Utrecht University
School of Economics

The engineering tools that shaped the rational expectations revolution

Marcel Boumans

U.S.E. Working Paper Series nr: 19-23

Utrecht University **School of Economics**

U.S.E. Research Institute
Kriekenpitplein 21-22, 3584 EC Utrecht, The
Netherlands Tel: +31 30 253 9800, e-mail:
use.ri@uu.nl www.uu.nl/use/research

Utrecht University School of Economics

U.S.E. Research Institute
Working Paper Series 19-23

The engineering tools that shaped the rational expectations revolution

Marcel Boumans
Utrecht School of Economics
Utrecht University

December 2019

Abstract

The rational expectations revolution was not only based on the introduction of Muth's idea of rational expectations to macroeconomics; the introduction of Muth's hypothesis cannot explain the more drastic change of the mathematical toolbox and concepts, research strategies, vocabulary, and questions since the 1980s. The main claim is that the shift from "Keynesian economics" to "new classical economics" is based on a shift from a control engineering approach to an information engineering methodology. The paper even shows that the "revolution" was more radical. The change of engineering tools has changed macroeconomics more deeply, not only its methodology but also its epistemology and ontology.

To show this shift in epistemology and ontology, the history of economics will be interwoven with the history of mathematics which cannot be detangled from the emergence of the digital computer and the influence of this emergence on the changed nature of mathematics: the adoption of a new concept of solution, no longer a number, a formula, or a function, but an algorithm. The result of this new concept of solution was a new approach to the analysis of processes.

Information engineering studies the fundamental limits in communication and finds its origins in Shannon's theory of communication, and incorporates the tools designed by Turing, Shannon, Kálmán, and Bellman. The resulting ontology of this kind of engineering is a world populated by machines that communicate with each other by exchanging information. This information does not, however, contain only signals about the system states but also noise that needs to be filtered out. It is not a deterministic world, but one governed by stochastic processes. The decisions these machines take is conditioned on the (noisy) information they have about the current state of the world but at the same time will affect future states. Policy in this world therefore means tracing an optimal trajectory taking all these issues into account.

Keywords: communication theory, control engineering, dynamic programming, information engineering, Kalman filter, rational expectations, Turing machine

JEL classification: B22, B23

Acknowledgements: I would like to thank the editors, Pedro G. Duarte and Yann Giraud, for encouraging me to revisit my earlier work on engineering and for nudging me to make it fit to this volume. I would also like to thank the other participants of the conference and the two anonymous referees for their thoughtful comments. Special thanks go to Mary S. Morgan who helped me enormously with the struggle I had to convey this paper's main claim.

Comments welcomed to: M.J.Boumans@uu.nl

The engineering tools that shaped the rational expectations revolution

Marcel Boumans[*]

1. Introduction

In 2011, a panel discussion was held marking the 50th anniversary of John Muth's "Rational Expectations and the Theory of Price Movements" (1961), with Michael Lovell, Robert Lucas, Dale Mortensen, Robert Shiller, and Neil Walace in the panel and Kevin D. Hoover and Warren Young (2013) as moderators. To open the discussion, Hoover remarked that: "Fifteen years later, it was commonplace to speak of a rational expectations revolution. And within another fifteen years, rational expectations had been fully integrated into macroeconomics" (Hoover and Young 2013: 1169-70). In his response, Lucas made clear that he did not like the term "revolution," because of its political connotation (see also Klamer 1984: 55-56), but when used in a Kuhnian sense, to mark a scientific change, however, he considered this term to be appropriate to denote what has happened since the 1970s: a change in scientific world view, research strategies, vocabulary, and questions.

At least at two occasions, Lucas explained what he saw as the nature of this revolution; it should be seen not as a change of theory, but as a technical development:

> One would expect developments to arise from two quite different kinds of forces outside the subdisciplines of monetary economics or business cycle theory. Of these forces the most important, I believe, in this area and in economic generally, consists of purely technical developments that enlarge our abilities to construct analogue economies. Here I would include both improvements in mathematical methods and improvements in computational capacity. (Lucas 1980: 697)
>
> We got that view from Smith and Ricardo, and there have never been any new paradigms or paradigm changes or shifts. Maybe there will be, but in two hundred years it hasn't happened yet. So you've got this kind of basic line of economic theory. And then I see the progressive – I don't want to say that everything is in Smith and Ricardo – the progressive element in economics as entirely technical: better mathematics, better mathematical formulation, better data, better data-processing methods, better statistical methods, better computational methods. (Lucas 2003: 22)

Considering the development of economics mainly driven by technical innovations, the rational expectations revolution should, however, be distinguished from an earlier technical development, called the "Keynesian," which was "the evolution of macroeconomics into a quantitative, scientific discipline, the development of explicit statistical descriptions of economic behavior, the increasing reliance of government officials on technical economic expertise, and the introduction of the use of mathematical control theory to manage an economy" (Lucas and Sargent 1979: 50).

---

[*] Correspondence may be addressed to Kriekenpitplein 21-22, 3584 EC Utrecht, Netherlands. E-mail: m.j.boumans@uu.nl. I would like to thank the editors, Pedro G. Duarte and Yann Giraud, for encouraging me to revisit my earlier work on engineering and for nudging me to make it fit to this volume. I would also like to thank the other participants of the conference and the two anonymous referees for their thoughtful comments. Special thanks go to Mary S. Morgan who helped me enormously with the struggle I had to convey this paper's main claim.

In two mainly reflective papers (1977 and 1980), Lucas discussed the outlines of the new methodology that he supposed to be needed for the kind of macroeconomics he was promoting. The first thing that stands out when reading both articles is the terminology Lucas used. Agents are "processing noisy signals" by "smoothing" them for which they use "information systems." Elsewhere (Boumans 1997, 2005), I have shown that Lucas introduced in these two articles Turing's approach by considering models as imitation games and hence should be validated by a Turing test.

When a description of the information system is seen as an essential element of a macro-economic model, the modeler should provide an explicit account of the nature of expectations-formation. It is only in this context that Lucas suggested that expectation is "rational" in the sense of Muth:

> To practice economics, we need *some* way (short of psychoanalysis, one hopes) of understanding *which* decision problem agents are solving. John Muth (1961) proposed to resolve this problem by identifying agents subjective probabilities with observed frequencies of the events to be forecast, or with "true" probabilities, calling the assumed coincidence of subjective and "true" probabilities *rational expectations*. (Lucas 1977: 15)

While one can question one part of Lucas's historical claim that scientific development in macroeconomics is only technical – which will not be discussed here –, the paper instead will show, and thereby confirm Lucas's claim that what he called the Keynesian revolution was based on a different technical development than the rational expectations revolution. This paper shows that the rational expectations revolution is based on the introduction of a different type of engineering mathematics. Control engineering was replaced by information engineering, that is to say, the framework of feedback loops was replaced by a framework of information processing. The rational expectations revolution was not only based on the introduction of Muth's idea of rational expectations to macroeconomics; the introduction of Muth's hypothesis cannot explain the more drastic change in macroeconomics, namely the change of the mathematical toolbox and concepts in macroeconomics since the 1980s.

The claim that the major shift from "Keynesian economics" to "new classical economics" is based on a shift from a control engineering approach to a different kind of engineering methodology is not new. Sent (1997: 57) makes a similar claim, "new classical economists criticized the use of control engineering in economics and adopted other engineering techniques instead … tools of optimal prediction and filtering." I, however, disagree with her view that "new classical economists realized that the use of control techniques was not only possible but essential under rational expectations" (58). Although some of the control techniques may still be used by the new classical economists, they were not "essential" anymore to their new approach. In my view, the "revolution" was even more radical. The change of tools has changed macroeconomics more deeply, not only its methodology but also its epistemology and ontology.

To show this 1970s shift in epistemology and ontology, the history of economics needs first to be interwoven with the history of engineering mathematics which cannot be detangled from the emergence of the digital computer and the influence of this emergence on the changed nature of mathematics: "The single most impressive thing that has happened is the arrival on the industrial, and the mathematical, scene of the large high speed digital computer" (McMillan 1962: 86). In his Presidential address given before the Society for Industrial and Applied Mathematics in 1961, Brockway McMillan presented an overview of what he considered to be the major trends since 1940.[1] It is in particular the rise of new domains such as statistical theory of communication, the theory of feedback and control, the theory of games, and linear programming which were relevant for the change of the nature of mathematics. Connected to this change, McMillan observed a new "attitude" to mathematical problems:

> Mathematicians, both the applied and some of the pure ones, have caught up with the logicians and have adopted a new concept of solution, a concept of what it means to solve a problem. A solution is no longer a number, a formula, or a function describing dependence upon a parameter. A solution may simply be an algorithm which is known to terminate in a reasonable number of steps and is available on cards or tape for one's own brand of computer. (McMillan 1962: 89)

The result of this new concept of solution is a new approach, which McMillan called the "operational method in human affairs,"

> the approach to a process in terms of its observables, its inputs, outputs, criteria, choices, without anthropocentric references to intentions, purposes, authority, or even "meaning". This is the way we have been forced to approach the mechanization of routine human data handling activities. (McMillan 1962: 90)

This method had a stronger influence on a new field called information engineering than control engineering. Control engineering maintained the anthropocentric references, by focusing on the human-machine relationship, and references to intentions and purposes.

The term "information engineering" will be used in this paper to distinguish a specific domain within mathematical engineering from control engineering. Because both engineering branches are not uniquely specified, my distinction is based on the differences between control theory and information theory. Control theory is the theory of feedback systems. According to Andrei (2005: 3), there are three fundamental concepts in control theory: feedback, need for fluctuations, and optimization. A textbook on information theory (Yeung 2008: 3) describes information theory as "the science of information" that "studies the fundamental limits in communication regardless of the technologies involved in the actual implementation of the communication systems" and finds its origins in Claude E. Shannon's (1948) "Mathematical Theory of Communication." To clarify the core concepts, tools and strategies of information engineering, they will be discussed as they were first introduced into mathematical engineering and computer science. In the subsequent section it will be shown how these tools have shaped Lucas's papers of the early 1970s. This article will focus only on five papers by Lucas (1971 with Prescott, 1972a, 1972b, 1973, 1975). This suffices for the main claim of this paper that the rational expectations revolution was based on a change of engineering tools. It does not mean to suggest that others did not co-create the rational expectations revolution and thus contributed to the development of the new tools, but it was Lucas alone who received the Nobel Prize in 1995, "for having developed and applied the hypothesis of rational expectations, and thereby having transformed macroeconomic analysis and deepened our understanding of economic policy" (NobelPrize.org).

---

[1] This paper was brought to my attention when reading Klein 2001.

2. The rise of the operational method in human affairs

> Our task as I see it … is to write a FORTRAN program that will accept specific economic policy rules as "input" and will generate as "output" statistics describing the operating characteristics of time series we care about.
> —Robert E. Lucas, Methods and problems in business cycle theory (1980)

To understand which tools Lucas used to shape a new ontology for macroeconomics, particularly in relation to the emergence of the computer, one can choose between two starting points: John von Neumann and Alan Turing. It has already been shown that for game theory and linear programming, von Neumann is the most appropriate starting point, but for statistical theories of communication and information it is better to start with Alan Turing.

In his 1936 paper, 'On computable numbers, with an application to the Entscheidungsproblem,' Turing introduced the idea of an idealized human computer. According to Soare (1996), this paper is "monumental" because:

> (1) Turing analyzed an idealized *human* computing agent (a "*computer*") …; (2) Turing specified a remarkably simple formal device (*Turing machine*) and proved the equivalence of (1) and (2); (3) Turing proved the unsolvability of Hilbert's *Entscheidungsproblem* …; (4) Turing proposed a *universal* Turing machine, one which carried within it the capacity to duplicate any other …. (Soare 1996: 291)

This idealized human computing agent, a "computing machine," which is only capable of a finite number of conditions, called "*m*-configurations," is supplied with a "tape," "the analogue of paper" (Turing 1936: 231), moving back and forth past the machine. The tape is divided into "squares," each capable of bearing a "symbol." At any moment there is just one square which is "in the machine," which is called the "scanned square." "The 'scanned symbol' is the only one of which the machine is, so to speak, 'directly aware'" (231). The possible behavior of the machine at any moment is determined by the *m*-configuration and the scanned symbol. The machine is able to erase the symbol on the scanned square, to print a symbol on the scanned square, and to move the tape to the left or right, one square at a time. In addition to these operations the *m*-configuration may be changed. This device of *m*-configurations functions as a simple memory. As Turing writes, "by altering its *m*-configuration the machine can effectively remember some of the symbols which it has 'seen' (scanned) previously" (231).

Turing (1936) emphasized that the behavior of the machine is similar to the computations of a human being. "Computing is normally done by writing certain symbols on paper. We may suppose this paper is divided into squares like a child's arithmetic book" (249). The behavior of a *human* computer then "at any moment is determined by the symbols which he is observing, and his 'state of mind' at that moment" (Turing 1936: 250).[2] Each "state of mind" corresponds to an *m*-configuration. Turing aimed to define "computing" not as a psychological process but to externalize it as operations on paper. He therefore wished to define the "state of mind" by a "more physical and definite counterpart":

---

[2] In the 1930s the term 'computer' always referred to a human making computations. In the late 1940s it became common to use this term as a reference to a machine, though Turing still used the term computer to refer to a human with paper (Soare 1996: 291).

> It is always possible for the computer to break off from his work, to go away and forget all about it, and later to come back and go on with it. If he does this he must leave a note of instructions (written in some standard form) explaining how the work is to be continued. (Turing 1936: 253)

This “note of instructions” is therefore the counterpart of the “state of mind.”

In Turing’s 1936 paper, machine operations were taken as being similar to a specific domain of human thinking, namely computations. With the building of real machine computers, the comparisons became more ambitious and were extended to “thinking” more generally. In the same year as Turing published his ‘Computing machinery and intelligence’ in which he introduced what later came to be known as the Turing test, Shannon published two papers on a chess-playing machine. In the one written for a more general audience, Shannon ([1950] 1988: 2099) explicitly discusses the question “Could a machine be designed that would be capable of ‘thinking’?” Chess-playing machines were considered to be “an ideal one to start with,” because “chess is generally considered to require ‘thinking’ for skillful play; a solution of this problem will force us either to admit the possibility of a mechanized thinking or to further restrict our concept of ‘thinking’” (Shannon 1950: 257).

A comparison of their discussions of the question whether machines can think shows how similar the views of Shannon and Turing were:

> From a behavioristic point of view, the machine acts as though it were thinking. It has always been considered that skillful chess play requires the reasoning faculty. If we regard thinking as a property of external actions rather than internal method the machine is surely thinking. The thinking process is considered by some psychologists to be essentially characterized by the following steps: various possible solutions of a problem are tried out mentally or symbolically without actually being carried out physically; the best solution is selected by a mental evaluation of the results of these trails; and the solution found in his way is then acted upon. It will be seen that this is almost an exact description of how a chess-playing computer operates, provide we substitute “within a machine” for “mentally”. (Shannon [1950] 1988: 2107)

Shannon is, however, better known for his ‘Mathematical theory of communication’ (1948), which aims to extent the theory of communication by including “the effect of noise in the channel” (1). The “fundamental” problem of communication, according to Shannon, is that of “reproducing at one point either exactly or approximately a message selected at another point,” but the “semantic aspects of communication are irrelevant to the engineering problem” (1).

Shannon (1948) discusses three categories of communication systems, discrete, continuous and mixed. A discrete system is one in which both the message and the signal are a sequence of discrete symbols. As an example of such a system, Shannon mentions telegraphy where the message is a sequence of letters and the signal a sequence of dots, dashes and spaces. A continuous system is one in which the message and signal are both treated as continuous functions, e.g., radio or television. A mixed system is one in which both discrete and continuous variables appear. The largest part of this 1948 paper discusses the discrete case, which has “applications not only in communication theory, but also in the theory of computing machines” (3).

The discrete information source is considered as a stochastic process, a sequence of symbols governed by a set of probabilities. In particular it is a Markov process, which can be described as follows: There exist a finite number of possible "states" of a system S_1, S_2, … S_n, and there is a set of transition probabilities; $p_i(j)$ the probability that if the system is in state S_i it will next go to state S_j. Among the possible discrete Markov processes there is a group with special properties of significance in communication theory, namely the "ergodic" processes. Ergodic processes have transition probabilities with the same constant values.

In his discussion of continuous systems, Shannon refers to Norbert Wiener as an important source for his own work:

> Communication theory is heavily indebted to Wiener for much of its basic philosophy and theory. His classic NDRC report, *The Interpolation, Extrapolation and Smoothing of Stationary Time Series* (Wiley, 1949), contains the first clear-cut formulation of communication theory as a statistical problem, the study of operations on time series. This work, although chiefly concerned with the linear prediction and filtering problem, is an important collateral reference in connection with the present paper. (Shannon 1948: 34)[3]

Wiener's *Interpolation, Extrapolation and Smoothing of Stationary Time Series* is restricted to continuous systems, and therefore Shannon discusses filters only in terms of these systems.

The purpose of Wiener's NDRC report (1949) is to "unite" time series in statistics and communication theory. Communication theory, according to Wiener is the "study of messages and their transmission," and because a "message to be transmitted is developed into time series" (2), fusing statistical techniques with communication engineering, such as Fourier analysis, into a "common technique" is "more effective than either existing technique alone" (9).

Wiener mentions three "things which we can do with time series or messages" (9): prediction, filtering and answering questions of policy. Prediction is the estimation of the continuation of a series that is most probable, and filtering or "purification" (9) is the estimation of what the data would have been without being corrupted or altered by mixture with other time series. While prediction and filtering operate on data without direct information as to how these data might have been altered, questions of policy require an "intrinsic" study of time series: "to ascertain how certain series dealing with the economics of a country might have changed if a different system of taxation had been adopted" (11).

To commemorate the twenty-fifth anniversary of the publication of Wiener's 1949 monograph, the *IEE Transactions on Information Theory* published a survey of "three decades of linear filtering" (Kailath 1974). According to Thomas Kailath, Wiener's work was "the direct cause for the great activity of the last three decades in signal estimation, but it was perhaps the greatest factor in bringing the statistical point of view clearly into communication theory and also control theory" (146).

The key problem of linear filtering is the determination of the linear least-squares (i.e. optimal) estimate of a signal process corrupted by additive white noise. For information

---

[3] In addition to this, Shannon (1948: 52) mentions in the Acknowledgements, that "Credit should also be given to Professor N. Wiener, whose elegant solution of the problems of filtering and prediction of stationary ensembles has considerably influenced the writer's thinking in this field."

theorists and communication engineers at that time, the conventional specification of the filtering problem was in terms of signal and noise covariance functions, and the Wiener filter, designed to tackle this kind of problems, was an obvious method to use. Rudolf Emil Kálmán replaced this approach by one in which state-space models – borrowed from Bellman – instead of covariance functions were specified for the signal and noise, and “it seemed to many that this difference in specification was the chief reason for the success of the Kalman filter” (153):

> The fact that the Kalman filter dealt with state-estimation made it comparatively easy to include it in books and courses in state-variable theory, without having to go very deeply into estimation theory or even into Wiener filtering. (Kailath 1974: 152)

The Kalman filter is not given by an explicit formula for the impulse response of the optimal filter – as in the case of Wiener filters – but as an algorithm suitable for direct evaluation by computers. As Kálmán (1960) noted, the Wiener filtering methods were subject to “a number of limitations which seriously curtail their practical usefulness”:

> (1) The optimal filter is specified by its impulse response. It is not a simple task to synthesize the filter from such data.
> (2) Numerical determination of the optimal impulse response is often quite involved and poorly suited to machine computation. The situation gets rapidly worse with increasing complexity of the problem.
> (3) Important generalizations (e.g., growing-memory filters, nonstationary prediction) require new derivations, frequently of considerable difficulty to the nonspecialists.
> (4) The mathematics of the derivations are not transparent. Fundamental assumptions and their consequences tend to be obscured. (Kálmán 1960: 35)

Kálmán’s “new approach” introduced “a new look at this whole assemblage of problems, sidestepping the difficulties just mentioned” (35): the filtering problem is approached from the point of view of conditional distributions and expectations, so all statistical calculations and results are based on first and second moments, no other statistical data are needed.

To clarify later in this article how much Lucas’s approach was based on Kálmán’s “new approach,” we will now discuss it in more detail: In this approach, the random signals are represented as the output of a linear dynamic system excited by independent or uncorrelated random signals (white noise). The behavior of this system is described through a quantity known as the system “state,” which is specified by what is called the “system equation”:

$$x_{t+1} = F_t x_t + \eta_t \tag{1}$$

where x_t is an $(n\times 1)$ vector of state variables, F_t an $(n\times n)$ state transition matrix, and η_t an $(n\times 1)$ vector of system noise. Associated with the system equation is what is called the “measurement equation”:

$$y_t = H_t x_t + \varepsilon_t \tag{2}$$

where y_t is an $(m\times 1)$ vector of measurements, H_t an $(m\times n)$ measurement matrix and ε_t an $(m\times 1)$ vector of measurement noise. To complete the model specification, a description of the noise terms η_t and ε_t is required; their means and variance-covariance matrices are given as $E[\eta_t] = E[\varepsilon_t] = 0$, $Q = \mathrm{Var}(\eta_t)$ and $R = \mathrm{Var}(\varepsilon_t)$.

Given the state-space representation of a discrete linear dynamic system and the measurement equation, the problem is to estimate the state x_t from the noisy measurements y_1, y_2, …, y_t. The estimation of the state will also be based on the estimate of x_{t-1}. Let $\hat{x}_{t|t-1}$ denote such an estimate, where the subscript $t|t$-1 denotes that the estimate at time t used the information available up to time t-1. The estimate of x_t using all available information at time t is then:

$$\hat{x}_{t|t} = L_t \hat{x}_{t|t-1} + K_t y_t \tag{3}$$

where L_t and K_t are time-varying weighting matrices to be specified by imposing on the filter the conditions that the estimate should be unbiased and of minimal variance.

The unbiasedness is ensured by having

$$L_t = I - K_t H_t \tag{4}$$

which, when substituted in (3) gives

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t[y_t - H_t \hat{x}_{t|t-1}] \tag{5}$$

To meet the requirement of minimal variance K_t, called the Kalman gain, should be of the following shape:

$$K_t = P_{t|t-1} H_t^T \left[H_t P_{t|t-1} H_t^T + R \right]^{-1} \tag{6}$$

where $P_{t|t-1} = Var(\hat{\varepsilon}_{t|t-1})$, that is to say, the variance-covariance matrix of the estimation error $\hat{\varepsilon}_{t|t-1} = \hat{x}_{t|t-1} - x_t$.

As mentioned above, Kalman's state-space models are based on Richard Bellman's theory of dynamic programming. Bellman had developed his theory "to treat the mathematical problems arising from the study of various multi-stage decision processes" (Bellman 1954: 1), which he described in the following way:

> We have a physical system whose state at any time *t* is determined by a set of quantities which we call state parameters, or state variables. At certain times, which may be prescribed in advance, or which may be determined by the process itself, we are called upon to make decision which will affect the state of the system. These decisions are equivalent to transformations of the state variables, the choice of a decision being identical with the choice of a transformation. The outcome of the preceding decisions is to be used to guide the choice of future ones, with the purpose of the whole process that of maximizing some function of the parameters describing the final state. (Bellman 1954: 1)

According to Bellman, the exceedingly difficulty and complexity of a multi-stage decision process require new methods instead of the conventional "enumerative" methods that only work well in a "computational nirvana." In the conventional formulation, the multi-stage decision process, that is an *N*-stage process where *M* decisions are to be made at each stage, is considered as an *MN*-dimensional single stage process. The fundamental problem is then: "How can we avoid this multiplication of dimension which stifles analysis and greatly impedes computation?" (Bellman 1957: xi). The answer is that it is the structure of the policy which is essential. By structure Bellman meant the characteristics of the system which determine the decision to be made at any particular stage of the process: "in place of determining the optimal sequence of decisions from some *fixed* state of the system, we wish to determine the optimal decision to be made at *any* state of the system" (xi). This approach makes the decision problem analytically more tractable and computationally vastly simpler.

Bellman (1957) outlines the "structure of dynamic programming processes" by first enumerating the features of these processes:

> a. In each case we have a physical system characterized at any stage by a small set of parameters, the *state variables*.
> b. At each stage of either process we have a choice of a number of decisions.
> c. The effect of a decision is a transformation of the state variable.
> d. The past history of the system is of no importance in determining future actions.
> e. The purpose of the process is to maximize some function of the state variables. (Bellman 1957: 81-82)

In this context, a "policy" is defined as "any rule for making decisions which yields an allowable sequence of decisions; and an optimal policy is a policy which maximizes a preassigned function of the final state variables" (82). An optimal policy is obtained by applying the "principle of optimality":

> An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision. (Bellman 1957: 83)

The resulting functional equation for optimal policy then is:

$$f(p) = \max_{q} f(T_q(p)) \tag{7}$$

where T_q is the transformation as a result of the first decision, obtaining in this way a new state vector $T_q(p)$.

Existence and uniqueness theorems were proven for this kind of functional equations, not by applying fixed-point theorems, but by Bellman's own method of successive approximation. He showed that a particular sequence of successive approximations

$$f_N(p) = \max_{q} f_{N-1}\big(T_q(p)\big) \tag{8}$$

converges to the unique solution.

The tools designed by Turing, Shannon, Wiener, Kálmán, and Bellman created a new world populated by machines that communicate with each other by exchanging information. This information does not, however, contain only signals about the system states but also noise that needs to be filtered out. It is not a deterministic world, but one governed by stochastic processes. The decisions these machines take is conditioned on the (noisy) information they have about the current state of the world but at the same time will affect future states. Policy in this world therefore means tracing an optimal trajectory taking all these issues into account.

3. What Lucas took from Carnegie Institute of Technology

> This is what I mean by the "mechanics" of economic development – the construction of a mechanical, artificial world, populated by the interacting robots that economics typically studies, that is capable of exhibiting behavior the gross features of which resemble those of the actual world that I have just described.
> —Robert E. Lucas, On the mechanics of economic development (1988)

As is well acknowledged by Lucas, he did not only borrow the tools from the above discussed field of information engineering, but also those that were developed at the Graduate School of Industrial Administration of the Carnegie Institute of Technology, in particular Muth's "rational expectations hypothesis." When Lucas introduced the concept of rational expectations in his early papers, he, however, also indicated that his interpretation of the rational expectations hypothesis differed from Muth's and in what way:

- "[Muth] applied it to the case where the expected and actual price (both random variables) have a common *mean value*. Since Muth's discussion of this concept applies equally well to our assumption of a common *distribution* for these random variables, it seems natural to adopt the term here." (Lucas and Prescott 1971: 660, fn. 4)
- "The main difference is in what superficially appears to be a fine mathematical point. Muth defines solutions to be elements of the space of sequences of realizations, as opposed to being elements of the space of functions of current state variables. The definition used here is much more restrictive." (Lucas 1972b: 55)

To clarify how much of Lucas's work of the early 1970s was based on the tools of information engineering and how much on Muth's rational expectations hypothesis, the background of the rational hypothesis will be briefly discussed. This background was shaped by the research project "Planning and Control of Industrial Operations" conducted by Charles C. Holt, Franco Modigliani, John F. Muth, and Herbert A. Simon. The project resulted into the volume *Planning Production, Inventories, and Work Force* (1960). Because the historical context of this research is excellently covered by Klein 2015, I will focus briefly on only one aspect of this research, namely adaptive expectations and Muth's own specific response to it.

According to Holt et al. (1960: 258), forecasts (of sales) must meet the following "tests":

> They must be made quickly, cheaply, and easily. The forecasting technique must be clearly spelled out, so that it can be followed routinely, either manually or using an electronic computer. The number of pieces of information required to make a single forecast must be small so that the total amount of informant required for all products will not be expensive to store and to maintain. It should also be possible to introduce current sales information easily.

The authors were quite explicit about the non-theoretical aspect of the forecasting techniques they were interested in: “do not ‘explain’ sales changes, but simply extrapolate a sales time-series” (258). The only relevant input is the past history of sales, “no direct information concerning the market, the industry, the economy, sales of competing and complementary products, price changes, advertising campaigns, and so on is used” (258).

The technique that is primarily discussed is the “exponentially-weighted moving-average forecast,” because it meets the above tests: it “extrapolates sales in the forthcoming period by correcting for observed error in the preceding forecast” (259). This technique can be described as follows:[4] Let y_t represent that part of a time-series which cannot be explained by trend, seasonal, or any other systematic factors; and let y_t^e represent the expectation of y_t on the basis of information available through the (t-1)st period. Then, the expectations are adapted from one period to the next proportional to the latest observed error:

$$y_t^e = y_{t-1}^e + \beta(y_{t-1} - y_{t-1}^e), 0 \leq \beta \leq 1 \qquad (9)$$

The solution of this equation gives the formula for the exponentially weighted forecast:

$$y_t^e = \beta \sum_{i=1}^{\infty} (1-\beta)^{i-1} y_{t-i} \qquad (10)$$

Muth 1960 studies the statistical properties of time series for which this forecasting method would work well, and concludes that this is for the case where y_t consist of a permanent component, $\bar{y}_t$, and a transitory component, η_t, so that: $y_t = \bar{y}_t + \eta_t$. The transitory components are assumed to be independently distributed with mean zero and variance σ_η^2, and the permanent components are defined by $\bar{y}_t = \bar{y}_{t-1} + \varepsilon_t = \sum_{i=1}^{t} \varepsilon_i$, where the ε’s are serially independent with mean zero and variance σ_ε^2. The optimality condition is to minimize the error variance $E(y_t - y_t^e)^2$. By applying this condition, Muth (1960) shows that the weight β has to be a specific combination of the variances σ_ε^2 and σ_η^2.

One year later, *Econometrica* published Muth’s paper on Rational Expectations (1961). Contrary to the aims of the *Planning* book (Holt et al. 1960) – namely forecasting and not explaining – the intentions of this paper was “to explain fairly simply how expectations are formed” (Muth 1961: 315). To do so, Muth suggested the “hypothesis” that “expectations, since they are informed predictions of future events, are essentially the same as the predictions of the relevant economic theory” (316). This hypothesis was “a little more precisely” rephrased as: “that expectations of firms (or, more generally, the subjective probability distribution of outcomes) tend to be distributed, for the same information set, about the prediction of the theory (or the ‘objective’ probability distributions of outcomes)” (316).[5]

Muth’s discussion of an isolated market shows nicely the difference and similarity of optimal adaptive expectations and rational expectations. Suppose the market equations take the form:

$$C_t = -\beta p_t \qquad \text{(Demand)}$$

---

[4] This description is based on Muth 1960.

[5] Hence, the interpretation of rational expectation in Lucas and Prescott 1971 is in this respect not different from Muth’s – contrary that what Lucas and Prescott 1971 claim (see quotation on page 10 above).

$$P_t = \gamma p_t^e + u_t \qquad \text{(Supply)}$$

$$P_t = C_t \qquad \text{(Market equilibrium)}$$

where P_t represents production, C_t consumption, p_t market price, p_t^e expected market price, and u_t an “error term.” If these latter error terms u can be represented by a linear combination of the past history of normally and independently distributed random variables ε_t: $u_t = \sum_{i=0}^{\infty} \varepsilon_{t-i}$, then the price and expected price will be linear functions of the same independent disturbances: $p_t = \sum_{i=0}^{\infty} W_i \varepsilon_{t-i}$ and $p_t^e = \sum_{i=1}^{\infty} W_i \varepsilon_{t-i}$. Because the aim is to write the expectations in terms of the history of observable variables, the result is that the expected price is a geometrically weighted moving average of past prices:

$$p_t^e = \frac{\beta}{\gamma} \sum_{j=1}^{\infty} \left(\frac{\gamma}{\beta + \gamma} \right)^j p_{t-j} \qquad (11)$$

Comparing this equation (11) with equation (10), the prediction formula is the exponentially weighted forecast, but with different weights: “The only difference is that our analysis states that the ‘coefficient of adjustment’ in the expectations formula should depend on the demand and the supply coefficients” (320). In the case of optimal adaptive expectations, the coefficient of adjustment depends on the variances σ_ε^2 and σ_η^2 of the disturbances, whereas for rational expectations the coefficients are determined by using knowledge of the structure of the system describing the economy, namely the model parameters, β and γ.

4. What kind of mathematical engineering tools did Lucas use?

> If Wesley Mitchell could view agents as “signal processors” in 1913, then I saw no reason to regard my own adoption of this viewpoint in 1972 as unduly speculative.
> — Robert E. Lucas, *Studies in Business-Cycle Theory* (1981)

Although Lucas made all kinds of references to the information engineering literature, as will be shown below, these references could be dismissed as rhetoric. This section will show that this is not the case. To apply the rational expectations hypothesis to macroeconomics, Lucas needed and therefore was in search for new mathematical tools that could enable this implementation.

In his first paper on rational expectations, ‘Investment under uncertainty’ (1971), Lucas and Prescott were very clear about their mathematical approach: “It is shown, first, that the equilibrium development for the industry solves a particular dynamic programming problem (maximization of ‘consumer surplus’)” (659). This was reconfirmed ten years later: “The idea of defining an equilibrium as a point in a space of functions of a few ‘state variables’ was one that Prescott and I had utilized in ‘Investment under Uncertainty’” (Lucas 1981a: 7).

To obtain a theory of the development of the industry through time, Lucas and Prescott considered it to be “natural” to define an anticipated price process as a sequence $\{p_t\}$ of functions of $(u_1, \ldots, u_t)$, where $\{u_t\}$is a Markov process (664). Similarly the investment-output plan $\{x_t, q_t\}$ is also considered to be as a sequence of functions of $(u_1, \ldots, u_t)$. To link the anticipated price sequence to the actual price sequence – also a sequence of functions of $(u_1, \ldots, u_t)$ – it was assumed that the expectations are “rational” in the sense that “the anticipated price at time t is the same function of $(u_1, \ldots, u_t)$ as is the actual price. That is, we assume that firms know the true distribution of prices for all future periods” (664).

Equilibrium then was defined as an element $\{q_t^0, x_t^0, p_t^0\}$ such that the market condition is satisfied for all $(t, u_1, \ldots, u_t)$ and such that

$$E\{\textstyle\sum_{t=0}^{\infty}\beta^t[p_t^0 q_t^0 - x_t^o]\} \geq E\{\textstyle\sum_{t=0}^{\infty}\beta^t[p_t^0 q_t - x_t]\} \qquad (12)$$

for all $\{q_t, x_t\}$ satisfying certain conditions. The largest and central section of their paper proved that this equilibrium is unique and exists by “using the techniques of dynamic programming” (665), that is to say, by utilizing the “method of successive approximation as applied in [Bellman 1957]” (668). Therefore an “operator” T was defined such that solutions to equation 12 coincides with solutions to $Tf = f$. Then, it was proven that $Tf = f$ has a unique solution, f^*, and for any g, $lim_{n\rightarrow 0}T^n g = f^*$.

This framework of dynamic programming was a year later again applied – “exactly the [same] way” (Lucas 1981a: 7) – in the (1972a) ‘Expectations and the neutrality of money’ paper.

> The substantive results developed below are based on a concept of equilibrium which is, I believe, new (although closely related to the principles underlying dynamic programming) and which may be of independent interest. In this paper, equilibrium prices and quantities will be characterized mathematically as *functions* defined on the space of possible states of the economy, which are in turn chracterized [*sic*] as finite dimensional vectors. (Lucas 1972a: 104)

In this 1972a article, the state of the economy in any period is entirely described by three variables m, x, and θ, representing respectively money supply, global money shocks and local allocation shocks. The “motion of the economy from state to state is independent of decisions made by individuals in the economy” (106), and is given by $m' = mx$, and the densities f and g of x and θ. As a result, the equilibrium price was considered to be a function $p(m, x, \theta)$ on the space of possible states and was defined to satisfy a specific functional equation. And again, the largest and central section (plus appendix) was spent on proving the existence and uniqueness of the equilibrium price as a solution to equation $Tf = f$.

In his 1972b, 1973 and 1975 papers, Lucas introduced also other mathematical tools in his analysis than those of dynamic programming used in the (1971 and 1972a) articles discussed above. Actually his 1972b paper, ‘Econometric testing of the natural rate hypothesis,’ is the only paper in which rational expectations are formed as “originally proposed by Muth” (1972b: 51) and hence were shaped by the same mathematical tools.

The aim of this ‘Econometric Testing’ paper is to compare two different models of expectations formation, adaptive and rational, to investigate which one can give an adequate formulation of the natural rate hypothesis. To show that the model of rational expectations formation in this paper is indeed the same as Muth’s (1960), we will discuss this model in more detail.

Letting y_t be the log of real output in t, P_t be the log of the price level, and P_t^* be the log of an index of expected future prices, the aggregate supply function is: $y_t = a(P_t - P_t^*)$, and aggregate demand is assumed to be $y_t + P_t = x_t$, where x_t, the log of nominal GNP, is viewed as a “shift parameter.” Policy is defined as a “rule” giving the current value of x_t as a function of the state of the system. Lucas considered the following particular rule: $x_t = \rho_1 x_{t-1} + \rho_2 x_{t-2} + \varepsilon_t$, where $\{\varepsilon_t\}$ is a sequence of independent random variables that are distributed identically and normally, each with zero mean. Expectations were defined as rational:

$$P_t^* = E\{P_{t+1} | x_t, x_{t-1}, \eta_t\} + \eta_t \tag{13}$$

where η_t is the forecast error, with the same properties as ε_t. To solve this system of four equations, Lucas used Muth’s solution method by considering P_t and P_t^* as linear combinations of the disturbances x_t and x_{t-1} and forecast error η_t.

Lucas’s (1973) ‘Some empirical evidence on output-inflation trade-offs’ was based on the main result of Lucas’s (1972a) ‘Expectations and the neutrality of money.’ This latter paper showed that an economic system in which “all prices are market clearing, all agents behave optimally in light of their objectives and expectations, and expectations are formed optimally” (1972a: 103) still could give rise to business cycles. This result was achieved by “the removal of the postulate that all transactions are made under complete information” (104). Information was incomplete because the economy was divided into two physically separated markets (“islands”), and

> information on the current state of these real and monetary disturbances is transmitted to agents only through prices in the market where each agent happens to be. In the particular framework presented below, prices convey this information only imperfectly, forcing agents to hedge on whether a particular price movement results from a relative demand shift or a nominal (monetary) one. (Lucas 1972a: 103)

This idea of incomplete information was taken over in his 1973 empirical study of real output-inflation trade-offs. To model “suppliers’ lack of information on some of the prices relevant to their decision” (326), “where agents are placed in this situation of imperfect information” (327), Lucas “imagined” suppliers as located in a large number of scattered competitive markets, indexed by z. The information available to suppliers in z at t comes from two sources. First traders enter period t with knowledge of the past course of demand and supply shifts. While this information does not permit exact inference of the (log of) current general price level, P_t, it does determine the distribution on P_t common to all markets. It is assumed to be normal with mean $\bar{P}_t$ – “depending in a known way on the above history” (328) – and a constant variance σ^2. It is supposed that the observed price in market z deviates from the general price level P_t with mean zero and variance τ^2. Then the observed price in z, $P_t(z)$ is the sum of independent, normal variates

$$P_t(z) = P_t + z \tag{14}$$

The information $I_t(z)$ relevant for estimation of the unobserved P_t consists then of the observed price $P_t(z)$ and the history summarized in $\bar{P}_t$. This information is used to infer the value of P_t:

$$E(P_t | I_t(z)) = E(P_t | P_t(z), \bar{P}_t) = (1-\theta)P_t(z) + \theta\bar{P}_t \tag{15}$$

where $\theta = \tau^2/(\sigma^2 + \tau^2)$, and variance $\theta\sigma^2$.

Although this result was apparently obtained by "straightforward calculation" (328), the same result would have been attained by framing this information processing as a Kalman filter. To see this, consider first equation (14) as the measurement equation (2), so that $H_t = I$, $y_t = P_t(z)$, $x_t = P_t$, $\varepsilon_t = z$, $R = \tau^2$. If one, subsequently, equates $\hat{x}_{t|t}$ with $E(P_t| I_t(z))$ and $\hat{x}_{t|t-1}$ with $\bar{P}_t$, then $P_{t|t-1} = \sigma^2$, and thus the Kalman gain $K = \sigma^2/(\sigma^2 + \tau^2) = 1 - \theta$ and $P_{t|t} = (1-K)\sigma^2 = \theta\sigma^2$. It can now be seen that the expectation of the general price level using information $I_t(z)$ as represented in equation (15) is equal to the Kalman-filtered signal as represented in equation (5).

Lucas's 1975 paper 'An equilibrium model of the business cycle' dealt with the same problem: "the expectations of agents are rational, given the information available to them; information is imperfect, not only in the sense that the future is unknown, but also in the sense that no agent is perfectly informed as to the current state of the economy" (1113). And similar to his 1972a and 1973 paper, he adopted "the device proposed by Phelps (1969) and, since utilized in similar contexts by Lucas (1972[a], 1973) …, of thinking of trading as occurring in distinct markets, or 'islands'" (1120).

The aggregate state of the economy is fully described by the variables k_t, m_t and x_t, representing capital stock, money, and nominal government spending. The situation of an individual market z is described by its capital relative to average, $u_t(z)$ and the government spending it receives relative to average, $\theta_t(z)$. Agents observe none of the variables directly. The only source of information is the history of market clearing prices $p_t(z)$, $p_{t-1}(z)$, $p_{t-2}(z)$, … of the markets in which traders happened to be currently and in the past. This information is used to form unbiased estimates of the current values of the aggregate state variables. On the basis of this information, the agents have a "well-informed opinion" of the relevant variables they cannot observe: $s_t = (k_t, m_t, x_t, \theta_t(z), u_t(z))$.

When new information is available, $p_t(z) - \hat{p}_t$, agents form a posteriori conditional mean on the state vector to be used in forecasting: $\tilde{s}_t$. And again, according to Lucas, "a straightforward calculation yields the conditional means" (1125)

$$\tilde{s}_t = \hat{s}_t + \sigma_p^{-2} V(p_t(z) - \hat{p}_t) \tag{16}$$

which again can be interpreted as a Kalman filter.

5. Evaluation

Lucas has never made clear and explicit what the larger background is from which his new methodology was developed. The early 1970s papers were written when he was at Carnegie-Mellon University, from 1963 till 1974. At several occasions, Lucas explained that the people who were around at Carnegie at that time had influenced him, but not in what way. When he, for example, clarified that he saw models as a "parallel or analogue system," he noted that "I do not know the background of this view of theory as physical analogue, nor do I have a clear idea as to how widely shared it is among economists. An immediate ancestor of my condensed statement is [Simon 1969]" (Lucas 1980: 697).

Also, with respect the kind of mathematics, he never made clear where it came from. In relation to Arjo Klamer's question about his collaboration with Prescott on the first paper in which rational expectations was introduced, Lucas answered: "We thought [the investment problem] was a pretty straightforward applied problem, but then we got in way over our heads technically. We didn't want to quit, so we read tons of difficult mathematical economics and mathematics, even though none of us had any prior familiarity with it" (Klamer 1984: 32-33).

The only exception with respect to more precise referencing to the mathematical literature is to be found in his 1975 paper. To clarify how he arrived at the above presented equation 16, Lucas refers in a footnote to a theorem in a textbook in statistics: Franklin Graybill's (1961) 'Introduction to Linear Statistical Models,' namely theorem 3.10 on page 63. To receive more clarification about the mathematical resources Lucas had used in his early 1970s papers, I wrote in March 9 2005 an email to Lucas with the following question:

> When carefully reading your work from the early 1970s, I found it striking that the way rational expectations is defined in these early papers differs from one paper to the other. While your papers written together with Prescott and your Neutrality paper is closely linked to dynamic programming, the 1973 paper 'Some International Evidence on Output-Inflation Tradeoffs' and your 1975 paper 'An Equilibrium Model of the Business Cycle' show a different approach – in the sense of mathematical concepts and techniques. In section 6, page 1125, of the 1975 paper, you refer to a theorem in a statistics textbook, which is an ordinary conditional expectations theorem. However, your specific application of it, namely taking expectations conditioned on information which give a noisy signal of unobserved prices, is in fact a Kalman filter.

On April 20th of 2005, Lucas responded as follows:

> I don't understand the distinction you are drawing here. I think of Kalman filtering as a recursive method for updating Bayesian posteriors as time passes that works in certain dynamic models. It, too, is an application of ordinary conditional expectations to a particular context. The linear set-up I used in these papers and some others is similar to Muth's, in his two papers on rational expectations. I suppose I just took it from him. But of course I knew about Kalman filtering at this time too.

The above references and answers show an instrumental relationship to mathematics. Consistent with his instrumental view on theory, namely to be considered as "an explicit set of instructions" (Lucas 1980: 697) for building a model, Lucas searched the mathematical literature for useful concepts and tools, and picked out the ones that solved his technical problems.

This paper shows that the mathematical "instructions" came from information engineering, which should be distinguished from those of control engineering. This could, however, not be shown by exhibits from the Lucas archive[6] or from his published work, neither from remarks made at interviews or in panel discussions. Evidence to show that the "rational expectations revolution" would have been better called the "information revolution" has to be constructed in another way. The reconstruction method applied in this paper is actually closer to forensic research, in the sense that I have looked for mathematical traces, that is to say, mathematical concepts and tools that I could trace back to original sources.[7]

---

[6] I searched the Lucas collection in the Economists' Papers Archive held in the David M. Rubenstein Rare Book & Manuscript Library at Duke University with no avail.

[7] The possibility of this kind of tracing is not limited to mathematics. In Boumans (2004) I was able to show that the Hodrick-Prescott is actually a Kalman filter, by deconstructing the FORTRAN subroutine which was

These original sources are the works where these concepts and tools for the first time were presented and discussed. But, the paper does not claim that Lucas used these original sources directly. By the time Lucas started to use these tools, they already had found their way into textbooks and other handbooks where they could be picked up for use.

But this forensic tracing is not sufficient to make any claim about the relevance of information engineering background for Lucas's new methodology. My e-mail exchange with Lucas shows that Lucas himself saw no sense in making any distinction between the various backgrounds of the tools he had used. In his view they are neutral with respect to the original context from which they were developed and neutral with respect to the context of application. Mathematical tools are like hammers and screwdrivers, the usage of them does not change one's world view. Contrary to this viewpoint, this article shows that this distinction is relevant because the tools of control engineering shape a different world, namely one consisting of human-machine relationships, than those of information engineering which is populated information-processing robots.

As Mary Morgan (2019) discusses in a paper on diagrams, but also in her earlier work on models (2012), a mathematical tool, such as a diagram or model, becomes a description or representation of the economic materials in the sense that economists cannot think of the materials without that description, the tool becomes wedded to the materials. What economists think is in the world, or in the way the world works – their ontology of what is in the world – can only be understood through this mathematically expressed projection. The mathematical form is essential to the way economists think of the world and how it works. Although originally the choice of tool depends on the economist seeing the problem in a particular way to which that tool fits that epistemological purpose, over time that choice has ontological implications as the tool and subject matter become more closely intertwined: Tools "have implications for the way they think about their materials and subject matters and so the objects that they think exist in the world" (Morgan 2019: 23). Due to the innovations of the 1970s, the (economic) world came to be seen as an "informational structure," providing "signal processors" with noisy information that needs therefore to be "filtered" (see e.g. Lucas 1977). With the rational expectations revolution, Lucas and the other new classical economist have changed the world such that in that new world, control techniques have lost their essence.

References

Andrei, Neculai. Modern Control Theory: A Historical perspective. Center for Advanced Modeling and Optimization. https://camo.ici.ro/neculai/history.pdf

Bellman, Richard. 1954. The Theory of Dynamic Programming. Working paper P-550. Santa Monica, CA: Rand.

Bellman, Richard. 1957. *Dynamic Programming*. Princeton, NJ: Princeton University Press.

Boumans, Marcel. 1997. "Lucas and Artificial Worlds." In *New Economics and Its History*, edited by John B. Davis. *History of Political Economy* 29 (supplement): 63-90.

Boumans, Marcel. 2004. "The Reliability of an Instrument." *Social Epistemology* 18, no. 2-3: 215-246.

Boumans, Marcel. 2005. *How Economists Model the World into Numbers*. London and New York: Routledge.

---

published in a research memorandum 'A Fortran subroutine for efficiently computing HP-filtered time series' by Kydland and Prescott in 1989.

Graybill, Franklin A. 1961. *An Introduction to Linear Statistical Models*. New York: McGraw-Hill.

Holt, Charles C., Franco Modigliani, John F. Muth and Herbert A. Simon. 1960. *Planning Production, Inventories, and Work Force*. Englewoods Cliffs, NJ: Prentice-Hall.

Kailath, Thomas. 1974. “A View of Three Decades of Linear Filtering Theory.” *IEEE Transactions on Information Theory* IT-20, no. 2: 146-181.

Kálmán, Rudolf Emil. 1960. “A New Approach to Linear Filtering and Predictions Problems.” *Journal of Basic Engineering* 82: 35-45.

Klein, Judy L. 2001. “Optimization and Recursive Residuals in the Space Age: Sputnik and the Kalman Filter.” Working paper Mary Baldwin College, USA.

Klein, Judy L. 2015. “The Cold War Hot House for Modeling Strategies at the Carnegie Institute of Technology.” Institute for New Economic Thinking Working Paper Series No. 19.

Lucas, Robert E. 1972a. “Expectations and the Neutrality of Money.” *Journal of Economic Theory* 4: 103-124.

Lucas, Robert E. 1972b. “Econometric Testing of the Natural Rate Hypothesis.” In *The Econometrics of Price Determination Conference,* edited by O. Eckstein, 50-59. Washington, DC: Board of Governors of the Federal Reserve System.

Lucas, Robert E. 1973. “Some International Evidence on Output-Inflation Tradeoffs.” *The American Economic Review* 63, no. 3: 326-334.

Lucas, Robert E. 1975. “An Equilibrium Model of the Business Cycle.” *Journal of Political Economy* 83, no. 6: 1113-1144.

Lucas, Robert E. 1977. “Understanding Business Cycle.” In *Stabilization of the Domestic and International Economy*, edited by Karl Brunner and Allan H. Meltzer, 7-29. Amsterdam: North-Holland.

Lucas, Robert E. 1980. “Methods and Problems in Business Cycle Theory.” *Journal of Money, Credit, and Banking* 12, no. 4: 696-715.

Lucas, Robert E. 1981a. *Studies in Business-Cycle Theory*. Oxford: Basil Blackwell.

Lucas, Robert E., ed. 1981b. *Rational Expectations and Econometric Practice*. University of Minnesota Press.

Lucas, Robert E. 1988. “On the Mechanics of Economic Development.” *Journal of Monetary Economics* 22: 3-42.

Lucas, Robert E. 2004. “Keynote Address to the 2003 HOPE Conference: My Keynesian education.” In *The IS-LM Model*, edited by Michel De Vroey and Kevin D. Hoover. *History of Political Economy* 36 (supplement): 12-24.

Lucas, Robert E., and Edward C. Prescott. 1971. “Investment under Uncertainty.” *Econometrica* 39, no. 5: 659-681.

Lucas, Robert, and Thomas Sargent. 1979. “After Keynesian Macroeconomics.” *Federal Reserve Bank of Minneapolis Quarterly Review* 3, no. 2: 1-16.

Hoover, Kevin D., and Warren Young. 2013. “Rational expectations: Retrospect ad Prospect.” *Macroeconomic Dynamics* 17: 1169-1192.

Klamer, Arjo. 1984. *The New Classical Macroeconomics; Conversations with the New Classical Economists and their Opponents*. Brighton, UK: Wheatsheaf.

McMillan, Brockway. 1962. “Mathematicians and Their Uses.” *SIAM Review* 4, no. 2: 79-90.

Morgan, Mary S. 2012. *The World in the Model*. Cambridge: Cambridge University Press.

Morgan, Mary S. 2019. Inducing Visibility and Visual Deduction. LSE Working Paper.

Muth, John F. 1960. “Optimal Properties of Exponentially Weighted Forecasts.” *Journal of the American Statistical Association* 55, no. 290: 299-306.

Muth, John F. 1961. “Rational Expectations and the Theory of Price Movements.” *Econometrica* 29, no. 3: 315-335.

Robert E. Lucas Jr. – Facts. NobelPrize.org. Nobel Media AB 2019. Fri. 23 Aug 2019. https://www.nobelprize.org/prizes/economic-sciences/1995/lucas/facts/

Sent, Esther-Mirjam 1997. "Engineering Dynamic Economics." In *New Economics and Its History*, edited by John B. Davis. *History of Political Economy* 29 (supplement): 41-62.

Shannon Claude E. 1948. "A Mathematical Theory of Communication." *The Bell System Technical Journal* 27: 379–423, 623–656.

Shannon. Claude E. 1950. "Programming a Computer for Playing Chess." *Philosophical Magazine* 41, no. 314: 256-275.

Shannon, Claude E. (1950) 1988. "A Chess-Playing Machine." In Vol. 4 of *The World of Mathematics*, edited by James R. Newman, 2099-2108. Washington: Microsoft Press.

Simon, Herbert A. 1969. *The Sciences of the Artificial*. Cambridge: MIT Press.

Soare, Robert I. 1996. "Computability and Recursion." *The Bulletin of Symbolic Logic* 2, no. 3: 284-321.

Turing, Alan M. 1936. "On Computable Numbers, with an Application to the Entscheidungsproblem." *Proceedings of the London Mathematical Society* 42: 230-265.

Wiener, Nobert. 1949. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series. With Engineering Applications*. New York: Wiley & London: Chapman & Hall & The Technology Press of the Massachusetts Institute of Technology.

Yeung, Raymond W. 2008. *Information Theory and Network Coding*. Springer US.