

## NOTES AND COMMUNICATIONS

### FOUR EXAMPLES OF A *TROMPE-L'OEIL* IN ECONOMICS

#### *Summary*

This essay explores the occurrence of illusionary images or *trompe-l'oeil* in economics and their role in economic policy design. In this context four specific examples are discussed, relating to money illusion, the bias in the consumer price index to measure purchasing power, the econometrics of hyperinflation and the measurement of scholarly productivity.

**Key words:** citation analysis, consumer price index bias, hyperinflation econometrics, money illusion

**JEL Code(s):** A11, C43, E40, Z13

Almost 70 years ago, the American sociologist of science Robert K. Merton asserted that ‘Theories which attempt to account for certain phenomena require facts, but not all facts are equally pertinent to the problem in hand. “Selection’, determined by the limits of the problem, is necessary’ (Merton (1938) but reprinted in (1973, p. 228)). In this *selection* Merton supposedly presumed that cultural variables matter as do values and sentiments in research programmes. In the following my working hypothesis is that in social sciences and particularly in economics the researcher’s convictions and beliefs play a major role, at least implicitly and often subconsciously. Perhaps some of these convictions and beliefs are influential and, according to Merton, worthy of further investigation. I have therefore opted for a critical assessment of four topics of research which, in my view, have in common the tenets of a misleading illusion or *trompe-l'oeil*. The topics considered, which are of either a positive or a normative nature, are discussed in the following four sections of this essay. They concern money illusion, consumer price measurement, the research on hyperinflation and the fashionable practice of citation analysis, i.e. counting the number of citations to assess scholarly productivity. The final section concludes.

#### 1 MONEY ILLUSION: NEGLECTED IN RESEARCH AND POLICY

The term money illusion was coined three quarters of a century ago by Irving Fisher, who defined it as ‘the failure to perceive that the dollar, or any other

unit of money, expands or shrinks in value' (Fisher(1928, p. 4)). As a matter of fact, money illusion refers to individual or aggregate economic behaviour that consists in failing to distinguish transactions in terms of nominal or real monetary values. This odd tendency is a corollary of the fact that money as a measure of value or unit of account, such as the dollar or the euro, differs fundamentally from physical yardsticks, like miles, kilos, or ohms, in that it is not an intrinsically fixed and invariable measure over time.

This failure of the public to recognize real and nominal monetary changes as distinguishable facts is primarily a psychological phenomenon. However, this psychological idiosyncrasy does not imply that it fails to have any economic effects in the real world, as will be indicated below. In neoclassical economic theory, with its remarkable bias towards monetary neutrality with its strong belief in a money veil, these monetary effects often have been assumed away. So money illusion has been regarded with some suspicion, because the basic assumption of its absence underlies the long-run neutrality property embraced by the quantity theory of money.

Authors such as Leontief (1936), and Haberler (1941) in a footnote on p. 460<sup>1</sup>, consider money illusion a violation of the homogeneity postulate of economic theory. This theory hypothesizes that for goods and services the demand and supply functions are homogeneous of degree zero in all nominal prices; i.e. that demand and supply depend on relative rather than absolute prices and thus are insensitive to absolute price changes. Patinkin (1949, 1965) extended this to include monetary assets as cash balances. Consequently, Patinkin postulates the absence of money illusion on the basis of the zero-homogeneity property of net-demand functions in all money prices and the money value of initial holdings of assets. Operationally this is equivalent to the assumption of rational economic behaviour predicting that a proportional change in all prices and monetary balances would leave money's purchasing power unaffected. It is this very property that offers an operational yardstick for the observation and perhaps measurement of money illusion in practice.

The absence of money illusion is the main assumption underlying neoclassical economic theory, which cherishes David Hume's famous but questionable money veil which, according to Pigou (1941, pp. 20–27), denotes that money, in itself not comprising any of the essentials of real economic life, is only useful as a means of exchange for other things that are of direct significance for economic welfare. Nevertheless, recognition of money illusion has a long tradition among heterodox and especially monetary economists. With his 1928 monograph *Money Illusion*, Irving Fisher devoted an entire book to this topic, attempting to illustrate the occurrence of money illusion in the real world on the basis of anecdotic, historical and statistical evidence. To him, money illusion was an important explanation for business cycle fluctuations.

1 Surprisingly the relevant footnote has disappeared in the later than 1941-editions of Haberler's monograph.

Nowadays, interest in the empirical validity of the assumption of money illusion is no longer anathema to the economics profession, firstly, because the presence of money illusion helps to account for price stickiness and less than perfect economic adjustment processes and, secondly, because a lot of empirical or quasi-empirical evidence seems to support the occurrence of money illusion in real life.

Two kinds of evidence seem to dominate. On the one hand, several well-designed psychological experiments at the individual level show a convincing bias towards nominal rather than real magnitudes, which according to these findings results in considerable inertia (see Shafir et al. (1997); Fehr and Tyran (1997)). On the other hand, recent experience also appears to provide clear-cut evidence for the existence of money illusion at the individual and aggregate levels. The most notable historical evidence is associated with the introduction of the euro in 2002. This gigantic currency operation offered a splendid opportunity for a real-life experiment to examine the occurrence of money illusion in the main western European countries then joining the European Monetary Union with its common monetary policy. From a purely monetary point of view, the replacement of the national currencies of the eurozone countries by a single currency merely amounted to a redefinition of prices through multiplication by a given and fixed number, e.g. 0.45 for the Dutch guilder. According to the homogeneity postulate -that is the absence of money illusion- demand and supply conditions remain unchanged. However, as a matter of fact this purely nominal operation actually resulted in an upward though hardly statistically registered pressure of prices for particular commodities and especially services, affecting household expenditure considerably, as national account statistics of eurozone countries unambiguously show. This statistical observation, combined with ad hoc information on expenditure in several sectors of the economy in the relevant countries, points to some degree of money illusion with the consumers. So, both economic experiments (in this case with questionnaires on hypothetical situations allowing either nominal or volume variations) and designed experiments with actually observed expenditure behaviour apparently violate the neutrality property of money. Hence, the experimental evidence indicates the existence of money illusion in day-to-day life. Moreover, the difficulty of distinguishing between real and nominal exchange rates in daily economic activity provides additional empirical support for this conclusion. Finally, recent statistics on remarkably increasing bank indebtedness of households in the Netherlands, which over the years since 2002 surprisingly coincides with the introduction of the euro, also provide fresh circumstantial evidence for the occurrence of money illusion among the public in recent time spans as indicated in Fase (2005). This however, does not mean that it offers a solid time series historical account of this interesting economic phenomenon, enabling economic historians to examine it further. Taken together, the occurrence of

money illusion is quite likely in the real world and is, in fact, nothing but a particular manifestation of either incomplete knowledge or market frictions in society that hamper the perfect market clearing. Both discredit the monetary neutrality assumption of neoclassical economic theory and the harmlessness of a nominal monetary reform for purchasing power such as the introduction of the euro in 2002. Perhaps the assumption of neutrality around 2002 was the *trompe-l'oeil* of monetary policy makers.

## 2 MEASUREMENT ILLUSION: BIASES IN THE CONSUMER PRICE INDEX

In almost every country the consumer price index plays a particular socio-economic role through the institutional arrangements of society. Firstly, it is an overall measure of consumer commodity inflation. In that role it is used to index social security benefits, other public security support and income tax brackets against inflation in order to compensate for the effect of price changes on purchasing power. Moreover and secondly, it serves as *the* or one of the indicators of monetary policy. Of the latter the ECB's monetary policy in the EMU is a splendid illustration. These two particular uses require an undisputable and transparent production process at the statistics agencies. However, this is beyond any doubt. Of course, this statistical competence also depends on the technical capabilities available to and the budgetary constraints imposed on the national statistical agencies. Furthermore, this presumes absolute integrity of the technical process of making the most accurate possible price index numbers to serve society. In western societies there is no reason to doubt this.

Assessment of the purchasing power of money is a classical topic in monetary economics, providing the yardstick for quantifying domestic inflation. While measuring prices and their rate of change, i.e. the inflation rate, is of great and vital importance for every economy, its accuracy is of a fundamentally different nature than the occurrence of money illusion. This is so in spite of the fact that both focus on prices. National statistical agencies of all over the world produce consumer price indices to establish a cost of living index to assess the purchasing power of money, which for obvious reasons has a high policy sensitivity in modern society. Among economists and index number statisticians there is agreement that the consumer price index is at best a close approximation of the cost of living. The accuracy of this approximation or the appropriateness of the consumer price index is the issue to be considered here.

According to economic theory the consumer price index, which tracks the cost of purchasing a fixed market based basket of goods and services, represents the upper bound on the change in the true cost of living, assuming stable consumer preferences. This is an ideal and for that reason in some way a misleading economic illusion or *trompe-l'oeil*, notably because under par-

ticular circumstances this specific presumption has important policy implications. In the case of many households, following Diewert (1998), the social cost of living index, which is more relevant than the index of the virtual representative economic agent in consumer theory, is the ratio of the total minimum expenditure required to enable each of the households present in two periods of time to attain their reference utility levels in both time periods. A Laspeyres index, which defines the cost of purchasing a fixed basket of goods and services in the base period and the cost of buying the same basket in the present, tends, according to standard theory as the upper bound of the cost of living index, to overstate the rise of the cost of living. It does so because it allows no substitution between goods and services. However, following Deaton and Muellbauer (1980, p. 170) and others, a single commodity bundle is an unnecessary restrictive interpretation of what is to represent a constant standard of living. An obvious alternative and certainly a more abstract theoretical avenue is to take a specific utility level or indifference curve as the reference to be held constant. According to this interpretation the cost of living index is the ratio of the minimum expenditures required to attain the reference indifference curve of two sets of prices. This reasoning, quite naturally, results in two possible index number devices. The first is the base quantity weighted, or Laspeyres price, index, representing the upper bound of the cost of living index. The other is the current quantity weighted index, known as the Paasche price index, representing the lower bound of the cost of living index. The Paasche index has the advantage of taking into account the occurrence of commodity substitution and the appearance of new consumer goods. Unfortunately, the practicality of this theoretical conclusion is limited because the required data on expenditure outlays obtained from consumer surveys unavoidably lag behind considerably. This all is well known among economists and consumer price statisticians, but becomes of vital importance in times of enhanced market dynamics such as shown in the last decades or so. In 1994 it was precisely because of these enhanced market dynamics of the economy that the then Fed chairman Alan Greenspan told US Congress at a hearing that the US inflation rate measured by the Bureau of Labor Statistics consumer price index, was likely to be annually at least one percentage point too high and thus over-rated the actual annual price increases. Greenspan's one percent upward bias in the measured consumer price index vis-à-vis the true cost of living was certainly an expert's guess he made in view of his policy perception and desire of monetary neutrality which, according to the Fed chairman, is the level of inflation that does not materially enter into the decisions of households and firms.<sup>2</sup> Two

2 This is an old idea going back to the Austrian School and notably Carl Menger. He proposed the so called inner value of money as the appropriate inflation measure, a concept closely related to today's core or headline inflation. See Fase and Folkertsma (2001) on this issue.

years later, in 1996, Greenspan's conjecture was substantiated empirically by a group of experts under the leadership of Michael J. Boskin, a former chairman of the Economic Counsel of the president of the US. This so-termed Boskin Commission, which for its findings relied solely on existing research drawn from the literature without performing original studies on index numbers by itself, concluded that the consumer price index as a measure of the cost of living overstated the increase in the true cost of living by 1.1% points a year. Moreover, it established a plausible range for this upward bias of 0.80–1.60% points per year with the reported 1.1 percentage points as the best estimate of its mean.<sup>3</sup> The policy implications of this quantitative finding for indexed government benefits and tax brackets as well as for the relative size of US government debt were enormous and provoked a lot of public and academic debate. One of the outcomes of this debate was a renewed interest among economists and official statistical agencies in making price index numbers. But this was not the only remarkable result. The impact of the Boskin report was profound and inversely proportional to the out-of-pocket cost of the commission's work, which amounted only to a slight 25,000 US dollars. Anyhow the Boskin report resulted in an *aggiornamento* of the theory on price index numbers, important initiatives on data collection, and an incentive for hopeful innovative research at the official statistical agencies in North America and Western Europe.

As to the sources of the established upward bias of the official consumer price index in the US – and presumably also in other western countries – three important elementary economic facts came to the fore. These are the goods and services substitution bias explaining 0.4% points of the discrepancy, the outlet substitution bias that accounts for 0.10% points of the upward bias and the quality *plus* new products bias accounting for 0.60% points of the estimated total bias of approximately 1.10% points. Presumably, this statistical inaccuracy is an annually recurring phenomenon and for that reason cumulative over time.

The substitution bias arises because the usual weighting with a fixed basket of goods and services does not take into account that consumers shift away from goods and services if their relative prices or household income change. Several commentators (Baker (1998); Diewert (1998); Abraham (2003); Schultze (2003)) consider the estimated substitution bias by the Boskin Commission in some detail by making a distinction between commodity and outlet substitution, i.e. shifting to other points of sale. Diewert e.g. accepts the Boskin estimate of commodity substitution bias but finds a much higher estimate for outlet substitution. Diewert's back-of-the envelope estimate is 0.4, that is 0.3 percentage points above Boskin's, noting that it may

3 The final report is included in *Getting Prices Right: the debate over the consumer price index*, edited by D. Baker, Washington/London, 1998, pp. 1–77.

even rise with the advent of discount selling of goods and services over the Internet.

The problem of quality change and the arrival of new products arise because in conventional price index number construction these are considered only implicitly by replacing items. The Boskin Commission obtained estimates for quality change by examining 27 separate final categories of goods and services, resulting in an estimate of 0.6 percentage points a year. Not surprisingly, the accuracy of this upward bias estimate was the most controversial issue in the debate, mainly because this assessment required quality imputation of thousands of goods and services including new products. For this, the economic statisticians used several methods such as hedonic regression and formal subjective assessment. Baker (1996) criticized this Boskin estimate on the ground that it neglected a great deal of the items in the consumer price index. He came up with another and substantially lower estimate of  $-0.4\%$  points, comprising an opposite and downward bias of 1% point for quality changes only. Deaton (1998), another commentator, was concerned with the rather arbitrary nature of Boskin's estimate for quality change and new products but nevertheless refrained from providing an alternative guess. Hausman (2003) also was skeptical on this score but recognizes that quality change and the arrival of new goods may amply affect the cost of living. He therefore concluded that the published consumer price index may overestimate inflation considerably. Together with Nordhaus (1998), Hausman believes the Boskin Commission estimate for quality change is persuasive but not yet conclusive. The latter is so because of omissions in areas, new products, asset prices and the impact of major tectonic -that is revolutionary new technology- shifts. These revolutions render the upward bias of the consumer price index vis-à-vis the true cost of living highly uncertain. Perhaps the inclusion of asset prices does not pose a special problem when treated analogously to durables such as occupational housing, taking their prices to be the rental prices. The Boskin Commission did not report on asset prices as an example of excluded goods in spite of the existence of thoughtful writings on this matter (see Bryan et al. (2001) for a review). This inference is provocative and suggests another *trompe-l'oeil*, now in the area of price indexes. However, what is certain is the fact that the overall upward bias is significant and at least 1.5% points a year. As Abraham et al. (1998) of the Bureau of Labor Statistics have pointed out it, is an illusion to produce a perfect cost-of-living measure although we can make important improvements when more resources for internal research are made available. A similar conclusion presumably holds for EMU and the ECB that by setting a medium term inflation target of 2% anticipated an upward bias in measured price changes.<sup>4</sup>

4 See also Fase (1999) for a discussion of this.

## 3 TOPIC ILLUSION: HYPERINFLATION ECONOMETRICS

Almost by definition empirical measurement of hyperinflation offers no problem at all. Neither does the mathematical analysis of its dynamics, although in practice hyperinflation is a great political problem, as European history shows. Hyperinflation typically occurs in a politically unstable environment with astronomical increases in prices and the nominal quantity of money that dwarf changes in real income, production, trade and similar real economic flows but drives up nominal assets and phenomena of the same kind. Following the seminal work of Cagan (1956), hyperinflation is conventionally defined by a monthly rise in prices of at least 50% over a prolonged period of time. The classical historical examples are the seven famous hyperinflations of the 1920s and mid-1940s, which Cagan examined over 50 years ago. These were the hyperinflation episodes in Austria (October 1921–August 1922), Germany (August 1922–October 1923), Greece (November 1943–November 1944), Hungary (March 1923–February 1924; August 1945–July 1946), Poland (January 1923–January 1924) and Russia (December 1921–January 1924). Of course, there exist other examples such as Belgium in the late 1920s, the Latin American countries in the 1970s and thereafter, Yugoslavia in the 1990s and Zimbabwe in the mid-2000s with in the second quarter of 2006 inflation hovering around 1000% a year, then the highest rate in the world. However, Cagan focused on the above six European countries presumably to illustrate Friedman's adage that inflation is a monetary phenomenon only.

Cagan studied the interwar historical episodes of hyperinflation as part of Milton Friedman's project to restore the quantity theory of money – in disrepute since the 1930s and 1940s – to professional respectability. The interesting point is that Cagan's courageous monetarist attempt provided interesting and, among economists of the time, unknown empirical information within a solid economic analytical framework accompanied with statistical verification. This was certainly novel and highly inspired by Friedman's restated quantity theory of money that goes back to Irving Fisher at Yale and the Chicago economists of the 1930s. However, later generations of economists ran off with Cagan's mathematical approach and perhaps lost sight of the genuine core of his work by solely focusing on econometric subtleties. These gave rise to another economic *trompe-l'oeil*. It seems warranted to consider this a particular and presumably innocuous example of illusion in economics which is entirely of an analytical nature. Perhaps, this is an exaggeration that has been encouraged by Baumol's (1959, p. 8) splendid monograph on the magnificent dynamics.

The starting point of Cagan's analysis of hyperinflation is a restatement of Irving Fisher's classical quantity theory of money. Working under the guidance of Milton Friedman, Cagan considered the quantity theory of money as a special case of the demand for money he postulated. His demand for money under hyperinflation assumes that the real money balances decline



with the rise of expected inflation in the next period, deliberately ignoring other influences like real income and the real rate of interest. He justifies this by hypothesizing that in a situation of hyperinflation these influences are negligible and could be ignored as a first approximation because inflation dwarfs these real changes. However, expected inflation is unobservable, which requires an additional assumption to model the expectation formation process. The expectation formation process Cagan assumed – a novelty at the time – was the adaptive expectation hypothesis. The idea behind this assumption is that the expected inflation rate is proportional to the discrepancy between the most recent inflation rate and its own expected previous value, with the factor of proportionality reflecting the speed of adjustment of expectations. This adaptive expectations hypothesis means that the expected rate of inflation is adjusted upward, relative to its previous value, when the most recent actual and observed inflation rate exceeds its own previously expected value. Correspondingly, when the most recent actual inflation rate is smaller than its own previously expected value, expected inflation will be lowered relative to its previous value. Rewriting the corresponding algebraic formulae of adaptive expectations, the unobservable expected inflation rate can be expressed as a weighted average of all current and past actual inflation rates, with exponentially declining weights dependent on the speed of adjustment. So, it becomes possible to rewrite Cagan's money demand equation for real balances in observable terms. Algebraic rearrangement results in an equation for inflation with the money stock and lagged inflation rates as explanatory variables. This equation has been estimated by Cagan and, in his footsteps, by numerous other researchers.<sup>5</sup>

In the perspective of the present essay on illusions three comments seem to be relevant. The first is the stability of the Cagan hyperinflation model and the numerical coefficients this requires. The second concerns the reliability of the estimates of Cagan's regressions. The third is the economic interpretation of Cagan's outcome and its relationship with possible alternative hypotheses on the formation of expectations.

To examine the dynamics of hyperinflation one should determine whether the price level increase results from the enormous money supply in the countries considered or whether the price increases are self-generating when started up. This means that the stability properties of the Cagan model are of crucial importance. As a matter of fact the Cagan model results in a first order difference (or differential) equation whose properties define the stability conditions. Decisive is the size of its root, which is determined by an amalgam of two model parameters: the price elasticity of money demand and the speed of adjustment of expectations. The numerical estimates Cagan obtained indicate that the hyperinflation processes in almost all of the seven countries

5 See e.g. Sargent (1993); Schuit and Winder (1992); Michael et al. (1944); Petrović and Vujošević (1996).

he studied were stable. This means that at any disturbance of equilibrium, monetary or not, the price level tends to return to its equilibrium level. By definition, an unstable dynamic process implies that any shock to the system leads to a departure of equilibrium that is self-generating. This conclusion, however, is based on a mechanistic view and mathematical properties that perhaps do not have much policy relevance. For that, another than the mathematical approach to interpret the Cagan model is wanted. A careful economic interpretation of the mathematical model solution shows that in the Cagan model an excessive money supply is the driving force of inflation. This means that policy design failures are the basic source of hyperinflations, leading to the conclusion that a change in the policy regime, which is sufficiently binding and for that reason widely believed, is required. However, to achieve that conclusion we do not need the Cagan model per se. Nevertheless this model is useful because it convincingly shows that money matters and does not behave erratically as was often believed at the time of Cagan's pioneering work. Though pathbreaking, and this regards my second point of comment, Cagan's statistical estimates were inaccurate because the estimation procedure he used was inappropriate, generating biased and inconsistent estimates. In the monetarist literature, several proposals have been made to improve them, resulting in a flow of novel and convincing research, mainly of an econometric nature. However, this fine econometric work has in no way altered the policy conclusion that the hyperinflations of many countries in the pre- and post-war-II episodes were mostly the simple consequence of the enormous money stock increases in the countries concerned and seldom the outcome of a self-generating process. Another novelty of Cagan's research was the use of adaptive expectations. From the late 1970s onwards this hypothesis came under heavy attack and was replaced by the idea of rational expectations. Sargent (1982; reprinted 1993), an influential proponent of this view, believes the rational expectations hypothesis is difficult to reconcile with Cagan's momentum model, while in an earlier article, co-authored with Wallace (1973), he shows that with feedback from price increases to subsequent rates of money creation adaptive expectations are fully rational and not ad hoc as was the criticism in the literature. Therefore, under specific circumstances, the rational expectations approach may be consistent with the policy measures that brought drastic inflations under control in the hyperinflation countries

Cagan examined because of its forward-lookingness. That behaviour takes into account all information available including the anticipated or pre-emptive policy design.<sup>6</sup> Leaving the extensive and often econometrically futile research

6 The empirical results of Sargent and Wallace (1973) indicate that to explain hyperinflation the monetary authorities seem to make money creation respond directly and systematically to inflation. Accordingly, this was probably the reason why hyperinflations developed rather than exogenous money creation.

aside, Cagan's study points out that extreme rises in prices depend almost entirely on changes in the quantity of money. Rises in wages and the price of foreign currency are in his view the result rather than the start. Therefore, hyperinflation can be explained fully in terms of money demand and supply, with monetizing government deficits as a dominant force. While for this policy conclusion the many critical scholarly comments on the original Cagan model may seem to be irrelevant, they are of course, as pointed out in Fase (2002), of great fun for econometricians and mathematical economists. However, for policy-making, this research, though intellectually certainly interesting, is overstated and deteriorated to *l'art pour l'art*, not doing justice to Cagan's pioneering monetarist work. This practice of elaborating Cagan is somewhat misleading and perhaps a genuine *trompe-l'oeil*. The work concerned is undoubtedly fine scholarly research but without any practical implication or usefulness.

#### 4 RESEARCH PERFORMANCE: MEASUREMENT OF SCIENTIFIC PRODUCTIVITY

Scholarly publications and evaluation of research are of great importance to society and an undeniable part of academic life. To assess academic output a variety of indicators have been developed to serve researchers and their management. However, many scholars are skeptical towards these indicators partly because of the increasing number of publications and, as noted by Laband and Piette (1994), the explosive growth of new economic journals. Many feel the fashion of citation counts resembles an alternative mode of hyperinflation. Of greater importance perhaps is the question whether counting publications or citations and ranking of economic journals is that relevant for society. Certainly, this practice has enhanced transparency in academia and encouraged the publish or perish adage as well as multiple authorship considerably in the face of increased competition among scholars. However, how to separate among journal publications the wheat from the chaff? This question gives rise to a variety of issues which all share a deep concern over the intrinsic value of academic productivity. These concerns have evoked an extensive literature and much debate. As to the latter, a recent critical reflection is offered by Moed's (2005) informative monograph on citation analysis. For economics and the other social sciences authors such as Anderson et al. (1989), Laband and Piette (1994) as well as Pieters and Baumgartner (2000) published profound studies too, while over the years a flow of articles for particular countries has appeared.<sup>7</sup> However, one may suspect that these attempts uncover another *trompe-l'oeil*, certainly in

<sup>7</sup> E.g. Dusansky and Vernon for the US, Kalaitzidakis et al. (1999, 2003) or Frey and Eichenberger (2000) for Europe while in the Netherlands since the early 1980s the bi-monthly *Economisch Statistische Berichten* and the weekly *Intermediar* have published rankings of individual scholars and departments.

economics. Citation analysis and bibliometric indicators to measure research performance and academic productivity heavily weigh on academia and individual scholars in basic science, social sciences and humanities. Economics is a prominent sister within the disciplines constituting the social sciences. So economic researchers have a genuine interest either to adhere to or to look critically at bibliometric performance indicators and question whether they should take these measurements at face value. Because of the many arbitrary elements involved, perhaps a more relativist view on this practice seems to be desirable. For this I see at least four reasons. First, citation analysis is used to assess past research and publication activities of individual scholars, i.e. the basis for evaluation of research groups, departments, universities and even countries. Evaluative bibliometrics count citations during a certain time period, constituting the database for the citation indexes. The starting point is that all papers, notes, reviews and corrections published in scholarly journals provide the documents that contain citations constituting the formal linkage between the publications expressing scientific production. A citation index reflects the information signals based on indexing. The pioneer of this registration activity was the Institute of Scientific Information or ISI, now named Thomson Scientific Inc., which in the 1950s started to generate all sorts of bibliographic information. Since then bibliometrics became a field of inquiry to serve the scientific community including economics. However, a scholarly publication is not only a piece of information but also a manifestation of research activity that seeks mutual recognition. Citation indexing serves this purpose. So coverage, accuracy and a sound methodology are important parts of this signalling device. In view of this, one could question the neglect of monographs and books as well as the role played by the time window chosen.

Second, an important practical use of citation analysis is the ranking of scholars and journals according to the so-called journal impact factor. For a particular year and journal this is defined as the ratio of the number of citations received by all documents published in the journal considered, and the total number of citable and published documents in the journal during the previous years. Citations are the scientific community's version of dollar – or for that matter euro – voting by the consumer. Consequently, the impact factor denotes quality, influence and mutual recognition which are of great interest to authors and research groups in image-making. Moreover, to the publishers the impact factor is a marketing tool. Focusing on this statistic itself, one may wonder how robust the impact factor is in view of the many more or less arbitrary decisions to be made in calculating it. In this context the most important parameter is the time window, which often is a time frame of 2–5 years, neglecting scores at later ages or in the very distant

past.<sup>8</sup> To enhance its robustness, calculation of the citation half-life may be appropriate to obtain an additional assessment tool.<sup>9</sup> This seemingly technical detail is important because impact factors are crucial for the ranking of journals, which as a matter of fact determines the weights of journal publications of individual scholars. On the other hand, this ranking also influences a potential author's choice of journal to which to submit his or her paper, which in turn to some extent determines citation behaviour and thus scholarly reputation. This all looks like the famous chicken and egg problem that wholly neglects a paper's content or the academic task to disseminate knowledge. In view of this, Kodrzycki and Yu (2005) of the Federal Reserve of Boston, proposed another approach in ranking scholarly journals by taking into account the broader intellectual influence of economic journals in social sciences, resulting in adjusted but flexible citation statistics. This is also desirable in view of the increasing importance of the Internet as a communication channel within the community of scholars. However, as Pieters and Baumgartner (2000) have argued, the frequency of most citations in social sciences also provides insight into the scholarly network through which academic knowledge diffuses in economics. They conclude that intra- and interdisciplinary communication of economic journals is limited and that only the general journals and the theory and methods journals entertain reciprocal citation relationships. Another point is that it is uncertain whether or not the impact factor of journals not covered by ISI are as good as or even better than those covered by ISI. Perhaps Moed (2005, p.3) was right when he noted that there is perhaps no single perfect indicator of journal performance. Undoubtedly journal impact is a performance aspect that reflects prestige and quality.<sup>10</sup>

Third, for evaluating scholarly performance use is made of the institution of peer review too. Peer review takes place both within the context of a journal's editorial policy of refereeing submissions, awarding grants, the allocation of scholarly prizes, and in evaluating research performance by committees. Generally speaking, the credibility of peer review is unquestioned. Neverthe-

8 Anderson et al. (1989, pp. 181–182) examined this statistically and found that in economics and other social sciences the decay rate is substantial, which supports the fact that also in economics certain research areas fall out of fashion and consequently disappear from current literature and citation performance.

9 For biology Moed (2005, p. 97) examines this and found a truncated exponentially declining pattern. Perhaps for economics, which according to Lazear (2000) is not only a social science but a genuine science, a similar pattern would result.

10 Quality and research performance is a multidimensional concept. Any attempt to express them in a single number resembles reducing a multidimensional Euclidean space to one dimension without any loss of relevant information incorporated in the whole picture. Social scientist have developed statistical methods such as factor and principal components analysis to implement this dimension reduction empirically (see Cramer (1966) and Fase (1973), for early applications in economics). So far I have not seen an attempt to use these techniques justifying impact factors as an indicator for overall research performance.

less, one may wonder whether citation scores and other bibliometric devices implicitly play a role in peer evaluation. Statistical correlation between peer judgements about research performances, on the one hand, and results of citation counts, on the other, reveals a strong relationship. However, this does not necessarily mean that the two approaches measure the same phenomenon. Strong correlation may validate the citation method because peer reviews often take a broader perspective on academic prestige than only publication or citation scores in top journals (see e.g. Fase (1996)). Perhaps this supports the desirability of a relativist approach to citation scores, which would make citation analysis less of a *trompe-l'oeil*.

Fourth, publishing in journals is mainly a supply driven process with a minor role, if any, for the demand side. Some, e.g. Laband and Piette (1994) in their industrial organisation approach to scientific publishing, even doubt, on the basis of careful quantitative analysis, whether journal articles are read at all. The main function of academic publishing in their view is to promote personal visibility rather than disseminating knowledge. Perhaps to this one may add that getting research published is sort of certification or scholarly award for its author. Research on this issue shows that the standards for getting published evolve, predicting a shift from developing new insights towards polishing submitted papers. Ellison (2000), who did extensive research on this feature of academic publishing, concludes that this shift may reflect slowly changing tastes that are likely to increase the gap between the preferences of authors and readers. In the end this will have subtle negative welfare effects. Authors such as Frey (2003) or Hodgson and Rothman (1999) even suggest that today's practice of referee's vetoing against publication of submissions are indicative of the existence of an institutional oligopoly of predominantly US-based economic journals, a suggestion that is strongly supported by the research of Kalaitzidakis et al. (2003) on academic journal ranking. This tendency may narrow theoretical pluralism in economics and is likely to constrain the scope of new ideas and approaches.

Definitive answers of how to evaluate research in a balanced manner are not yet available. Acknowledgement of the multifacetedness of scholarly performance may help to recognize the danger of a one-dimensional approach by way of citation analysis to express academic productivity. Perhaps an analogy with performance measurement in business and government, now popular in the management literature, may be of some help. This literature (see e.g. Kaplan and Norton (1996); Chang and Morgan (2000)) proposes the so-called balanced scorecard as a suitable instrument to consider, besides productivity, other performance dimensions such as professionalism, personal skills, policy relevance, influence and intellectual independency. Taking into account such a broader view may strengthen the evaluation of academic research in general and in economics in particular. This will reduce the likelihood of citation scores to be a *trompe-l'oeil* in the economics. At the end

of the day in academia the adage *do right without expecting thanks* should perhaps be the rule to assess performance. A forward looking rather than the backward looking attitude of counting citations of past work might be more fruitful for genuine scholarship.

## 5 CONCLUSIONS

The maintained working hypothesis in this essay was that in research convictions and beliefs matter, introducing the danger of a *trompe-l'oeil* in assessing research efforts. In art history a *trompe-l'oeil*, which literally means deception of the eye, is a tour de force of illusionism, prominent in the Baroque but also popular with seventeenth century Dutch painters like Gerard Dou or Frans van Mieris. Both artists liked to deceive their spectators with artistic jokes (see e.g. Kleiner et al. (2001, p. 628)). To examine illusions in economics I considered four examples of what may be a *trompe-l'oeil*. These concern suggestive theoretical work or facts that deceive the economic observer and preclude sound policy judgement. In some way, the chosen examples have in common the notion of inflation in the form of either price changes or excessive growth in journal publications or multiple authorship. Moreover, all explored cases have subtle social welfare effects which are likely to be negative. This is most obvious in the neglect by policy makers of money illusion and the measurement bias in price indexes. The one-dimensional emphasis on certain types of research and performance indicators may result in sub-optimal allocation of scarce resources too. We all know that this is inefficient and causes welfare losses. From a social welfare point of view, this is unforgivable and ought to be avoided.

M.M.G. Fase\*

## REFERENCES

- Abraham, K.G. (2003), 'Toward a Cost-of-living Index: Progress and Prospects,' *Journal of Economic Perspectives* 17, pp. 45–58.
- Abraham, K.G., J.S. Greenless and B.R. Moulton (1998), 'Working to Improve the Consumer Price Index,' *Journal of Economic Perspectives* 12, pp. 27–36.
- Anderson, G.M., D.M. Levy and R.D. Tollison (1989), 'The Half-life of Dead Economists,' *The Canadian Journal of Economics* 22, pp. 174–183.
- Baker, D. (1996), 'Does the CPI Overstate Inflation? An Analysis of the Boskin Commission Report,' in: D. Baker (ed), *Getting Prices Right; the Debate Over the Consumer Price Index*, Washington/London, pp. 79–155.

\* Fellow Tinbergen Institute and the Managing editor of *De Economist*. Until 2001 deputy executive director as well as head of research with De Nederlandsche Bank and from January 2003 professor emeritus in monetary economics, University of Amsterdam. I am grateful to professor J.S. Cramer and dr Peter van Els for their useful comments. Corresponding author: e-mail: mmg.fase@planet.nl

- Baumol, W.J. (1959), *Economic Dynamics; An Introduction*, New York.
- Bryan, M.F., S.G. Cecchetti and R. O'Sullivan (2001), 'Asset prices in the measurement of inflation,' *De Economist* 149, pp. 404–431.
- Cagan, Ph. (1956), 'The Monetary Dynamics of Hyperinflation,' in: M.Friedman (ed.), *Studies in the Quantity Theory of Money*, Chicago, pp. 23–117.
- Chang, R.Y. and M.W. Morgan (2000), *Performance Scorecards, Measuring the Right Things in The Real World*, San Francisco.
- Cramer, J.S. (1966), 'Une analyse de budget de famille par composantes principales,' *Economie Appliquée* 19, pp. 250–268.
- Deaton, A. (1998), 'Getting prices right: What Should Be Done?,' *Journal of Economic Perspectives* 12, pp. 37–46.
- Deaton, A. and J. Muellbauer (1980), *Economics and Consumer Behavior*, Cambridge.
- Diewert, W.E. (1998), 'Index Number Issues in the Consumer Price Index,' *Journal of Economic Perspectives* 12, pp. 47–58.
- Dusansky, R. and C.J. Vernon (1998), 'Ranking of U.S. Economics Departments,' *Journal of Economic Perspectives* 12, pp. 157–170.
- Ellison, G. (2000), 'Evolving Standards for Academic Publishing: A q-r Theory,' *NBER Working Paper 7805*.
- Fase, M.M.G. (1973), 'A Principal Components Analysis of Market Interest Rates in the Netherlands, 1962–1970,' *European Economic Review* 4, pp. 107–134.
- Fase, M.M.G. (1996), 'Erkenning Van Onderzoek: Toetsing van de Nieuwe Spelers op het Onderzoekveld,' *Beleid en Maatschappij* 23, nr. 5, pp. 246–253.
- Fase, M.M.G. (1999), 'Prijsstabiliteit als Monetaire Doelstelling: Het Belang van een Volmaakt Prijsindexcijfer,' *Tijdschrift voor Politieke Economie* 22, pp. 73–81.
- Fase, M.M.G. (2002), *Van Droom naar Daad: Veertig Jaar Monetaire Economie in Vogelvlucht*, Amsterdam.
- Fase, M.M.G.(2005), 'Zet de Prijzen voor lange Tijd Vast,' *Het Financieele Dagblad* 1 december, p.13.
- Fase, M.M.G. and C.K. Folkertsma (2001), 'Measuring Inflation: testing Carl Menger's concept of the Inner Value of Money,' *Kredit und Kapital* 34, pp. 197–222.
- Fehr, E. and J.-R. Tyran (1997), 'Does Money Illusion Matter?,' *American Economic Review* 112, pp. 239–262.
- Fisher, I. (1928), *The Money Illusion*, New York.
- Frey, B.S. (2003), 'Publishing as Prostitution? Choosing Between one's Own Ideas and Academic Success,' *Public Choice* 116, pp. 205–223.
- Frey, B.S. and R. Eichenberger (2000), 'The Ranking of Economists and Management Scientists in Europe: A Quantitative Analysis,' *Journal des Economistes et des Etudes Humaines* 10, pp. 575–581.
- Haberler, G. (1941), *Prosperity and Depression*, 3rd edn. London.
- Hausman, J. (2003), 'Sources of Bias and Solutions to Bias in the Consumer Price Index,' *Journal of Economic Perspectives* 17, pp. 23–44.
- Hodgson, G.M. and H. Rothman (1999), 'The Editors and Authors of Economics Journals: A Case of Institutional Oligopoly?' *Economic Journal* 109, pp. 165–189.



- Kalaitzidakis, P., Th. P. Mamuneas and Th. Stengos (1999), 'European Economics; An Analysis Based on Publication in Core Journals,' *European Economic Review* 43, pp. 1150–1168.
- Kalaitzidakis, P., Th. P. Mamuneas and Th. Stengos (2003), 'Rankings of academic journals and institutions in economics,' *Journal of the European Economic Association* 1, pp. 1346–1366.
- Kaplan, R. and D. Norton (1993), 'Putting Balanced Scorecard to Work,' *Harvard Business Review* 71 no 9/10, pp. 73–87.
- Kleiner, F.S., C.J. Mamiya and R.G. Tansey (2001), *GARDNER's Art through the Ages*, London, 11th edn.
- Kodrzycki, Y.K. and P.D. Yu (2005), 'New Approaches to Ranking Economic Journals,' *Research Review Federal Reserve Bank of Boston*, No. 4, pp. 13–15 (and <http://www.bos.frb.org/economic/wp/wp2005/wp0512.pdf>).
- Laband, D.N. and M.J. Piette (1994), 'The Relative Impacts of Economics Journals: 1970–1990,' *Journal of Economic Literature* 32, pp. 640–666.
- Lazear, E.P. (2000), 'Economic Imperialism,' *The Quarterly Journal of Economics* 115, pp. 99–46.
- Leontief, W. (1936), 'The Fundamental Assumptions of Mr. Keynes' Monetary Theory of Unemployment,' *Quarterly Journal of Economics* 51, pp. 192–197.
- Merton, R.K. (1973), 'The Puritan Spur to Science,' chapter in his 1938: *Science, Technology and Society in Seventeenth-Century England*, reprinted in: R.K. Merton, *The Sociology of Science*, Chicago/London, pp. 228–253.
- Michael, P., A.R. Nobay and D.A. Peel (1944), 'The German hyperinflation and the Demand for Money Revisited,' *International Economic Review* 35, pp. 1–22.
- Moed, H.F. (2005), *Citation Analysis in Research Evaluation*, Dordrecht.
- Nordhaus, W.D. (1998), 'Quality Change in Prices Indexes,' *Journal of Economic Perspectives* 12, pp. 59–68.
- Patinkin, D. (1949), 'The Undeterminacy of Absolute Prices in Classical Economic Theory,' *Econometrica* 17, pp. 1–27.
- Patinkin, D. (1965), *Money, Interest and Prices*, 2nd edn, New York.
- Petrović, P. and Z. Vujošević (1996), 'The Monetary Dynamics in the Yugoslav Hyperinflation of 1991–1993: the Cagan Money Demand,' *European Journal of Political Economy* 12, pp. 467–483.
- Pieters, R. and H. Baumgartner (2000), 'Who Talks to Whom? Intra- and Interdisciplinary Communication of Economic Journals,' *Journal of Economic Literature* 40, pp. 483–509.
- Pigou, A.C. (1941), *The Veil of Money*, London.
- Sargent, T.J. (1993), 'The Ends of Four Big Inflations,' reprinted in T.J. Sargent: *Rational Expectations and Inflation*, 2nd edn New York, pp. 43–115.
- Sargent, T.J. and N. Wallace (1973), 'Rational expectations and the dynamics of hyperinflation,' *International Economic Review* 14, pp. 328–350.
- Schuit, M.E.J. and C.C.A. Winder (1992), 'Analyse van Hyperinflatieprocessen: Cagan's Visie in Theorie en Praktijk,' *Maandschrift Economie* 56, pp. 372–385.

- Schultze, Ch. L. (2003), 'The Consumer Price Index: Conceptual Issues and Practical Suggestions,' *Journal of Economic Perspectives* 17, pp. 3–22.
- Shafir, E., P. Diamond and A. Tversky (1997), 'Money Illusion,' *Quarterly Journal of Economics* 112, pp. 341–374.