3]. From the frequentist perspective (popularized in the phrase "the law of large numbers"), quantification is presented as the paradigmatic tool of "objectivity", "common language" and "shared diagnosis" between observers or agents, who, through quantification, can communicate with each other, overcoming subjective differences, and consequently, substituting rational language for passionate speech. At several points, Cournot evokes the central question of interpretation, which underpins simultaneously the construction, format and conclusions of quantitative arguments. In this way, he places this mode of reasoning within an open reflection on the means of knowledge, whereas, more commonly, these arguments are advanced in order to close a debate. In particular, he applies this process to two questions. On the one hand, he introduces the distinction between so-called ***objective*** and ***subjective*** probabilities. On the other hand, he discusses the interpretation of ***segmentation*** or ***cuts***: the latter are the categories organized by statisticians according to equivalence conventions, with the aim of ordering and comparing the objects of study. Following a brief review of the legacy of these two intuitions of Cournot, I shall examine how they can be implemented in three contexts where quantitative arguments are often mobilized: risk management, macroeconomic policies, and the evaluation of public policies. The essay will revisit, although with different tools, certain questions that Cournot raises in his ***Exposition***. How should

[3] Although some historians see in Cournot a proponent of the frequentist interpretation, Thierry Martin (1994) shows that "if the concept of mathematical probability is not univocally determined […], the reason is that for Cournot, it is a matter of classifying the different possible meanings, in order to appreciate the value of the results that calculation yields". In the same spirit, the aim in the present paper is to explore the multiplicity of meanings and interpretations that probabilistic and statistical assertions may carry.

probabilistic tools and statistical methods be integrated with other means of knowledge and action? What conventions does this integration imply? How should one "compare incomparables"? Can one complete Cournot's rigorous logical examination with an empirical study of the social uses of these instruments, as they have developed over two centuries?

Cournot was the first to place strong insistence on the *dual* character of the calculus of probabilities, which, on the one hand, quantifies *reasons for belief*, and, on the other hand, often (but not always) relies on *observed frequencies*. Since its earliest appearance in the 1660s (Hacking 1975), this duality has been described in different ways. Condorcet distinguished "reasons to believe" and "facility". Cournot spoke of "chance" and "probability". Carnap contrasted "inductive" with "statistical" probabilities. In the 18th century, the decision-theoretic aspect, based on "reasons to believe", predominated (Daston 1989), particularly in procedures stemming from Bayes's theorem. This theorem set out a way of taking account of partial information about unknown situations, in order to estimate a "causal probability" enabling one to support a decision. The 19th century frequentist perspective contested this way of reasoning. It distinguished radically decisions based on non-quantifiable judgements (for example, those of a trial jury), from those that relied on repeated observations, in particular those provided by the new statistical offices advocated by Quetelet. For frequentists, Bayesian procedures, combining a small number of observations with a purely conjectural "*a priori* probability" to infer a stronger "*a posteriori* probability", seemed like a fantasy. As the choice of *a priori* probabilities was often arbitrary, the reasoning appeared built on sand. In the 20th century, by contrast, the question of decision-making under uncertainty attracted new interest with the work of Keynes, de Finetti and Savage. Discussions of Bayesianism and its interpretation assumed primary importance. Yet, in 1843, Cournot had already sensed the significance of Bayesian reasoning, at the very moment when it was being discredited:

> *A rule, first formulated by the Englishman Bayes, and on which Condorcet and Laplace wished to build the doctrine of* a posteriori *probabilities, has become the source of numerous equivocations which must first be clarified, and of serious errors that must be rectified, and which disappear as soon as one is made aware of the fundamental distinction between those probabilities that have an objective existence, which give a measure of the possibility of things, and subjective probabilities, relative in part to our knowledge, in part to our ignorance, and which* vary from one individual to another[4], *according to their capacities and the data provided to them.*
> (Cournot 1843, p. 155)

In the 20th century, the idea that subjective probabilities "vary from one individual to another" became a topic of research in experimental psychology, in particular in the works of Kahneman and Tversky (1973), who showed that the human mind does not function according to Bayesian assumptions. In turn, these results were contested by Gigerenzer and Murray (1987), who criticized the poor experimental framework and the weakness of the interpretations (Amossé, Andrieux, Muller 2001). The interest of these controversies is to centre the debate on an empirical question: how does the human mind combine quantitative information with *other information* that is temporally prior or of another kind? This approach differs from research that was limited to studies of putatively comprehensive data files. The advances of the latter reflect the fact that since the 1930s and the work of Ronald Fisher, William Gosset (*alias* Student), Jerzy Neyman and Egon Pearson, inferential statistics had made remarkable progress in parameter estimation and hypothesis testing using data files.

[4] My emphasis.

Exploring the contact zones between propositions involving different registers

Mathematical statisticians have handled less well the question of interpretation, that is to say, the articulation of knowledge produced in terms of *what one knows (or thinks one knows) from other sources*. Yet, this type of question was suggested by Cournot in 1843. Of interest here are the ***contact zones***, or mediation points, between the rhetoric of statistics and other rhetorics[5]. Progress in mathematical statistics and econometric models in particular[6] has enlarged the space within which interlinked and mutually reinforcing statistical styles of reasoning come to appear self-sufficient, and have less and less contact with other types of argumentation. This development accompanied the increased professionalization of statisticians and econometrists. Matters were different in the 19th century, when methods of analysis were less sophisticated, and the cognitive and professional divisions of labour were less clearly delineated. The questions raised by zone border crossings were more visible, and hence often more discussed and rendered more explicit.

The point is neither to criticize or denigrate current uses of quantitative methods, nor to deepen the epistemological questions that these uses throw up, nor to make normative proposals for an improved methodology. I wish simply to provide some pointers for empirical studies of statistical practices

[5] The word "rhetoric" is not intended to have the pejorative meaning that it sometimes possesses, but the neutral meaning of a form of argumentation, or, as Hacking says, a "style of reasoning".

[6] This reading has links with the work on the history and sociology of modelling and on the role of models, undertaken in the Anglo-American world by Mary Morgan and Margaret Morrison (1999), under the suggestive title *Models as Mediators*, and in France by Michel Armatte and Amy Dahan-Dalmedico (2004) at the Alexandre Koyré Centre. In this perspective, a *model* is a *mediator* in two ways: on the one hand, it mediates between formalism and a non-formalized world, and on the other hand, it serves as a *common language* between agents. Quantification procedures can be viewed in the same manner.

and, more precisely, of transactions in the contact zones, in which quantified assertions inscribed in more or less formal models replace non-quantified assertions formulated in natural language. How is the world altered by the production and circulation of these quantified formal assertions? To what extent do they enable or prevent the production of unified incontestable interpretations, as their producers and users wish and even claim? If this is not the case, how should variations of interpretation be analysed?

When Cournot observed that subjective probabilities "vary from one individual to another", he did not venture to explore the forms and causes of this variability: are they psychological, cultural, or biographical? Subsequent research has expanded these questions. Nevertheless, Cournot did see an aspect of statistical work for which the question of articulation in terms of pre-existing knowledge is crucial: what he called the "choice of segments", that is to say, nomenclature. Remarking that if one classifies French administrative departments according to some variable (alphabetical order, crime rate), he asks whether the "top classes" and "bottom classes" are the result of random chance, or, on the contrary, of some relevant feature? Today's means of rapid calculation enable the savvy statistician (or, in the eyes of some, the less-than-honest data-miner) to calculate all possible correlations in a file, to choose the "best", and *then* to formulate hypotheses which are miraculously confirmed by the data in the file. Cournot explicated this precise question in 1843, when he spoke of the "prior judgement that orients the gaze towards the segments" (i.e., the nomenclature), and the interpretation of the "spreads observed":

> *A further element lies in the prior judgement, through which we perceive the nomenclature giving rise to the spreads observed, as one that is natural to employ out of the multitude of possible divisions, and not as one that catches our attention merely on account of the spreads observed. This prior judgement, by which statistical experience appears obliged to fix on one nomenclature*

> *rather than another, results from motives whose significance cannot be rigorously estimated and* may be very differently estimated by different minds[7]. *It is a conjectural judgement, itself based on probabilities, but on probabilities that cannot be resolved into an enumeration of chances, the discussion of which does not properly belong to the doctrine of mathematical probability.*
> (Cournot 1843, p. 196)

According to this reasoning, "segmentation" (in this case the French administrative departments) is a given, but one cannot pretend prior complete ignorance about the specificities of the departments of the Seine, which includes Paris, or Corsica, which is an island. Cournot, however, does not raise the question of the genesis of the "segments". Yet the construction, coding and interpretation of statistical nomenclature constitute privileged moments in the study of the contact zones mentioned above. A statistical category is the result of an equivalence convention. The verb ***convene***[8], from which the word ***convention*** is derived, evokes the social procedure which yields the category. This procedure is a key element of the contact zone[9].

The methodological doubt that Cournot articulated was not, however, the product of an arbitrary relativist scepticism, claiming (as is sometimes done in polemical contexts) that "statistics can be made to say anything… ". On the contrary, in a perspective that finally brings him close to Quetelet, he considered that belief in certain truths rested on a rational order, above individual subjectivities. Thus, the variability of perceptions and interpretations is imputed to individual error, as was the case in the language

[7] My emphasis.

[8] The French verb *convenir* (translated above as *convene* and subsequently as *agree*) has the double sense of to agree and to convene [translator's note].

[9] Various studies of these phenomena have been undertaken since the 1970s. Several are reviewed in Chapter 8 of Desrosières (2000).

of the 18th-century astronomers. Objectivity is induced through averaging, which, through the magic of the law of large numbers, enabled one to base contingent individual observations on "chains of interlinking truths", closer to the "rational order"[10]:

> *Our belief in certain truths is therefore founded solely neither on the repetition of the same judgements nor on unanimous or near unanimous assent: it rests principally on the perception of a rational order according to which these truths are interlinked, and on the conviction that the causes of error are abnormal, irregular and subjective causes, which could not give rise to such a regular and objective coordination.*
> (Cournot 1843, p. 421)

Cournot's idea combines, on the one hand, a methodological doubt justifying his attention to individual subjectivity, and, on the other hand, the conviction that there exists a rational order transcending individual subjectivity. This ambivalence is echoed in the controversies raised in recent sociology of science, polarized between different forms of *realism* and *constructivism* (Hacking 1999). Taking these observations as the point of departure, I shall study in what sense quantification can be said to create objectivity. In one fell swoop, quantification appears to constrain, reduce and delimit the space of possible interpretations of the world, but at the same time, it *creates another world*, with new possibilities of interpretation and action. Quantification reconfigures the world, creating new objects that enter human social circulation.

[10] This brings to mind the references to a divine order in Quetelet, or, more recently, in the work of the French statistician Jean-Paul Benzécri, an advocate of correspondence analysis.

Quantify = convene + measure

In an experimental spirit, I shall suggest a framework in which to examine quantification procedures and their cognitive and social effects. The framework differs slightly from realist epistemology coming from the natural sciences, which often prevails in the social sciences. I shall test the framework using examples of problems drawn from current debates: risk management, macroeconomic analysis, and public management indicators. In order to carry out this experiment, it will be indispensable to distinguish two commonly confused ideas: the idea of *quantification*, and the idea of *measurement*. The verb *quantify* is used here in a broad sense: *to express and realize in numerical form that which was previously expressed in words and not in numbers*. By contrast, the idea of *measurement*, drawn from the natural sciences, implies that something already exists in a form that is measurable using a realist metrology, for example, the height of Mont Blanc. In the case of the social sciences or the evaluation of public services, profligate use of the term *measure* and its cognates leads to error, by leaving in the shadows the *conventions of quantification*. The verb *quantify*, in its transitive form (*make* into a number, *put a figure on*, numericize), presupposes that a series of prior equivalence conventions has been developed and made explicit[11], involving comparisons, negotiations, compromises, translations, registrations, encodings, codifiable and replicable procedures, and calculations leading to numericization. Measurement, strictly understood, comes afterwards, as the rule-based implementation of these conventions. From this viewpoint, quantification splits into two moments: *convention* and *measurement*.

The use of the verb *quantify* draws attention to the socially and cognitively creative dimension of the activity. This activity does not just provide a

[11] This social and logical notion of equivalence convention owes much to the early work of Bruno Latour (1984) in the supplement *Irréductions* to his book on Pasteur, and the paper of Laurent Thévenot (1983).

reflection of the world (the common viewpoint), but it transforms the world, by reconfiguring it differently. The distinction between *quantify* and *measure* is not "relativist" in the pejorative sense occasionally attributed to the word. It aims at separating analytically two moments that are historically and socially distinct. The distinction is convincingly illustrated by examples such as "intelligence", when the "intelligence quotient (IQ)" was initially conceived, "public opinion", when "Gallup"-style polls first appeared (introduced into France by Jean Stoetzel), or the more recent debates about the quantification of the effects of public programmes. The invention of the notion of *probability* in the 17th century, in order to *quantify the uncertain* by means of a number lying between 0 and 1, is an illustrious precedent in this domain. The "reality" and the ontological status of the concept of probability were discussed at length, in particular by Cournot, whose distinction between objective and subjective probability was a cunning way of dealing with the epistemological challenge.

The suspicion of relativism may arise when the real existence of the object, prior to its being measured, is put in doubt by those for whom the measure actually creates the object. Intelligence is "what is measured by IQ tests". Opinion is "what is measured by opinion polls". The standing hypothesis of this paper is that quantification, understood as the totality of socially agreed conventions and mensuration operations, creates a new way of thinking, representing and expressing the world, and of acting upon it. The recurring question whether "a statistic reflects reality more or less well" is deceptive shorthand, contaminated by the metrological realism of the natural sciences. Statistics, and more generally all forms of quantification (for example, probabilistic quantification, accounting quantification), transform the world, through their very existence, by their diffusion and use in argumentation, whether in science, politics, or journalism. Once the procedures of quantification have been coded and programmed, their results are reified. They tend to become "reality", by an irreversible "ratchet effect". The initial

conventions are forgotten, the quantified object is naturalized, so to speak, and the use of the verb "measure" automatically springs to mind and into ink on the page. This naturalization remains in force until, for reasons that require case-by-case analysis, controversies erupt and the "black boxes" are reopened. An example is provided by the recent contestation of the "volume–price split" in the economic growth rate.

The question of the objectivity and univocity of statements formulated in quantitative terms has already been raised above, inspired by Cournot's remarks on the variability in the ways in which human minds navigate the contact zone between, on the one hand, a non-formalized world, and on the other hand, the world of formalisms, via what is referred to as a "model". The definition of the verb *quantify* that I propose, distinct from the definition of the verb *measure*, allows one to raise this question in a different way. Quantification provides a specific language, endowed with remarkable properties of transferability, standardized computational manipulations, and programmable systems of interpretation. Thus, it makes available to researchers and policy makers "coherent objects", in the triple meaning of intrinsic coherence (resistance to criticism), combinatorial cohesiveness, and power of social cohesion, keeping people together by encouraging (and sometimes forcing) them to use this universalizing language rather than some other language. This perspective, which differs from the common received standpoint advocated by the quantitative social sciences and, more generally, by users of statistical and accounting tools, is advanced as an hypothesis that I shall attempt to apply in exploring three areas where quantitative arguments are widely invoked: risk management, macroeconomic planning, and the evaluation of public management. Three types of equivalence space will be deployed: the space of probabilities, which Cournot studied in 1843, the space of value and wealth, which he examined in 1838, and finally, the space of means and ends of public management, which, as a theoretician, Cournot scarcely envisaged at all.

When people agree to set up equivalence spaces

The three examples mentioned above are intentionally disparate. They have been chosen because, in each of the three cases, something that, *a priori*, was expressed in verbal form, ended up in numerical form, despite the fact that this transition was far from evident and the translation was (and often still is) debated in various ways. What price was paid to achieve this conversion from words to numbers? The historical moments when these "numericizations" (just as one says "dramatizations") occurred, are entirely different: risk has been "probabilized" since the middle of the 18th century, the aggregates of national accounting have been quantified since the middle of the 20th century, and discussion of the quantified evaluation of the performance and quality of public service (also called *benchmarking*) started in the 1980s. Other examples have already been mentioned: the (highly contested) identification of intelligence with IQ, and of public opinion with Gallup polls. The decisive stage is the negotiation of the conventions that make things *commensurable*, that is to say, comparable according to a numerical scale, whereas, *a priori*, this comparison was judged "impossible" by many: "You are comparing things that are not comparable, it *cannot be quantified*". These criticisms are frequently heard from those who contest the commensurability in question. The objections are centuries old; they invariably surface, at some moment or other, in relation to the cases mentioned above.

The ambivalence of these objections lies in the French infinitive *pouvoir*. The French verb has two meanings: "to be physically possible...", and "to be permitted...". In English, the ambivalence is expressed by two distinct verbs: *can* and *may*. The former is quasi-physical: it appears to stem from the nature of the thing in question. By contrast, the latter relates to the moral, social or political order. To compare (that is to say, *see together*) is a political act: in certain societies, one could not (in the sense of "it was inconceivable to...")

compare slaves and free men, women and men, commoners and nobles, black people and white people. "Social inequalities", as this expression is understood today, in terms of a reference to a common equivalence space, were thought of in this way at the end of the 19th century only in connection with inequalities of income, and (with rare exceptions) in the middle of the 20th century in connection with other types of inequality, such as consumption, access to education, or social mobility. To postulate and construct an equivalence space enabling quantification and hence mensuration, is an act that is at once both political and technical. It is political in that it *changes the world*: to compare commoners and nobles required the night of August 4, 1789[12], to compare blacks and whites required the abolition of slavery, to compare women and men required truly universal suffrage including women.

American sociologists have put forward the related idea of *commensuration*. Under the title *Commensuration as a Social Process*, the idea of which is close to what I am proposing here, Wendy Espeland and Mitchell Stevens (1998) analyse the social processes that aim increasingly to *monetize* human acts, as an effect of the extension of market mechanisms. In this case, the equivalence space is money, the antiquity and generality of which should not be ignored. From a closely related perspective, Viviana Zelizer (2001) describes how, in divorce cases in the United States, previous amorous relationships, which it would have been inconceivable to valorize, suddenly become the object of bitter negotiations with the aim of quantifying them in dollar terms, in order to fix levels of compensation, generally for women injured by the separation. In these diverse cases, the authors study the *resistances* of all sorts that these commensurations encounter and must overcome. Their case studies are interesting for the proposal I am advancing, but they have nevertheless the disadvantage of restricting commensuration to *monetization* (within a

[12] The night on which seigneurial rights and prerogatives were renounced in a session of the National Assembly, bringing to the end the *Ancien Régime* in France.

perspective that is perhaps unsurprising in the US context)[13]. The passage to a cash equivalent is one case (certainly historically important) among other constructions of equivalence spaces that have marked the history of humankind. Of the three cases presented here, the second (the evaluation of macroeconomic aggregates) involves the question of monetization, of course, but this is not necessarily so for the two others: risk management, and public service performance evaluation, using benchmarking, where the quantifications employed may or may not be monetary.

Probability in the 18th century: a daring intellectual construction

Many of the problems raised by contact zone border crossings had already catalysed in the 18th century with the use of the calculus of probabilities. At that time, probability seemed an astonishing construction, uniting in a single cognitive space, quantified by a number between 0 and 1, three forms of "degrees of belief" that were *a priori* very different (Daston 1989): (1) forms issuing from geometric constructions such as games of coin tossing or dice; (2) forms deduced from regularities observed in a large number of events, such as the sex ratio or mortality; and (3) forms resulting from a bundle of clues and conjectures about a unique event that was not comparable to any other, such as the guilt of a crime suspect. The fact of collecting and indexing in this way, within the same equivalence space, three entirely heterogeneous ways of knowing or believing, appeared a daring intellectual feat. It is true that in his *Ars Conjectandi*, published in 1713 eight years after his death, Jacob Bernoulli, using his model of successive drawings of black and white balls from an urn, had proposed an ingenious way of linking the first and second of the three forms of degrees of belief. His "law of large numbers" suggested a

[13] This approach is also related to the concern of certain sociologists to position themselves relative to economists, for whom money is *the* reference variable.

convergence of the frequencies observed of black and white balls, as the number of drawings increased. It created the possibility of assimilating the sex ratio or the suicide rate to the drawing of a ball from a Bernoulli urn. However, the same was not true of the subjective probability of a unique event, of which Cournot could say that it "varied from one intelligence to another". This brought about the relative discredit that both this type of probability and Bayesian reasoning experienced for almost a century, from the 1830s until 1930, to the advantage of the frequentist interpretation of probability.

Frequentist reasoning, originating from the model of Bernoulli urns, enjoyed great success in the 19th century, with the spread of so-called moral statistics, collected by the new offices of statistics promoted by Quetelet. Observed regularities induced a form of statistical determinism (even fatalism), and thereby, the idea that it is possible to ***predict*** at the collective macro-social level, phenomena that are unpredictable at the individual level, such as crime or suicide. This manner of reasoning, dubbed ***the taming of chance*** by Ian Hacking (1990), is at the origin of several types of activity. On one hand, the ***quantitative social sciences*** (sociology since Durkheim, econometrics since Frisch and Tinbergen) could share the ambition of being able to predict the future course of the world, like their big sisters astronomy and physics. On the other hand, ***insurance*** could now ground its rates (premia) and future repayments on the basis of the frequency of accidents observed in the past. But to do this, it is necessary to ***agree*** on the definition and scope of the Bernoulli urn (the risks to be covered), the identity of the balls (the elementary events), the nomenclature of the colours of the balls (the accident categories) and the coding procedures (once an accident is reported and a repayment made). Hence, prior to any risk measurement, risk quantification involves a complex game of conventions, negotiated in the contact zone. Recurrent controversies are normal, because the choices enumerated above involve ***judgements*** that are variable not only "according to each intelligence"

as Cournot thought, but also according to the interests of the actors. In this way, quantification is not sufficient to unite the various diagnoses around a univocal objectivity, although this may be the aim, in all good faith, of some (but not all) of the actors involved.

The notion of risk, associated with frequentist reasoning, has become essential in medical fields, on the one hand, in epidemiology (preventative intervention), and, on the other hand, in clinical medicine (therapeutic effectiveness). Some of the research on the subject is contemporaneous with Cournot: in France, the work of Docteur Louis, whose "numerical method" aimed at comparing the effects of various treatments of typhoid, and the studies of the English epidemiologist William Farr on the prevention of cholera epidemics (Desrosières 2000). The criticisms encountered by these applied quantitative methods are typical of what is played out in the contact zone. Resistance was of two sorts. The first "traditionalist" criticism invoked the singularity of the patient–doctor consultation, and the impossibility of reducing the complexity of a person to a family of "equivalence classes" by "segmenting the former into slices"[14]. The other "modernist" criticism was more interested in "the" precise direct cause of a symptom or treatment effect, and not in statistical regularities or average causes. This was the position of Claude Bernard, and later of Pasteurian microbiologists. The latter sought "the" cholera bacterium, or "the" AIDS virus. The same battle was replayed (and quickly resolved) during the AIDS epidemic at the beginning of the 1980s, when blinkered epidemiology led to talk of an ill that, statistically, hit "the 4 Hs" (Haitians, haemophiliacs, heroine addicts, and homosexuals), before the HIV virus was identified. If, in our time, the two perspectives – the first "macro" and statistical, and the second "micro" (in the sense of an individual case, but also in the sense of microscope) – are perceived as

[14] This perspective remains very much alive, in particular in the idea of the *patient–doctor private consultation* (*colloque singulier*) in general medical practice, and in psychoanalysis and homeopathic medicine.

complementary, the opposition, reflected in the history of medicine, refers back to a more general question, central in the analysis of what happens in the contact zone, concerning the kind of "causality" that quantitative methods suggest.

Statistical regularities and causality

Karl Pearson (1857–1936), one of the founders of mathematical statistics, was the first to formulate the ideas of correlation and regression. Drawing inspiration from the theories of the German physicist and anti-realist philosopher of science Ernst Mach (1838–1916), he emphasized the fact that statistics merely showed distributions, co-occurrences, regularities and "contingency tables" (that is to say, joint distributions), but *in no case causes* (Pearson 1912). Pearson thought causality was a "metaphysical notion". Even if, at a philosophical level, this position is conceivable, it certainly does not work for a man of action. Chased out the door, causality simply slips in again by the window, under a different, or even the same, name. Modern uses of the notions of "risk", "risk factor" and "risk category", in epidemiology or the treatment of delinquency or drug dependency, provide examples of these metamorphoses of causality, torn between an anodyne epistemology and sets of practices, which, in these fields, grab any means available to integrate economic, social and political observations and objectives of all sorts more or less coherently. By an irony of history, despite his anti-causalist *credo*, Karl Pearson himself furnished a formalism that through its very terminology induced an apparently causalist interpretation. Linear regression models, which put the "dependent variable" on the left-hand side of the equals sign and the "explanatory variables" on the right-hand side, lend themselves to such a reading, despite whatever possible precautionary admonitions the

statistician may utter[15]. The verb ***explain*** is sufficiently ambiguous to suggest a causality without explicitly affirming its existence. This lies at the core of questions about contact zone crossings. The contact zone is a translation area, like a canal lock, a decompression chamber or a corridor between two cognitive universes.

In linear regression models, the notion of a ***variable*** constitutes the core of the transformation that takes place in the transition from one world to another. It works like "The Purloined Letter" of Edgar Allan Poe, that no one can see, although it is clearly visible on the chimney. The subjects of verbs, and thus of actions, cease to be persons or social groups, and become variables, which are new entities, resulting from a series of equivalence conventions, taxonomies, codings, and evaluations according to various frameworks. People are decomposed into ***items***, which are recomposed into ***variables***. The crucible of this transformation is the ***table***, which crosses ***rows*** containing persons (or any other kind of beings, be they individuals or groups), against ***columns*** containing normalized coded items concerning each of these beings. In the first world, the table is read horizontally across the rows, and the individuals or groups are the subjects of verbs. Stories are told. In the second world, that of statistics, the gaze undergoes a perpendicular swivel: the table is vertically read down the columns, the variables become the actors. They are now the verb subjects. They are related, explained, and positively or negatively correlated. Each variable acts in a uniform way, provided that all the other variables are held constant. Thus, one seeks to separate and isolate their ***pure effects*** (under a ***ceteris paribus*** assumption) using econometric methods, involving logistic regression, that generalize the rationale of linear models. The coefficients of these regressions are assumed to provide the man of action

[15] The question of the absence of any automatic causal link is most often raised in connection with *correlation,* although the formula for the correlation coefficient is symmetric. By contrast, in virtue of their asymmetric and hence oriented form, ***regression equations*** invite causal readings even more strongly.

with the means of quantifying the marginal effects of the different levers that he controls.

This statistical language has two related properties. On the one hand, it is inspired by the natural sciences, which are ahistorical, and in which putative universal substances or concepts interact according to equally universal mechanisms. On the other hand, it lends itself well to the rationalization and optimization of action sought by executives in administration, politics and economics. For the latter, a *variable* defines the brief of a ministerial office, an objective to be attained, an indicator, a dial on a control panel. Linear economic models relate, on one side of the equation, those variables, often expressed in terms of *risks*, on which the executive *wishes to act* (the rate of unemployment, delinquency, road accidents, alcoholism) and, on the other side of the equation, other variables, expressed as *risk factors* (an alcohol limit, a speed limit), on which the executive *can act* through regulation, taxes or (a more recent solution) mechanisms of judicious incentive. The two properties are related. They are suitable for engineering models of intervention, which look for experimental regularities of a general scope, to orient, optimize and evaluate interventions.

The separation of *risks* and *risk factors*, a defining characteristic of causal linear models, results from explicitly discussed conventions. Patrick Peretti-Watel (2004) speaks of the "porousness" of the equations of these models, in the sense that there may be some hesitation regarding the status of certain variables. In investigations into hard drug use or teenage suicide, are alcohol consumption, nicotine ingestion or hashish smoking merely "risk factors" or in fact "risks"? The problem is more complicated, when, in so-called "multi-factor" models, the aim is to isolate the "pure effect, *ceteris paribus*," of certain factors, using econometric methods, the results of which depend crucially on the sets of dependent and explanatory variables chosen. The idea of separating the two categories of variable is less obvious in the first world

than in the second world of assessable effective intervention, conceived according to the mechanistic model of cause and effect. Several controversies concerning the use of quantitative methods feed on this tension. They start in the system of concepts and conventions according to which the problem is defined, and in terms of which probability estimates can be made. Very frequently, the protagonists do not share a consensus about the appropriate system. Insoluble dialogues of the deaf are the result. Three recent examples (among many) are the notion of the "precautionary principle", the potential dangerousness of genetically modified organisms (GMOs), and the possibility of evaluating various sorts of psychotherapy. Protagonists in these controversies understand and interpret the very notions of risk and uncertainty in different ways according to their positions.

Risk, uncertainty and the precautionary principle

The fact that, despite the wishes of 18th century philosophers, not all situations of uncertainty can be probabilized, was emphasized by Knight (1921). He introduced the distinction, frequently taken up by others (notably by Keynes), between *risk*, which is probabilizable, and ***uncertainty***, which is not. I have already drawn attention to the ambiguity in the word "possibility" corresponding to the senses of can and may: is "possibility" to be seen as a technical eventuality or a social agreement? Some insurance companies pride themselves on covering the most exceptional risks. Indeed, the business of reinsurance is to cover such risks[16]. Knight's distinction has been very useful in subsequent economic reflection. In practice, however, it assumes the aspect of

[16] In theory, one should distinguish *non-probabilizable* events from events of *very small probability*. Cournot was especially interested in the latter from a philosophical viewpoint, and a so-called "Cournot principle" for events of very small probability became the subject of several subsequent debates (Martin 1994).

a convention[17]. In recent history, marked by so-called exceptional catastrophes (9/11 in 2001, the heat wave in France in 2003, the Asian *tsunami* in 2004, Hurricane Katrina in 2005), the delimitation into risk and uncertainty in Knight's sense is once more called into question. Some commentators, such as Ulrich Beck, make risk, understood in both meanings, an essential characteristic of the current period[18]. In this context, the publication in 2005 of a report for the French General Plan Commission on the relations between "Uncertainty, precaution and insurability", shows that the distinction between risk and uncertainty is the result, if not of assessments "*which vary from one individual to another*", but rather of conventions relating to the argumentative and political use that is made of the distinction. The report puts forward an "economic theory of insurability under uncertainty" (Chemarin 2005).

The three debates mentioned above (the precautionary principle, GMOs, and psychotherapy) have each been the object of an abundant literature. Within the perspective I am proposing here, these texts can be re-read, with an emphasis on examining and comparing the place and role of probabilistic and statistical argument in each case, keeping in mind the notion of a "style of reasoning" developed by Alistair Crombie (1994) and Ian Hacking (1992). Of course, the controversies bear on very different questions. Nevertheless, positional homologies are discernible. In each case, the styles of reasoning of the two adversarial camps are quasi-incommensurable. Yet, at a transversal

[17] The distinction is often used by economists, but infrequently adopted by statisticians, for example. The standard reference text by Stephen Stigler (1986) on the history of statistics is called: *The History of Statistics: The Measurement of Uncertainty Before 1900*. In this case, uncertainty is most certainly "probabilized".

[18] This judgement requires some qualification since warnings of catastrophe were already expressed in comparable terms in earlier periods. For example, in the period 1820 to 1850, alarms were sounded concerning the then very new and spectacular accidents involving steam engines, gasometers and railways (Jean-Baptiste Fressoz, thesis in preparation at the Alexandre Koyré Centre under the supervision of Dominique Pestre).

level, analogies can be observed between the respective ways of arguing of the homological poles. On one side, that of the adversaries of the "precautionary principle", the GMO partisans and the advocates of behavioural cognitive therapy (BCT), the probabilistic argument is seen as decisive and directed towards ending the debate. The problem is assumed to be sufficiently well-defined in order that hypotheses about the probabilities of risk, or of therapeutic success, may be advanced and serve as evidence. At best, the methods of quantification and its results may be debated, but not the idea that the particular quantification brings an answer to the problem. In the three cases, this way of seeing things has the support of important scientific institutions, with arguments that are convincing within the canon or style of statistical reasoning. The French Academy of Sciences criticized the adoption of the precautionary principle, on the grounds that, in its view, the principle outlaws all risk taking. The Academy's criticism was founded on its assimilation of the principle to an outright proscription of any venture that involves potential danger. It intervened thus in the name of what it took to be freedom of research. The National Institute for Agronomic Research (INRA) endorsed experimental GMO cultivation, arguing that research was needed precisely in order to evaluate and quantify the potential risk of these crops. The National Institute for Medical Research (INSERM) compared BCTs to psychoanalysis by means of statistical "meta-analyses" of previous evaluations of these therapies. In the three cases, quantification and the expression of risk in terms of probabilities aim to unify and aggregate radically different, even antagonistic, viewpoints within a commensurable space. They seek to gain the status of a *common language*.

In each of the three cases, the opposing side questioned the equivalences bolstering the arguments of the first camp. They returned to debating within the contact zone intermediate between the complex world of words and the world modelled by numbers and probabilities. Advocates of the precautionary principle obtained its inclusion in the Charter of the Environment attached to

the Constitution. They deduced and took seriously the consequences of the notion of ***non-probabilizable uncertainty***, in Knight's sense. Far from proscribing all risk-taking research, on the contrary they asked that research and consultation should take place ***as far upstream as possible*** when new techniques or industries emerge (Godard 1997). They did not forbid the quantification or estimation of probabilities, but they did wish that it be done in a pluralist manner, within the framework of an enlarged universe of possibles, resulting from the confrontation of the most varied viewpoints and interests. They suggested transforming the relations between science, expertise and political decision, by including the doubts and uncertainties at the centre of democratic debate, instead of confining them within the work of experts required to deliver ready-made certainties into the hands of reassured decision makers. Procedures of this kind, such as "consensus conferencing", gathering together diverse categories of experts and the people involved, enabling them to express their points of view, have been tested. Some have attempted to theorize (occasionally in an idyllic or even utopian fashion) these practices, under the name of "hybrid forum", drawing attention to the diversity of actors involved in these new ways of conjugating expertise and society (Callon, Lascoumes, Barthe 2001). Nonetheless, until very recently, probabilistic and statistical tools have rarely been unfolded and discussed in these forums, except in the framework of associations[19].

One of the difficulties in the quantification of the problems raised relating to the precautionary principle is that the confrontation frequently involves *two* equivalence spaces, which for moral reasons are judged incommensurable. The first space is public health risks, probabilized or not, which are concerned

[19] Some associations set themselves this precise goal. In France, Pénombre, http://www.penombre.org, founded in 1993, "offers a public space for reflection and exchanges on the use of figures in society's debates: justice, sociology, the media, statistics". In Great Britain, official statistics are vigorously discussed by the association RadStats, http://www.radstats.org.uk, in existence since the 1970s: "We believe that statistics can be used to support radical campaigns for progressive social change. Statistics should inform, not drive policies. Social problems should not be disguised by technical language".

with life and death. The other space is the economy, for which, according to the analysis of Espeland and Stevens, "commensuration" is guaranteed by monetary evaluation. Although economists have long incorporated the "price of life" into their calculations, for example in choosing roadwork projects, the conjunction of the two spaces of quantification remains problematic. The inventors of the calculus of probabilities had, in theory, formalized a common space and a decisional criterion combining money and uncertainty (even beyond questions of life and death): mathematical expectation as the product of a potential loss or gain, multiplied by a probability. However, despite three centuries of debate and reflection on the criterion of mathematical expectation, strong reasons subsist to contest, reject or ignore it, precisely because it predicates an equivalence between beings that, for right or wrong, some people refuse to "co-measure". We are here at the heart of the contact zone alluded to above. The bitter debates about GMOs have to do with the difficulties in agreeing on a common equivalence space, insofar as the interests and issues, real or imagined, of farmers, seed producers and consumers are *simultaneously* uncertain and contradictory.

The reticence of psychoanalysts concerning comparative evaluations of the effectiveness of psychotherapeutic methods can also be read in terms of a refusal to accept the definition of effectiveness used in the meta-analyses undertaken at INSERM by the specialists in the field of BCTs, who concluded that the latter were superior. Psychoanalysts since Freud posit the singularity of the personal relationship that is constituted in an analysis. They refuse to circumscribe this interaction within the categories of the disappearance (in their view often momentary) of symptoms duly coded in a system of pre-established equivalences[20]. Their adversaries deplore "this typically Gallic

[20] Debates around the "numerical method" proposed by Docteur Louis in the 19th century set in opposition similar arguments on the theme of the private patient–doctor consultation, in which the patient "must be treated in his or her unique wholeness". The tension is the same between, on the one hand, the appeal to singularity and, on the other, the classification into nosographical categories. It lies at the heart of the turbulent history of relations between medicine and statistics: the "comparison of

refusal of the culture of assessment". Moreover, it is true that to this debate on the definition and ultimate purpose of the various methods, several economic arguments are more or less explicitly added, in terms of competition or public health economics[21].

The project of exploring the contact zone between the worlds of words and numbers encounters along the way several controversies on the equivalence conventions necessary for quantification. After emphasizing, in memory of Cournot, some of the controversies bearing on probability and mathematical expectation, I shall outline the recent debates involving, on the one hand, the evaluation and interpretation of macroeconomic aggregates and, on the other hand, the indicators promoted by so-called benchmarking techniques.

Controversies involving the volume–price split of the economic growth rate

Up to this point, I have followed the lead on the equivalence convention that is provided by the calculus of probabilities, which Cournot expounded in 1843. This convention bundles together a multiplicity of conjectures into one real number lying between 0 and 1. Yet Cournot is above all best known for his work in economics. His first book, *Recherches sur les principes mathématiques de la théorie des richesses* (1838), bears on another equivalence convention, one on which economics is founded, elaborating the notions of value and wealth through the general equivalent term of money (Aglietta and Orléan 2002). The co-existence of these two books, one from 1838, the other from 1843, by the same author, reveals a paradox. Whereas nowadays the idea of

incomparables" is always involved.

[21] The tension between an approach that is centred on the individual person and an approach that *compares and aggregates* within a perspective of the collective good, is subtly analysed from an ethical viewpoint by Anne Fagot-Largeault (1991). In her study of the "notion of the quality of life", she describes these two approaches as "deontological" and "teleological" respectively.

the *mathematization* of the economy seems synonymous with *quantification*, the effective synthesis of the two ways of doing economics, notably in the form of econometrics, dates only from the 1930s. The two books of 1838 and 1843 appear independent of each other, as though Cournot the economist and Cournot the probability theorist were unacquainted. It is true that the 1838 text of the *Recherches* is principally devoted to the analysis of partial equilibria. Schumpeter (1983) attributes Cournot's reticence regarding more global analyses to the fact that, in his view, global analyses would exceed "practical methods of calculation", and this leads him to envisage the use of a "small number of aggregates" and a "social income", which brings to mind modern national accounting:

> *Cournot recognized that* "in the complete rigorous solution of problems relating to some components of the economic system, it is indispensable to take the whole system into consideration" (*Mathematical Principles...*, p. 127; *Recherches...*, pp. 191–192). *This is exactly what Walras was to do. However, just like the Keynesian group of economists post Marshall, Cournot believed that* "this would exceed the power of mathematical analysis and our practical methods of calculation" (*Mathematical Principles...*, p. 127; *Recherches…*, p. 192). *Instead, he envisaged the possibility of dealing with the problems in terms of a small number of aggregates, in which the social income and its variants would occupy the place of honour.* (Schumpeter 1983, Vol. III, p. 281)

In his partial analyses, Cournot distinguishes carefully "real" and "nominal" wealth and variations in quantity and price. When, a century later, public accountants *quantified* (in the sense defined above) the aggregates used to express economic growth, they ran into the problem of splitting this growth into "volume" and "price" (Vanoli 2002). The ensuing controversies illustrate

the irreversible ratchet effect that quantification produces. Once quantification has been programmed, debates take place in realist terms, which are the only coin plausible according to the rationale of the practical and political uses of national accounts. The question of the volume–price split has consequences on current debates on the elusive equilibrium between stability and growth. The growth rate in *volume* terms (in constant money) of an economy from one period to another is calculated by deflation (division) of the progression in *value* terms (in current money) by a price index, itself also the result of a calculation. The price index thereby plays a key role in the calculation of the growth rate.

The volume–price split provoked lively debate in the 1990s because of the difficulties of taking into account a "quality effect" in the measurement of price progression, particularly in the case of computers, whose power had increased rapidly. How should agreement be reached on what constitutes *constant quality*? This question stimulated controversy in the United States, following a report by Michael Boskin (1996) to the US Senate. The report argued that price increases were over-estimated because the quality effect was insufficiently reflected. As a result, volume increases were under-estimated, a fact that, according to the report, had major political and economic consequences. In this debate, all the participants, who may very well have disagreed on the report's methodology and its conclusions, were implicit realists, since the notions of "over-estimation" and "under-estimation", accepted by everyone involved, presuppose that a "bias" exists relative to some reality that pre-exists any measurements. The language of realism was never in doubt. Incorporation of the quality effect implies a *judgement* and *conventions* (precisely concerning the said "qualities") and is not a simple matter of a purely realist metrology. Yet, this fact is rarely mentioned, even with sophisticated mathematical methods of the "hedonic price" type.
Assessments of the consequences of European stability policies are founded on measures of the inflation rate and the volume growth rate. The European

Central Bank (ECB) is frequently criticized for dealing only with the first of these rates to the detriment of the second (Fitoussi 2002). In debate, evocation of the volume–price split problem would cloud a politically important message. Would it be possible to re-endogenize these questions of measurement conventions within scientific and social debate? In which "hybrid forum" (in the sense of Callon, Lascoumes and Barthe, 2001) could this be done? Whatever one's view on such a thorny problem, it is clear that the ***social division of labour*** between statisticians, national accountants, university economists, ECB directors, political executives, journalists and citizens, plays an essential role in the distribution of realist and conventionalist rhetorics. This suggests a programme of research and public debate, rather than abstract normative responses. It does not involve ***relativizing*** the work of national accountants by exhibiting their conventional, and hence judgemental, character, but rather to suggest an analogy with legal rules, decided by common agreement, with the aim of creating a ***common language*** between the actors.

Since the 1980s, evaluations of national accounting are taken into account in indexation procedures, European regulations and treaties, in the Growth and Stability Pact and in the decision-making processes of the ECB. The ***constitutive*** (even constitutional) character of national accounting is thereby accentuated[22]. The horizons of national account use have changed. Some wish to "include in GDP" the quantification of new questions: the domestic work of women, externalities relating to destructions of the natural environment (Gadrey and Jany-Catrice 2005). In these different cases, quantification fashions and re-fashions society, and does not just measure or reflect it. National accounts seem spread between increasingly different uses, from their appearance in the context of their initial employment to enthrone Keynesian

[22] The fact that the Boskin report was commissioned and published by the US Senate shows well how the measurement conventions of the national accounts contribute to *institute* society, and not simply to *describe* it. For an update on the Boskin controversy after a decade, see: http://www.csls.ca/ipm/ipm12.asp.

policies or to guide indicative economic planning. These slippages of use prompt one to look again at a contact zone that, pre-1960, had been studied by economists such as Frisch and Hicks, but since then has been left to small teams of international expert specialists and little visited by the best-known economists (Vanoli 2002).

Tensions such as these, resulting from the multiplicity of uses, are also visible in the case of *corporate accounting*. Thus, in order to "value" balance sheet assets, three conventions may be employed, corresponding to three rationales of use. The *original cost* (or *historic value*) is used by the *manager* who is seeking to distribute depreciation annuities. The *resale value* is of concern to the *creditor* of the firm, who wonders what its assets are actually worth. Finally, the *sum of discounted future earnings* interests the *investor*, who wishes to allocate his or her financial assets. A comparable diversity exists in the different manners of calculating the profit of a firm, according to the objectives of the calculation. The *active* form of the verb "value" used by accountants signifies a procedure that is implicitly more constructivist than realist. Whereas economists debate the "foundations of value", accountants "value", that is to say, fabricate a value according to conventions. Within the legal rules and conventions of auditing, firms have degrees of freedom that allow them to show higher or lower profits, depending on whether their concern is to convey a message to their shareholders, potential acquirers, the State, or other actors in the economy. The parameters and the effects of the techniques (sometimes called window dressing) by which firms optimize their accounting decisions in light of various constraints are the object of an elaborate mathematical branch of accounting research, which draws support from the assumptions of microeconomic theory, *Positive Accounting Theory* (Casta 2000; Chiapello and Desrosières 2003).

Qualification, comparison, evaluation, and classification: the politics of statistical indicators

Unlike market activities, public policies, be they national, European or local, do not have available accounting criteria such as "market share" or profitability in order to judge their capacity to satisfy users' needs, or simply their efficiency. Traditional notions of public service and rational administration presuppose strong commitment by their members, monitored through structures of hierarchical subordination, of which the French and German states have long been the prime examples. Since the 1980s, however, this civic sense of public service has been widely judged insufficient to monitor democratically and efficiently activities that themselves are financed by the public purse. Quantified indicators were sought that could play a role more or less similar to the cost accounting, operating accounts and balance sheets of commercial firms. National accounting had exercised this role only partially, because its place was at the macroeconomic level, in a Keynesian or central planning perspective, without entering into the detail of public interventions. In this new perspective, indicators cannot be simply monetary, because the effects of interventions (schools, public health, security, foreign affairs, defence...) are not in general expressible in the familiar equivalence space provided by money. Thus the efforts undertaken by both the French state and the European Union can be seen as vast tentative experiments in the construction and negotiation of *new equivalence spaces*, by *agreeing* procedures for the quantification of the means and ends of intervention, using different units, amongst which money may be included but is not the sole unit. I shall discuss two examples: the *Constitutional Bylaw on Budget Acts* (CBBA)[23], unanimously adopted by the French Parliament in 2001, and the

[23] In French, this is referred to as the *Loi organique relative aux lois de finances* (LOLF).

Open Method of Coordination (OMC) employed by the European Union[24]. The political and historical contexts of these instruments (one French, the other European) of government public policy are different, but they share the common feature of giving a central role to *statistical indicators*, that is to say, tools little discussed in public debates, although these tools constitute the actual spaces and languages that delimit and structure these debates[25].

The CBBA is a new way of structuring the State Budget, according to objectives to be achieved and not according to the means allocated. It entails that these objectives be made explicit and *quantified*, so that Parliament no longer just approves expenditure but verifies the achievement of objectives and the performance of services. This idea of the quantification of the means and ends of public intervention seems evident if Parliament is to play fully its constitutional role in voting and executing the implementation of the Budget. Nevertheless, it entails important processes of objectivation and the "equivalencing" of disparate activities within the contact zone already discussed. These activities must be articulated, discussed, named, qualified, compared, classified and evaluated. The right indicator is never evident. A pre-existing institutional and social order is often described and made explicit. In theory, this can only happen through the deep involvement of the persons concerned. Often, however, the very notion of a quantitative indicator arouses reticence, comparable to those described above in connection with medicine. The idea resurfaces that these procedures lead to the comparison of incomparables. Sometimes this idea appears absurd, and all the more so the

[24] There are historical precedents, that despite some differences, could be studied within this perspective: the economic planning experiments of socialist countries, and the "rationalization of budget choices" undertaken in France in the 1970s and subsequently pursued under the name of "public policy evaluation" (Spenlehauer 1998). In these different cases, non-monetary quantified indicators were implemented.

[25] On the instrumentation of public intervention via "the choice and use of tools (techniques, operating methods, rules) that bring into effect and operationalize government action", see the collection *Gouverner par les instruments*, edited by Pierre Lascoumes and Patrick Le Galès (2004).

more deeply the actors are involved in their tasks. The fact of creating categories, designed in principle to simplify the world and render it intelligible, at the same time modifies it, and makes it a different world. Actors, by changing the system of reference, are no longer the same actors, since their actions are henceforth directed by these indicators and classifications, which become criteria for intervention and evaluation.

The CBBA was presented as enabling Parliament to know better and evaluate public service interventions, within the perspective of a rebalancing of legislative and executive powers. In this context, the fact that this entails the invention and installation of a large number of quantified indicators does not seem to have attracted much attention from commentators, at least between 2001 and 2003. It seemed a technical question, to be resolved by technicians. Yet, the ever more detailed discussions starting from 2004 (the law is due to come into force from 2006) show that this moment of quantification (in the sense of the ***action of quantifying***) is decisive for the course of events, although this did not precipitate any more general study of the questions mentioned above. The difficulties and perverse effects appeared ***one by one***, occasionally becoming the object of denunciations or jokes. Thus, for example, the police force (under the Ministry of the Interior) and the ***gendarmerie*** (under the Ministry of Defence) responsible for road safety chose as indicator of their performance the percentage of positive alcohol tests out of the total of all tests effected. However, the police initially wished to evaluate its performance by an ***increase*** in this proportion, whereas the ***gendarmerie*** sought to ***diminish*** the proportion. Each choice had its logic. The example shows what sort of issue a political sociology of quantification could treat in a study of the methods and effects of "indicator politics" entailed by the CBBA, or, at the European level, by the OMC.

In other certainly different contexts, similar effects have been observed. Thus the centralized planning of former communist countries failed in part because

it proved impossible to fix reliable indicators for the achievement of the objectives of the Plan. The problem was caused by the perverse effects of retroaction, induced by these indicators, on the behaviour of actors. In the American context, in a study on the installation of a system of professional classification in hospitals, Bowker and Star (1999) showed how the formal explicitation of previously implicit activities resulted in their transformation. Indicators and classification are simultaneously *constraints* and *resources* which, by their very existence, change the world. Further, these management instruments, which the authors describe as boundary objects, are sufficiently ambiguous and polysemiotic to circulate from one world to another with partially different interpretations and uses. This is a sociological way of understanding the multivocity of statistical assertions, discussed in the introduction starting from Cournot's observations. The metaphor of *boundary objects* is close to what I have referred to as the *contact zone*, on condition that the latter is interpreted not just cognitively (contacts between more and less formalized languages), but also sociologically (contacts between more and less expert actors, using different languages):

> *Boundary objects are those objects that both inhabit several communities of practice and satisfy the informational requirements of each of them. Boundary objects are thus both plastic enough to adapt to local needs and constraints of the several parties employing them, yet robust enough to maintain a common identity across sites. They are weakly structured in common use and become strongly structured in individual-site use.* (Bowker and Star 1999, p. 297)

The OMC is used by the European Union to try to harmonize social policies (employment, education, welfare) that do not involve monetary and economic domains falling explicitly within its sphere of competence. The first example of the method was the European Strategy for Employment (ESE) proposed at the

Amsterdam Summit in 1997. The name and the procedure of the OMC were decided at the Lisbon Summit in 2000. The principle underlying the OMC is that, in an intergovernmental way, states set themselves common objectives, expressed in terms of quantified indicators, relative to which states are then classed and evaluated, as in a prize list. In theory, the results of this benchmarking exercise are purely indicative, but the simple fact that they are published serves as a powerful stimulant to guide national policies in the directions indicated at Summits (Dehousse 2004). For example, an employment rate of 70% was fixed at the Lisbon Summit in 2000. Thus, just like the CBBA, the OMC gives a key role to statistical indicators, the former for the presentation and monitoring of the state budget, the latter for the indirect guidance of European social policies.

The way in which member states of the European Union agree on methods for this quantification is therefore essential, although it is poorly known. Technically, the work is divided into two parts. The political authorities decide on the choice of indicators and define them in a succinct verbal manner. Then they transmit an order to quantify to the statisticians at Eurostat (the Office of Statistics of the European Union) and the National Institutes of Statistics. The expression "agree" is therefore itself shared, since the political executives leave to the statisticians the business of sorting out the "details", as for example in the precise definitions of the notions of *employment rate* (Salais 2004), *disposable household income* (Nivière 2005), and *homeless person* (Brousse 2005). These three studies show that, at this stage, given the great institutional differences between countries, statisticians cannot avoid leaving vague certain sometimes-important specifications in measurement procedures, and cannot harmonize them completely. The method is called "open" because it is not binding and leaves states free to adapt it to their institutional specificities, notably by choosing as sources direct enquiries or administrative registers (Desrosières 2005).

Indicators produced in this way can be seen as boundary objects in the sense mentioned above. These objects have a vague, non-exhaustively defined character, which allows them to serve in several universes that were previously unrelated but now become comparable. In this way, the idea of boundary objects comes close to the idea of common language. Natural language has analogous properties: it is because interlocutors do not spend their time making explicit the meaning and content of the words uttered that communication is possible. The objects produced by public statistics, the unemployment rate, the price index, GDP, share this sub-explicitness to some extent. A complete explicitation of the method of their construction and their content would risk weakening their argumentative effectiveness, not just because it would uncover the conventions and approximations that the user had not suspected, but simply for reasons of economy in the course of the exchanges, debates and demonstrations in which statistical arguments find their place. Except in the case of controversy (such as that resulting from the Boskin report), all this remains implicit. However, the idea of *vagueness* can only shock, and with perfect justification, those professionals concerned with the definition and standardization of their objects. They are caught between two contradictory requirements. On the one hand, as good engineers, they wish to specify their procedures completely, but, on the other hand, negotiations encourage them to tolerate compromises without which the indicators necessary for benchmarking would be simply impossible to provide. The equilibrium that they actually seek to maintain between these two requirements has received little formalization[26].

[26] Some formalization is nevertheless sometimes perceptible at the level of the *meta-data* (data about data). These are sought and provided, but giving too many details would introduce an undesirable element of insidious doubt. A statistical argument is more effective if it can be invoked naked, without footnotes.

Social conditions for recourse to the Bernoulli urn model

The "resistance to statistics" attributed to Cournot can be interpreted in the light of the controversies discussed here: "how should one compare data collected in places and human contexts, in such heterogeneous environments?" The question can be juxtaposed with the criticism formulated by Baron de Keverberg in 1827 to invalidate the probabilistic sampling methods (the future "polling") employed by Laplace since the end of the 18th century to estimate the French population. How, Keverberg asked, can one justify the assumption of the unicity of the Bernoulli urn when the French territory is so heterogeneous? Does the procedure of replacing the whole by a (small) part allow one to extrapolate a result obtained under these conditions? The equivalence convention of balls contained in an urn was thereby thrown into question. This criticism had such an impact (in particular on Quetelet) that the probabilistic polling method was disbarred from use until the beginning of the 20th century (Desrosières 2000, chap. 7).

The notion of equivalence convention articulates the social (convene – agree) and logical (the mathematical relation of equivalence) dimensions of the process of quantification. It enables one to show how the objections of Keverberg and Cournot were overcome not just logically but also socially, in a way that posed and resolved practical problems. From this standpoint, the social uses of probability and statistics have been ill served by their juxtaposition with the natural sciences, which the pioneers of the quantification of the social sciences, and more generally of the guidance and evaluation of public intervention, frequently made. They hoped to benefit from the putative objectivity of the natural sciences, according to which in the 19th century "the only science is the science of the measurable". A different and less banal rapprochement could be made with the constructions issuing from law and the political sciences. A society cannot exist without constitutive conventions that are negotiated and inscribed in stable texts. The adjective

"conventional" is not synonymous with "arbitrary". Since the 18th century, probabilistic and statistical tools have been included in the panoply of the common languages and instruments in terms of which human societies think of themselves, act, and express their projects and disagreements. Cournot was perhaps too much of a logician and an epistemologist to venture further in his fertile intuitions on the multiplicity of meanings of the probabilistic and statistical tools of his time. But then, it is true that their social uses were far less numerous than is currently the case. The questions that Cournot the philosopher treated can be dealt with today in terms of the sociology of the diverse forms of quantification and modelling that suffuse the world of action and the economic and social sciences.

References

Aglietta, M. and A. Orléan (2002), *La monnaie entre violence et confiance*, Paris, France: Odile Jacob.

Amosse, T., Y.V. Andrieux and L. Muller (2001), "L'esprit humain est-il bayésien? ", *Courrier des statistiques*, n°100, 25-28; http://www.insee.fr/fr/ffc/docs_ffc/cs100g.pdf.

Armatte, M. (1995), *Histoire du modèle linéaire. Formes et usages en statistique et économétrie jusqu'en 1945*, thesis, EHESS, Paris.

Armatte, M. and A. Dahan-Dalmedico (2004), "Modèles et modélisations, 1950-2000 : Nouvelles pratiques, nouveaux enjeux", *Revue d'histoire des sciences*, **57** (2), 245–305.

Béraud, A. (2005), *Richesse et valeur : la contribution des économistes français du début du 19ème siècle*. Economies et Sociétés, Cahiers de l'ISMEA, Œconomia, *Histoire de la pensée économique*, n° 36, 1009-32; http://www.u-cergy.fr/rech/pages/beraud/valeur.pdf.

Boskin, M., E. Dulberger, R. Gordon, Z. Griliches and D. Jorgenson (1996), *Toward a More Accurate Measure of the Cost of Living: Final Report to the Senate Finance Committee,* Washington, DC, U.S. Government Printing Office.

Bowker, G. and L.S. Star (1999), *Sorting Things Out. Classification and Its Consequences*, Cambridge, MA: MIT Press.

Brousse, C. (2005), "Définir et compter les sans-abri en Europe : enjeux et controverses", *Genèses*, **58** (March), 48–71.

Callon, M., P. Lascoumes and Y. Barthe (2001), *Agir dans un monde incertain. Essai sur la démocratie technique*, Paris: Seuil.

Casta, J.F. (2000): "Théorie positive de la comptabilité", in B. Colasse (ed.) (2000), *Encyclopédie de la comptabilité, Contrôle de gestion et Audit*, Paris: Economica, pp. 1223–1232.

Chemarin, S. (2005), "Vers une théorie économique de l'assurabilité en incertitude", in C. Henry, S. Chemarin and E. Michel-Kerjean (eds.) (2005), *Incertitude, Précaution et Assurabilité*, Report to the General Plan Commission; http://ceco.polytechnique.fr/CDD/PDF/2005-005.pdf.

Chiapello, E. and A. Desrosières (2003), "La quantification de l'économie et la recherche en sciences sociales : paradoxes, contradictions et omissions. Le cas exemplaire de la *Positive Accounting Theory*", Contribution to the Colloquium "Conventions et Institutions : approfondissements théoriques et contributions au débat politique"; http://forum.u-paris10.fr/cd/fr/seminaires/coll_convention.

Cournot, A. (1838), *Recherches sur les principes mathématiques de la théorie des richesses*, re-edited in G. Jorland (ed.) (1980), *Œuvres complètes de Cournot*, tome VIII, Paris: Vrin; English translation (1898/1971), *Researches into the mathematical principles of the theory of wealth*, translated by N. Bacon, 6th ed., New York: A.-M. Kelley.

Cournot, A. (1843), *Exposition de la théorie des chances et des probabilités*, re-edited in B. Bru (ed.) (1984), *Œuvres complètes de Cournot*, tome I, Paris: Vrin.

Crombie, A. (1994), *Styles of Scientific Thinking in the European Tradition: The History of Argument and Explanation Especially in the Mathematical and Biomedical Sciences and Arts*, London: Duckworth.

Daston, L (1989), "L'interprétation classique du calcul des probabilités", *Annales des Economies, Sociétés et Civilisations*, **3**, 715–731.

Dehousse, R. (2004), "La méthode ouverte de coordination. Quand l'instrument tient lieu de politique", in P. Lascoumes and P. Le Galès (eds.) (2004), *Gouverner par les instruments*, Paris: Presses de Sciences Po, pp. 331–356.

Desrosières, A. (2000), *La politique des grands nombres. Histoire de la raison statistique*, Paris: La Découverte /Poche; English translation (2002): *The Politics of Large Numbers: A History of Statistical Reasoning*, Cambridge, MA: Harvard University Press.

Desrosières A., 2005: "Décrire l'Etat ou explorer la société: les deux sources de la statistique publique", *Genèses*, **58** (March), 4–27.

Detienne, M. (2000), *Comparer l'incomparable*, Paris: Seuil.

Espeland, W. and M. Stevens (1998), "Commensuration as a Social Process", *Annual Review of Sociology*, **24**, 313–343.

Fagot-Largeault, A. (1991), "Réflexions sur la notion de qualité de la vie", *Archives de philosophie du droit*, tome 36, volume "Droit et science", 135–153, reprinted in R. Launois and F. Régnier (eds.) (1992), *Décision thérapeutique*

et qualité de vie, Paris: John Libbey, pp. 83-100; English translation in L. Nordenfelt (ed.) *Concepts and Measurement of Quality of Life in Health Care*, Philosophy of Medicine Series, n° 47, European Subseries n° 1, Dordrecht: Kluwer, pp. 135–160.

Fitoussi, J.P. (2002), *La Règle et le Choix : De la souveraineté économique en Europe*, Paris: Seuil.

Gadrey, J. and F. Jany-Catrice (2005), *Les nouveaux indicateurs de richesse*, Paris: La Découverte.

Gigerenzer, G. and D.J. Murray (1987), *Cognition as Intuitive Statistics*, Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Godard, O. (ed.) (1997), *Le principe de précaution dans la conduite des affaires humaines*, MSH and INRA.

Hacking, I. (1975), *The Emergence of Probability*, Cambridge: Cambridge University Press.

Hacking, I. (1990), *The Taming of Chance*, Cambridge: Cambridge University Press.

Hacking, I. (1992), "Statistical Language, Statistical Truth, Statistical Reason: The Self-Authentification of a Style of Reasoning", in H. MacMullin (ed.) (1992), *Social Dimensions of Science*, Indiana: Notre Dame University Press, pp. 130–157.

Hacking, I. (1999), *The Social Construction of What?*, Cambridge, MA: Harvard University Press.

Jorland, G. (1987), "The Saint-Petersburg Paradox, 1713-1937", in L. Kruger, L. Daston and L. Heidelberger (eds.) (1987), *The Probabilistic Revolution*, Vol. 1, *Ideas in History*, Cambridge, MA: MIT Press, pp. 157–190.

Kahneman, D. and A. Tversky (1973), "Availability: A heuristic for judging frequency and probability", *Cognitive Psychology*, **5**, 207–232.

Knight, F. (1921), *Risk, Uncertainty and Profit*, Boston: Houghton Mifflin Company.

Lascoumes, P. and P. Le Gales (eds.) (2004), *Gouverner par les instruments*, Paris: Presses de Sciences Po.

Latour, B. (1984), *Les microbes. Guerre et paix*, followed by *Irréductions*, Paris: Métailié; English translation by A. Sheridan and J. Law (1988), *The Pasteurization of France*, Cambridge, MA: Harvard University Press.

Martin, T. (1994), "La valeur objective du calcul des probabilités selon Cournot", *Mathématiques, Informatique et Sciences humaines*, CAMS-EHESS, n° 127, 5–17.

Ménard, C. (1977), "Trois formes de résistance aux statistiques : Say, Cournot, Walras", in *Pour une histoire de la statistique*, tome 1: *Contributions*, Paris: Economica-INSEE; reprint (1987), pp. 417–429.

Morgan, M. and M. Morrison (1999), *Models as Mediators: Perspectives on Natural and Social Science*, Cambridge: Cambridge University Press.

Nivière, D. (2005), "Négocier une statistique européenne : le cas de la pauvreté", *Genèses*, **58** (March), 28–47.

Pearson, K. (1912), La *Grammaire de la science*, Paris: Alcan; English original (1911), *The Grammar of Science*, London: A. and C. Black.

Peretti-Watel, P. (2004), "Du recours au paradigme épidémiologique pour l'étude des conduites à risque", *Revue Française de Sociologie*, **45** (1), 103–132.

Salais, R. (2004), "La Politique des indicateurs. Du taux de chômage au taux d'emploi dans la stratégie européenne pour l'emploi (SEE)", in B. Zimmermann (ed.) (2004), *Action publique et sciences sociales*, MSH, Paris; http://www.insee.fr/fr/nom_def_met/colloques/acn/colloque_10/Salais.pdf.

Schumpeter, J. (1983), *Histoire de l'analyse économique*, Vol. III, *L'âge de la science*, Paris: Gallimard.

Spenlehauer, V. (1998), *L'évaluation des politiques publiques, avatar de la planification*, thesis, Université Grenoble II - Pierre Mendès-France, Institut d'études politiques de Grenoble.

Stigler, S. (1986), *The History of Statistics. The Measurement of Uncertainty Before 1900*, Cambridge, MA: Harvard University Press.

Thevenot, L. (1983), "L'économie du codage social", *Critiques de l'économie politique*, 23-24, 188–222.

Vanoli, A. (2002), *Une histoire de la comptabilité nationale*, Paris: La Découverte.

Zelizer, V. (2001), "Transactions intimes", *Genèses*, **42** (March), 121–144.