Economics
The Open-Access, Open-Assessment E-Journal

**Discussion Paper**
No. 2015-51 | July 15, 2015 | http://www.economics-ejournal.org/economics/discussionpapers/2015-51

# An Application of Econophysics to the History of Economic Thought: The Analysis of Texts from the Frequency of Appearance of Key Words

*Estrella Trincado and José María Vindel*

**Abstract**
This article poses a new methodology applying the statistical analysis to the economic literature. This analysis has never been used in the history of economic thought, albeit it may open up new possibilities and provide us with further explanations so as to reconsider theoretical issues. With that purpose in mind, the article applies the intermittency of the turbulence in different economic texts, and specifically in three important authors: William Stanley Jevons, Adam Smith, and Karl Marx.

**JEL** B16 B40 C18 Z11
**Keywords** Econophysics; history of economic thought; bibliometrics; applications of statistical analysis

**Authors**
*Estrella Trincado*, ✉ Universidad Complutense de Madrid, Spain, estrinaz@ccee.ucm.es
*José María Vindel*, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas, Madrid, Spain

**Citation** Estrella Trincado and José María Vindel (2015). An Application of Econophysics to the History of Economic Thought: The Analysis of Texts from the Frequency of Appearance of Key Words. Economics Discussion Papers, No 2015-51, Kiel Institute for the World Economy. http://www.economics-ejournal.org/economics/discussionpapers/2015-51

**Received** July 7, 2015 **Accepted as Economics Discussion Paper** July 13, 2015 **Published** July 15, 2015

## 1. INTRODUCTION

The most obvious and usual way to deal with economic texts is to try to understand in an analytical way the theories put forward by the authors providing a scientific, social and economic context for that specific literature. Using analytical tools, we can assess what the author defends, something that will furnish us with some of his "good-will statements". But, can the wording of the texts provide us with further knowledge about the authors' deep thoughts? This study draws a methodology to achieve such an objective, the application of the statistical analysis used in physics to the economic literature. An analogy between physics and economics has been extensively developed in recent years (Mantegna and Stanley 1996, 1997, 2000, Loffredo 2004, Lux, 2001); for instance, fluid dynamics and turbulent behavior have been used to predict the behavior of the stock market (Vindel & Trincado 2010; Trincado & Vindel 2013). However, this analysis has never been used in the history of economic thought, albeit it may open up new possibilities and provide us with further explanations so as to reconsider theoretical issues.

Actually, rarely is the statistics, one of the most widespread methodological tools in economics, used in the comparative contents of economic literature and their rhetoric.[1] However, in scientometric and informetric fields content analysis and mapping are widely developed such as, for example, in the word cloud literature (Schütze 1998). Certainly, the study of literature using mathematical statistics is not new. The American linguist Zipf (1935, 1949) stated that in any meaningful text, the frequency of any word is inversely proportional to its rank in the frequency table. Thus, according to the Zipf's law, the most frequent word will occur approximately twice as

1 Such analysis is the purview of "bibliometrics", but it is generally within the applied sciences that it is framed, exploring the impact of papers in the field of library and information science or examining the sociology of academia in the field of economic doctrines (Nicholas and Ritchie 1978). Bibliometrics does not usually have a theoretical or philosophical aim.

often as the second most frequent word, three times as often as the third most frequent word, etc. The remarkable question is that Zipf's law holds for most languages. Zipf (1949, 1) provided an explanation: the principle of the least effort used by speakers and listeners to reach an understanding.[2] This is not the last time that mathematics will be used to analyse the arts. Actually, fractals were observed by artists before scientists (Edgar, 2004). Some scientists believe that this kind of analysis can give deep insights into how mathematical structures guide our visual perception and how they shape our appreciation of art and artists' instinctive ability to pick out structures (Taylor *et al*, 2010). 3

**2. Objective of the article**

This article abounds in this scholarship by studying the intermittency of the turbulence in different economic texts, and specifically in three important authors: William Stanley Jevons, Adam Smith, and Karl Marx. If turbulence is characterised by a chaotic and stochastic regime and property changes, and by the existence of a cascade of energy from larger turbulent eddies to smaller ones (Richardson, 1922), intermittency is a specific property of a turbulent regime, usually explained as the irregular alternation of data which are not continuously available due to some factor outside direct control.[4]

Thus, the first hypothesis that the article aims at addressing is that economic texts are literature. As McCloskey (1983) said, being economics heavily metaphorical, the rhetoric of economics is a literary matter. We need to study how economists really argue and we have to

2 This law has been empirically verified in many types of data in the physical and social sciences, which can be approximated with a Zipfian distribution part of the family of relate discrete power law probability distributions. For instance, city population ranking in various countries, corporation sizes, income rankings, and so on, follow Zipf's law, as first shown by Auerbach already in 1913 (see Gabaix, 1999 and Kali, 2003). Actually, it was Pareto (1897) the one who firstly discovered a power law tail in the distribution of individual wealth.

3 For instance, by using statistical analysis some researchers have discovered that there are turbulent structures in paintings. Recently Aragón et al. (2006) showed that the curls and swirls of Van Gogh's later paintings adhere to the same physical laws as turbulent fluids found in nature. In effect, the so-called probability density function (PDF) of luminance fluctuations in some Van Gogh paintings compares notable well with the PDF of the velocity differences in a turbulent flow as predicted by Kolmogorov's statistical theory (Kolmogorov, 1941).

4 Although intermittency analysis was first applied in physics, at present it is frequently used in other disciplines, e.g., to assess the volatility of financial markets (Plerou *et al*, 2000).

examine the workaday rhetoric that leads to the prior beliefs. Ricoeur (1984, 1986, and 2004) said that literature has to do with the revealing, through writing and within the textual structure, of unconscious links, not made explicit and never surfacing at the conceptual or theoretical level. Finally as Adam Smith said "The best method of explaining and illustrating the various powers of the human mind, the most useful part of metaphysics, arises from an examination of the several ways of communicating our thoughts by speech, and from an attention to the principles of those literary compositions which contribute to persuasions or entertainment" (cited in Jermolowicz, 2004, 202). [5]

### 3. Analysed data

In this paper, three very different authors in terms of theoretical construct and cultural and analytical background have been compared. As mentioned above, the authors selected are William Stanley Jevons, Adam Smith and Karl Marx. We will study Jevons' theory by applying a statistical study of one of his most important works, *The theory of political economy*[6]. To study Marx's theory, we have used volume I of the English version of *Capital*[7] - using the German version could make our comparison unequal and our results less robust. Finally, in the study of Adam Smith's theory we have used *The Wealth of Nations*[8].

In an attempt to choose some significant words that are present in all of the texts, we have come across a sentence by Jevons in the preface to the fifth edition of the *Theory of political economy.* There, he specifies that the most difficult notions in political economy are "Utility, Value, Labour and Capital" which can be subject to mathematical analysis and expression. We have considered this sentence by Jevons a good starting point: these words seem to be important notions

5 If we may interpret in a psychological way the appearance of an approximate turbulent structure in the products human creativity, such as a painting, then, why do not do the same in a text? They are both the representation of a conscious desire of communication by the author and a representation of the author's subconscious.
6 This can be found at http://mises.org/books/political_economy_jevons.pdf .
7 This can be found at http://www.marxists.org/archive/marx/works/download/pdf/Capital-Volume-I.pdf.
8 Which can be found at http://www2.hn.psu.edu/faculty/jmanis/adam-smith/Wealth-Nations.pdf

for Jevons, and it seemed appropriate and relevant to analyse them in all the authors[9].

## 4. Methodology

We will now apply the statistical analysis to the thinking of these economists by looking at the flow of thought within the manuscript which talks about the evolution or melody of the text. To assess a melody, we need to look at the full sequence or arrangement of words, the harmony that tries to combine chords in a certain way.

The idea is to study the uni-scaling or multi-scaling behaviour of those key words found. Our study variable will be the number of times that a certain word appears on every page of the book. The corresponding scale will be the distance between pages. Thus, the value of the variable for a certain scale will be the difference between the number of times that the word appears on one page compared to the frequency of the word on another page with a page distance corresponding to that scale. For example, at scale 1, the variable represents the number of times that the word appears on a page minus the number of times that the word appears on the previous page. At scale 5, the variable is the difference between the number of times that the word appears on a specific page and the number of times it appears 5 pages before. Obviously, the word appearance frequency shows the importance that the author gives to the concept represented by the word. Given a certain scale, in this case a certain distance between pages, a high probability that the frequency of appearance of the word on both pages is similar implies that there is a high persistence at this scale. That implies a relationship between the ideas expressed on one page and the ones expressed on another page at the distance of the scale in question.[10]

Then we may represent the Probability Density Functions (PDFs) of the variable, who

9 By not selecting the words on purpose we claim not to be charged with trying to draw some specific results from our study. For that same reason, we have not chosen synonyms or association of words, but only registered those words themselves.

10 The higher the scale, the lesser the transmission of information that we may expect from one page to the other.

describe the likelihood for a random variable to take on a given value (inevitably, the addition of all the probabilities equals 1). These PDFs can be represented, in general, as Lévy distributions (Lévy, 1925), characterized by the presence of fat tails. A PDF with a sharper peak implies that the tail is fatter or more stretched out. It also implies that it is likely that the variable will have high values at a certain scale (see Figure 1).

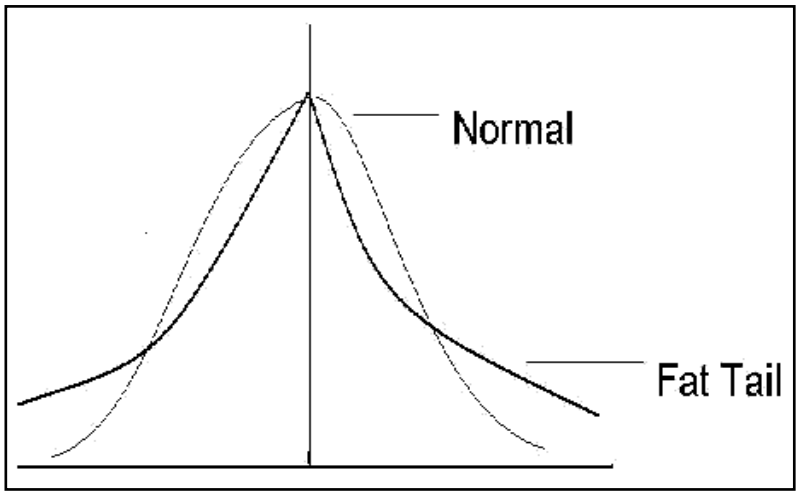


**Figure 1.** PDFs with different shapes (with more or less stretched tails).

Thus, in case of fatter tails, there are more events or values that stray widely from the average. Fatter tails show high and low values more frequently than when the shape of the tail is different. These are the unexpected results. Taking the case of our variable (differences between frequencies), high persistence (high values in the PDF for x=0) should mean, in general, more transmission of information. On the other hand, previous information can also yield very large jumps in the value of the variable (rare events).

On the other hand, if the text shows multi-scaling (different behaviour at different scales), it may mean that the words are represented as a turbulent regime characterized by intermittency. An intermittent source of data may be quite predictable or persistent but it is unintentionally stopped or unavailable. Thus, intermittency is sometimes considered to be synonymous with variability of the

variability. It indicates the existence of undesired or uncontrolled changes and of a greater - undesired - loss of information rate in the short-run than in the long run (fatter PDFs for the short-term and thus, greater probability of rare events in the small scales).

**5. Distribution analysis along the text**

Although Jevons makes a plea for considering the words "Utility, Value, Labour and Capital" as the most important notions in political economy, data processing shows that Jevons only made significant use of three of those words in the text analyzed: utility, labour and capital. Utility is cited 457 times, labour 411 times and capital 403 times, but value is only cited 218 times in a text of 182,628 words. Although 400 is not a very high number of words for a statistical analysis, we have found it to be sufficient, but value had to be ruled out due to statistical constraints. However, Marx only cites the word utility 35 times in his *Capital* (337,280 words), although the frequency of the other words is large enough: capital is cited 1,284 times, and labour, 3,401. Smith also frequently cites labour (1,148 times) and capital (769 times), but does not often cite utility (18 times) in his *Wealth of Nations* (383,766 words). The fact that Adam Smith uses the word utility infrequently is much more surprising since he has been considered a utilitarian for decades and even now a "contemplative utilitarian" (Ross, 1995). Did history or the historians label Adam Smith as a utilitarian as a way to disclose his real hidden purposes? If so, we may contribute to the process of disclosing Adam Smith's "statements of good will" in a more positivistic way than the merely bringing charges against him.¨

The differences in the author's theories are made evident in the statistical analysis. For doing this analysis, first we have made the table on sections and chapters of the three books (table 1, table 2 and table 3) and then we have represented the number of times that each of the words labour and capital appear in each page of the text, including the word utility for Jevons as he is the only one to mention that word with significant frequency (Figure 2).

## TABLE 1 (WEALTH OF NATIONS)

### BOOK I. OF THE CAUSES OF IMPROVEMENT IN THE PRODUCTIVE POWERS OF LABOUR, AND OF THE ORDER ACCORDING TO WHICH ITS PRODUCE IS NATURALLY DISTRIBUTED AMONG THE DIFFERENT RANKS OF THE PEOPLE



### BOOK II. OF THE NATURE, ACCUMULATION, AND EMPLOYMENT OF STOCK



### BOOK III. OF THE DIFFERENT PROGRESS OF OPULENCE IN DIFFERENT NATIONS



### BOOK IV. OF SYSTEMS OF POLITICAL ECONOMY



### BOOK V. OF THE REVENUE OF THE SOVEREIGN OR COMMONWEALTH

## TABLE 2 (CAPITAL)

### Part 1: Commodities and Money



### Part 2: Transformation of Money into Capital



### Part 3: The Production of Absolute Surplus Value



### Part 4: Production of Relative Surplus Value

**Part 5: Production of Absolute and Relative Surplus Value**



**Part 6: Wages**



Part 7: The Accumulation of Capital

**Part 8: Primitive Accumulation**

## TABLE 3 (THE THEORY POLITICAL ECONOMY)

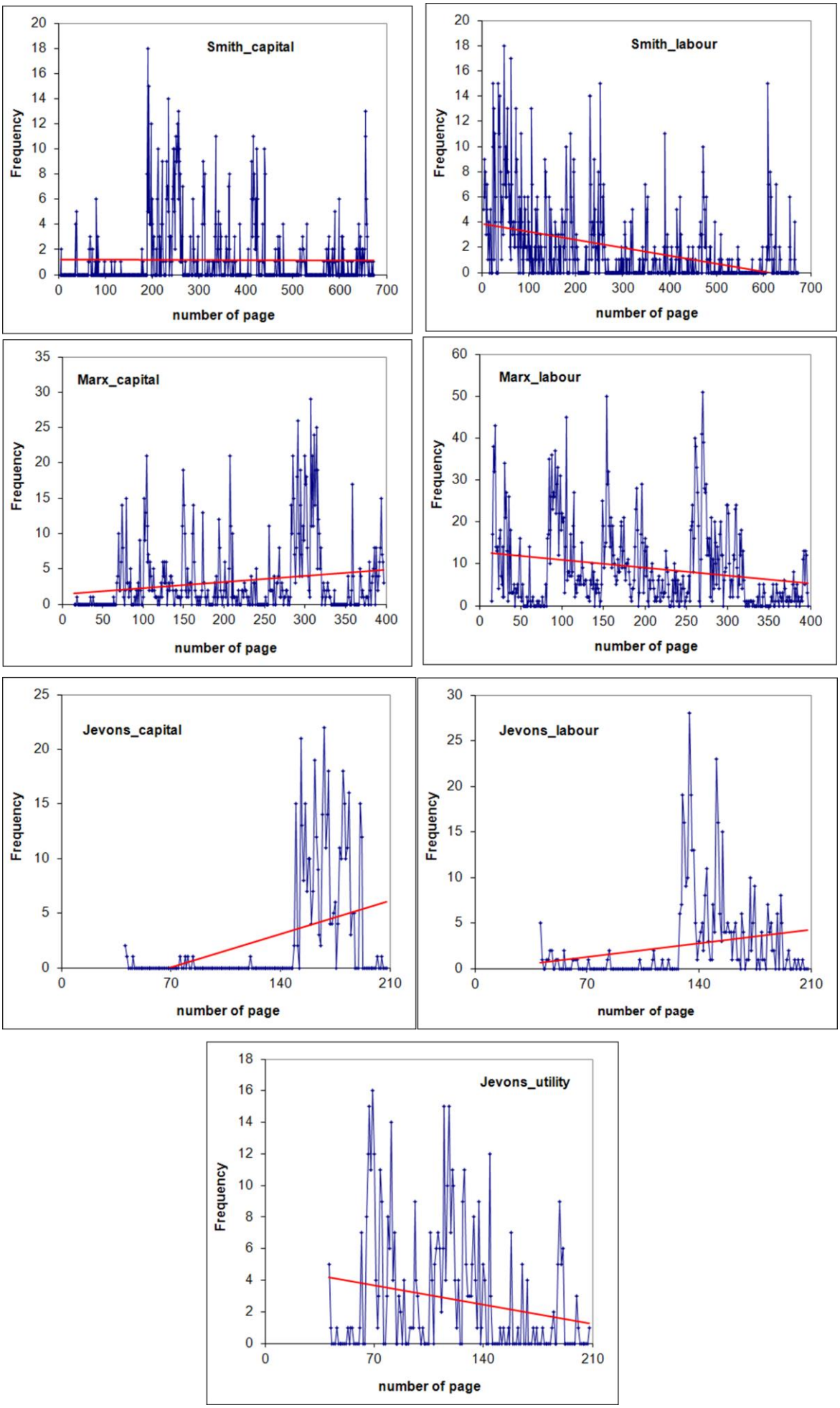


**Figure 2.** Frequency of appearance of the words analyzed.

From these, we may draw some initial conclusions about the theories. Actually, these graphs are versions of what in corpus linguistics/stylistics would be called dispersion plots[11]. Firstly, Smith uses labour and capital frequently throughout the *Wealth of Nations*. Since he uses those words to discuss all economic issues, it seems that both words are important for him. However, the trends of the frequencies show a slight decrease in the importance given to the word labour. The word capital is also cited throughout the book, and, although Book II on stock cites it more frequently, the steady trend shows that it is given practically equal importance in all sections. Certainly, although the first chapters of the *Wealth of nations* defend that the value of commodities depends on the labour commanded, Smith soon recognizes that the value reflects wages, rent of land and profit of capital. Once explained the concept of labour commanded, capital acquires greater importance in Smith's theory. This seems to be true in special within a growing economy and, so, in the course of Book III on the progress of opulence of different nations. Not in vain the accumulation of stock determines the average wage of labour. Besides, as shown in the low frequency of the word labour in Book IV Smith criticism of the previous systems of political economy seem to be related to their disregard of labour. Finally, as we may imagine and we observe in the frequencies of the word labour in Book V, labour seems to be a source more than a reason of expense for the sovereign. However, as a general conclusion, Smith has a quite steady trend of frequencies.

Marx also talks about labour and capital throughout *Capital*. Although 'labour', the basis of value according to Marx, is actually cited twice as often as 'capital', the importance of labour declines as the importance of capital rises, and the greatest variability can be found around the page 300. Capital begins to dominate and acquires ever greater importance as we run through the text. This seems to be logical as in the first chapters Marx explains the character of commodities, of exchange and money in the more classical way and it is not until money is transformed into Capital,

11 See Hearst (1997) for one early approach in computational linguistic. The distribution of the words are quantified (and rendered comparable) by means of dispersion statistics (Gries 2008).

in part 2, that the word is more frequently named. Then, it is in part 7, The Accumulation of Capital, when Marx cites the most frequently the word capital - in special when denouncing the degree of exploitation of labour – power.

In contrast, the data show that labor and capital do not have the same value for Jevons in all sections of the text. The two words do not even appear until they suddenly burst onto the scene, when the number of words jumps significantly. Curiously, the peaks for both words are quite similar. Moreover, the trend in the frequency for both words slopes slightly upwards, as the importance given to the words is greater in the late chapters. In the case of the word utility, however, the variability throughout the text is more stable, although Jevons does not cite this word more frequently than the other ones. In this case, the relative frequencies of the key terms are related to their place in the overall structure of exposition. Capital and labor are only cited when trying to explain the classical theory of value determined by the cost of production and also in the chapters that deals with the theory of labor and the theory of rent and capital. Certainly, labor understood in merely utilitarian terms, as "the amount of work done by certain muscles and the rate of fatigue" (Jevons, 1871, 207) is only one more example of pleasure-pain over a period of time, something that creates a useful effort in exchange for a wage, but cannot be key in Jevons' theory. As Jevons says, labor determines supply, supply determines final degree of utility and final degree of utility determines value. In the word capital, Jevons himself explains the lack of information transmission within the text saying that "in considering the nature and principles of Capital, we enter a distinct branch of our subject. There is no close or necessary connection between the employment of capital and the processes of exchange" (Jevons, 1871, 223). It is with good reason that, as he says, capital is concerned with time - the period of time for which it is invested that amounts to a lengthening of the period of production or the time elapsing between the beginning and end of a work. A theory based on a subjective and successive perception of time can run into difficulties when dealing with time itself.

## 6. Variability and intermittency

These remarks being made, we go on to analyse the scaling behaviour displayed by the different study cases. With this purpose in mind, Figure 3 represents the probability mass functions (the equivalent in discrete probability distributions to the PDFs that define continuous distributions) corresponding to each case for the two scales considered. One page, naturally the smaller scale that we may study, will furnish us with information about the variability between one page and the next, and five pages, that reports on the longer run variations, but yet at a distance small enough to maintain a relationship between the frequencies of presence of the word. In this case, it is not so much the theoretical considerations that we have to take into account; this statistical analysis does not show the frequency, but the "variation of the frequency" of some words, which may result from a certain "state of mind" of the author. Frequency may have to do with the conscious desire of using intensively a word; “variation of frequency” is somewhat an unconscious product as it has to do with a certain path of the melody. Frequency may be interpreted as cognition; “variation of frequency” as metacognition.

In Jevons case, we have represented the probability mass functions of the part of the book in which the words are actually present; a practically null citing of the words until the page 120 would have distorted the results giving Jevons a higher persistence than his state of mind would suggest. Obviously, not citing the word on one page or on the next implies persistence, but it does not mean anything in terms of the information flow or transmission. Quite simply, there is no information available. Besides, we have only represented the probability mass functions of the word “utility” in this author as he is the only one to use this word significantly.

The first remarkable result is that there is a hierarchy in the persistence of both labour and capital for different authors, a pattern of variation that seems to signal a certain authors' state of mind. Let us remember that the high persistence or probability that the number of words in one page

and the next in the scale is the same implies that the probability mass functions are taller in x=0. Smith is clearly the most persistent author. Marx wording is less persistent, and then, Jevons displays the lesser persistence in both words, as shown by the low 0,4 probability of persistence in the y axis of the graphs. Besides, for both Marx and Smith capital is more persistent than labour – this can be related with the above mentioned downward slope of the frequency maps of the word labour and the lesser importance given to it as the texts progress.

Our interpretation of this result is the following: utilitarian authors may be, as their theories, goal-oriented and, so, they are constrained by an image embedded in the authors' psyche. They take aim at something unreal that leads them to an anxious and turbulent perception of time. The image of utility, being the product of imagination, is moody and volatile (see Trincado 2003, 2006, 2009). However, persistence may imply that the author is being guided by something which is not volatile, maybe reality.

In this sense, Smith seems to be the more persistent author. He seems to seek a clear objective with his wording. His line of thought is more coherent, meaning that what he discusses on a given page will probably be explained on the next one. In the *Wealth of Nations* there is a continuous use of the words capital and labour and when Smith uses those words, he does so with the clear purpose of providing understanding to the reader. His coherence was also displayed in the first analysis of frequencies: he talks about labour more frequently in the first chapters on the productive powers of labour, and about capital on Book II on the nature and employment of stock. Indeed, according to Smith, language must aim at the natural precision and simplicity of both the spoken and written word (Smith, 1983). According to his theory, value is objective and based on reciprocity.

Marx however seems to be less persistent. According to our hypothesis, this implies that his

mind is more “utilitarian” than that of Smith, although less than that of Jevons. Marx did not much value utilitarian theory (Marx, 1867, 426). Marx affirms that a use value “has value only because human labour in the abstract has been embodied or materialised in it” (Marx, 1867, 28). However, he writes about a system that he thinks that will collapse – a breakdown that Marx actually desired[12] He, then, seems to be led by his image of the system criticized and maybe with the image of the desired system underlying his thought. He wrote his works with an eye on a new free foreseeable world. Then, his perception of time may have led Marx to already have in mind that a rupture would take place, like a mirror image that gives the theory a desire to find discontinuity.[13]

Finally, Jevons clearly is the less persistent author, and the information flow seems to be the least coherent one. Capital and labour, words considered at first by him so important, are only cited in his work in a burst, being obliterated by the continuous use of the word utility that, however, declines in importance over the text. Jevons explicitly recognizes that his theory is linked to the founder of the utilitarian radical philosopher, Jeremy Bentham (Jevons 1871, 29).This search for utility is a volatile activity that depends on changing feelings - actually based on a concept of perception by impressions sent from “outside” to the inner self. Science is based on the summing up of the subjective feeling of pleasure of different people. [14]

Besides, the probability mass functions corresponding to Smith have longer tails and, so, in Smith's case there are more rare events than in Marx's case, and Marx also displays longer probability mass functions than Jevons. Actually, persistence is usually related to these large jumps

12 In his doom watcher theory, a destructive embryo of capital keeps growing until it breaks the shell and the system's equilibrium. Men are controlled by the inevitability of the future. They are moved by forces that will lead them to violence and eventually to communism. Marx tries to give an objective definition of capital and capitalism, but his own contemptuous tone is his undoing (Marx, 1867, 532).

13 Marx accepted that utilitarianism is the basis for capitalistic behaviour, although eventually, his desire to change the system may lead him to blow out of proportion the possible utilitarian bias of capitalism.

14 According to Jevons, economics should be regarded as a calculus of pleasure and pain or, more correctly, of positive and negative feeling. The idea is that we need to study the psychological basis of economics as one branch of the science of human behaviour to arrive at sound conclusions that will maximise human happiness when they are applied.

and large jumps cannot be explained by the information transmitted – information must have been lost or there has been an exogenous shock affecting. In one page the same word appears a lot but in the next there are very few of that word: maybe inattention or a change of mind makes the author not use that information continuously in his line of thought. Then, a persistence of the objective is also accompanied by sudden movements, maybe produced by hard reality or maybe due to the search by the author himself to achieve clarity for the understanding of the reader and for a full comprehension of reality. Jumps are the sign of an attempt not to understand the mind itself, but reality.

Finally, another striking result is that all the probability mass functions for all the three authors and words display a similar behavior regarding the phenomenon intermittency of the turbulence. That is shown by the different shape of the probability mass functions in scale 1 and scale 5 (Figure 3). This intermittency means that the rhythm at which the information is lost is different in different scales. The information loss rate is greater in the short run than in the long run. When we compare a page with the one immediately after it, we see that there is a greater persistence in the number of times that a word appears (more pointed probability mass functions) than when comparing the same page to one five pages after. At the same time, it means that the tails are more stretched for the scale 1 and, there is thus a higher probability of finding rare events or unexplained jumps which are not linked to the previous information. In the long run (in scale 5) the jumps in the number of apparitions of the word is less important. Consequently, the long run displays the author's basic line of thought as there is less loss of information and information in one page seems to be drawn from information in the other page. Although it is more probable that the number of words in two correlative pages is similar than the number of words 5 pages apart, in the short run we may find jumps that may not be found in the long run. Those jumps do not transmit information as they come back to the general trend: they are, in some sense, nervous movements of the mind. The jumps in information in the short run do not take into account the previous

information and probably, they do not transmit information immediately to the next page.[15] These jumps may be part of the effort of communication. They are comings and goings, struggles for coming to the idea and for making others understand what we are trying to convey. In short, a signal of the difficulties of communication becomes apparent. In this sense, all the texts might have some level of this phenomenon, but the logic of the analysis leads to variations on the frequency of apparition of crucial concepts. In this case, we propose that the cause of variation is the fact that the author try to stick to reality or not. But, the greater or lesser intermittency may also be understood in terms of other issue: the novelty of the theory. Let us explain this further.

In financial markets analysis, jumps represent great increases or decreases of the indexes which can be associated to explosive pieces of information that affect the index in the short term (Vindel and Trincado, 2010). Therefore, these jumps usually appear at small time scales and their presence is associated to sporadic and rare events; if those events became usual, then the market would foresee their appearance, making them less explosive. Besides when seeing a jump or large price movements, it is reasonable to assume that people might stop their trading activities to reevaluate their investing options (Amatyakul, 2009). The traditional companies that are a great source of common knowledge tend to have less sharp jumps than the high-tech innovate firms. Although the context is different, an analogy could be made with the theories of the three authors studied. In both cases we are talking about a certain flow of information. New theories are not being based on common knowledge. So, there must create jumps in information in the short run. When the jump happens, the author stops its line of thought and re-evaluates his ideas. Conversely, words understood in a traditional way must have a lesser degree of intermittency. Indeed, that happens with the word utility, the probability mass functions of which we have represented in Figure 3. Jevons considered utility in the traditional way, and actually he defines utility as Bentham did, as

15 If there were indubitable statistical evidence of persistence - e.g. of uniform or full probability of zero differences in the frequency of these terms on every page, this will mean that the information flow is very steady; however, the concept of intermittency of the turbulence tries to assess the different path of frequencies in consecutive pages or in 1 to 5 pages and, so, if it is possible that the transmission is made in the long run or in the short run.

Bentham “put forward the Utilitarian theory in the most uncompromising manner (…) The words of Bentham on this subject may require some explanation and qualification, but they are too grand and too full of truth to be omitted" (Jevons, 1965, 23).

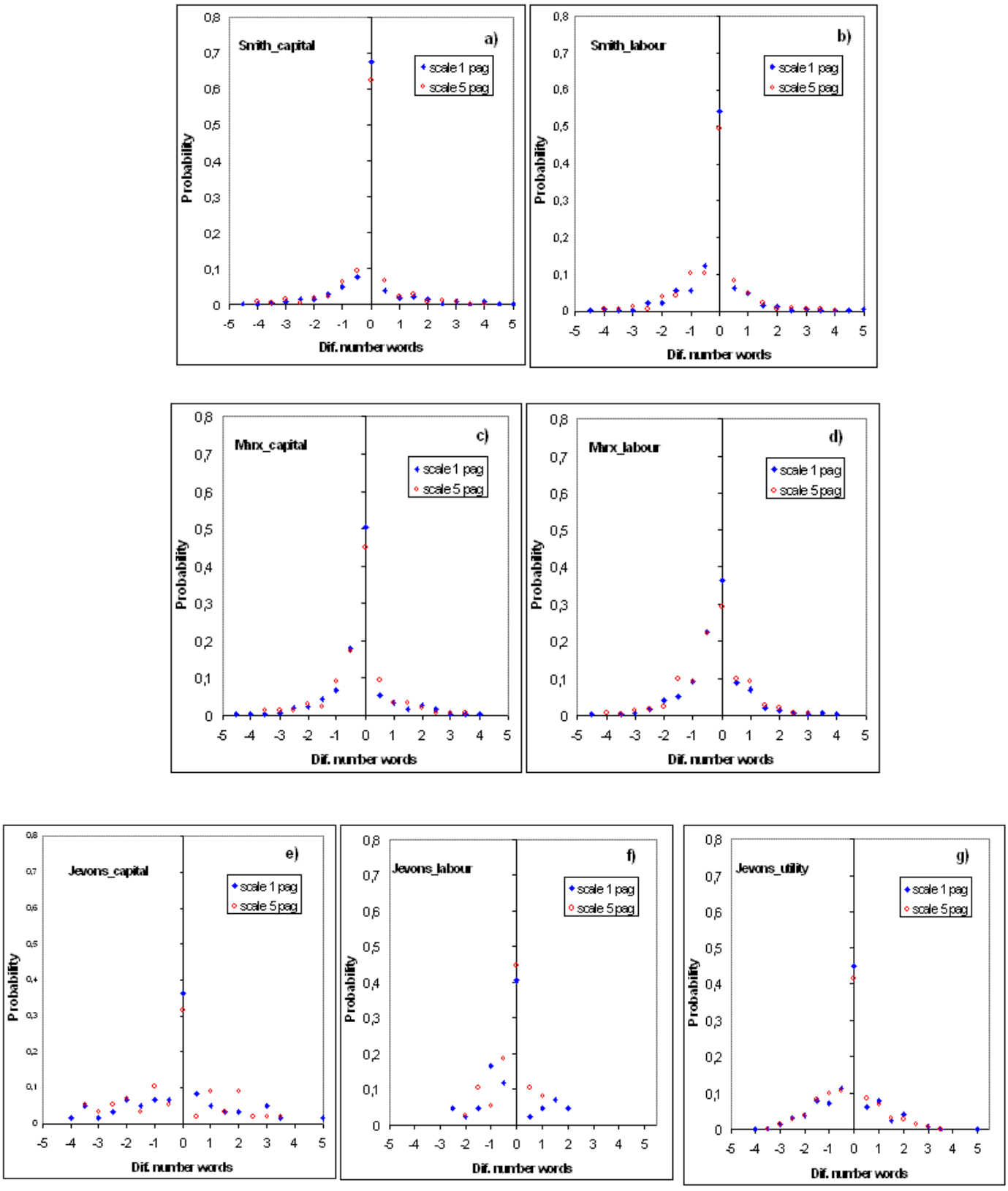


**Figure 3.** Probability mass functions corresponding to scales 1 and 5 for the words analysed

## CONCLUSIONS

This article intends to be a programmatic study. It does some spadework by using some statistical tools for the study of texts in the history of economic thought. This analysis provides us with further explanations of some theoretical issues. In this respect, the first conclusion to be drawn is that we may use a mathematical analysis for the study of the texts. Besides, this scientific research programme brings closer the methodology currently used in economics and in other scientific disciples to the one used in the history of economic thought. This article, then, calls for inter-disciplinarily and for the interest of different disciples – history of economic thought, statistics, econometrics, philology, psychology, mathematics or physics.

This study could even be understood as a tool of the experimental field of neuropsychology: only from the recording of electrical activity in the brain functioning of living creatures may we draw information; but words used in a text may provide us with at least some information about the continuity of the thought of a dead man.

In any case, the results obtained could not have been drawn from a classical textual analysis. In particular, we have explored the possibility that the perception of time by different authors may affect the wording of their texts. The analysis made provides empirical support for such a general psychological theory and is an inferential study with an empirical test using frequency of keyword analysis. The conclusions have been drawn from the analysis of the frequency of appearance of the words “capital, labour and utility” and from the variation of this frequency of appearance at different scales in Jevons, Smith and Marx's masterpieces.

In particular, Jevons gives a lot of importance to the word utility and the variability throughout the text in the use of the word is quite stable. Conversely, the importance that Jevons gives to labor and capital is greater in the late chapters. For those words there is a lack of

information transmission within the text. In contrast, Marx and Smith do not give any importance to utility, something that may be evidence of their anti-utilitarian basis. Smith can deal with the accumulation in time – in the present - more easily than a subjectivist theory based on the succession of pleasures. For Marx the importance of labour also declines as the importance of capital raises, the same as in his analytical conclusions.

Besides, there is a hierarchy in the persistence of both labour and capital for different authors. This hierarchy seems to signal the authors' state of mind. Smith is clearly the most persistent author, Marx wording is less persistent and Jevons displays the lesser persistence in both words. Besides, for both Marx and Smith capital is more persistent than labour. If we consider persistence synonymous with the author not being guided by an image, Smith seems to be the less utilitarian author.

Finally, the probability mass functions for all the three authors and words display a similar behavior regarding the intermittency of the turbulence. So, the information loss rate is greater in the short run than in the long run for the three authors, and there is thus a higher probability of finding in the short run unexplained jumps which are not linked to the previous information. Those jumps do not transmit information as they come back to the general trend: they are a signal of the difficulties of communication. The long run displays the author's basic line of thought.

However, as we have seen, the greater or lesser intermittency may be understood also in terms of the novelty of the analysis of the different authors. Words understood in a traditional way must have a lesser degree of intermittency. Indeed, the words labour and capital – with a high variability between scales - are analysed in an innovative way by all of the authors; but the word utility – with a low variability between scales - is considered by Jevons in the traditional way. This analysis, then, may provide a way to assess originality in literature and, thus, furnish us with a way

to defend the authors from the charge of plagiarism (the, it is a new defending argument in the case Salim v. Smith! see Rashid, 1990).

We cannot give up to the temptation of addressing possible lines of study for improving the analysis in future research. Certainly, the quantitative framework hereby proposed could be expanded by providing statistical analysis of text wording for different distances or for more words or group of words. Besides, more similar studies may be studied to see if we can draw some new conclusions from the analysis. These conclusions should have to do with the flow of though: it may be, then, the attention or inattention put by the author on the passage and, so, the emphasis placed on it; or the anxiousness of the author in terms of the hostility he thinks he will find in his audience towards his theory. In short, there are many other possibilities of interpreting the results. Our interpretation of the data is only our option made; an option that we hope will be of interest to the academia.

**REFERENCES**

Amatyakul, P. (2009): *The relationship between trading volume and jump processes in financial markets,* Durham: Duke University.

Aragón, J. L, Naumis, G. Bai, M., Torres M. and Maini, P. K. (2006): "Turbulent luminance in impassioned van Gogh paintings" arXiv:physics/0606246v2 [physics.flu-dyn] for this version).

Barndorff-Nielsen, O. E., Blaesild, P. and Schmiegel. J. (2004): "A parsimonious and universal description of turbulent velocity increments" *European Physical Journal B* 41, 345-363.

Bentham, J. (1843): "Table of the springs of actions" *The Works of Jeremy Bentham, vol. 1 (Principles of Morals and Legislation, Fragment on Government, Civil Code, Penal Law),* published under the Superintendence of his Executor, John Bowring, Edinburgh: William Tait, 11 vols. Vol. 1.

Castaing, B., Gagne, Y. and Hopfinger, E J. (1990): "Velocity probability density fluctuations of high Reynolds number turbulence" *Physica D* 46**:** 177-200.

Cattoni, D., Ozu M., and Chara. O. (2004): "Ruidos en la naturaleza" *ANALES AFA. Tandil, Argentina:*

*CRECIC* 16: 294-299.

Chevillard, L, Castaing B. and Lévêque E. (2005): "On the rapid increase of intermittency in the near-dissipation range of fully developed turbulence" *European Physical Journal B* 45**:** 561-567.

Di Matteo, T. (2007): "Multi-scaling in finance" *Quantitative Finance* 7(1): 21-36.

Edgar, G. (2004): *Classics on Fractals*. Boulder, CO: Westview Press.

Frisch, U. (1995): *Turbulence*. Cambridge: Cambridge University Press.

Gabaix, X. (1999): "Zipf's Law for Cities: An Explanation" *Quarterly Journal of Economics* 114(3): 739–767.

Gagne, Y, Marchand M. and Castaing. B. (1994): "Conditional velocity pdf in 3-D turbulence" *Journal de Physique II France* 4: 1-8.

Gries, S. (2008): "Dispersions and adjusted frequencies in corpora" *International Journal of Corpus Linguistics*, 13:4 , 403–437.

Hearst, M. (1997): "TextTiling: Segmenting Text into Multi-Paragraph Subtopic Passages" *Computational Linguistics*, 23 (1), pp. 33-64, March.

Hurst, H. E. (1951): "Long-term storage capacity of reservoirs" *Transactions of the American Society of Civil Engineers*. 116: 770–808.

Jermolowicz, R. (2004): "Remarks on Adam Smith Lectures on Rhetoric and Belles Lettres" *Studies in Logic, Grammar and Rhetoric* 7 (20): 199- 204.

Jevons, W. S. (1965): *The theory of political economy.* Fifth edition, New York: Agustus M Kelley: Bookseller.

Kali, R. (2003): "The city as a giant component: a random graph approach to Zipf's law" *Applied Economics Letters* 10 (4): 717-720.

Kolmogorov, A. N. (1941): **"**Dissipation of energy in locally isotropic turbulence" *C. R. Academia of Sciences USSR* 32**:** 16-18.

Kolmogorov, A. N. (1962): "A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number" *Journal Fluid Mechanics*. 13**:** 82-85.

Lévy, P. P. (1925): *Calcul des probabilités.* Paris: Gauthier-Villars.

Loffredo, M. I. (2004): "On the statistical physics contribution to quantitative finance" *International Journal of modern Physics B* 18: 705-713.

Lux, T. (2001): “Turbulence in financial markets: the surprising explanatory power of simple cascade models” *Quantitative Finance* 1: 632-640.

Mandelbrot, B. B. (1972): “Possible refinements of the lognormal hypothesis concerning the distribution of energy dissipation in intermittent turbulence” In M Rosenblatt and C Van Atta (eds), *Statistical Models and Turbulence*. New York: Springer Verlag.

Mantegna, R. N. and Eugene, S. H. (1996): “Turbulence and financial markets” *Nature* 383: 587-588.

Mantegna, R N. and Stanley, H. E. (1997): “Stock Market Dynamics and Turbulence: Parallel Analysis of Fluctuation Phenomena” *Physica A* 239**:** 255-266.

Mantegna, R. N. and H. E. Stanley, 2000, *An Introduction to Econophysics*, Cambridge University Press, Cambridge.

Marx, K. (1887): *Capital A Critique of Political Economy Volume I Book One: The Process of Production* of Capital. Moscow: Progress Publishers.

McCloskey, D. N. (1983): “The Rhetoric of Economics” *Journal of Economic Literature* 21(2): 481–517.

Meneveau, C. and Sreenivasan. K. R. (1987): “Simple multifractal cascade model for fully developed turbulence” *Physical Review Letters.* 59**:** 1424-1427.

Mordant, N. Metz, P., Michel, O. and Pinton. J. F. (2001): “Measurement of lagrangian velocity in fully developed turbulence” *Physical Review Letters* 87 (21): 214501-214505.

Nicholas, D. and Ritchie. M. (1978): *Literature and Bibliometrics.* London: Clive Bingley.

Pareto, V. (1897): *Cours d'Economie Politique.* Lausanne: Rouge.

Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, A. and Stanley. H. E. (2000): “Econophysics: financial time series from a statistical physics point of view” *Physica A*, 279, 443–456.

Rashid, S. (1990): “Adam Smith's Acknowledgements: Neo-Plagiarism and the Wealth of Nations” *Journal of Libertarian Studies*, 9, no. 2 (Fall): 1-24.

Richardson, L. F., (1922): *Weather prediction by numerical process*. Cambridge: Cambridge University Press.

Ricoeur, P. (1984): *Time and Narrative,* Chicago: University of Chicago Press.

Ricoeur, P. (1986): *Fallible Man*, New York: Fordham University Press.

Ricoeur, P, (2004): *Memory, History, Forgetting.* Chicago: University of Chicago Press.

Ross, I. S., (1995): *The Life of Adam Smith.* Oxford: Clarendon Press.

Schütze, H., (1998). "Automatic word sense discrimination". *Computational Linguistics*, 24 (1): 97-123.

She, Z. S. and Levêque. E. (1994): "Universal scaling laws in fully developed turbulence" *Physical Review Letters.* 72**:** 336.

Smith, A. (1790): *The Theory of Moral Sentiments.* London: A. Millar [1759].

Smith, A. (1978): *Lectures on Jurisprudence* (The Glasgow Edition of the Works and correspondence of Adam Smith), R. L. Meek, D.D. Raphael and P.G. Stein (eds). Oxford: Clarendon Press [1762-3; 1766].

Smith, A. (1979): *An Inquiry into the Nature and Causes of the Wealth of Nations* (The Glasgow Edition of the Works and Correspondence of Adam Smith). R. H. Campbell, A. S. Skinner and W. B. Todd (eds). Oxford: Clarendon Press [1776].

Smith, A. (1983): *Lectures on Rhetoric and Belles Lettres* (The Glasgow Edition of the Works and Correspondence of Adam Smith). J.C. Bryce and A.S. Skinner (eds). Oxford: Clarendon Press [1759].

Sorriso-Valvo L., Carbone, V., Veltri, P., Politano, H., and Pouquet, A. (2000): "Non-Gaussian probability distribution functions in two dimensional magnetohydrodinamic turbulence" *Europhysics Letters* 51: 520-526.

Taylor, R., Micolich A. P. and Jonas D. (2010): "Fractal Expressionism: Can science be used to further our understanding of art?" *Phys.unsw.edu.au.* http://www.phys.unsw.edu.au/PHYSICS_!/FRACTAL_EXPRESSIONISM/fractal_taylor.html, Retrieved 2010-10-17.

Trincado, E. (2003): *Crítica a la doctrina de la Utilidad y Revisión de las teorías de Hume, Smith y Bentham*, Madrid: E-PrintsUCM.

Trincado, E. (2004): "Equity, utility and transaction costs: On the origin of judicial power in Adam Smith" *Storia del Pensiero Economico*, anno 1: 33-52.

Trincado, E. (2005): "Utility, Money and transaction costs: Authoritarian vs libertarian monetary policies" *History of Economic Ideas* XIII(1): 57-77.

Trincado, E. (2006): "Adam Smith criticism of the doctrine of utility: a theory of the Creative Present" In Montes, Leonidas y Schliesser, Eric. *New Voices on Adam Smith*, New York and London: Routledge: 313-327.

Trincado, E. (2007): "Costes de Transacción vs Costes de Jerarquía en la Provisión de Bienes: una interpretación de la teoría Smithiana" *Revista de Historia Económica. Journal of Iberian and Latin American*

*Economic History*, nº 2, año XXV: 261-292.

Trincado, E. (2009): *Crítica al Utilitarismo*, Madrid: Maia ediciones.

Trincado, Estrella; Vindel, Jose María (2013), "Not Fully Developed Turbulence in the Dow Jones Index", *Brazilian Journal of Physics*, vol. 43, n. 4, pp. 260-267.

Urrutia, J. (1983): *Economía Neoclásica*, Madrid, Pirámide.

Vindel, J. M. and Trincado, E. (2010): "The timing of information transmission in financial markets" *Physica A* 389(24): 5749-5758.

Vindel, J. M. and Yagüe, C., (2011): "Intermittency of turbulence in the atmospheric boundary layer: scaling exponents and stratification influence" *Boundary-Layer Meteorology* 140(1): 73-85.

Zipf, G. K. (1935): *The Psychobiology of Language*. Houghton-Mifflin.

Zipf, G. K. (1949): *Human Behavior and the Principle of Least Effort*. Cambridge, Massachusetts: Addison-Wesley.