

---

## 45 Econometrics: the Keynes–Tinbergen controversy

*Marcel Boumans*

---

No one could be more frank, more painstaking, more free from subjective bias or *parti pris* than Professor Tinbergen. There is no one, therefore, so far as human qualities go, whom it would be safer to trust with black magic. That there is anyone I would trust with it at the present stage or that this brand of statistical alchemy is ripe to become a branch of science, I am not yet persuaded. But Newton, Boyle and Locke all played with alchemy. So let him continue. (Keynes 1940, p. 156)

The first two macroeconomic models were constructed by Jan Tinbergen. Tinbergen's first model was of the Dutch economy, published in 1936. In the same year Tinbergen was commissioned by the League of Nations to perform statistical tests on business-cycle theories. The results of this later study were published in a two-volume work, *Statistical Testing of Business-Cycle Theories* (1939). The first contained an explanation of the "method" as well as an "illustration" of this method by demonstrating its "application" to three examples relating to fluctuations in total investment, residential building and net investment in railway rolling stock (Tinbergen 1939a, p. 9). The second volume developed a model of the United States; the second macroeconomic model in the history of economics (Tinbergen 1939b).

As Mary Morgan (1990, p. 121) notes, the first volume on the method "proved highly controversial". The manuscript of this volume was circulated in 1938; it was sent to a number of statisticians in different countries for comment, and two meetings of economists and statisticians were held at which the assumptions made and methods adopted were discussed (Tinbergen 1939a, p. 10). One of the economists who was approached for comment by Alexander Loveday, the Director of the Financial Sector and Economic Intelligence Service of the League of Nations, was Keynes. Keynes's comments and Tinbergen's responses, first expressed in an exchange of letters between Keynes, Loveday and Tinbergen, and later worked out in a book review by Keynes (1939) with a reply from Tinbergen (1940) and a final comment by Keynes (1940) came to be known as the "Keynes–Tinbergen debate".

This debate can be best described as a *Methodenstreit* in the sense that it is not only a discussion about the most appropriate methods for empirical economics, but also an epistemological debate about the most appropriate economic methodology. Since this kind of debate takes place at these two different levels, we find different assessments of its outcome. On the one hand, as Morgan (1990, p. 121) concluded, Keynes "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 technique by which trends are eliminated, Tinbergen (1940, p. 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".<sup>1</sup> On the other hand, David Hendry, in an inaugural lecture of 1980, "Econometrics – alchemy or science?" admitted that "it is difficult to provide a convincing case for the defence against Keynes's accusation almost 40 years ago that econometrics is *statistical alchemy* since many of his criticisms remain apposite" (Hendry 1980, p. 402, original emphasis).

To gain some understanding of what the debate was about, we need a brief account of their different backgrounds. Tinbergen was trained as a physicist in Leiden and Keynes had studied philosophy and economics in Cambridge. Their backgrounds were so different that the debate at times can be best characterized as being incommensurable in the Kuhnian sense: each world view was so different that they misunderstood each other profoundly.

## TINBERGEN'S METHOD

By describing his method as “a synthesis of *statistical business cycle research* and *quantitative economic theory*” (1939a, p. 11, original emphasis), Tinbergen made clear that his method belonged to the new field of econometrics. Its main founder, Ragnar Frisch (1933, p. 1), had defined the aim of econometrics as: “to promote studies that aim at a unification of the theoretical-quantitative and the empirical-quantitative approach to economic problems and that are penetrated by constructive and rigorous thinking similar to that which has come to dominate in the natural sciences”.

The aim of Tinbergen's method was to identify and test “direct causal relations” (1939a, p. 12). These relations are suggested by various economic theories, and it is the task of the “statistician” to test them. Since “no statistical test can prove a theory to be correct” (Tinbergen 1939a, p. 12), this task is restricted. A statistical test can only “prove that theory to be incorrect, or at least incomplete, by showing that it does not cover a particular set of facts” (Tinbergen 1939a, p. 12). However, according to Tinbergen, the task of the statistician is not restricted to testing; it includes also measurement: “what we want to discover is, not merely what causes are operative, but also *with what strength each of them operates*” (1939a, p. 12).

The statistical part of Tinbergen's method is what today is called regression analysis, but what he called “correlation analysis”. So, the main task of the statistician is to estimate regression equations and to assess the statistical significance of the results. However, the economic part was no less important: “The reliability of results may be judged by statistical as well as economic criteria” (Tinbergen 1939a, p. 27). Take, for example, the following relationship:  $x_1 = b_2x_2 + b_3x_3 + \dots$ , of which the causal factors,  $x_2, x_3, \dots$ , are “provided” by economy theory. Tinbergen emphasized that “this expression is justified only so far as the economic theory which has prompted the calculation is accepted as valid” 1939a, (p. 22). The term  $b_2x_2$  was called the “influence” of  $x_2$  and the “strength of the influence” was determined by standard deviation of the influence:  $b_2\sigma_2$ . In the analysis of economic significance it was taken into account whether the regression coefficients  $b_i$  have the right sign, and statistical significance was, among other things, based on the measurement of the strength of the influences.

According to this method, whether an influence, suggested by theory, should be considered as having explanatory power was tested by using the results of regression analysis. The assessment of the “relevance” of the variables, or “variates” as Tinbergen called them, was determined by three tests:

1. whether or not the variate in question increases the correlation coefficient to any considerable extent,

2. whether or not the sign of a fairly stable regression coefficient is right, and
3. whether or not the influence of that variate is perceptible. (Tinbergen 1939a, p. 50).

In addition to these three tests, we can distinguish seven more in the first volume (see Hendry and Morgan 1995, pp. 53–4):<sup>2</sup>

4. tests of lag length on the independent variables,
5. causality testing using the structure of correlation,
6. the linearity assumption,
7. serial correlation of the residuals,
8. significance, using standard errors of the regression coefficients,
9. multicollinearity, by bunch maps and simple correlations, and
10. impact of measurement errors by calculating limits on regression coefficients by bunch map methods and by Koopmans's test.

## THE DEBATE IN LETTERS

During the 1938 summer holidays, Keynes received manuscripts of both volumes with a request from Royall Tyler, expert in the Economic and Financial Section of the League of Nations, to comment upon them. In a letter dated 23 August 1938, Keynes replied crustily: “I confess that I have the utmost difficulty in making head or tail of them. No doubt this is partly due to my lack of familiarity with the matter, but partly also, I think, to the author's cryptic method of exposition” (Keynes 1973, p. 285).

Keynes's first concern was “the central question of methodology”, namely, the applicability of regression analysis to “economic material, which we know to be non-homogeneous through time” (1973, p. 286). In his view, regression analysis was only applicable to “the action of numerically measurable, independent forces . . . on material constant and homogenous through time” (Keynes 1973, p. 286). Related to this main concern was a list of questions that needs to be addressed, the first of which was the assumption that the coefficients were supposed to be constant for ten years or a longer period: “There is no reason at all why they should not be different every year” (Keynes 1973, p. 286). Another issue was the selection of factors being investigated, whether this selection is comprehensive or merely a partial selection? Since the equations are based on past statistics, the question is whether they also hold for the future.<sup>3</sup> Also, what about “non-numerical factors, such as inventions, politics, labour troubles, wars, earthquakes, financial crises?” and “if a factor has in fact not varied much during the period in question, it therefore necessarily emerges as unimportant?” (Keynes 1973, p. 287). Note that statistical significance (that is, influence and strength of influence) is determined by the standard deviation. So factors that did not vary during a specific period are considered to be statistically insignificant, though that does not mean they are also economically insignificant.<sup>4</sup> Keynes's final question was about the uniqueness of the results: “Would someone else, that is to say, faced with the same problem and using the same method and the same statistics, but without having seen these calculations, necessarily bring out the same result?” (1973, p. 287).

In his response to Keynes's letter to Tyler, Loveday (quoted in Keynes 1973, p. 290) acknowledged the importance of Keynes's central question: “You begin with the point

which always seemed to me the most vital of all, namely, that of the non-homogeneity of the material through time”, and therefore he made sure that Tinbergen would reply to this and the other questions. Keynes (1938) replied to Loveday by summarizing the problems he saw in Tinbergen’s method:

He has taken (i) periods too long for a constant value of his coefficients to be at all probable, (ii) problems where the factors he takes account of are most unlikely to be comprehensive, (iii) examples where the available statistics are not sufficiently complete and accurate for the weight he throws on them, whilst the intervals at which they have been collected are too long to allow adequate inferences as to time lags.

Tinbergen responded to Keynes’s criticisms in a letter to him, dated 12 September 1938. He did not address Keynes’s central question concerning non-homogeneity, but only responded to the subsequent questions. As regards the assumption that the coefficients are constant for ten years or more, Tinbergen noted that “in general this is true”, in the sense of a “first approximation” (Tinbergen, quoted in Keynes 1973, p. 291). He agreed with Keynes’s remark that “there is no reason at all why (coefficients) should not be different every year”, but that would only mean that another step had to be taken to find the underlying stable causal relationship that made this coefficient change: “coefficients changing just by chance would, of course, render this method as well as the whole of quantitative economic science impossible” (Tinbergen, quoted in Keynes 1973, p. 292).

Concerning the comprehensiveness of the factors, Tinbergen remarked that this is actually dealt with by his statistical tests to see whether their omission of factors “seems to be serious or not” (Tinbergen, quoted in Keynes 1973, p. 292). Also, he emphasized that he started “from the principle that economic theory should tell what the variables to be included are, and if one requires that lags shall be reasonable and the signs of the coefficients in accordance with economic theory” (Tinbergen, quoted in Keynes 1973, p. 292).

Tinbergen’s response was, according to Keynes, in a letter dated 20 September 1938, too much about the method, and too little on its methodological justification: “I do emphasise the consideration that very practical weight ought to be given to your provisional conclusions pending a justification of the application of your general method to statistics of the character and quality in question” (Keynes 1973, p. 294). To follow up his remarks about the constancy of the coefficients, Keynes suggested that, for a period of 20 years, these should be divided into shorter periods of five years each, and “calculate a proper equation for each period separately, and then consider what concordance appears between the different results”, otherwise the equation covering 20 years “can have very little significance” (Keynes 1976, p. 294). “My expectation would be that the broad problem of the credit cycle is just about the worst case to select to which to apply the method, owing to its complexity, its variability, and the fact [that] there are such important influences which cannot reduced to statistical form” (Keynes 1973, p. 294).

The debate in the exchange of letters stopped with this last reply from Keynes, and was continued in *The Economic Journal*. The reason was that Jacques Polak, employed to finish the statistical testing project at Geneva after Tinbergen had left, did not consider it worthwhile to continue this exchange. To Loveday he wrote (23 September 1938): “As Mr. Keynes’ letter once more shows, I think, [he is unwilling] to read carefully Tinbergen’s books or letter . . . I hope [therefore] you will approve this answer, which cuts off any further correspondence, which would be useless” (Polak 1938).

While Keynes's remarks were understood by Tinbergen and Polak as addressing only method, Keynes's main concern was methodological, the non-homogeneity of economic phenomena. The point he was trying to make was expressed more explicitly in an exchange of letters Keynes had with Roy Harrod: "economics is essentially a moral science and not a natural science. That is to say, it employs introspection and judgments of value" (Keynes 1973, p. 297). In a subsequent letter (dated 16 July 1938) he explicated what he meant by this in relation to Tinbergen's approach:

In chemistry and physics and other natural sciences the object of experiment is to fill in the actual values of the various quantities and factors appearing in an equation or a formula; and the work when done is once and for all. In economics that is not the case, and to convert a model into a quantitative formula is to destroy its usefulness as an instrument of thought. Tinbergen endeavours to work out the variable quantities in a particular case, or perhaps in the average of several particular cases, and he then suggest that the quantitative formula so obtained has general validity. Yet in fact, by filling in figures, which one can be quite sure will not apply next time, so far from increasing the value his instrument, he has destroyed it. . . . I also want to emphasize strongly the point about economics being a moral science. I mentioned before that it deals with introspection and with values. I might have added that it deals with motives, expectations, psychological uncertainties. One has to be constantly on guard against treating the material as constant and homogeneous. (Keynes 1973, pp. 299–300)

## THE DEBATE IN *THE ECONOMIC JOURNAL*

The debate in *The Economic Journal* was a public debate, and hence the style was different, though not the points that were already discussed in the earlier letters exchanged between Keynes and Tinbergen. Keynes's main concern was not the discussion of the statistical method, which he judged "very good indeed", but Tinbergen's discussion of the underlying methodology, which was "grievously disappointing" in its failure to explain "fully and carefully the conditions which the economic material must satisfy if the application of this method to it is to be fruitful" (Keynes 1939, p. 559).

Keynes's criticisms were presented as six conditions that the economic material must satisfy. The first was that Tinbergen's method "is only applicable where the economist is able to provide beforehand a correct and indubitably complete analysis of the significant factors. The method is neither of discovery nor of criticism" (Keynes 1939, p. 560). It is a method of measurement. Secondly, the method presumed that all significant factors are measurable. The third condition was the assumed independence of the factors towards each other, and fourth is the assumption that the correlations are linear. Keynes (1939, p. 564) summed up to this point: "But it is a very drastic and usually improbable postulate to suppose that all economic forces are of this character, producing independent changes in the phenomenon under investigation which are directly proportional to the changes in themselves; indeed, it is ridiculous." His fifth point was about the choice of trends and lags, which were, according to him, chosen arbitrarily.

The sixth and final point was about the epistemological reach of Tinbergen's method: "How far are these curves and equations meant to be no more than a piece of historical curve-fitting and description, and how far do they make inductive claims with reference to the future as well as the past?" (Keynes 1939, p. 566). Referring to his own study on

inductive inferences “thirty years ago” (Keynes 1939, p.566), Keynes emphasized the condition for inductive inferences:

the environment in all relevant respects, other than the fluctuations in those factors of which we take particular account, should be uniform and homogeneous over a period of time. We cannot be sure that such conditions will persist in the future, even if we find them in the past. But if we find them in the past, we have at any rate some basis for an inductive argument. The first step, therefore, is to break up the period under examination into a series of sub-periods, with a view of discovering whether the results of applying our method to the various sub-periods taken separately are reasonably uniform. If they are, then we have some ground for projecting our results into the future. (Keynes 1939, pp.566–67)

Anna Carabelli (1988) discusses in detail how Keynes’s (1921) *A Treatise on Probability* provides the theoretical background of his criticism of the applicability of Tinbergen’s method to “economic material”. As Keynes argues in the *Treatise*, for inductive inference based on statistical analysis it is of relevance to show whether the statistical series show some “stability”. According to Keynes, the method developed by Wilhelm Lexis is a “valuable aid to inductive correlation”: “This method consists in breaking up a statistical series, according to appropriate principles, into a number of sub-series, with a view to analysing and measuring, not merely the frequency of a given characteristic overt the aggregate series, but the stability of this frequency amongst the subseries” (Keynes 1921, p. 392).

The verification of whether a statistical series has such a stability is required because the condition for economic material to be applicable for inductive inference is that it shows sufficient homogeneity, or “uniformity”, as Keynes called it in his *Treatise*. By uniformity, Keynes (1921, p. 226) meant that “mere differences of position in time and space are treated as irrelevant”.

## CONCLUSIONS

The controversy between Keynes and Tinbergen was about the applicability of regression analysis to economic material. According to Keynes, a necessary requirement was that the economic material is homogeneous, in time and place, and that therefore Tinbergen should test his material for this condition. Tinbergen never addressed this problem, because he assumed that economic material was like the material studied in natural science, namely, homogeneous. He therefore could not understand Keynes’s criticisms, and so assumed that Keynes did not understand his method.

## NOTES

1. This view was reiterated in an interview with Tinbergen by Jan Magnus and Mary Morgan (1987, p.129), held in 1986: “From Keynes’ criticisms of your League of Nations first report, it seems fairly clear that he knew very little of the developments in econometrics over the 1920s and 1930s, despite being on the editorial board of *Econometrica*. . . T: Indeed, I did feel that, at least on certain points, he was badly informed. . . . It was a bit strange to me because he had written the *Treatise on Probability*, so I thought he was somewhat familiar with statistics.”
2. Hendry and Morgan actually arrive at 17 tests, an “impressive range of tests by the standard of any period” (Hendry and Morgan 1995, p.54). But this “battery of tests” also includes those of the second volume (Tinbergen 1939b).

3. This problem was the main point of critique in Milton Friedman's (1940) review of Tinbergen's second volume (1939b).
4. This problem of Tinbergen's methodology was analysed in great detail by Trygve Haavelmo (see Boumans 2014).

## REFERENCES

- Boumans, M. (2014), 'Haavelmo's epistemology for an inexact science', *History of Political Economy*, **46** (2), 211–29.
- Carabelli, A.M. (1988), *On Keynes's Method*, Basingstoke: Macmillan.
- Friedman, M. (1940), 'Review of "Business Cycles in the United States of America, 1919–1932" by J. Tinbergen', *American Economic Review*, **30** (3), 657–60.
- Frisch, R. (1933), 'Editor's note', *Econometrica*, **1** (1), 1–4.
- Hendry, D.F. (1980), 'Econometrics – alchemy or science?', *Economica*, **47** (188), 387–406.
- Hendry, D.F. and M.S. Morgan (eds) (1995), *The Foundations of Econometric Analysis*, Cambridge: Cambridge University Press.
- Keynes, J.M. (1921), *A Treatise on Probability*, London: Macmillan.
- Keynes, J.M. (1938), 'Letter to Loveday', dated 5 September 1938, Loveday Paper 1938/39, Archive of the League of Nations, Geneva.
- Keynes, J.M. (1939), 'Professor Tinbergen's method', *The Economic Journal*, **49** (195), 558–68.
- Keynes, J.M. (1940), 'Comment', *The Economic Journal*, **50** (197), 154–6.
- Keynes, J.M. (1973), *The General Theory and After: Part II Defence and Development*, in D. Moggridge (ed.), *The Collected Writings of John Maynard Keynes, XIV*, London: Macmillan
- Magnus, J.R. and M.S. Morgan (1987), 'The ET interview: Professor J. Tinbergen', *Econometric Theory*, **3** (1), 117–42.
- Morgan, M.S. (1990), *The History of Econometric Ideas*, Cambridge: Cambridge University Press.
- Polak, J. (1938), 'Minute sheet to Loveday', dated 23 September 1938, 10B.33994.12653, Archives 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 Economic Intelligence Service.
- Tinbergen, J. (1939b), *Business Cycles in the United States of America, 1919–1932. Statistical Testing of Business-Cycle Theories II*, Geneva: League of Nations Economic Intelligence Service.
- Tinbergen, J. (1940), 'On a method of statistical business-cycle research: a reply', *The Economic Journal*, **50** (197), 141–54.