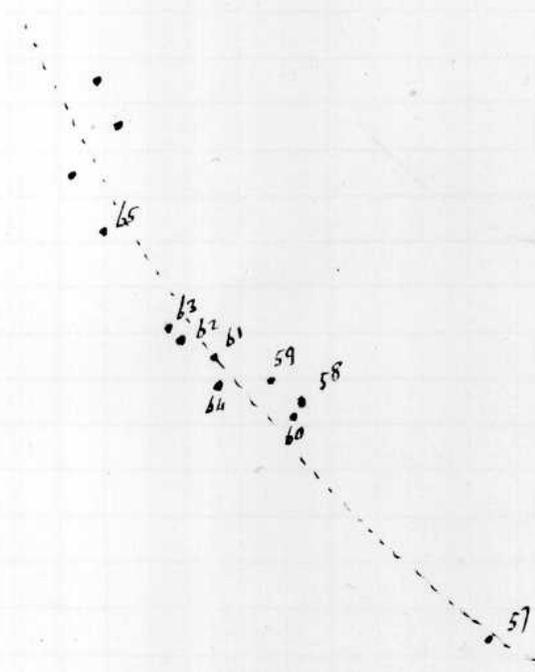


189

# Observations of an Expert

Marcel Boumans

41  
40  
39  
38  
37  
36  
35  
34  
33  
32  
31  
30  
29  
28  
27  
26  
25



University of Amsterdam  
2009



1.5 1.6

2.0

2.5

# Observations of an Expert

Marcel Boumans

Dept. of Economics

University of Amsterdam

## 1. Introduction

Maurice Kendall and Alan Stuart (1963) define statistics as “the branch of scientific method which deals with the data obtained by counting or measuring the properties of populations of natural phenomena. In this definition ‘natural phenomena’ includes all the happenings of the external world, whether human or not” (p. 2). The same word ‘statistics’ is also applied to the numerical material with which the method operates, so the data obtained by counting and measuring. More generally, statistics in the sense of data is in this paper defined as quantitatively registered history. To avoid misunderstanding in the interpretation of which meaning is meant, statistics as a method takes the singular, the word statistics as data takes the plural.

For various reasons, statistics, sometimes, are not enough, complete or available to legitimate policy decision-making. Then recourse has to be taken to other kinds of observations. These are the observations made by human beings, so they are not the registrations of a measuring device, nor mechanical observations, nor ‘objective’ observations. The standard view (in science) is that these non-statistical observations run the risk of being more biased than

statistics, therefore often denoted as ‘subjective’ observations. Some people, however, are assumed to be less biased than others, because of training, because of experience, or because they have better intuitions. These people are called experts.

This paper discusses the role of expert’s observations in different practices of decision making. In these practices it is never the case that the observations of one sole expert is being used, so discussing the role of expert’s observations implies a discussion of how these observations are combined.

The practices that will be discussed are cases from economics and econometrics, with the exception of a case in risk analysis, which provides a rather lucid example of a practice in between statistics and the intuitive expert. These practices will be discussed in their historical context, starting with a notorious debate in econometrics on the role of statistics.

## **2. Pitfalls Debate**

The ‘Pitfalls debate’<sup>1</sup> was a debate of the early 1930s between “two young men from the emerging econometrics group” (Hendry and Morgan 1995: 38), Wassily Leontief and Ragnar Frisch. In 1929, Leontief had published an article in which he tried to treat demand and supply curves together by working out a method of simultaneously estimating these two curves from one set of observations on price and quantities. His method depended upon splitting his data set into two subsets and imposing common

---

<sup>1</sup> This naming is from Morgan (1990). This “acrimonious debate” is discussed in Morgan 1990, chapter 6.3 and Hendry and Morgan 1995: 38-40.

supply and demand curves on the two data sets. This method appeared to give him sufficient information to estimate both the demand and supply elasticities. This method also depended upon a number of other assumptions relating to the structure of demand and supply and the changes to which the demand and supply relations were subject. Frisch (1933) objected “almost violently” (Hendry and Morgan 1995: 38) to Leontief’s method. Subsequently, a ‘Reply’ by Leontief (1934a), a counter-reply by Frisch (1934), Leontief’s ‘Final word’ (1934b) and finally ‘Some comments’ by Jakob Marschak (1934) appeared in *The Quarterly Journal of Economics*, which together drew up the Debate.<sup>2</sup>

As a member of the staff of the *Institut für Weltwirtschaft und Seeverkehr* (Institute for World Economics and Sea Traffic) at the University of Kiel, Germany, from 1927 to 1931<sup>3</sup>, Leontief was engaged in research on the derivation of statistical demand and supply curves. He published two articles based on this research in *Weltwirtschaftliches Archiv*: ‘Ein Versuch zur statistischen Analyse von Angebot und Nachfrage’ (1929; An attempt to statistical analysis of supply and demand) and ‘Studien über die Elastizität des Angebots’ (1932; Studies on supply elasticity).

The subjects of derivation of statistical demand curves was chosen by me rather than suggested by Löwe<sup>4</sup> or anybody else. Methodological

---

<sup>2</sup> This part of the debate published in *The Quarterly Journal of Economics* is reprinted in Hendry and Morgan 1995: 257-270.

<sup>3</sup> To be more precise, Leontief worked at the Kiel institute from May 1927 till May 1929, and, after a year working in China as economic adviser of the Nanking/China railroad ministry, from May 1930 till April 1931 (Beckman 2000: 65-66).

<sup>4</sup> From 1926-1930, Adolf Löwe (1893-1995) was the leader of *Abteilung für Statistische Weltwirtschaftskunde und Internationale Konjunkturforschung* (ASTWIK: Department of statistical world economics and international business cycle research), the department where Leontief was employed.

problems arising in connection with the possible use of theoretical approach in concrete empirical analysis always interested me.

May I add that it was disappointment in the possibility of basing concrete factual explanation of economic phenomena by traditional theory of supply and demand that led me to development of the “input-output” approach (letter of Leontief to Beckmann 03.05.1993, published in Beckmann 2000: 531-532).

Only the first paper did play a role in the ‘Pitfalls debate’, so for a clarification of Leontief’s position we will focus only on this text. The largest part of this 1929 paper was a discussion about the assumptions that has to be made with respect to the demand and supply curves. Therefore, Leontief made a distinction between the ‘structure’ of demand and supply and their ‘conditions’ (*Bedingungen*).

Structure was understood to be production technology and all psychological, biological or social determined preference orderings – today recapped as ‘tastes and technology’. Structure was, according to Leontief, represented by the elasticities of both supply and demand. The conditions were understood to be the totality of all other factors influencing equally the whole range of supply and demand. The levels of the curves represented these conditions. With respect to the stability of both elasticities and conditions, Leontief argued for the following assumptions (pp. 15-17):

1. *Perturbations of price-quantity changes*: Changes of elasticities and conditions will mainly affect the conditions of other goods and only

in rare cases and to a small extent lead to changes of elasticities of other goods.

2. *Coincidence and relative rareness of changes of elasticities:* Price and quantity fluctuations are more often caused by shifts of the levels of the demand and supply curves than by changes of elasticities.
3. *Independence of changes of the levels:* For each good, there is no relation between changes in its demand structure and changes in its supply structure.

The rest of the paper aimed at the development of a statistical framework to determine both supply and demand elasticities and both supply and demand levels. The theoretical assumptions above were, therefore, translated into the following more specific assumptions:

1. Supply and demand curves have for each point an equal and constant elasticity. That means that for a logarithmic scale both curves can be drawn as a straight line of which the slope represents elasticity, and changes in supply and demand are only due to parallel shifts of these curves.
2. Shifts of both curves are independent of each other.

Both elasticities were determined on the basis of the times series of quantities and of prices, and by dividing both series in two (and numbered as 1 and 2). In an appendix, written by Robert Schmidt<sup>5</sup>, the formulas for the

---

<sup>5</sup> The year after, Schmidt (1930) discussed in a separate paper the “Prägnanz” (pithiness, sententiousness) of the calculated elasticities. This is however a mathematical characteristic of elasticities and not a statistical one, and therefore is not relevant for the discussion of this paper.

calculation of the elasticities were derived using various mathematical techniques.

Four years after Leontief's 'statistical attempt', the *Frankfurter Gesellschaft für Konjunkturforschung* (Frankfurt's society for business cycle research<sup>6</sup>) published a critique on this approach by Ragnar Frisch. Frisch's critique appeared in a series in which already two publications on demand and supply analysis were published: Hans Staehle's 'Die Analyse von Nachfragekurven in ihrer Bedeutung für die Konjunkturforschung' (1929, issue 2; The analysis of demand curves in relation to their meaning for business cycle research), Henry Schultz's 'Der Sinn der statistischen Nachfragekurven' (1930, issue 10; 'The meaning of statistical demand curves').

In many cases the coefficients obtained by [Leontief's] method are, I believe, entirely meaningless, their magnitude being determined essentially by the random disturbances in the material. And in those cases where they have a sense, they do not as a rule express demand and supply elasticities, but simply express the historical trend connection between price and quantity. (Frisch 1933: 10)

To discuss Leontief's method, Frisch first set up a more "general" framework in which he (only) "granted" Leontief's first assumption of constant elasticities. He presented the demand and supply functions as:

---

<sup>6</sup> This society was led by Eugen Altschul (1887-1959) and existed from 1926 till 1938 (de facto only active till Altschul's emigration in 1933 to the US). Altschul was very critical on pure statistical-empirical methods; he advocated a synthesis of economic theory and statistics (see Beckmann 2000: 456-457).

$$\text{Demand:} \quad x = u + \alpha p \quad (1)$$

$$\text{Supply:} \quad x = v + \beta p \quad (2)$$

where  $x$  stands for the log of the quantity demanded and supplied,  $p$  for the log of price,  $u$  and  $v$  are the shifts,  $\alpha$  and  $\beta$  are the demand and supply elasticities respectively. Next he defined the “relative violences”:

$$l = \frac{\sigma_x}{\sigma_p} \quad \text{and} \quad \lambda = \frac{\sigma_u}{\sigma_v}.$$

From equations (1) and (2), Frisch derived the “fundamental equations”:

$$(\alpha\beta - \rho\lambda\beta^2) - (\alpha + \beta - 2\rho\lambda\beta)rl + (1 - \rho\lambda)l^2 = 0 \quad (3a)$$

$$(\alpha^2 - \lambda^2\beta^2) - 2(\alpha - \lambda^2\beta)rl + (1 - \lambda^2)l^2 = 0 \quad (3b)$$

where  $r$  is  $\sigma_{xp}$ , and  $\rho$  is  $\sigma_{uv}$ .

These fundamental equations were, thereupon, used to discuss different cases of the nature of the shift distributions:

1. Demand curve stability:  $\lambda = 0$ , which means that  $r^2 = 1$ . As a result  $\alpha = \pm l$  and  $\beta$  is indeterminate, which he labeled as the Cournot effect on the demand side.
2. Supply curve stability:  $\lambda = \infty$ . Similar to the case of demand stability,  $r^2 = 1$ . As a result  $\beta = \pm l$  and  $\alpha$  is indeterminate, which is labeled as the Cournot effect on the supply side.
3. Bilateral and uncorrelated shifts:  $0 < \lambda < \infty$  and  $\rho = 0$ . Then, equation (3a) reduces to:

$$\alpha\beta - (\alpha + \beta)rl + l^2 = 0$$

This is, according to Frisch, the case Leontief discussed.

4. Bilateral and highly correlated shifts:  $0 < \lambda < \infty$  and  $\rho = \pm 1$ , which means that in case of  $\lambda \neq \pm 1$  and  $\lambda \neq \pm \alpha/\beta$ ,  $r^2 = 1$ . This correlation is, however, not a Cournot effect (cases 1 and 2), but a trend effect. “The slopes of the  $(x, p)$  regression therefore no expresses neither the demand elasticity, nor the supply elasticity, but the historical trend relation between  $x$  and  $p$ ” (p. 18).

Next, Frisch showed that Leontief’s formulas for the calculation of the elasticities were equivalent to the system of three equations ( $\rho = 0$ ):

$$\alpha\beta - (\alpha + \beta)r_1l_1 + l_1^2 = 0 \quad (4a)$$

$$\alpha\beta - (\alpha + \beta)r_2l_2 + l_2^2 = 0 \quad (4b)$$

$$\alpha\beta - (\alpha + \beta)\eta + \eta^2 = 0 \quad (4c)$$

And this can be reduced to an even more simple equation, which Frisch called the “two-material equation”:

$$\begin{vmatrix} 1 & r_1l_1 & l_1^2 \\ 1 & r_2l_2 & l_2^2 \\ 1 & \eta & \eta^2 \end{vmatrix} = 0 \quad (5)$$

In other words, Frisch was able to “strip Leontief’s problem of irrelevant complications” (p. 21), and to reduce the mathematics to the solution of one equation (5).

In a Table, ‘Nature of the solution of the two-material equation’ (p. 30), Frisch discussed the roots of the “two-material equation” (5) for the various classes determined by the parameters  $l_1$ ,  $l_2$ ,  $r_1$  and  $r_2$ . Thereupon, Frisch discussed for each class, whether the roots, when they are determine, are

meaningful with respect to correlations and violences of the shifts,  $\rho$ , and  $\lambda$ . He arrived at the following conclusions:

Thus the only situation in which there might be a meaning in using Leontief's method, is the case where *all* the following three conditions are fulfilled:

1. The relative  $(x, p)$  violence must be *significantly* different in the two materials on which the computations are built.
2. The  $(x, p)$  correlation must be *significantly* different in these two materials.
3. The shifts must be uncorrelated in both materials.

[...]

Is there a great likelihood that we shall meet such a situation in practice? I think it is safe to say that it would be a veritable miracle if we should ever find a material satisfying all these conditions and having nevertheless the same demand and supply elasticities. It would even be a miracle, I think, if the two observable criteria 1) and 2) should be satisfied. In virtually all practical cases where it is plausible to assume that the elasticities have been constant I believe we shall have the situation where at least one, if not both of the conditions 1) and 2) are violated. If Leontief had discussed the conditions 1) and 2), I believe he would have found that they are not fulfilled in any of his data. But Leontief has not gone into any analysis of these conditions.

It is true that there passages where he expresses the necessity of using two materials between which there exists some sort of difference, but he does not seem to understand what sort of difference this must be. (Frisch 1933: 37)

The year after this publication, the debate moved to the pages of the *Quarterly Journal of Economics*, starting with “a reply” by Leontief (1934a).<sup>7,8</sup> First of all, Leontief emphasized the difference in approach. While the larger part of his (1929) paper was devoted to “a lengthy analysis of the economic forces concealed behind the smooth shapes of Marshallian supply and demand curves, and leads to the formulation of a few fundamental propositions concerning the economic properties of supply and demand relations”, Frisch, instead, in the very beginning of his essay, “reproduces the two fundamental demand and supply equations, but from this point on never mentions their economic significance” (pp. 355-356). With respect to the three conditions for which Leontief’s method is applicable, Leontief responded they are

essentially identical with the fundamental properties of the supply and demand relations which I have derived from a detailed discussion of the economic aspect of the problem. Being interested only in this side of the problem, I had no reason whatever to discuss any mathematical set-ups other than those which do strictly fulfill the economic premises. This means that in the larger part of his criticism, devoted to an elaborate

---

<sup>7</sup> This does not mean that Leontief’s method was not already been discussed in this journal. Elizabeth Waterman Gilboy (1931) gives a critical review of both Schultz’s and Leontief’s method.

<sup>8</sup> In the meantime, in 1931 Leontief had moved to the United States to join staff of the National Bureau of Economic Research, but after only a few months accepted an appointment at Harvard University, where he remained for the following 44 years.

analysis of different mathematical configurations which do *not* comply with the fundamental economic assumptions, Professor Frisch is tilting at windmills. (Leontief 1934a: 357)

When discussing these three conditions, Leontief could only but agree with the third one: if there exists an interdependence between both shifts “then all the theoretically possible price-quantity combinations necessarily have the tendency to be distributed along a definite single curve. This, however, would be neither the supply nor the demand curve” (p. 358).

Although to Frisch the simultaneous satisfaction of the first two conditions would be a “miracle”, Leontief instead claims that it is a “direct mathematical necessity” (p. 359). If the two equations (4a) and (4b) are fulfilled, we may write

$$-(\alpha + \beta)r_1l_1 + l_1^2 = -(\alpha + \beta)r_2l_2 + l_2^2 \quad (6)$$

According to Leontief, it is “evident” that if  $l_1 = l_2$  it follows that  $r_1 = r_2$  and, on the other hand, if  $r_1 \neq r_2$ ,  $l_1 \neq l_2$ . This relation holds so long as  $\alpha + \beta \neq 0$ . If  $\alpha + \beta = 0$ , then  $l_1 = l_2$ . In other words, equation (6) shows that the “significance” of the (in)equality of  $r_1$  and  $r_2$  is related to the “significance” of the (in)equality of  $l_1$  and  $l_2$ : “any discrepancy in the fulfillment of Professor Frisch’s two conditions is mathematically impossible” (p. 360).

Frisch’s rebound “More Pittfalls” (1934) consists of two parts. The first part about whether the assumption of uncorrelated shifts is admissible showed how this debate increasingly displayed signs of Kuhnian linguistic incommensurability. According to Frisch, “the nature of the data at hand” will

contradict the assumption of independent shifts, and independent shifts is only one of Leontief's theoretical premises: "Apparently he is not aware of the possibility that there may be things in his data that contradict such premises" (p. 750). Leontief's mentioning that he is only interested in those situations "which do strictly fulfill the economic premises" (see Leontief's quotation above), was considered by Frisch as "a form of abstractism which, fortunately, is very seldom encountered." (p. 750).

The features of the directly observable ( $xp$ ) distribution he calls "assumptions" and "mathematical set-ups", while his own abstract assumption about the independence of the shifts (which are not directly observable) he calls the "economic aspect". (Frisch 1934: 750)

By first discussing extensively and detailed how various economics factors are related to each other, Leontief had arrived at his premise of independent shifts, and for this specific situation he had developed his estimation method. Frisch approach is clearly different; he starts with a mathematical framework and discusses for all possible situations which ones are realistic, and subsequently assess for which realistic situation Leontief's method is appropriate. 'Direct observation' had a different meaning to both men. To Frisch it referred to statistical properties of the economic variables, to Leontief it referred to structural features of an economic system: how economics factors were connected. Frisch's approach is formal in its treatment, using statistical models and the statistical categories of correlation and variance

relations to investigate Leontief's claims, while Leontief uses economic observations to justify his theoretical assumptions.

After this linguistic confusion, the second part focuses on the mathematical relationship between  $r_1$ ,  $r_2$ ,  $l_1$  and  $l_2$ . First, Frisch discusses Leontief's "discovery" that  $l_1 = l_2$  entails  $r_1 = r_2$ . Therefore, he first rewrote equation (6) in the form

$$l_1^2 - l_2^2 = (\alpha + \beta) (r_1 l_1 - r_2 l_2) \quad (7)$$

Frisch agrees with Leontief that if  $l_1 = l_2$  and  $\alpha + \beta \neq 0$  we must have  $r_1 = r_2$  (since of course we may assume  $l_1 \neq 0$  and  $l_2 \neq 0$ ). In other words  $r_1 \neq r_2$  entails  $l_1 \neq l_2$ , when  $\alpha + \beta \neq 0$ . As a result, for the situation for which  $r_1 \neq r_2$  and  $l_1 = l_2$ , there are two possibilities: (1) if we assume uncorrelated shifts,  $\alpha = -\beta$ ; or (2) the shifts are correlated. This means that we can only have uncorrelated shifts,  $r_1 \neq r_2$  and  $l_1 = l_2$ , when  $\alpha = -\beta$ . This situation was also mentioned in his first (1933) Pitfalls article.

This also means that, when shifts are uncorrelated,  $r_1 = r_2$  does not entail that  $l_1 = l_2$ , as Leontief, according to Frisch, seems to believe. By putting  $r_1 = r_2 = r$  in equation (7) we have:

$$l_1^2 - l_2^2 = (\alpha + \beta) (l_1 - l_2)r$$

This means that we have  $l_1 = l_2$  or  $l_1 + l_2 = (\alpha + \beta)r$ . This shows that we may have  $r_1 = r_2$  and  $l_1 \neq l_2$ .

In his "final word", Leontief (1934b) responded to both parts of Frisch's critique. With respect to the first part related to the assumption of independent shifts, Leontief claimed again that Frisch not really challenged that part of his

analysis: Frisch expressed “with great emphasis the opinion that ‘the nature’ of the empirical data must frequently contradict my fundamental assumption but neglects to specify a single instance of such a contradiction. And it still seems to me impossible that he should do so” (p. 756). Moreover, Leontief assumed that without this assumption the determination of the corresponding supply and demand curves would be impossible.

All the existing methods of statistical supply and demand analysis, whatever the specific character of each of them may be, are based on the Marshallian theory of particular equilibrium. The fundamental postulate of this theory is the independence of supply and demand factors. Any other assumption concerning this point would constitute an open contradiction in objecto. (Leontief 1934b: 757)

About the critique of part two, Leontief could only but say that Frisch failed to reproduce the statements in which he claimed that  $r_1 = r_2$  entails  $l_1 = l_2$ . “I find myself unable to make it more clear that I was not and am not ‘clearly under the impression’ that ‘ $r_1 = r_2$  entails  $l_1 = l_2$ ’” (p. 757).

The acrimony of the debate shows that Leontief and Frisch could not agree and “it was left to Marschak to try and calm the debate and suggest compromises” (Morgan 1990: 187). Whoever made this decision, Jakob<sup>9</sup> Marschak seemed to be the right mediator. Leontief and Marschak shared a similar background. Like Leontief, Marschak was born in Russia, and worked from 1928 to 1930 at the ASTWIK of the Institute for World Economics and Sea Traffic (Beckmann 2000). He was forced to flee Germany in 1933, and

---

<sup>9</sup> After 1933, he turned his name into Jacob.

became director of the Oxford Institute of Economics and Statistics in 1935. After his move to the USA in 1939, he met his earlier ASTWIK colleagues, Adolf Löwe, Gerhard Colm, Hans Neisser and Alfred Kähler, again at the New School for Social Research. He directed the Cowles Commission in 1943-1948, “the period of its intense and influential work on theoretical econometric problems” (Morgan 1990: 153n).

Marschak was already familiar with Leontief’s statistical work on supply and demand elasticities. In his *Elastizität der Nachfrage: Zur empirischen Feststellung relativer Marktkonstanten durch Beobachtung von Haushalt, Betrieb and Markt* (1931; Elasticity of demand: On the empirical determination of relative market constants by observing household, business and market), he discusses Leontief’s 1929 article.<sup>10</sup>

Comparing the “complicated” mathematical derivation of the method by Schmidt in the appendix of Leontief (1929) with Frisch’s exposition of Leontief’s method, one cannot but agree with Marschak that Frisch “succeeded in giving to this method a elementary mathematical exposition which is considerably simpler and at the same time more general” (p. 760). It is therefore that Marschak took Frisch’s “simple way of exposition” to discuss Frisch’s criticisms and Leontief’s replies. Marschak’s exposition clarified the key issues of this debate by making the method’s assumptions more explicit. He arrived at five assumptions necessary to apply Leontief’s method:

---

<sup>10</sup> In the preface of this book, Marschak thanks Colm, Leontief, Löwe and Neisser of the Institut für Weltwirtschaft in Kiel, who saw this work in its different versions, for their encouragements.

Assumption I: Assume the elasticities  $\alpha$  and  $\beta$  of the demand and supply curve respectively to be constant all along the curves.

Assumption II: Assume the elasticities  $\alpha$  and  $\beta$  of the demand and supply curve respectively to be constant over time.

Assumption III: Assume that demand shifts are noncorrelated with supply shifts:  $\rho_1 = \rho_2 = 0$ .

Assumption IV: The price-quantity-correlations  $r_1$  and  $r_2$  must be “significantly different” in both materials.

Assumption V: The relative volences  $l_1$  and  $l_2$  must also be “significantly different” in both materials.

And the discussion of Frisch and Leontief about the relation between  $r_1$ ,  $r_2$ ,  $l_1$  and  $l_2$  (7), shows that Assumption IV and Assumption V are independent and so both are needed.

To meet Assumption III, Marschak suggested that one could adjust “the material so as to eliminate some of the more obvious causes which influence simultaneously both supply and demand (population, price level)” (p. 762).

And to meet the last two assumptions, “why not select *deliberately* such ‘two materials’ as would reasonably satisfy the Assumption IV and V, instead of splitting the material *arbitrarily*?” (p. 762). In other words, according to Marschak, “the criticism of R. Frisch thus could be so far adequately met with” (p. 763).

Marschak considered, however, Assumption II as the vulnerable point of Leontief’s method. It is here that Marschak directed the focus of the discussion

to an equally important part of the Pitfalls debate, namely the nature of the “economic premises”.

At this point the main economic reasoning of W. Leontief becomes important. As W. Leontief rightly observes in his reply, this part of his work has been somewhat neglected in R. Frisch’s criticism. W. Leontief claims for the parameters “elasticity” and “level” more than a purely mathematical significance. (Marschak 1934: 764)

### **3. Autonomy**

The Pitfalls debate shows that Leontief worried much more about invariance and independence and Frisch more about correlations and spurious results. It is only two years later that Frisch started to take publicly note of this problem of invariance and independence, and again two years before he gave this requirement of independence and invariance a name: ‘autonomy’. At the 1936 Oxford meeting of the Econometric Society<sup>11</sup>, a student of Frisch, Trygve Haavelmo, presented a paper on ‘Confluent relations as a means of connecting a macrodynamic subsystem with the total system’. In a discussion following Haavelmo’s presentation of his paper, as a reply to a question by Marschak about the distinction between ‘structural’ and ‘confluent’ relations, Frisch answered by expounding the meaning of autonomy but without introducing the term (Aldrich 1989: 22; Bjerkholt 2005: 510): “any coefficient in a structural

---

<sup>11</sup> The Econometric Society was founded in 1930, at the initiative of Irving Fisher (the Society's first president) and Ragnar Frisch. The first meetings of the Society were held in September, 1931, at the University of Lausanne, Switzerland, and in December, 1931, in Washington D.C.

relation might be changed *institutionally* without necessarily entailing a change in the other structural relations” (Phelps Brown 1937: 374). Haavelmo’s paper was published in the July 1938 issue of *Econometrica*, where he gave a more precise account of the difference between structural and confluent relations: A general linear dynamic relation between the  $x_i$  variables is a structural relation if the (lagged) variables in the right-hand side of the equation “may take on different values *independent of each other*, giving *alternatively* different values to  $x_i$  [...]. The reason for the existence of such structural relations may be certain technical conditions, psychological laws, etc” (Haavelmo 1938: 203).

Frisch (1938) introduced the term ‘autonomy’ for the first time in his famous Autonomy Memorandum written to discuss Jan Tinbergen’s work for the League of Nations (see section 5). Autonomous equations were the equations that “maintained unaltered while other features of the structure were changed” (p. 17).

The higher this degree of autonomy, the more *fundamental* is the equation, the deeper is the insight which it gives us into the way in which the system functions, in short, the nearer it comes to being a *real explanation*. Such relations form the essence of “theory”. (Frisch 1938: 17)

Autonomy, as Frisch admitted, is not a “mathematical property of a closed system”, but is built on “some sort of knowledge outside this system” (p. 15). Moreover, statistics only leads to “confluent equations”, and generally speaking, these relations are far from able to give information about the

autonomous structural relations. Therefore, it is necessary to employ experiments, as Frisch recommended, or its substitute in economics, the “interview method”. This problem was expressed more explicitly and publicly in a paper published in *The American Economic Review*:

It is very seldom indeed that we have a clear case where the statistical data can actually determine numerically an autonomous structural equation. In most cases we only get a covariational equation with a low degree of autonomy. [...] We must look for some other means of getting information about the numerical character of our structural equations. The only possible way seems to be to utilize to a much larger extent than we have done so far the interview method, *i.e.*, we must ask persons or groups what they would do under such and such circumstances. (Frisch 1948: 370)

Frisch had argued in favor of using interview information since 1926 and applied it primarily in his investigations of methods for estimating marginal utility. Frisch, however, claimed that he had developed such a method already in 1922 (“and have since tried it out occasionally on friends”) as an alternative to “a lack of reliable statistical data” (Frisch 1932: 140).

The term autonomy was discussed more elaborately in Haavelmo’s (1944) Probability Approach paper when addressing the problem of “judging the degree of persistence over time of relations between economic variables”, or more generally speaking, “whether or not we might hope to find elements of invariance in economic life, upon which to establish permanent ‘laws’” (p. 13).

Like Frisch, Haavelmo understood that for dealing with this problem, statistics is not sufficient, it is “a problem of actually *knowing something* about real phenomena, and of making realistic assumptions about them” (p. 29):

The construction of systems of autonomous relations is, therefore, a matter of intuition and factual knowledge; it is an art. (Haavelmo 1944: 29)

Nevertheless, to turn art into science, Haavelmo suggested “to *find* such a basic system of highly autonomous relations in an actual case [...] it is a task of making fruitful hypotheses as to how reality actually is” (p. 31), and subsequently testing these hypotheses. Therefore he introduced to econometrics Neyman and Pearson’s theory of statistical testing, for which this paper became to be renowned.

Haavelmo’s suggested method for econometrics was adopted by the researchers at the Cowles Commission<sup>12</sup>. The Cowles Commission view became to be that to understand a particular aspect of economic behavior, it is necessary to have a system of descriptive equations. These equations should contain relevant observable variables, be of a known form (preferably linear), and have estimatable coefficients. The Cowles Commission programme aimed to provide an appropriate method to choose the variables relevant to a particular problem so as to obtain a suitable system of equations and estimate the value of the parameters. However, “little attention was given how to choose

---

<sup>12</sup> The Cowles Commission for Research in Economics was set up in 1932 to undertake econometric research. The journal *Econometrica*, in which Haavelmo’s paper appeared, was run from the Cowles Commission.

the variables and the form of the equations; it was thought that economic theory would provide this information in each case” (Christ 1994: 33).

#### **4. Ask the engineer**

While the Cowles Commission took the direction of relying on theory to solve the problem of autonomy, Leontief took another route. His programme was made most explicit in his Presidential address delivered at the meeting of The American Economic Association in 1970:

In contrast to most physical sciences, we study a system that [is] in a state of constant flux. I have in mind [...] the basic structural relationships described by the form and the parameters of these equations. In order to know what the shape of these structural relationships actually are at any given time, we have to keep them under continuous surveillance. By sinking the foundations of our analytical system deeper and deeper, by reducing, for example, cost functions to production functions and the production functions to some still more basic relationships eventually capable of explaining the technological change itself, we should be able to reduce this drift. It would, nevertheless, be quite unrealistic to expect to reach, in this way, the bedrock of invariant structural relationships (measurable parameters) which, once having been observed and described, could be used year after year, decade after decade, without revisions based on repeated observation. (Leontief 1971: 3-4)

Besides emphasizing that this requires a “steady flow of new data”, it also shows the need to look beyond the traditional domain of economic phenomena:

The pursuit of a more fundamental understanding of the process of production inevitably leads into the area of engineering sciences. To penetrate below the skin-thin surface of conventional consumption to develop a systematic study of the structural characteristics and of the functioning of households, an area in which description and analysis of social, anthropological and demographic factors must obviously occupy the center of the stage. (Leontief 1971: 4)

To obtain this new kind of information, direct observation was considered to be more appropriate than what he called “indirect statistical inference”:

Establishment of systematic cooperative relationships across the traditional frontiers now separating economics from these adjoining fields is hampered by the sense of self-sufficiency resulting from what I have already characterized as undue reliance on indirect statistical inference as the principal method of empirical research. (Leontief 1971: 4)

Indirect statistical inference would just be “circular”, not widening and deepening the empirical foundations of economic analysis, because we then construct models in which prices, outputs, rates of saving and investment are explained in terms of production functions, consumption functions and other

structural relationships; but to measure the parameters of these relationships we use the magnitudes of these prices, outputs and other variables. According to Leontief it would be much better to use “exogenous information”, to cross “the conventional line separating ours from the adjoining fields” (p. 5).

An alternative and more direct way of determining, for example, the amount of coke required to produce a ton of pig iron or the amount of corn feed required per hundredweight of live hogs is that of asking the ironmaster in the first and a specialist in animal husbandry in the second case. As a matter of fact one can easily visualize the possibility of assembling a complete set of input coefficients describing the structural characteristics of all branches of the national economy entirely on the basis of such direct information without recourse to any actual statistical input or output figures. (Leontief 1949: 213)

Leontief considered what he called “statistical econometrics” developed by the Cowles Commission as providing only “indirect inference”. His proposed approach of input-output analysis was based on “direct observation”. This latter term should be taken almost literally.<sup>13</sup> Since the 1930s at Harvard, to collect data necessary to fill the input-output table, Leontief, with his assistants, wrote letters, phoned, asked engineers, firms, statistical bureaus and unions in order to get these data:

When I constructed the first input-output table, which was very early, I often used the telephone. I called up industries, particularly firms which

---

<sup>13</sup> According to Silk (1976: 160), having interviewed Leontief, it meant “to be gained with eyes and nose and hands and ears and measuring instruments of all types”.

were engaged in the distribution of commodities, and got the data from them. (Leontief interviewed by Foley 1998: 121)

It should be noted here that not just anyone was asked for their observations, they only asked engineers, technicians and other experts on a relevant sector or component of the economic system, like the ironmaster and the specialist in animal husbandry as mentioned in the quotation above.

In analyzing the changing structure of the steel industry, we must get our information from the technical literature, from ironmasters and from rolling mill managers. To study the changing pattern of consumer behavior, we have to develop practical co-operation with psychologists and sociologists. (Leontief 1949: 225)

This was also worded by Leontief in about similar terms at the first conference on input-output systems held in the Netherlands:

Such empirical description requires many months of works by a large staff of experienced economic statisticians and experts intimately acquainted with the various branches of manufacturing, mining, agriculture, transportation, etc. (Leontief 1953: 7-8)

But it makes no sense to ask these experts for information on a too abstract level, or too high level of aggregation. For example, a theoretical aggregate production function, intended to describe the relationship between, say, the amount of steel produced,  $y_1$ , and the quantities of two different inputs,  $y_2$  and  $y_3$ , needed to produce it, is typically described as a CES function:

$$y_1^{\rho} = ay_2^{\rho} + (1-a)y_3^{\rho}$$

To ask a manager of a steel plant or a metallurgical expert for information on the magnitude of the [...] parameters appearing in the [equation] would make no sense. Hence, while the labels attached to symbolic variables and parameters of the theoretical equations tend to suggest that they could be identified with those directly observable in the real world, any attempt to do so is bound to fail: the problem of “identification” of aggregative equations after they have been reduced – that is, transformed, as they often are – for purposes of the curve-fitting process, was raised many years ago but still has not found a satisfactory solution. (Leontief 1982: 104)

Leontief’s attempts to achieve the above described aims for empirical work can be characterized by two pervasive concerns: his disapproval of aggregate variables and his emphasis on enlarging the primary data base for economic analysis with engineering and technical data (Carter and Petri 1989: 17)

You see, I was some what skeptical of the whole curve-fitting notion. I thought of technological information. The people who know the structure of the economy are not statisticians but technologists, but of course to model technological information is very difficult. My idea was not to infer the structure indirectly from econometric or statistical techniques, but to go directly to technological and engineering sources. (Leontief interviewed by Foley 1998: 123)

As we have seen above, Frisch, Leontief's initial fierce opponent in the Pitfalls debate and most outspoken advocate of 'statistical econometrics', moved gradually in the direction of Leontief, while the institutions he created and his students and sympathizers put the probabilistic approach in the centre of the econometric program. This is probably what Leontief meant in an interview by DeBresson (2004: 139): "we had our disagreements, but in retrospect our views are quite close"?

On various occasions Frisch claimed that he had, prior to Leontief, invented the principles of input-output analysis in an article he published in 1934 in *Econometrica* entitled 'Circulation planning: proposal for a national organization of a commodity and service exchange'. Bjerkholt and Knell (2006) refute this claim by extensively discussing the (only formal) similarity and many differences between both approaches. They, however, point at a closer and more interesting similarity:

His promotion of methods for getting behavioral information directly from economic agents by interview method, is matched by another idea he also pursued, namely to gather production information directly as engineering data. Frisch pioneered the estimation of 'engineering production functions' in Frisch (1935). (Bjerkholt and Knell 2006: 406-407)

Bjerkholt and Knell (2006) refer to Frisch's paper, 'The principle of substitution', in which Frisch gives a presentation of how the principle of substitution works in practice, to be more precise for the Freia Chocolate

Fabrik, Oslo, with data provided by the Managing Director and the Chief-Engineer.

Many of the quantitative data which are being utilized in this work are in themselves not novel. Some of them have, to a smaller or larger extent, already been utilized a long time by engineers and cost accountants. The particular way in which they are now being utilized is, however, novel. By being interpreted in the light of modern economic theory they receive a new significance and throw new light on the many problems which are of special concern to the industry or to the firm and also on problems which are of a much more general economic interest. On the whole this is a field of study in which the engineer, the cost accountant and the economist have much to learn from each other. (Frisch 1935: 12).

## **5. The number two is from Keynes**

In 1969, Ragnar Frisch received the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel together with another pioneer of econometrics, Jan Tinbergen. Tinbergen was the first to succeed in modeling a real economy on the basis of the then new econometrical techniques. In 1936, he presented his very first macro-econometric model of the Dutch economy to the Dutch Society of Economics and Statistics. The paper was read and published in Dutch, but in the same year Tinbergen was commissioned by the League of Nations to perform statistical tests on business-cycle theories. The

results were published in a two-volume work, *Statistical Testing of Business-Cycle Theories* (1939a, 1939b). The first contained an explanation of this new method of econometric testing as well as a demonstration based on three case studies to show what could be achieved. The second volume developed a model of the United States, the second macro-econometric model in the history of economics.

To explain and justify this new econometric method, Tinbergen wrote several reports on his work at the *League*. They provide us with an explicit account of what early modeling practice entailed. The ‘method’ Tinbergen employed to understand the causation of business-cycle phenomena “essentially starts with *a priori* considerations about what explanatory variables are to be included. This choice must be based on economic theory or common sense” (Tinbergen 1939b, 10). Tinbergen was quite aware of the fact that economists did not agree upon which were the most important causes of the business-cycle phenomenon. From his PhD supervisor and mentor Paul Ehrenfest he had learned that:

to formulate differences of opinion in a ‘nobler’ way than merely as conflicts. His favourite formulation was cast in the general form: if  $a > b$ , scholar A is right, but if  $a < b$ , then scholar B is right. The statement applied to a well-defined problem, and both  $a$  and  $b$  would generally be sets of values of elements relevant to the problem treated, with possibly a number of components of qualitative nature. (Tinbergen 1988: 67)

This method was exactly the method he would adopt in his work for the *League of Nations*:

It is rather rare that of two opinions only one is correct, the other wrong.

In most cases both form part of the truth [...] The two opinions, as a rule, do not exclude each other. Then the question arises in what ‘degree each is correct’; or, how these two opinions have to be ‘combined’ to have the best picture of reality.

[We can] combine these different views, viz. by assuming that the movements [...] can be explained by some *mathematical function* of all the variables mentioned. We then have not a combination in the physical sense – an addition of two quantities or of two amounts – but a combination of influences. In many cases the mathematical function just mentioned may be approximated by a linear expression. (Tinbergen 1936, 1-3)

The equations that were chosen were linear with parameters that remain constant over time. The values of the parameters were found by multiple regression analysis.

This view on models and on how to deal with differences of opinions would never leave him. He considered it as a scientific responsibility to synthesize: “If we, economists, continue to oppose each other, we fail in our duty as scientists” (Tinbergen 2003: 303). In a 1982 article ‘The Need of a Synthesis’, Tinbergen repeated his lifelong credo:

In quantitative terms one can also say that in certain regression equations some of the coefficients indicate what the weights are of the explanatory variables, as these are put forward in the competing theories, in the explanation of the independent variable. In the search for a synthesis what matters is that, as has been stated by Klein: “It is less important that the effort be labeled Keynesian, monetarist, neoclassical or anything else, than that we get good approximation to explanation of this complete system ...”.

Indeed, the criterion by which we test the various competing views, has to be in the best possible explanation of the developments observed [...].

The point with which I want to end this argument is that the synthesis is only completed when such partial studies – the usefulness of which I accept fully – are made part of a complete model. The reason for that I gave earlier already: consistency with the other ‘blocs’ of a complete model.

That is why, we cannot do without our largest model factory, the Central Planning Bureau, in establishing the synthesis intended.

(Tinbergen 2003: 303, 305-306)

According to him, models are “an order of thinking. They make it possible to localize differences of opinion: to indicate the equation about one disagrees, the term of that equation, or the term that is lacking, or the variable that is lacking” (Tinbergen 1987a: 106, trans by the author).

Tinbergen's League of Nations study proved highly controversial. It was circulated in 1938 prior to publication, and provoked interesting discussions about the role of econometrics in theory testing. It was John Maynard Keynes' (1939) critique of Tinbergen's first volume, 'Professor Tinbergen's Method,' that sparked off the debate about the role of econometrics and what might be achieved by it. Keynes' attack, with his usual rhetorical flourish, was such that Morgan (1990: 121) concluded that "he had clearly not read the volume with any great care" and revealed ignorance about the technical aspects of econometrics (In his reply to Keynes, when discussing the method by which trends are eliminated, Tinbergen (1940: 151) even remarked: "Mr. Keynes does not seem to be well informed. ... A glance ... at any elementary text-book on these matters could have helped him").

It is remarkable, however, that in various interviews, whenever Tinbergen was asked to give his view on this debate, he provides two standard answers. One is that as result of this debate he received from Keynes a free lifelong subscription of *The Economic Journal*, the journal which published the debate and of which Keynes was the editor. And the other is about a conversation they had in 1946. They only met once, but this conversation shows that they would never move closer to each other, in spite of Tinbergen's attempt to synthesize:

Indeed, I did feel that, at least on certain points, he [Keynes] was badly informed. The best illustration is that he thought that a trend was determined by connecting the first and the last observation. It was a bit

strange to me because he had written the *Treatise of Probability*, so I thought he was somewhat familiar with statistics.

At first I was a bit disappointed, because I thought that he would be especially happy with my work, since we had very largely followed his main macro-theories. But all that seemed not to impress him very much. I had the privilege of meeting him later, just once in 1946. On that occasion I told him we had done quite a bit of research on the price elasticity of exports and that we had really found that the elasticity is about 2, the figure that he uses in his famous book about German reparation payments. I thought that he would be very glad that we had found that figure, and “that he had been right”. But he only said: “How nice for you that you found the right figure”. That was a most funny experience. (Magnus and Morgan 1987: 129-130)

In the Appendix, other versions are given. These are oral histories and should be treated differently than other sources of historical evidence (Mata 2008). They fulfill a specific role. Tinbergen’s stories about the number two is an epistemological parable, telling the only exceptional case one can rely on intuition, instead of measurement, namely in the case that it is the intuition of a genius<sup>14</sup>:

---

<sup>14</sup> This parable is rather similar to one about another 20th century genius, Albert Einstein. Like Tinbergen’s parable, it has different versions, but as an example I take the one from *Time* February 19, 1979, ‘The Year of Dr. Einstein’:

General relativity indicated that when light from a distant star passes very close to the sun on its way to earth, it should be deflected by solar gravity, thereby shifting the star’s position in the sky. The amount of shift, Einstein calculated, should be 1.75 seconds of arc—a small variation, but one discernible by astronomers of the day. But how could astronomers photograph a star nearly in line with the sun when it would certainly be obscured by sunlight? Answer: during a total eclipse. On May 29, 1919, during an eclipse expedition to the island of Principe off the West African coast, the British astronomer Arthur Eddington found

Sometimes, indeed, intuition constitutes a basis for new scientific results. I should be the intuition of a genius, however. For simpler souls, intuition may be less reliable! (Tinbergen 1988)

To understand Keynes's response to Tinbergen's announcement that he had verified the number 2, it is helpful to look at the passage in the book where this number is produced.<sup>15</sup>

Let us put our guess as high as we can without being foolish, and suppose that after a time Germany will be able, in spite of the reduction of her resources, her facilities, her markets, and her productive power, to increase her exports and diminish her imports so as to improve her trade balance altogether by \$500,000,000 annually, measured in pre-war prices. This adjustment is first required to liquidate the adverse trade balance, which in the five years before the war averaged \$370,000,000; but we will assume that after allowing for this, she is left with a favorable trade balance of \$250,000,000 a year. Doubling this to allow for the rise in pre-war prices, we have a figure of \$500,000,000. (Keynes 1920: 105-6)

This number 2 appeared in the first macro-model of the Centraal Plan Bureau (Central Planning Bureau), in 1955 (this model, therefore, became to be known as the '1955 model'). Tinbergen was the first director of this new Netherlands Bureau for Economic Policy Analysis, as its more appropriate

---

deflections in starlight that almost matched Einstein's prediction. Later, when Einstein was asked what he would have concluded if no bending had been detected, he replied: "Then I would have been sorry for the dear Lord—the theory is correct."

<sup>15</sup> I am not really sure whether this is the spot Tinbergen is referring to, direct references are never given, but this is the only part of the text where export and prices are discussed in relation to each other.

name in English is.<sup>16</sup> The Bureau was never used for planning. At CPB, the export elasticity became to be called ‘Tinbergen Two’ (Theeuwes 1987; see also Don and Verbruggen 2006: 166), reason for Tinbergen (1987) to write a note, ‘The number two is from Keynes’, to give Keynes the honor of being its source (see Appendix).<sup>17</sup>

Today, economic policy analysis at the CPB is still very much in the tradition of its first director:

Many policy measures in the macroeconomic sphere can only be understood and discussed properly with the help of a model which sets out the key relationships between macroeconomic variables. Such a model is an important instrument in considering relevant relationships.

(Don and Verbruggen 2006: 146)

In a panel discussion to explore policy makers’ perspectives on their experiences of the modeling-policy interaction, Henk Don, director of CPB from 1994 to 2006, explains the unique role models have in Dutch policymaking, when compared to other countries:

Perhaps the most important one is to use models as an information processing device: to monitor the economy, to monitor the budget outlook in particular, and to provide information about different scenarios that the near future might bring.

---

<sup>16</sup> CPB was founded in September 1945.

<sup>17</sup> The Dutch title of this note is ‘Het getal twee is van Keynes’ in which the preposition ‘van’ has the double meaning of being ‘of’ (property indication) and ‘from’ (reference).

In Dutch policy-making there is still another use of economic models, which is to use it as a tool in consensus building. [...] Using the model may help to locate exactly where the political differences are and whether these are differences of preferences in what people would like the economy to produce or whether these are difference in analysis of what the economic trade-offs really are. The model helps very much in assessing all these difference and in getting as much common ground as you can get. (Don quoted in Morgan 2000: 264-5)

## **6. Monetary Policy Committee**

The Dutch case of Tinbergenian policy decision-making where a model frames the discussion and where regression estimates function as arbitrator is quite the opposite of a case of policy decision-making in the UK: the Bank of England's Monetary Policy Committee (MPC) process. According to Downward and Mearman (2008), who have investigated this process and suggest to call it 'triangulation', decision-making by the MPC neither follows procedural rules (section 7), nor relies on the strict applications of econometric techniques of analysis (section 5), but "appears to reflect a more pragmatic and pluralist approach that draws upon a variety of sources of argument and evidence" (p. 385). Moreover when it comes to empirical evidence, it will appear that this rather similar to the kind of observations Leontief took as evidence (section 4).

In 1997, the Chancellor of the Exchequer, Gordon Brown, gave operational independence to the Bank of England. The Bank of England Act (1998) legislated that the Bank was to pursue price stability, and subject to that, to support government economic policy. The Bank's main policy instrument is the interest rate. In setting the appropriate interest rate, the Bank utilizes the for that reason established MPC, comprising the Governor, two Deputy Governors and six other members. The decisions are made by a vote of the MPC, with each member having one vote. Two internal members take management responsibility for monetary policy and market operations, respectively. They are appointed by the Governor, after consultation with the Chancellor, for three year terms. The remaining four members will be appointed by the Chancellor, for three year terms. "They will be recognised experts" (Brown quoted in Rodgers 1997: 244). These external members are typically academic or professional economists.

To build up credibility to influence public expectations, the MPC's policy is being as open and transparent as possible about the way it comes to its decisions. For this reason, the *Bank of England Quarterly Bulletin* published an article by Lambert (2005) to explain "what life is like as a member of the MPC". This article gives a coloured, but even so nice account about the role and character of evidence being used in the decision-making process. The monthly meetings to set interest rates take place on the Wednesday and Thursday following the first Monday of every month, and a pre-MPC meeting is held on the preceding Friday. The idea of the pre-MPC meeting is to draw

out all the important economic news of the previous month and to put it into context. All MPC members attend, so that they can prepare for the following week's policy meeting on an equal footing. Throughout the weeks since the previous meetings, MPC members have received scores of emails from the Bank's staff, analyzing the latest economic news from around the world. They also have been sent studies from Bank analysts on topical issues: the outlook for growth and employment; what's happening to wages; growth in the euro area and so on. "On the Thursday night before pre-MPC, they will have received a pack of around 500 charts and tables which are updated every month to give a consistent picture of the economic world" (p. 57).

It is early on a Friday morning, and the Bank's economists are gathering for coffee along with the members of the MPC and some of the Bank's regional Agents from around the United Kingdom for the big briefing session which is known as the pre-MPC meeting. In all, there may be as many as 100 people in the room. [...] The meeting takes the form of a series of set-piece presentations by senior Bank staff, each illustrated by dozens of graphics which are projected on to large screens around the room. Each presentation covers a different aspect of the economic landscape, building up a broad picture of the big economic and financial developments over the previous month and concentrating on those elements which are most important to the UK economy. For a newcomer to the MPC, this is often the first exposure to the scale and quality of the Bank's economic engine room. Graphics fly across the screens.

Occasional questions by Committee members are fielded by the presenter or, if he or she wishes to bring in a colleague, a turn of the presenter's head will bring in a swift response from the back row.

(Lambert 2005: 57)

The pre-MPC meeting provides the Bank's regional Agents with an opportunity to report on what they have picked up in recent weeks from their business contacts around the country. There are twelve Agents in all, each covering an area, who are in contact with a total of roughly 8,000 businesses large and small. The Agents usually have two slots at the Friday meeting. In one, they give an overview of their discussions with hundreds of business people over the previous month. Key regional and sector differences are highlighted, and comparisons are drawn with what the official data are showing. In the other, the Agents report back on the special survey which they undertake most months at the request of the Committee. At the end of each rate-setting meeting, the MPC identifies a topical issue about which it would like to learn more. The Agents discuss the month's topic with their business contacts in the next few weeks, and report their findings back to the pre-MPC meeting.

These Agents are the eyes and ears of the Bank. Willem Buiter, a former MPC member, once remarked: "In the years I was there, when there was a lot of smoke, any kind of ground troops, such as the agents, were helpful" (Financial Times, January 12, 2007). The role of each Agency<sup>18</sup> is to maintain

---

<sup>18</sup> Each Agency consists of an Agent, one or two Deputy Agents and up to two additional team members providing support (Beverly 1997: 424)

contact with industry and commerce and to report on the economy as seen by those based in its area:

The Agencies carry out most of their economic intelligence work through face-to-face contact with individual companies, who know that the confidentiality of sensitive information will be protected. [...] Every month each Agency makes direct contact with around 50 firms, either through company visits or at various functions, in addition to numerous telephone calls. Agencies maintain standing panels of business people whose views about the economy are regularly canvassed. These can serve as control groups, whose discussions help to identify changes in trends. (Beverly 1997: 425)

For the Bank, the distinctive contribution that the Agencies make comes from their first-hand contact with a wide range of business people. These contacts provide the Bank with a regular flow of up-to-date economic news that complements the published statistics. “The Agencies are able to pick up developments of local significance and, by comparing these local reports, Monetary Analysis is able to form a balanced picture of what is happening in the economy as a whole” (Beverly 1997: 427).

But also the MPC members spend time travelling around the United Kingdom, visiting companies and talking to business people and others about how the economy is doing. These visits are set up by the Bank’s regional Agents. “In all, about 60 or so of them are undertaken in any given year” (Lambert 2005: 63). These visits serve two purposes:

The first is to find out what is actually going on at the coal face of economic life. You can read all the data in the world and still not fully understand the issues which are keeping business people awake at night. The second is to explain why the Committee has acted in the way that it has, and to give people a chance to express their views about the way that monetary policy is being managed. To build their credibility, its members need to get out and about and pay serious attention to the views of as wide a slice of the community as possible. (Lambert 2005: 63-64)

Being fully updated by the Bank staff, the Wednesday MPC meeting is spend to allow the MPC members “talk about what has happened in the previous month that might have changed their views about the outlook for inflation. They discuss the latest economic data and business surveys, and they report what they have heard from their own contacts about business conditions around the country” (Lambert 2005: 59). At the Wednesday meeting it is not allowed for actually discussing the interest rate decision itself. At the Thursday meeting each member is given 10 minutes to highlight the issues that have been thought most relevant in the previous weeks and explaining the thinking that has led the decision on the rate, which is then announced.

The decision goes to the majority and there is no attempt to arrive at a consensus: members are individually accountable for their decisions. (Lambert 2005: 59)

The MPC is an “individualistic committee” and is not required to reach a consensual decision but a conclusion of the majority (Downward and Mearman 2008: 391).

## **7. Procedures for structured expert judgment**

There is a field in which you wish not to build up a historical record to quantify a decision model, namely risk modeling of consequences of nuclear accidents, but also those of chemical weapons disposal, nitrogen oxide emissions, and microbiological risk. To build up such risk models, nevertheless, one needs data to quantify these models. By lack of observations, one has to turn to experts for their judgments.

To explore a case of expert observations in between measurement and professional intuition, we now discuss an approach to deal with expertise outside economics, namely engineering. Over the last fifteen years, at Delft University of Technology, Roger M. Cooke has developed procedures to support the formal application of expert judgment. He is the author of *Experts in Uncertainty: Opinions and Subjective Probability in Science* (1991), and since September 2005 appointed to the Chauncey Starr Chair in Risk Analysis of Resources for the Future (RFF).

For the kind of cases Cooke has worked on, expert judgment is used to obtain results from experiments and/or measurements, which are physically possible, but not performable in practice. Such experiments are “out of scale” financially, morally, or physically in terms of time, energy, distance, etc. Since these experiments cannot in fact be performed, experts are uncertain about the

outcomes, and this uncertainty is quantified in a formal expert judgment exercise. In a video interview made at RFF, he expressed this as follows:<sup>19</sup>

... so we never built up historical data to quantify our risk models by the nature of the case. So in building a risk model and quantifying this model we must have recourse to expert judgment, and this has been a theme throughout my work and a lot of risk analysis is directed to this. We really want to look at expert judgment as a new form of scientific data which we can use in a methodological proper way to quantify our models and make this whole process transparent. (Cooke 2009, transcription by the author)

Instead of performing a physical experiment, experts are asked to do a “hypothetical experiment”:

First you have to be very clear what you want to ask, and this is a very difficult part of an expert judgment exercise to formulate a protocol of the questions you exactly ... exactly what you want to know. I like to think of it as follows: that expert judgment is just a different way of doing experiments. And you should have questions that you could in principle ask the nature, but for various reasons practically you cannot do so. So you ask these questions to experts who are familiar with the whole field and who can tell you what they think what will probably happen, and what their uncertainty is on what is going to happen if you could do such experiment. (Cooke 2009, transcription by the author)

---

<sup>19</sup> This video can be watched at <http://www.rff.org/Researchers/Pages/ResearcherSpotlightCooke.aspx>.

If a parameter is uncertain, and if the uncertainty cannot be quantified with historical and/or measurement data, then the analyst must ask the expert how the values *would* be determined if suitable measurements could be performed. Although these experiments are hypothetical, i.e. they cannot be performed in practice, they must be physically possible. The values are known to depend on a large number of physical parameters which cannot all be measured or controlled on any given experiment. Moreover, the functional form of the dependence is not known. Hence, if a controlled experiment is repeated many times, different values will be found reflecting different values of uncontrolled and unknown physical parameters. If a measurement set-up is described to an expert, (s)he can express his/her uncertainty via a subjective distribution over possible outcomes of the measurement. In such cases the experts are questioned directly about uncertainty with respect to model parameters.

So, while Cooke does not define explicitly what he means by hypothetical experiments, the way he discusses them is rather close to the “experiments-in-imagination” as described by Carl Hempel (1965). Hempel makes a distinction between two kinds of experiments-in-imagination: the intuitive and the theoretical. An intuitive experiment-in-imagination is aimed at anticipating the outcome of an experimental procedure which is just imagined, but which may well be capable of being actually performed. Prediction is guided here by past experience concerning particular phenomena and their regularities, and occasionally by belief in certain general principles which are accepted as if they were *a priori* truths. Imaginary experiments of this kind are

intuitive in the sense that the assumptions and data underlying the predictions are not made explicit and indeed may not even enter into the conscious process of anticipation at all: past experience and the belief in certain general principles function here as suggestive guides for imaginative anticipation rather than as a theoretical basis for systematic prediction. The theoretical kind of imaginary experiment, on the other hand, presupposes a set of explicitly stated general principles – such as laws of nature – and it anticipates the outcome of the experiment by deductive or probabilistic inference from those principles in combination with suitable boundary conditions representing the relevant aspects of the imagined experimental situation. Sometimes the latter is not actually realizable. The question what *would* happen *if* is answered here but by rigorous deduction from available theoretical principles. Imagination does not enter here; the experiment is imaginary only in the sense that the situation it refers to is not actually realized and may indeed be technically incapable of realization.

In my view, for theoretical imaginary experiments, one does not need an expert, so I assume that Cooke's usage of thought experiment refers to Hempel's intuitive experiment-in-imagination. It is interesting to note to both Hempel and Cooke emphasize that resource to thought experiments are only legitimate if inevitable: "intuitive experiments-in-imagination are no substitute for the collection of empirical data by actual experimental or observational procedures" (Hempel 1965: 165).

A formal expert judgment exercise is called an ‘elicitation’. And because this exercise is to reveal and quantify the expert’s uncertainties, Cooke considers the preparation for elicitation as “really nothing more than carefully designing these hypothetical experiments, so as to obtain the information that we require” (Cooke and Goossens 1999: 24).

And then we look at experts really as statistical hypotheses. When an expert says he has a certain uncertainty distribution over the range of outcomes of some possible experiments: that is a statistical hypothesis. And that is how we look at it. (Cooke 2009, transcription by the author)

In describing this hypothetical experiment to the expert, the physical factors which may influence the outcome of the experiment are first identified by the analyst. Each relevant physical factor will fall into one of the two classes: (1) The case structure assumptions; and (2) the uncertainty set. Some relevant factors will have their values stipulated by the assumptions of the study, as reflected in the case structure. Other factors may influence the outcome of the hypothetical experiment, but their values are not stipulated by the case structure. These factors belong to the uncertainty set. The expert is made aware that these factors are uncertain, and should fold this uncertainty into their distributions on the outcome of the hypothetical experiment. The general format for elicitation is given as the following figure (Figure 5, Cooke and Goossens 1999: 27):

*Conditional on*  
< values of factors in the case structure assumptions >  
*Please give the 5%, 50% and 95% quantiles of your uncertainty in*  
< Hypothetical experiment >  
*taking into account that values of*  
< uncertainty set >  
*are unknown*

Although it is explicitly noted that expert assessments should not be treated “as if they were physical measurements in the normal sense, which they are not” (p. 10), they are assessed as if they are measuring instruments:

Empirical control is built in the elicitation procedure by asking experts to assess calibration or seed variables. Seed variables are variables whose values are or will be known to the analyst within the frame of the exercise but not to the expert. Seed variables are important for assessing the performance of the combined experts’ assessments. Seed variables also form an important part of the feedback to experts, helping them to gauge their subjective sense of uncertainty against quantitative measure of performance. (Cooke and Goossens 1999: 28).

Calibration and gauging are typical techniques for increasing the reliability of a measuring instrument, but here they are applied the expert judgments: “expert judgement is recognized as just another type of scientific data, and methods are developed for treating it as such” (p. 10).

An “expert for a given subject” is described as a “person whose present or past field contains the subject in question, and who is regarded by others as being one of the more knowledgeable about the subject” (p. 29-30). The following general selection criteria are recommended (p. 30):

- Reputation in the field of interest
- Experimental experience in the field of interest
- Number and quality of publications in the field of interest
- Diversity in background
- Awards received
- Balance of views
- Interest in and availability for the project.

To take decisions in a rational manner, it is not rational to base a decision on the judgment of one single expert, because they are partial and will be favorable to the interest of some stakeholders:

An appeal to ‘impartial’ or ‘disinterested’ experts will fail for two reasons. First, experts have interests; they have jobs, mortgages and professional reputations. Second, even if an expert interest could somehow be quarantined, even then the experts would disagree. (Cooke and Goossens 1999: 15)

As a result, the views of a diverse set of experts must be taken into account. Simply choosing a maximally feasible pool of experts and combining their views by some method of equal representation might achieve a form of “political consensus” among experts involved, but will, according to Cooke and

Goossens, not achieve “rational consensus”. “If expert viewpoints are related to the institutions at which the experts are employed, then numerical representation of viewpoints in the pool may be, and/or may be perceived to be influenced by the size of the interests funding the institutes” (p. 15).

In a survey article discussing fifteen years of research in expert judgement at Delft, other considerations are given for preferring a rational consensus, that is a “mathematical aggregation” above a political consensus, that is “agreement among experts”:

a group of experts tends to perform better than the average solitary expert, but the best individual in the group often outperforms the group as whole [...]. This motivates the elicitation of the assessments of individual experts without any interaction, followed by mathematical aggregation in order to obtain a single assessment per variable, thereby weighting the individual experts’ assessments based on their quality [...]. (Goossens et al. 2008: 234-235)

Rational consensus is perceived to be attainable if stakeholders commit in advance to the method by which expert views are selected and combined. This method should be constrained by the following “principles for rational consensus” (p. 15):

- Scrutability/accountability: all data, including experts' names and assessments, and all processing tools are open to peer review and results must be reproducible by competent reviewers.
- Fairness: experts are not pre-judged.

- Neutrality: methods of elicitation and processing must not bias results.
- Empirical control: quantitative assessments are subjected to empirical quality controls.

## **Conclusions**

Rational decision making must be based on models quantified by data. Data are usually provided by statistics or experiments. There are however many situations for which no statistics are available because the situation is new, unique or not replicable, or for which no experiments are possible because of ethical, political or practical reasons. For these situations, an appeal is made to experts. This paper discusses how the need for expert observations first was acknowledged by the early econometricians who by exploring the use of statistics for economics soon found its epistemological borders. The most pronounced advocate of expert observations in economics became Wassily Leontief. His concept of input-output tables can be considered as a forerunner of expert systems. While in the laboratory no distinction is being made between an 'expert' observation and a 'subjective' observation', outside the laboratory it matters a lot. But unlike Leontief's case, in more recent economic decision-making, only one single expert observation is not considered to be reliable. It is the combination of several expert observations that creates trust. This trust-generating combination is created by a model, arguments or by procedures. To explore model-based trust, modeling at the CPB Netherlands Bureau for

Economic Policy Analysis in the 1940s and 1950s has been taken as a case. To explore argumentation-based trust, the case of decision-making at the Bank of England was discussed. To obtain a procedural-based trust, we discussed a case of risk modeling in engineering.

## **Appendix**

Gesprek met prof. dr. J. Tinbergen over: *Zending en Ontwikkeling Maandblad Reünisten Organisatie Societas Studiosorum Reformatorum* October 1971: 1-4.

Ik heb [Keynes] nog één keer in persoon ontmoet, kort voor zijn dood, en heb van die ontmoeting zeer genoten. Zijn zekerheid omtrent zijn eigen visie op het vraagstuk van de Duitse herstelbetalingen bleek op humoristische wijze als volgt. In die visie had hij aangenomen dat een prijsverlaging van 1% door een exportland een vermeerdering van 2% van het volume van de uitvoer tot gevolg zou hebben. Niemand had dat tot dan toe geverifyeerd. Toen ik hem vertelde dat ik daartoe pogingen had gedaan en inderdaad op het getal 2 was gekomen, dacht ik dat hij daarmee wel blij zou zijn. Hij reageerde echter slechts met een: “Dat is leuk voor u.” Ik ben het eens met de vele, vooral Engelse economen, die zeggen dat hij werkelijk een man met visie was. (p. 3)

[I have met [Keynes] once, shortly before his death, and have very much enjoyed this meeting. His certainty about his own vision about the problem of the German reparations was shown in a humorous way as

follows. In that vision he had assumed that a price reduction of 1% by an export nation would lead to an increase of 2% of the export volume. Until then nobody had verified that. When I told him about my attempts and that I indeed arrived at the number 2, I expected that he would be pleased with it. He only responded with a “That is nice for you”. I agree with the many, mainly English economists, who say that he was really a man with vision.

Het getal twee is van Keynes

J. Tinbergen (1987b)

*Economisch Statistische Berichten* 72.3632 (18-11-1987): 1092.

Er is nog een grappige discussie tussen Keynes en mij aan verbonden. Ik vertelde hem dat wij op het CBS regressies berekend hadden, om deze elasticiteit te schatten en dat wij inderdaad in de buurt van -2 waren uitgekomen. Ik dacht dat dat voor hem welkom nieuws zou zijn. Hij vond echter dat het goed nieuws voor ons was, omdat wij het juiste getal hadden gevonden. Zijn eigen intuïtie vond hij aanmerkelijk betrouwbaarder dan econometrische schattingen. Misschien terecht.

[This is connected to a funny discussion between Keynes and me. I told him that at the CBS we had calculated regressions to estimate this elasticity, and that we had indeed come close to the figure of -2. I

thought that he would welcome that news. He responded that this was good news for us, because we had found the right figure. He considered his own intuition considerable more reliable than econometric estimates. Perhaps rightly so.]

Jan Tinbergen over zijn jaren op het CPB

A.H.M. de Jong, C.W.A.M. van Paridon, J. Passenier (1988)

*Economisch Statistische Berichten* 13-7-1988: 652-657, 662.

Wel dat mooie verhaal dat ook in de ESB heeft gestaan over die -2; dat was erg leuk. Ik weet niet precies meer waar we verder over spraken; ik heb hem ook niet zo lang gesproken. De kijk van deze man was misschien meer waard dan de uitkomsten van de regressie. (p. 656)

[But that nice story that was also published in the ESB about that -2; that was really nice. I do not recall exactly where we talked about; I haven't talked with him that long. The vision of this man was probably more valuable than the results of the regression.]

Recollections of professional experiences.

J. Tinbergen (1988) *Recollections of Eminent Economists*, (ed.) J.E.

Kregel. Macmillan Press.

A famous book by J.M. Keynes (1919) tried to show that the war indemnity required by the Allied Nations from Germany after the First World War was completely unrealistic, since the export possibilities of Germany were limited as a consequence of the limited price elasticity of the demand for Germany's (and any country's) export goods. In his study, Keynes took this elasticity to be -2. His argument would have been weakened considerably if this elasticity had been assumed to be, say, -10, or to be equal to the theoretical value of minus infinity. Hence, some of my collaborators in Holland and myself undertook a series of econometric studies in order to estimate the elasticity's value. We actually found values around -2, and I told Keynes so, expecting that he would consider this to be a strengthening of his position. His reaction was different; "how nice for you to have found the correct figure!" Sometimes, indeed, intuition constitutes a basis for new scientific results. I should be the intuition of a genius, however. For simpler souls, intuition may be less reliable!

## **References**

- Aldrich, J. (1989). Autonomy. In *History and Methodology of Econometrics* (pp. 15-34), N. de Marchi & C. Gilbert (eds.). Clarendon Press.
- Beckmann, Ulf (2000). *Von Löwe bis Leontief: Pioniere der Konjunkturforschung am Kieler Institut für Welwirtschaft*. Marburg: Metropolis-Verlag.

- Beverly, John (1997). The Bank's regional Agencies. *Bank of England Quarterly Bulletin*.
- Bjerkholt, Olav (2005). Frisch's econometric laboratory and the rise of Trygve Haavelmo's probability approach. *Econometric Theory* 21: 491-533.
- Bjerkholt, Olav and Mark Knell (2006). Ragnar Frisch and the origin of input-output analysis. *Economic Systems Research* 18.4: 391-410.
- Carter, Anne P. and Peter A. Petri (1989). Leontief's contributions to economics. *Journal of Policy Modeling* 11.1: 7-30.
- Christ, C.F. (1994). The Cowles Commission's contributions to econometrics at Chicago, 1939-1955. *Journal of Economic Literature* 32: 30-59.
- Cooke, Roger M. (1991). *Experts in Uncertainty: Opinions and Subjective Probability in Science*. New York and Oxford: Oxford University Press.
- Cooke, R.M. (2009). Resources for the future. Researcher Spotlight Roger Cooke.
- Cooke, R.M. and L.H.J. Goosens (1999). *Procedures guide for structured expert judgment*. Report EUR 18820. Luxembourg/Brussels: European Commission.
- DeBresson, Christian (2004). Some highlights in the life of Wassily Leontief – an interview with Estelle and Wassily Leontief. In *Wassily Leontief and Input-Output Economics* (pp. 135-147), Erik Dietzenbacher and Michael L. Lahr (eds.). Cambridge: Cambridge University Press.

- Don, F.J.H. and J.P. Verbruggen (2006). Models and methods for economic policy: 60 years of evolution at CPB. *Statistica Neerlandica* 60.2: 145-170.
- Downward, Paul and Andrew Mearman (2008). Decision-making at the Bank of England: a critical appraisal. *Oxford Economic Papers* 60: 385-409.
- Foley, Duncan K. (1998). An interview with Wassily Leontief. *Macroeconomic Dynamics* 2: 116-140.
- Frisch, Ragnar (1932). *New Methods of Measuring Marginal Utility*. Tübingen: Mohr.
- Frisch, R. (1933). Pitfalls in the statistical construction of demand and supply curves. *Veröffentlichungen der Frankfurter Gesellschaft für Konjunkturforschung*, New series 5. Leipzig: Hans Buske Verlag.
- Frisch, R. (1934). More pitfalls in demand and supply curve analysis. *The Quarterly Journal of Economics* 48.4: 749-755.
- Frisch, R. (1935). The principle of substitution. An example of its application in the chocolate industri. *Nordisk Tidsskrift for Teknisk Økonomi* 1: 12-27.
- Frisch, R. (1938). Statistical versus theoretical relations in economic macrodynamics. Memorandum prepared for the Business Cycle Conference at Cambridge, England, July 18<sup>th</sup> – 20<sup>th</sup>, 1938. Dated July 17, 1938. Reproduced as *Memorandum fra Universitetets Socialøkonomiske Institutt* (September 10, 1954), Oslo. Reprinted in Hendry and Morgan (1995), pp. 407-419.

- Frisch, R. (1948). Repercussion studies at Oslo. *The American Economic Review* 38: 367-372.
- Gilboy, Elizabeth Waterman (1931). The Leontief's and Schultz methods of deriving "demand" curves. *Quarterly Journal of Economics* 45: 218-261.
- Goossens, L.H.J., R.M. Cooke, A.R. Hale, Lj. Rodić-Wiersmae (2008). Fifteen years of expert judgement at TUDelft. *Safety Science* 46: 234-244.
- Haavelmo, T. (1938). The method of supplementary confluent relations, illustrated by a study of stock prices. *Econometrica* 6: 203-218.
- Haavelmo, T. (1944). The probability approach in econometrics. Supplement to *Econometrica* 12.
- Hempel, Carl G. (1965). Typological Methods in the Natural and Social Sciences. In: *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science* (pp. 155-171). New York: The Free Press.
- Hendry, David F. and Mary S. Morgan (1995). *The Foundations of Econometric Analysis*. Cambridge: Cambridge University Press.
- Kendall, Maurice G. and Alan Stuart (1963). *The Advanced Theory of Statistics, Volume 1*, 2<sup>nd</sup> edition. London: Charles Griffin.
- Keynes, John Maynard (1920). *The Economic Consequences of the Peace*. New York: Harcourt, Brace and Howe.
- Keynes, J.M. (1939). Professor Tinbergen's method. *The Economic Journal* 49: 558-568.

- Lambert, Richard (2005). Inside the MPC. *Bank of England Quarterly Bulletin*: 56-65.
- Leontief, Wassily W. (1929). Ein Versuch zur statistischen Analyse von Angebot und Nachfrage. *Weltwirtschaftliches Archiv* 30: 1-53.
- Leontief, W. (1932). Studien über die Elastizität des Angebots. *Weltwirtschaftliches Archiv* 35.1: 66-115.
- Leontief, W. (1934a). Pitfalls in the construction of demand and supply curves: A Reply. *Quarterly Journal of Economics* 48.2: 355-361.
- Leontief, W. (1934b). A final word. *The Quarterly Journal of Economics* 48.4: 755-759.
- Leontief, W. (1949). Recent developments in the study of interindustrial relationships. *The American Economic Review* 39.3: 211-225.
- Leontief, W. (1953). The input-output approach in economic analysis. In *Input-Output Relations. Proceedings of a Conference on Inter-Industrial Relations Held at Driebergen, Holland* (pp. 1-23), The Netherlands Economic Institute (ed.). Leiden: Stenfert Kroese.
- Leontief, W. (1971). Theoretical assumptions and nonobserved facts. *The American Economic Review* 61.1: 1-7.
- Leontief, W. (1982). Academic economics. *Science* 217 (9 July): 104, 107.
- Magnus, Jan R. and Mary S. Morgan (1987). The ET interview: Professor J. Tinbergen. *Econometric Theory* 3: 117-142.

- Marschak, Jakob (1931). *Elastizität der Nachfrage: Zur empirischen Feststellung relativer Marktkonstanten durch Beobachtung von Haushalt, Betrieb and Markt*. Tübingen: J.C.B. Mohr.
- Marschak, Jakob (1934). Some comments. *The Quarterly Journal of Economics* 48.4: 759-766.
- Mata, Tiago (2008). Naming crisis: oral history of financial media, 2007-2008.
- Morgan, Mary S. (1990). *The History of Econometric Ideas*. Cambridge: Cambridge University Press.
- Morgan, M. (2000). The relevance of economic modelling for policy-making: a panel discussion. In *Empirical Models and Policy-Making: Interaction and Institutions* (259-78), Frank A.G. den Butter and Mary S. Morgan (eds.). London and New York: Routledge.
- Phelps Brown, E.H. (1937). Report of the Oxford meeting, September 25-29, 1936. *Econometrica* 5: 361-383.
- Rodgers, Peter (1997). Changes at the Bank of England. *Bank of England Quarterly Bulletin*.
- Schmidt, Robert (1930). Die Prägnanz der Elastizitätskoeffizienten. *Weltwirtschaftliches Archiv* 32.1: 264-273.
- Silk, Leonard (1976). *The Economists*. New York: Basic Books.
- Theeuwes, J.J.M. (1987). Van dingen die stabiel zijn. *Economisch Statistische Berichten* 21-11-1987: 987.

- Tinbergen, Jan (1936) Memorandum on the continuation of the League's business cycle research in a statistical direction, unpublished memorandum. Archive of the League of Nations, Geneva.
- Tinbergen, J. (1939a). *A Method and its Application to Investment Activity. Statistical Testing of Business-Cycle Theories I*. Geneva: League of Nations.
- Tinbergen, Jan (1939b). *Business Cycles in the United States of America 1919-1932. Statistical Testing of Business-Cycle Theories II*. Geneva: League of Nations.
- Tinbergen, J. (1940). On a method of statistical business-cycle research. A reply. *The Economic Journal* 50.197: 141-154.
- Tinbergen, J. (1987a). Over modellen. In *Lessen uit het verleden. 125 jaar vereniging voor de Staathuishoudkunde* (pp. 99-112), A. Knoester (ed.). Leiden and Antwerpen: Stenfert Kroese.
- Tinbergen, J. (1987b). Het getal twee is van Keynes. *Economisch Statistische Berichten* 72.3632 (18-11-1987): 1092.
- Tinbergen, J. (1988). Recollections of professional experiences. *Recollections of Eminent Economists*, J.E. Kregel (ed.). Basingstoke: Macmillan.
- Tinbergen, J. (2003). The need of a synthesis. In *Jan Tinbergen. The Centennial Volume* (pp. 303-306), Jacob Kol (ed.). Rotterdam University. Translation of De noodzaak van een synthese. *Economisch Statistische Berichten* 1-12-1982: 1284-5.

**HISTORY OF OBSERVATION IN ECONOMICS WORKING PAPER  
SERIES**

1. Mary S. Morgan: Seeking Parts, Looking For Wholes. 2009
2. Theodore M. Porter: The Engineer and the Sage: Le Play's Quest for Social Renewal and the Reconstruction of Observation. 2009
3. Harro Maas: Sorting Things Out: The Economist as an Armchair Observer. 2009
4. Marcel Boumans: Observation of an Expert. 2009