THE RATIONALITY OF RATIONAL EXPECTATIONS

Cloda Lane
Junior Sophister

> *The advent of rational expectations in econometric models has marked a revolution in economic thinking that is comparable in the magnitude of its impact on the economics profession to the Keynesian revolution of a half century ago.*[6]

The purpose of this essay is not to defend the numerous models and policy prescriptions which are based on the rational expectations hypothesis but to provide an insight into why the hypothesis coincides with optimising economic behaviour. As Simon (1978, p 12) outlines, Economics, whether normative or positive, has not simply been the study of the allocation of scarce resources, it has been the study of the *rational* allocation of scarce resources. Expectations often form a major part of the decisions which are made in the economy and as such they should also come under the doctrine of rationality. In line with other rational economic behaviour[7]we must see if the rational expectations hypothesis is the best (the objective) available (the constraint) method of modelling expectations. To show that this hypothesis is rational I will outline the theory itself and then I will show how it can withstand both theoretical and empirical criticisms.

The Rational Expectations Hypothesis

The premise of the rational expectations hypothesis is that economic variables are generated by systematic processes. Over time, economic agents learn what the process determining a variable is and they will use this knowledge to form expectations of that variable. Individuals learn about the variable generating process by using all the information available to them that is related to the variable. The end result is that the expectations of firms (or, more generally, the subjective probability distribution of outcomes) tend to be distributed, for some information set, about the prediction of the theory (or the objective probability distribution of outcomes).[8]

To see how the hypothesis works imagine an economic variable, Y, whose value is determined by its own lagged value, by the lagged value of two other variables, X and Z, and by a random variable U. This provides us with the simple linear process:

$$Yt = 0 + 1Yt\text{-}1 + 2Xt\text{-}1 + 3Z\,t\text{-}1 + Ut$$

The expected values of Yt is found by finding the mathematical expectation of Yt . Since Yt-1, Xt-1 and Z t-1 are lagged, their values are known at the end of period t-1 (when the forecast is being made). The values of Ut, however, only becomes known at the end of period t so the rational forecaster must form some expectation of its value at the end of period t-1. This means that:

$$Et\text{-}1(Yt) = 0 + 1Yt\text{-}1 + 2\,Xt\text{-}1 + 3Z\,t\text{-}1 + Et\text{-}1\,(Ut)$$

The random variable is assumed to be distributed with mean zero and variance ^{2}. The best estimate that can be made of the expected value of Ut is to use its mean value, zero. This leaves us with a formula for the expected value of Y as:

$$Et\text{-}1(Yt) = 0 + 1Yt\text{-}1 + 2\,Xt\text{-}1 + 3Zt\text{-}1$$

Thus, the rational expectation of the variable Y in period t is its mathematical expectation given the available information. Thus, as Muth (1961) explained, rational expectations should be generated by the same (stochastic) process that generates the variable to be forecast.[9]

The rational expectations hypothesis does not argue that agents are always right in their expectations of future variables. In fact, the forecast error is exactly equal to the random variable that determines Yt. This random variable is uncorrelated with the other variables in the process and with the information set available to the agent. This makes sense because if such correlation existed, it would logically be included in arriving at the initial expectation.[10] These random variables, and hence any forecast errors, are surprises or news in the system. They are random, they exhibit no definite pattern, they have a mean value of zero and they have a variance less than that associated with any other model of forecasting.[11] This means that, *on average*, rational expectations will be correct because the mean value of the forecast error is zero and it also means that they are the most efficient (in a statistical sense) means of forming expectations because their forecast errors have the property of minimum variance.

The rational expectations hypothesis thus puts forward a means of forming expectations which is based on agents taking account of all necessary available information to make their forecasts. The information is used efficiently to determine the process which generates the variable in question and the process is then used to formulate an expected value of that variable. The end result is that, Rational expectations, by Muths definition, yield predictions of future events which differ from the corresponding eventual outcomes only by errors which are themselves independent of the variables used to generate the predictions.[12]

Rational expectations are best in a statistical sense but do they stand up to theoretical and empirical scrutiny? In answer to this question the following discussion will focus on some of the major theoretical criticisms of the hypothesis and some of the empirical evidence concerned with the formation of expectations.

Theoretical Analysis

One of the main criticisms of the rational expectations hypothesis is that, as Arrow (1978) outlines, Economic agents are required to be superior statisticians, capable of analysing the future general equilibria of the economy.[13] This criticism stems from the mis-conception that Muth was proposing that economic agents use the exact model used by economists.[14] The fact of the matter, however, is that the rational expectations hypothesis argues that trained economists and economic agents produce the same expectations but it does not argue that they come to that conclusion by using the exact same method. In fact for the hypothesis to hold it is sufficient that in the light of past observation and experience they possess some concept of a reduced-form approach to economic modelling to permit them to make reasonable predictions.[15] This criticism of economic agents needing to be qualified economists for the hypothesis to be viable does not, therefore, withstand closer scrutiny of Muth's thesis.

A closely related criticism is the one that argues that the idea of rationality is implausible in itself. Can we really assume that all decision-makers are intelligent enough to use and fully understand all the available information? Once again, this criticism is based on a misconception of what the hypothesis is saying. The hypothesis does not apply to every individual in the economy. Rather, it claims that on the average expectations are rational. Thus, some agents may irrationally overpredict and some may underpredict but this does not mean that on average the expectations in the market cant be rational.

We must also remember that the hypothesis doesnt require that every single agent in the market gathers and formulates the information themselves and makes the expectations for themselves. In many cases individuals let other people form their expectations for them. For example, peoples expectations of inflation are often based on the expectations that have been carefully constructed by economists in the Central Bank, the ESRI or the Department of Finance. These expectations are based on full information and are rational. Thus, the market as a whole has rational expectations even though these expectations have been formed by only a subset of society.[16] Another situation where individuals allow others to form their expectations for them is in the labour market: Here, many such agents are perfectly willing to delegate the model-analysing role of their elected or appointed trade union representatives who do indeed invoke former models and often employ expert financial specialists and consultants to assist them.[17] Thus, the expectations of the market as a whole can be rational without making the highly unlikely assumption that every single individual forms rational expectations.

Finally, the criticism of the hypothesis on the grounds of rationality undermines the basis of economics: The idea that the typical individual is capable of making the best of the opportunities open to him is a common

one in economics.[18] Thus, to claim that agents are not rational when making forecasts is equivalent to claiming that the core of economic argument, that economic agents are rational decision-makers, is incorrect.

A third criticism of the rational expectations hypothesis is that the information necessary to form expectations is not always available and when it is it may be very costly to use it. It is true that individuals can not automatically know which variables are important in the variable generating process or know what the size of the coefficients in that process are but it is also true that the rational expectations hypothesis doesnt claim that they do. What the hypothesis argues is that on average and after a period of time, economic agents will learn from past experience what the process is. They will combine this developed knowledge with current available information to form their expectations. This is why the rational expectations hypothesis is best seen as a long-run argument. It is based on a learning process which takes time but once the necessary knowledge is acquired the process determining a variable will be known. It must be noted, however, that Friedman (1979, p 24) is right to point out that what is typically missing in Rational Expectations models, however, is a clear outline of the way in which economic agents derive the knowledge which they then use to formulate expectations meeting the requirement. This is something which needs to be developed in models that are based on rational expectations but we must remember that the models are based on the hypothesis, and are not the same as it. Thus, although a knowledge of the learning process of economic agents would make economic models more concise, the absence of it does not take away from the hypothesis itself.

The other side of this information criticism is that even when information is available it is costly to use. This criticism, however, does not take away from the rational expectations hypothesis. The crux of the hypothesis is not that a rational agent should simply use all the available information but that he should use all the available information in an efficient manner. That is, an efficient and rational individual will carry out a form of cost benefit analysis on the information, using only that which is of net benefit to him. Thus, in fact, the limitation imposed by costly information coincides with the efficiency standpoint of the hypothesis. There is also the argument that when information, that is absolutely necessary, gets too costly, agents can pool together to obtain that information or the government can obtain it and provide it to the public. Both of these methods ensure that agents still get the information and thus they can still form rational expectations.

A fourth criticism of the hypothesis is that it has limited applicability. As it is not always easy to determine the process by which a variable is generated it may not always be possible to form rational expectations. However, as Attfield, Demery and Duck (1985, p 28) outline a rational expectation can still be formed without knowing the exact process. In fact, we can still form expectations from an intelligent appraisal of circumstances, though the process behind such circumstances may be a bit harder to discern. Thus, rational expectations can be made even when variables are generated by unique and unusual processes because the economic agent will have enough information to make an intelligent estimate of the process.

The final criticism of the rational expectations hypothesis is the argument that the hypothesis is not testable. The obvious retort to this criticism is that although expectations are inherently immeasurable, there have been numerous attempts made to incorporate them into econometric models and to test their validity in these models. This fact holds for all proposed means of expectations formation (including adaptive and rational). Although the attempts made to test these hypotheses are not perfect, they are no worse for the rational expectations hypothesis than they are for any other expectations hypothesis. To see how the hypothesis holds up under these tests we must look at some of the empirical work that has been carried out on it.

Empirical Analysis

Since Muths seminal article was published, numerous attempts have been made to prove and disprove the rational expectations hypothesis. The hypothesis has been supported by much empirical work in the financial markets and commodity exchanges.[19] For example, Mishkin (1983, p 157) found that on balance the results justify using the assumption of rational expectations in empirical work, especially when financial markets are studied. The results for these specialised markets are very robust but there is much truth in the argument that the hypothesis holding true in these markets does not prove that it is the ideal way of forming expectations across the economy. While no major favourable insights of rational expectations in other markets have abounded, those empirical studies that have claimed to disprove the hypothesis have not been technically strong.

For example, many of these empirical tests of the rational expectations hypothesis have used survey data to proxy expectations.[20] Using survey data, however, presupposes that for market expectations to be rational all agents surveyed must be forming rational expectations. We know, however, that the hypothesis is based on the market, on average, having rational expectations and not on every individual forming such expectations. There is thus a data identification problem here and we must conclude that survey forecasts do not necessarily describe the forecasts inherent in market behaviour, and irrationality of survey forecasts does not in itself imply that market forecasts are also irrational.[21] Thus, any empirical attack on the rational expectations hypothesis which is based on survey data can not be taken as absolute.

Another problem with many of the empirical criticisms of the rational expectations hypothesis is that they have been based on tests which do not satisfy important statistical criterion. The most notable of such tests is the Chow Test.[22] This test has been used by many economists to test the rational expectations hypothesis but often their data fail to be consistent.[23] When alternative testing methods are used what the Chow Test showed to be irrational has often been shown as rational.[24] Thus, any conclusion about the irrationality of the rational expectations hypothesis based on these tests can not be assumed to be accurate.

These are just two examples which highlight some of the many problems that arise in empirical work concerned with expectations. No overall conclusion about whether expectations in the market are rational can be obtained from empirical work as it is so imperfect. All that we can do is realise that in the case of the financial markets there is widespread evidence to support the rational expectations hypothesis while in the case of other markets there is not sufficiently strong evidence to completely disprove this theory. Until this hypothesis has been empirically falsified it can not be dismissed as irrational.

Conclusion

Rational expectations are the best available models that economics has for modelling economic expectations. At a statistical level, they are efficient because they have an error term whose variance is less than that of any other method of modelling expectations and which has a mean value that predicts that on average, the error variable will be zero. The rational expectations hypothesis is also best because, unlike other hypotheses, it coincides perfectly with the concept of *homo economicus* and of the utility-maximising individual.[25] Finally, there is more information available that discredits other expectations models than there is to disprove the rational expectations hypothesis. Thus, the hypothesis that expectations are rational must be taken seriously, if only because its alternatives, for example various fixed-weight autoregressive models, are subject to so many objections.[26] The main point to be made is that the rational expectations hypothesis is not perfect but, given that expectations need to be incorporated into economic models, it is the best available method that we have for modelling these expectations. It therefore fits the loose economic criterion of rationality.

Bibliography

Arrow, K (1987) The Future and The Present in Economic Life in *The Economic Enquiry*

Attfield G, Demery D & Duck N (1985) *Rational Expectations in Macroeconomics*

Begg, D (1982) *The Rational Expectations Revolution in Macroeconomics*

Friedman, B (1979) Optimal Expectations and the Extreme Information Assumptions of Rational Expectations Macromodels in the *Journal of Monetary Economics*

Friedman, B (1980) Survey Evidence on the Rationality of Interest Rate Expectations in the *Journal of Monetary Economics*

Greenaway, D & Shaw, G (1983) *Macroeconomics: Theory and Policy in the UK*

Mishkin, F (1981) Are Market Forecasts Rational ? in the *American Economic Review*

Mishkin, F (1983) *A Rational Expectations Approach to Macroeconomics*

Muth, J (1961) Rational Expectations and the Theory of Price Movements in *Econometrica*

Mullineaux, D (1978) On Testing for Rationality: Another look at the Livingston Price Expectations Data in the *Journal of Political Economy*

Sargent, T & Wallace, N (1975) Rational Expectations, The Optimal Monetary Instrument and the Optimal Money Supply Rule in the *Journal of Political Economy*

Schiller, R (1987) Expectations in *The New Palgrave: A Dictionary of Economics* by Eatwell, J, Milgate, M and Newman, P (eds)

Shaw, G (1984) *Rational Expectations*

Shaw, G (1989) Expectations in Macroeconomics in *Current Issues in Macroeconomics* by Greenaway, D (ed)

Simon, H (1978) Rationality as Process and as Product of Thought in the *American Economic Review*

INDETERMINATE BARGAINING: A HINDRANCE TO THEORY AND AN INDICATION OF REALITY.

Mark Wall
Senior Sophister

Prices are central to economic co-ordination in the neo-classical market model. Standard economic theory makes a good case for market (i.e. price and quantity) determination when there are many agents on both the demand and supply side. This process is assumed to be flawless and efficient. The theory is not nearly as sufficient in the case of bilateral monopoly bargaining, where just one buyer and one seller must decide between them the terms of trade. The division of the core is the bargaining problem and gives rise to the indeterminacy in the theory.

This essay is an examination of the issues at stake in bargaining and of the contribution by modern literature to the solution of the indeterminacy. I shall begin by examining simultaneous bargaining (Edgeworth and Nash) and then move to sequential bargaining with Rubinstein. Following will be an analysis of the contingencies in bargaining models i.e. commitment, time, and information. Recent literature will be reviewed, and examples of current applications of bargaining in the field of industrial organisation presented.

model to which the agents would eventually end up after bargaining. He proposed that the two would trade to a Pareto optimal outcome that left each at least as well off as he would be in the absence of trade.

The problem with Edgeworths analysis is that the final solution to the terms of trade is indeterminate [10] One claim made is that the solution or outcome will depend on relative bargaining strengths. The question is, how do you model these psychological factors?

Trade will occur i.e. each player can potentially be made better off by moving to a point on the core. But, what parameters are causal in determining the outcome, ability to commit? Outside options if bargaining collapses? Patience? As we shall see, all these issues and more, are of importance. Already there is a sense that if we could be more precise about the protocols and institutional arrangements of bargaining then the solution would become more determinate.

The greatest development for bargaining theory came from Nash (1950, 1951, 1953).[11] Nashs intention with his axiomatic approach was to refine away the noise that surrounds bargaining to leave the bones of what actually occurs within all bargaining contexts. The desire was not to formulate the bargaining procedure but to identify common characteristics in bargaining and in its solutions. Nashs approach was successful and he claimed a unique outcome.

The Edgeworth and Nash analyses are almost equivalent. Nashs diagram is a reformulation of Edgeworths in a utility space rather than a quantity space. The endowment point, the Pareto frontier (contract curve), the set of feasible bundles, and the core of the model are common to both.

Nashs analysis is co-operative in flavour and based on fixed outcome bargaining. The initial assumptions are that X is the set of feasible payoffs, û is the disagreement point, and U* is the Nash bargaining solution. The axioms are:

1. Invariance: the solution is invariant to equivalent utility transformation (affine only) i.e. the solution is independent of the units in which we measure utility.

2. Efficiency: the solution is Pareto optimal i.e. at least one player is made better off.

3. Independence of Irrelevant Alternatives: if we drop some set of possible utility combinations from X leaving a smaller set Y, then if U* was not one of the points eliminated, U* does not change.

4. Anonymity (Symmetry): switching the labels on the axes does not alter the solution. In some sense this implies an equal bargaining power assumption.

Nash states that:

i.e. U* is that value or outcome which satisfies the maximisation problem as stated, subject to the underlying constraints. It is equivalent to a Lagrangean constrained optimisation problem.

Nashs theory implies a unique outcome, U*. The uniqueness of the solution is guaranteed by the convexity properties of X, much like the convexity conditions on preferences in the Edgeworth analysis.

It is possible to analyse changes in bargaining power in the Nash context by introducing a mathematical and diagrammatic approach that uses marginal utility reinterpreted as an interest in the good at the margin. A similar analysis can be evoked to show that risk aversion imposes a penalty on the bargainer - ones outcome diminishes. The risk aversion effect can be obtained by applying a transformation to the risk averse players utility function.

A more recent and considerable development in bargaining theory was by Rubinstein (1982). Whereas the Edgeworth and Nash models were static, the Rubinstein model introduces the concept of dynamism to bargaining. In addition to its dynamic nature, the Rubinstein model is built on non-co-operative game theory, i.e., where commitments are not binding. Game theory is interactive, rational economic decision-making, utilising mathematical and intuitive models to analyse instances of co-operation and conflict between economic agents. Bargaining requires exclusively dynamic extensive-form analysis rather than strategic normal-form analysis.

The Rubinstein model is as follows: Assume two players. One makes an offer. The other can either accept or reject. As a tie break rule assume that an indifferent offer is accepted. Utility is discounted. The discounting of payoffs implies delay is costly. Assume different discount factors for the two players, representative of different degrees of impatience. Assume time passes between offers and counter-offers.

The conclusion of the model is that players realise the others degree of impatience and make an optimal offer, taking account of relative impatience. It can be shown that if player 1 is to make the offer, the outcome is:

If player 2 is to make the offer

if the discount factors were equal

The outcome of the sequential, alternating offer/counteroffer game is immediate and efficient. Rubinstein shows that the equilibrium is a subgame perfect equilibrium i.e. a Nash equilibrium based on credible threats.[12]

The outcome of the game exhibits the first mover advantage concept: he who moves first has the advantage in that he can determine what the other is allowed to take in the fixed output bargaining model. The follower will wish to avoid delay and so may accept an unreasonable offer. The eventual outcome will depend on relative discount factors, i.e. on to whom delay is more costly. The larger output will go to he who can remain unaffected by the passage of time.[13]

The Rubinstein model is susceptible to changes in bargaining protocols. From Kreps (1990) and Sutton (1986):

1. The existence of an outside option for either of the players does not affect the Rubinstein subgame perfect equilibrium. This may be an unintuitive result. You may expect an outside option to improve bargaining power.

2. If player 1s discount factor is less than 2s, then player1 has a better bargaining position as he is more patient.[14]

3. The quicker a player can respond with another offer the greater their share from fixed bargaining.

The deficiency of each is that the outcome is immediate and efficient. Bargaining is not delayed. Conflict does not occur. This is in contrast to what we observe in reality (See below). The main criticisms of Rubinstein concern its unrealistic solution (immediate and efficient) and its very specific bargaining procedure (perfect information alternating offer or counteroffer).

Rubinsteins analysis makes extensive use of game theory. There are numerous problems with games. How do you define a game? When is it completely presented? What is the appropriate equilibrium / solution technique?[15] It is left to the model builder to decide these options. He must distil from reality the necessary details of commitment, time and information upon which the solution of the model is contingent. For slight variations in the specifics, the institutional protocols, one will be defining a different game with a different solution. The problem is allowing enough of the relevant economic factors into the formal, mathematical game analysis. But, how do you model judgement and intuition?

The contention of the Nash Program[16] and the course of action recommended by Sutton (1986) to further develop bargaining, is the examination of contingent-specific models, i.e. analyse the effect of time, commitment, and information on the outcome of bargaining problems. The hope is to refine Nashs bargaining axioms. Time, in the form of delay, is costly and ill be avoided by rational players[17]. Commitment refers to the ability to commit to bargained over outcomes. It represents trust[18]. Information is an asymmetrically distributed scarce resource. Unevenly distributed information is a cause of anomalies (e.g. conflict) in the models. There is the dual problem of imperfect and incomplete information[19].

Lyons and Varoufakes (1989) introduce a reputation argument to explain the occurrence of conflict. One would expect rational and well-informed agents to avoid costly delay, but unions, by delaying agreement can in some sense signal a tougher reputation in bargaining through willingness to accept the costly delay. Conflict is thus exhibited by relaxing the perfect and complete information assumptions [20]. This is where the literature has proliferated. See Crawford (1982).

Cramton (1984) examines issues in the urgency of bargaining in the incomplete information setting. There is the need to acquire information, which is achieved through analysing your opponents bids. Waiting is costly, so there is an incentive to optimally communicate to resolve bargaining. Cramton shows that the outcome depends on whose information set is restricted. There may be an incentive to lie or manipulate the situation with signal bids. The revelation principle does however rectify this[21].

Fudenberg and Tirole (1983) examine how changes in bargaining costs, the size of the contract zone, and the length of the bargaining process can effect aspects of the solution such as the probability of impasse and the likelihood of concession.

One can also consider the implications of the outcome for industrial organisations of the debate on bilateral monopoly determination. Is the Walrasian equilibrium of perfect competition robust to how we define the bargaining process between the agents in the model? Markets are assumed to be frictionless, and trade occurs flawlessly. What, if like Gale (1986), one assumes that agents matched randomly cannot agree on terms of trade thus incuring a search cost in looking for their next match? The indeterminacy of bargaining imposes a negative externality on the operation of a market with perfectly competitive characteristics.The Walrasian equilibrium no longer holds. Could vertical restraints affect bargaining positions and powers? Would the cost of delays from incomplete information be enough to encourage vertical integration to avoid these costs? For an analysis of the incentives of firms to merge because of the bargaining problem, see Horn and Wolinsky (1988). The bargaining problem may not be eliminated by vertical integration. In the vertically integrated firm the inputs of one division are the outputs of another. If management was decentralized then the interdivisional costs would be determined by bargaining. These are the transfer prices. These transfer prices are subject to influence activities and costs[22] indicating that bargaining can still impose a negative externality on firms, even after vertical integration to remove them.

Bargaining has been central to economic theory throughout its history. Its apparent indeterminacy has been its deficiency. The resolution of the (in)determinacy problem has been addressed in static and now dynamic-interactive frameworks. The literature has proliferated and is coalescing on a contingent-specific approach. The indeterminacy is now being appreciated as the of real-life anomalies such as the absence of Walrasian prices, conflict, mergers, transfer prices etc.. Indeterminate bargaining - a hindrance to theory and an indication of reality.

Bibliography

Binmore, KG (1990) *Essays on the Foundation of Game Theory*

Binmore, KG & Dasgupta P (1987) *The Economics of Bargaining*

Cramton, P (1984) Bargaining with Incomplete Information: An Infinite Horizon Model with Two-sided Uncertainty in *Review of Economic Studies*

Crawford, V (1982) A Theory of Disagreement in Bargaining in *Econometrica*

Foldes, L (1964) A Determinate Model of Bilateral Monopoly in *Economica*

Friedman, J (1989) Bilateral Monopoly, in *The New Palgrave*: Game Theory.

Fudenberg, D & Tirole J (1983) Sequential Bargaining with Incomplete Information in *Review of Economic Studies*

Gale, D (1986) Bargaining And Competition Part I: Characterizationin *Econometrica*.

Gibbons, R (1991) *A Primer in Game Theory*

Gul, F & Sonnenschain, H (1985) *Uncertainty Does Not Cause Delay*

Harsanyi, JC (1989) Bargaining in *The New Palgrave*: Game Theory.

Horn, H.& Wolinsky, A (1988) *Bilateral Monopolies and Incentives for Merger*

Kreps, D.M (1990a) *A Course in Microeconomic Theory*

Kreps, D.M (1990b) *Game Theory and Economic Modelling*

Lyons, B & Varoufakes, Y (1989) Game Theory, Oligopoly, and Bargaining in *Current Issues in Microeconomics*

Milgrom, P & Roberts, J (1992) *Economics, Organization, and Management*

Nash, J (1950) The Bargaining Problem in *Econometrica*.

Nash, J (1951) Non-cooperative Games in *The Annals Of Mathematics*

Nash, J (1953) Two-person Cooperative Games in *Econometrica*

Osborne, M & Rubinstein, A (1990) *Bargaining and Markets*

Rasmusen, E (1989) *Games and Information*

Rubinstein, A (1982) Perfect Equilibrium in a Bargaining Model in *Econometrica*

Sutton, J (1986) Non-cooperative Bargaining Theory: An Introduction in *Review of Economic Studies*

Von Neumann, J & Morgenstern, O (1944) *The Theory of Games and Economic Behaviour*.

Wilson, R (1987) Game Theoretic Analyses of Trading Processes in Bewleys *Advances in Economic Theory*

Zeuthen, F (1930) *Problems of Monopoly and Economic Welfare*

IS GAME THEORY USEFUL FOR THE ANALYSIS AND UNDERSTANDING OF DECISION MAKING IN ECONOMIC SETTINGS?

Fiona Hayes
Senior Sophister

Game theory is best described as cold blooded rational interactive decision making.[4] It is the extension of individual rational decision making to the behaviour of rational decision makers whose decisions affect each other. In recent decades much progress has been made in applying game theoretic models to a wide range of economic problems. Indeed, Varian claims that most economic behaviour can be viewed as special cases of game theory. Game theory is divided into two branches, cooperative and non-cooperative. In non-cooperative game theory, individuals cannot make binding agreements and the unit of analysis is the individual who is concerned with doing as well as possible for himself, subject to clearly defined rules and possibilities. In cooperative game theory, binding agreements are allowed and the unit of analysis is the group or coalition. This discussion will concern itself exclusively with non-cooperative game theory. Furthermore, I will avoid complex mathematical equations and detailed analysis of the fundamentals of the theory, choosing instead to outline the most important concepts in simple terms and to see how the theory applies to specific economic examples.

One preliminary issue needs to be addressed in order to approach the discussion in context: by what standard are we to judge the usefulness of a theory? Kreps (1990) suggests that a useful theory is one which helps us to understand or predict behaviour in concrete economic situations. Hence studying the interactions of ideally rational people should aim to help our understanding of the behaviour of real individuals in real economic situations.

Formulation of the Game

Von Neumann and Morgenstern (1944) first recognised explicitly that economic agents must take into account the interactive nature of situations when making their decisions. Their methodology consisted of taking an economic problem, formulating it as a game, finding the game theoretic solution, and translating the solution back into economic terms.

So how do we formulate an economic problem as a game? The essential elements of a game are players, actions, information, strategies, payoffs and equilibria. The sets of players, strategies and payoffs combine to give us the rules of the game. There are two models the strategic and extensive forms. Indeed, the former can be viewed as a summarized description of the latter. Two equilibrium concepts commonly used are dominant strategy equilibrium and Nash equilibrium. A dominant strategy must be the best response to all possible strategy combinations by other players. A dominant strategy equilibrium is a strategy combination consisting of all players dominant strategies. It can be arrived at by iterative deletion of dominated strategies, i.e. eliminating all actions that the players will *not* choose. This requires common knowledge of rationality. A Nash equilibrium requires that si*[5] be the best response to a particular strategy combination by other players. The idea here is that no single player has the incentive to deviate. We can have equilibria in pure or mixed strategies. Nash equilibrium is the single game theoretic concept most frequently applied in economic examples. In all the examples which follow we will concern ourselves with the search for a Nash equilibrium (or some refinement thereof).

The Games

Consider this first example, known as the Prisoners Dilemma. We have two prisoners in separate cells, faced with the dilemma of whether or not to inform. Each player has defect as best response to any action by the other. The only Nash equilibrium is (defect,defect), despite the fact that both would be better off if they were to cooperate, because neither player acting unilaterally has the incentive to deviate. Note this is also a dominant strategy equilibrium (DSE).

The Prisoners Dilemma

		Player 1	
		Cooperate (Advertise)	Defect (Do Not)
Player 2	Cooperate (Advertise)	(-1,-1)	(-5,0)
	Defect (Do Not)	(0,-5)	(-4,-4)

This general sort of situation arises in many contexts in economics. For example, where we have two firms selling the same product and deciding whether to advertise or not. The Nash and dominant strategy equilibrium is (do not, do not).

Next let us turn to a Battle of the Sexes game. This game is representative of many situations in which two or more parties seek to coordinate their actions, although they have different preferences concerning how to coordinate. In this simple example, player 1 would prefer to go to a Chinese restaurant, while player 2 would prefer Italian, but both would prefer dining together to dining alone. We have two pure strategy Nash equilibria at (Ch,Ch) and at (It,It). We do not know which of these will be selected. Perhaps there is a focal point an equilibrium which for psychological and other reasons is particularly attractive. For example, the couple in question may have eaten Chinese food the previous night.

The Battle of the Sexes

		Player 1	
		Chinese	Italian
Player 2	Chinese	(6,3)	(0,0)
	Italian	(0,0)	(3,6)

There are many such coordination problems in macroeconomics. Kreps cites the example of two adjacent tax authorities who wish to coordinate on the tax system they employ in order to prevent taxpayers benefiting from any differences.

Application to Oligopoly Theory

One of the most successful applications of game theory has been to oligopoly theory. Oligopoly is concerned with market structures in which the actions/payoffs of the individual firms affect and are affected by the actions/payoffs of other firms. I will discuss three models of oligopoly, restricting our attention to duopoly for the sake of simplicity. We assume the firms are producing an homogeneous product.

Cournots Analysis

In Cournots (1838) model, firms choose quantities *(q*l,*q*2) from strategy space, S=(0,). Firms payoff functions are their respective profit functions:

where *q* = *q*l + *q*2 and *p(q)* is the inverse demand curve.

There is interdependence i.e. 1 depends not only on *q*l but also on *q*2. A Cournot Nash equilibrium occurs at a pair of output levels *(q*l*,*q*2*) such that neither firm could have obtained higher profit by having chosen some other output i.e. no player has incentive to deviate. Let us consider the simple example of a linear demand curve and constant marginal cost:

and, by symmetry:

These are the reaction curves for firm 1 and firm 2 respectively, and show the optimum reaction for each firm given how the other has reacted. Substituting, we obtain the result that:

So our equilibrium pair of output levels is

Stackelbergs Analysis

Stackelberg (1934) proposed a basic derivative of the Cournot model: the leader-follower duopoly, in which one firm, say firm 1 chooses *q*1 and, after that choice is communicated to firm 2, *q*2 is chosen. Firm 2 will always choose *q*2 according to *q*2 = *f(q*l) and this is known to firm 1 who maximizes *q*1. The output vector Q = (ql,q2) is a Stackelberg equilibrium with firm 1 as the leader and firm 2 as the follower if firm 2 maximizes profit subject to the constraint that firm 2 chooses according to his reaction function. Once again, neither firm can increase his profit by a unilateral decision to alter its output. Refering back to our simple example, the Stackelberg equilibrium is

We see that the profits of the leader will be higher than those of the follower because of first mover advantage. We also note that more is produced in this model than in Cournots model.

Bertrands Analysis

Bertrand (1883) realised that the focus on quantity competition was unrealistic. He proposed an alternative model whereby firms compete via price. Firms choose prices *(p*1, *p*2) from strategy space. The firms payoff functions are given by:

(Assuming a linear demand curve and constant marginal cost.)

The only Bertrand Nash equilibrium is where *p*l=*p*2=*c* and no firm can make a (higher) profit by altering its own price decision, i.e. even with only two firms we obtain competitive results. This is the Bertrand paradox: we know that firms do compete in prices and that they do make positive profits in the real world.

Returning once again to our simple example, we see that the Bertrand equilibrium outcome is the same as in the Cournot model. Despite the fact that the three models differ, they can all be seen as the application of the Nash equilibrium concept to games which differ with respect to the choice of strategic variables and the timing of moves. Cournot and Bertrand equilibria are the Nash equilibria of simultaneous move games where the strategic variables are quantities and prices respectively. The Stackelberg equilibrium is the subgame perfect (explained below) Nash equilibrium of a game where quantities are chosen, but the leader moves before the follower. We can conclude that, despite the fact that the same equilibrium concept is applied to all three models, the different outcomes suggest that oligopoly theory results are very sensitive to the details of the model.[6]

Dynamic Games

Thus far we have concentrated almost exclusively on Nash equilibria in their purest sense. The Stackelberg model hinted at further issues to be addressed. The first of these is that of time in games: what happens if we repeat a prisoners dilemma type game infinitely often? The Folk Theorem tells us that players acting non-

cooperatively can attain the cooperative outcome via trigger strategies, under certain conditions. In this case, it is in the long run interest of prisoners to cooperate, but it is in their short run interest to defect; only the (Defect,Defect) outcome can occur in equilibrium, but if the discount rate is sufficiently high, we may get a cooperative outcome.

Repeated games model the psychological, informational side of ongoing relationships.[7] Phenomena like altruism, trust, revenge and punishment are all predicted by the theory. In repeated games, payoffs in each period are determined solely by actions in that period, yet strategies are a function of the entire history of the game up to that period. Dynamic games reflect the fact that current actions affect not only current payoffs, but also opportunities in the future - we learn from others, we also teach.

Refinements of Nash Equilibrium

		Incumbant	
		Accommodate	Fight
Entrant	Enter	(-1,-1)	(-10,0)
	Stay out	(0,-10)	(-8,-8)

The entrant-incumbent example can be viewed as a two period game where the entrant makes his decision in the first period and the incumbent responds in the second period. This is our starting point for a brief look at some of the refinements of Nash equilibrium. If this is a once off game, (5,5) is a Nash equilibrium. So too is (0,10). But in the latter case the problem is that the threat of fight, if not enter, is not credible. Equilibria supported by incredible threats are not subgame perfect. (A subgame perfect Nash equilibrium must be the Nash equilibrium of every subgame within the game in question). Hence the only subgame perfect Nash equilibrium is (5,5). This concept can be applied to the problem of whether or not government plans are sustainable.

A related refinement is the concept of trembling hand perfection. Suppose there exists a small probability that players dont play their dominant strategies. A trembling hand perfect equilibrium must be robust against slight perturbations in strategy: (5,5) is trembling hand perfect; (0,10) is not. Issues surrounding this concept are similar to the problems of applying rational expectations models to study reforms and regime changes.[8]

No discussion of game theory is complete without reference to the crucial issue of information in games, i.e. what do players know about each other? There are four separate categories of information, but here we consider only incomplete information models i.e. where nature chooses a type for one player and this goes unobserved by at least one player. The solution concept we use is known as Bayesian Nash equilibrium, a simple variation of the concept we have used in all models to date. A specific category is signalling games. The classic example is Spences (1975) model of education as a signaling device in the job market. Other examples include price as a signal of quality in the goods market. Further examples of games with incomplete information include auctions (where parties hold proprietary information).

A Final Analysis

At this stage we have seen that the game theoretic approach can be applied to virtually all interactive decisions in economics. But what about the deficiencies? Firstly, game theory requires clearly defined rules of the game, and there is a tendency to take these rules too much for granted, without asking from where they come. Secondly, in many games we have multiple equilibria and no way to choose, e.g. Edgeworths bilateral bargaining problem in cooperative game theory. Thirdly, we have inefficient outcomes in many cases, because of payoff dependence. This can be seen as a type of externality. For example, in the Cournot duopoly model, both firms would do better to restrict output to half the monopoly output, but this is not a sustainable equilibrium as both have the incentive to cheat and produce more.

Despite these shortcomings, non-cooperative game theory has provided a unified and flexible language for analysing interactive decision-making in a wide variety of economic contexts. Often the results obtained are closely related to those from the more conventional approaches; in other cases, game theoretic models lead to

new insights. Furthermore, game theoretic models enable us to identify similarity in superficially different situations, and to move insights from one context to another. In addition, game theory rightly stresses the importance of mechanics i.e. who moves when and with what information. However, we must keep a sense of proportion about the usefulness of the theory. The problems mentioned in the previous paragraph suggest that we must supplement game theory with considerations about experience and memory[9] that cannot be incorporated into the formal structure of the theory. In conclusion, I echo the words of Kreps who maintains that we should be happily dissatisfied overall - happy with the progress that has been made, but dissatisfied with our inadequate knowledge about how individuals actually behave in a complex and dynamic world.

Bibliography

Aumann, R (1989) Game Theory in *The New Palgrave: Game Theory*

Cournot, A (1898) *Recherches sur les Principes Mathematiques de la*

Theorie des Richesses

Friedman, J (1989) Duopoly in *The New Palgrave: Game Theory*

Kreps, D (1990) *A Course in Microeconomic Theory*

Kreps, D (1989) Nash Equilibrium in *The New Palgrave: Game Theory*

Kreps, D (1989) *Game Theory and Economic Modeling*

von Neumann, J & Morgenstern, O (1944) *Theory of Games and Economic Behaviour*

Rasmusen, E (1989) *Games and Information*

Sargent, T (1992) Book Review in the *Journal of Political Economy*

Sonnenschein, H (1989) Oligopoly and Game Theory, in *The New Palgrave: Game Theory*

Spence, M (1974) *Market Signalling*

Tirole, J (1989) *The Theory of Industrial Organisation*

Varian, H (1987) *Microeconomic Analysis*

A DIFFERENTIAL APPROACH TO GENERAL EQUILIBRIUM

Alan G. White
Postgraduate

General equilibrium theory aims at studying the equilibrium price vector as a function of the parameters defining the economy; it describes those states in which the independent plans of many agents with conflicting interests are compatible such a state is called an equilibrium.

Ideally, it is desirable to have some general principles concerning the equilibrium concept. Naturally, the first question is existence and the existence theorem has been proved by Arrow-Debreu (1954). As equilibrium has no connections with well-being, welfare theorems are particularly interesting. The first and second theorems of welfare economics, which make explicit the connection between equilibrium and Pareto optimality are well understood. Uniqueness of equilibrium is desirable so that policy measures and outcomes can be equivocally compared. Finally, one would like to have a dynamic theory according to which equilibria are approached over time.

One can think of many examples of economies with an infinite number of equilibria, the two good, two agent Edgeworth box is a good example. It is an arduous task to provide conditions guaranteeing (global) uniqueness of equilibrium. Assumptions guaranteeing global uniqueness have been seen to be excessively strong.[1] Without results guaranteeing uniqueness of equilibrium comparative statics becomes devoid of meaning, and a dynamic theory becomes particularly difficult to develop.

The introduction of differential techniques into economics was brought about by the study of several basic questions stemming from models proporting equilibrium concepts, one such being the uniqueness issue. The theory of regular and singular economies can be seen as an attempt to progress general equilibrium theory in the absence of a definitive uniqueness result. Debreu (1970, p. 387) recognises the possibility of multiplicity of equilibria but requires them to be locally unique. He notes ...if the set of equilibria is compact (a common situation), local uniqueness is equivalent to finiteness. One is thus lead to investigate conditions under which an economy has a finite set of equilibria. The equilibria are well defined and are not destroyed by small perturbations in parameters. Indeed, the data of an economy (such as endowments) cannot be exactly observed. If the equilibrium correspondence *W()* is not continuous at the economy, a small observational error will yield entirely different sets of predicted equilibria: in this case the explanatory power of the model is limited. Thus the continuity principle is a desirable property. The differential approach to general equilibrium theory thus attempts to go beyond the often overstudied existence question, endowing the equilibrium set with a more regular structure and with differentiability assumptions it permits a greater examination of the properties of equilibria that are economically interesting (such as uniqueness).

The following exposition of the differential approach examines the nature and charac-teristics of the differential approach for a simple pure exchange economy the emphasis is not on intricacies of the mathematics *per se*. Section two outlines some of the mathematical concepts and definitions, while Section three presents the methodology of the differential approach. Section four considers some extensions that have been developed concurrently with the differential approach.

Notation

This section and the next draw heavily on Balasko (1988)

Consider a pure exchange economy where there are *l* commodities and *m* consumers where each consumer is endowed with an amount of each of the *l* commodities i *l*. Individual demand is *fi* and is assumed to be fixed.

The space of economies is denoted by = *lm.* Let *r l* , be the vector of total resources for the economy. Then the space (*r*) = { *lm* | i i = *r* } *is* the space of economies with totat resources *r* and *S* is the set of prices.

The equilibrium correspondence or the Walras correspondence (Balasko 1975a, p 907). *W()* associates with each economy the set of prices for which the economy is in equilibrium

W() = {*p S* | i *f*i (*p*,*p*.) = i i } *S*

The pair (*p*,) *S* x is said to be an equilibrium if supply is equated with demand at (*p*,). The set ***E*** *S* x denotes the set of equilibria. The natural projection (Balasko 1988) : ***E*** is the restriction to the equilibrium set ***E*** of the projection : *S* x : (*p*,) X Note that the preimage of the natural projection

-1() = {(*p*,) : (*p*,)} =

and thus -1() = *W*() x

It is easily seen from this that the study of the correspondence *W* is thus formally equivalent to the study of the function . Indeed the cardinal number of this set #-1 () is the number of equilibria associated with the economy . Thus #-1 () = 1 corresponds to a unique euqilibrium, and #-1 () = is the case of an infinte number of equilibria.

The program of study suggested by the above has become apparant. Firstly, the structure and properties of the equilibrium set *E* must be explored. Secondly, the relationship #-1 () = *W* () x indicates that studying the characteristics and behaviour of the natural projection will provide information on the equilibrium correspondance *W*.

The Nature of the Differential Approach

The standard axioms of preference (such as transitivity and convexity) do not guarantee the differentiability of the demand function. In order to use the differentiablility properties of demand functions it is necessary to strengthen the set of axioms without weakening their economic relevance. The solution to this lies in the mathematical concept of *approximation*, the best example of one such approximation technique being that depicted by the *Stone - Weierstrass Theorem.* One special case of this theorem states that any continuous function can be approximated arbitrarily close by a *C* polynomial (one whose derivatives exist up to any order exist and are continuous).

Approximation techniques are widespread in mathematics. Taylor expansions can be used to approximate more complex functions. Such techniques permit the original functions to be replaced by ones that are simpler to analyse, characterise and manipulate. The focus here is on the consumer and similar approximation results exist in the context of analysing preferences and consumer choice (see Mas-Colell, 1974). It can thus be assumed that functions in economics exhibit the property of differentiability.

The following definition provides the link with the differential approach.

Definition 1 *Let M V W N be a function where M,N are finite dimensional real vector spaces. The function f is called a Cr diffeomorphism if*

(a) V is open in M, W is open in N.

(b) V W is bijective

(c) f and f-1 are of class Cr, that is the derivatives up to the rth order exist and are continuous.

Theorem 1.

*The demand fuction of consumer i, f*i*: S* x *l is a C diffeomorphism. Put simply, this means that f*i *is smooth and bijective and has a well defined smooth inverse g*i *where*

*g*i *: l S* x : *g*i (*x*i) = (*grad*n *u*i (*x*i) , *x*i *grad*n *u*i (*x*i))

Structure of the Equilibrium Set

The desirable properties of discreteness of equilibria and continuity of the equilibrium correspondence can be examined once the theoretical framework has been established. The conditions for the solution of the problems posed are differentiability assumptions, and the main mathematical tools for the solution of the problem of discreteness of the set of equilibria have been provided by differential topology.

To appreciate the concept of smooth manifold it is necessary to have an understanding of the notion of a coordinate system which is used to parameterise (label) the elements belonging to a topological space. The search for structure then, is one of finding appropriate parameterisations.

Definition 2 *Let y = (y*1 *, ... , y*n *) be a sequence of real valued functions on an open subset V of a topological space. Then y is called an n-dimensional coordinate system on* n *with domain V if*

X V y(V) n

is a homeomorphism of open V onto open y(V) in n*. If z = (z*1*, . . . ,z*n*) is another n-dimensional coordinate system on X, then y is Cr compatible with z if F = z y*-1 *is a Cr diffeomorphism of y (V W) onto z(V W).*

Coordinate systems and compatible coordinate systems are brought together in the manifold concept.

Definition 3 *Let y = (y*1 *, ... , y*n *) be a sequence of real valued functions on an open subset V of a topological space. Then y is called an n-dimensional co-ordinate system on X with domain V if*

X V y(V) n

is a homeomorphism of open V onto open y(V) in n*. If z = (z*1 *, ... ,z*n *) is another n-dimensional co-ordinate system on X, then y is Cr compatible with z if F = z y*-1 *is a Cr diffeomorphism of y (*V W*) onto z(V W)*

Definition 4 *A topological space X is called a Cr manifold if a family Y, A of mutually Cr compatible coordinate systems is given, each Y has domain V and*

As a first approximation a manifold is a topological space that can be identified with a Euclidean space, that is, for any point belonging to the manifold, there exists a neighbourhood of this space that is hoeomorphic to a Euclidean space. More precisely, and in line with the formal definition, the smooth manifold structure consists of a collection of local coordinate systems covering the whole space, where functions defining coordinate substitutions are of class C.

Theorem 2

The equilibrium set E is a smooth manifold of dimension lm embedded in S x *and thus the equilibrium (p,) E can be parameterised by lm local coordinates.*

The natural projection : *E* : (*p,*) is a mapping from the equilibrium set *E S* x into the space of economies. It has been proved that the equilibrium set *E* has the structure to a smooth manifold diffeomorphic to *lm* via a diffeomorphism. This enables one to use the notion of smooth mappings, that is mappings of class *C*, whose natural projection is smooth, then it puts at ones disposal many powerful tools of differential topology.

Figure 1. Regular and Singular Values

Regular and Singular Economics

Definition 5: *Let f: X Y be a smooth mapping between two smooth manifolds X and Y. A point x X is said to be a critical point of f: X Y if the linear tangent mapping at x, Tx f is not surjective. The image f(x) is called a singular value. It is said that y Y is a regular value of f: X Y if y is not the image of a critical point.*

The way out of the difficulty with uniqueness and continuity issues is provided by differential topology. A well-behaved economy (with locally unique equilibria and continuous dependence with respect to the parameters defining the economy) is called a *regular economy* such that

(a) there are not many non-regular economies

(b) each regular economy has a finite set of equilibria.

(c) the equilibrium set depends continuously on the economy.

To illustrate informally the idea of singular and regular values outlined above, consider a function *f:* such that *f* is of class *C* (Debreu 1976). The critical points are those where the linear tangent mapping is not surjective (at that point the derivative is zero). The three critical points are *a,b,* and *c*. Critical values are, by definition the images of critical points. Thus the critical values associated with the critical points *a,b* and *c* are respectively *d,e* and *f*. Sards theorem guarantees that, in some sense there are not many critical values.

These concepts can be extended to examples involving smooth manifolds and diffeomorphisms. The equilibrium set of an economy defined by finitely many parameters can be represented by a finite-dimensional smooth manifold. To show that economies are well-behaved in the sense of criteria (a), (b) and (c), introduce a *C* funtion *F* from a manifold *L* to *M* such that a regular economy is defined as a regular value of *F* satisfying (a), (b) and (c). Sards theorem shows that most economies are of this type.

To cast the issues in the framework that has been constructed, precise definitions of *regular economies* and *singular economies* are now introduced.

Definition 6. *The economy is regular (resp. singular) if it is a regular value (resp. singular) value of the natural projection* : *E .. A singular value is the image of a critical point.*

Denote by **R,** the set of regular economies and by the set of singular economies.[2] As these are the only two classifications of economy types that will be used, then = **R** . For a regular economy the projection of the tangent space of ***E*** at *(p,)* covers .

Sards theorem states that the set of singular values of a smooth mapping has Lebesgue measure zero. Thus the set of singular values of is small from a measure theoretic point of view. It is possible to attach a probabilistic interpretation to Lebesgue measure. Essentially, the probability that a randomly chosen economy will be singular is zero, in the sense that singular economies are quite exceptional. Of course, having measure zero is not equivalent to smallness from the topological point of view.

Balasko (1975b) derives a criterion for ascertaining the type of a particular economy.

Theorem 3 (Balasko)

The equilibrium (p,) ***E*** *is critical if and only if*

det J (p,) = 0

Where J (p,) is the Jacobian matrix of aggregate excess demand z(p,), where z* is the vector of the first l*-1 *coordinates of z l .*

This theorem gives an alternate description of a regular economy, and while less convenient from a diagrammatic point of view, is useful for computational purposes.

Uniqueness and Multiplicity of Equilibria

The desirable property of uniqueness of equilibria has been extensively surveyed: see Arrow & Hahn (1971). It is seen that global uniqueness, while desirable is a property which is excessively strong: the global uniqueness of equilibria can only be derived under highly restrictive assumptions. Clearly then, uniqueness of equilibria must be established by diluting or weakening some of the more restrictive assumptions

Debreus 1970 paper made the break. He showed that the number of equilibria for almost all exchange economies with continuously differentiable demand funtions was finite, and moreover, that the equilibria were *locally* unique. Furthermore, by introducing the notion of regular and singular economies he established that the set of economies with an infinite number of equilibria has a closure whose Lebesgue measure is zero.

To explore the issues associated with the uniqueness and multiplicity of equilibria. It is necessary to exploit fully the structure of the equilibrium manifold and its associated topological properties. The following outlines some of the ideas of *connectedness* which has already been mentioned briefly. It can be shown that uniqueness of equilibrium can be established for Pareto optimal economies. For an alternate proof of Debreus finiteness theorem see Dierker & Dierker (1972).

Extensions to the Differential Framework

It must be emphasised that only one member of a very large family of general equilibium models has been examined namely the simple pure exchange economy with a fixed number of l commodities and a fixed number of m consumers. Needless to say, there is ample room for further extensions. Rather, the initial intention was to introduce the nature of the differential approach, and give a characterisation of some of its features. Some of the aims and results of the differential approach have been outlined in a pure exchange setting. Indeed, any investigation of these issues would require several volumes - instead some of the more important developments and extensions will be mentioned.

There are several aspects of economics that are impossible to ignore. Production is one such feature and its inclusion in the general equilibrium model is so obvious as to obviate the need for justification. Arrow-Debreu (1954) included it in their model and its inclusion in the differential framework not only makes sense, but is entirely logical and consistent with earlier developments in the subject. Taking uncertainty into account is another natural extension. Debreus (Debreu 1954, Ch 7) concept of a contingent commodity embodies in its definition the state of nature, the realisation of which is necessary and sufficient for the contingent commodity to actually be delivered.

The Core of an Economy

One very important theme in the development of mathematical economics in recent years has been the study of economies with large numbers of agents. This has been helped by the introduction into economics of game theoretic concepts such as the *core* and by measure theory. The core was first named by Gillies and Shapley (Aumann 1964) who studied properties of the von Neumann-Morgentern solutions. The core consists of that set of allocations which can be improved upon by no coalition of agents.

The relationship between the set of Walras (competitive) allocations $W()$ and the set of core allocations $\boldsymbol{C()}$ has been the subject of much research. The usual rough statement is that the core approaches the set of equilibrium allocations as the number of agents tends to infinity. The two sets satisify the mathematical trivial but economically important relationship (Debreu & Scarf 1963)

$W() \; C()$

Thus a state of an economy decentralised by an equilibrium price system cannot be improved upon by co-operation of individual agents. This strengthens the well known first theorem of welfare economics which states that every equilibrium allocation is Pareto optimal.

The equivalence of the set of Walras allocations and the core has been established by Debreu & Scarf (1963) who consider m types of each of r consumers - the equivalence is proved as r tends to infinity (m fixed) given the equal treatment property that allocation in the core assign the same consumption to all consumer of the same type.

The notion of perfect competition is fundamental in the study of economic equilibrium. A mathematical model appropriate to the intuitive notion of perfect competition is one which has a continuum of agents. Aumann (1964) notes that the reason for adopting this model is that one can integrate over a continuum, and that is, the actions of a single individual agent are negligible. Aumann introduces a atomless measure space

of agents[3] and the chief mathematical tools used are those of Lebesgue measure and integration, but only their most elementary properties are needed. Aumann proves that for a continuum of traders

W() C()

Basing his proof on one given by McKenzie (1959), he proves the existence of equilibrium for a market with a continuum of traders, thus establishing the non-emptiness of the core (Aumann 1966). Here, however, no convexity assumptions are needed to prove the existence in some sense the individually insignificant traders have a convexifying effect on the aggregate.

The mathematical notion that an economy becomes more competitive as the number of agents increases has led to questions concerning the rate of this convergence. Debreu (1975) considers an economy , whose n-replica *n* has each agent repeated *n* times. *Cn* is the core of *En* and *W* are the Walras allocations of and therefore of ***En***. Debreu proves

Theorem 4 (Debreu)

For a regular economy , as n tends to infinty the Hausdorff distance

(Cn,W) 0 (1/n)

That is to say that ***n(Cn, W)*** is bounded. In other words, the distance between the core of ***n ,*** and its set of Walras allocations converges to zero at least as fast as then inverse of the number of agents. Aumann (1979) notes that the generic nature of Debreus convergence theorem leaves open the possibility of cores converging as slowly as one wishes.

Global Analysis: Smales approach

The approach adopted by Smale to the study of general economic equilibrium was to put the main results of classical general equilibrium theory in a global calculus context. There has been a movement away from calculus during the 1950s recall that Debreu (1959) in his preface to *Theory of Value* set out to free mathematical economics from the traditions of differential calculus. One limitation of subsequent developments in the theory was in the inadequate treatment of how equilibrium was reached. Smale (1976, p 290) notes that one step in the liberation from the static point of view would be to use a mathematics of a different kind. The liberation that Smale speaks of is the liberation from the mathematics of fixed point theorems and such like. Interestingly, Debreu returned to calculus foundations in his 1970 paper on the finiteness of the number of equilibria. (Debreu 1970)

Smale (1976) enumerates a number of reasons for this reversion to calculus foundations. Dynamic questions are more accessible via calculus formulations. Comparative statics are integrated in the framework in a natural way as they depend on first derivatives. Mathematical approximations by differentiable *(C)* functions gives further justification to the use of calculus. It must be remembered that this particular approach is closest to the oldest traditions of the subject found in Hicks (1939) and Samuelson (1947). The goal of Smale was thus to approach equilibrium theory with mathematics with dynamic and algorithimic overtones.

Computation of Equilibrium

Although the first proof of existence is now close to sixty years old it is only in the past thirty years that attempts have been made to develop algorithims for the computation of Walras equilibria. Scarf (1967 & 1973) showed that how to compute an approximate Walras equilibrium and proposed a general algorithm for the calculation of a fixed point of a correspondence. This algorithm has been surprisingly efficient, though does not permit a gradual improvement in the degree of approximation of the solution. This problem was overcome by Eaves (1974)

Conclusion

The results of the differential approach have been extended and modified in different directions: the above represents only a small sample of the research that has been conducted in equilibrium analysis. Demand functions, which have been assumed to be fixed have been weakened and the number of agents has been increased. The computation of approximate economic equilibria, and the introduction of dynamics into the theory have progressed the mathematical development of the theoryof general equilibrium, providing a richer theoretical foundation for examining issues concerning the properties of equilibria.

Bibliography

Arrow, K & Debreu, G (1954) Existence of an equilibrium for a competitive economy in *Econometrica* Vol. 22

Arrow, K & Hahn, F (1971) *General Competitive Analysis*

Aumann, R (1964) Markets with a continuum of traders in *Econometrica* Vol 32

Aumann, R (1966) Existence of competitive equilibria in markets with a continuum of traders in *Econometrica* Vol. 34

Aumann, R (1979) On the rate of convergence of the core in *International Economic Review* Vol 20

Balasko, Y (1975a) The graph of the Walras correspondence in *Econometrica* Vol 43

Balasko, Y (1975b) Some results on uniqueness and on stability of equilibrium in general equilibrium theory in *Journal of Mathematical Economics* Vol 2

Balasko, Y (1988) *Foundations of the Theory of General Economic Equilibrium*

Debreu, G (1954) *Theory of Value, an Axiomatic Analysis of Economic Equilibria*

Debreu, G (1970) Economies with a finite set of equilibria in *Econometrica* Vol 38

Debreu, G (1975) The rate of convergence of the core of an economy in *Journal of Mathematical Economics* Vol 2

Debreu, G (1976) Regular differentiable economies in *American Economic Review Proceedings* Vol 66

Debreu, G & Scarf, H (1963) A limit theorem on the core of an economy in *International Economic Review* Vol 4

Dierker, E & Dierker, H (1972) The local uniqueness of equilibria in *Econometrica* Vol 40

Eaves, B (1974) *Properly labelled simplices* in Dantzig, G & Eaves, B eds *Studies in Optimisation, MAA Studies in Mathematics*

Hicks, J (1939) *Value and Capital, an enquiry into some fundamental principles of economic theory*

Mas-Colell, A (1974) Continuous and smooth preferences in *Journal of Economic Literature* Vol. 8

McKenzie, L (1959) On the existence of general equilibrium for a competitive market in *Econometrica*

NASH EQUILIBRIUM AS AN IDEAL SOLUTION CONCEPT

Diarmuid O'Conghaile
Senior Sophister

What interests economists is to model the strategies on which equilibrium is based, to understand and explain human behaviour. Traditional microeconomic theory regards the individual as a rational and utility-maximizing price-taker, which is to say that the individuals actions have no effect on the aggregate economy. Game theory, on the other hand looks at situations in which the aggregate outcome is dependent on the individuals actions. The individuals in the model know this and then form and pursue rational and utility-maximising strategies on the basis of that knowledge. This is what we call strategic interaction.

In the real world, we can observe an economy moving, over time, from one state to another, from one equilibrium to another. This movement takes place continuously, rather than in discrete jumps, but we do nonetheless speak of the economy as moving from state a to b. If we then employ the concept of strategic interaction we can say that the economy arrives at state b as a result of the interacting behaviour of the individuals involved, and that each individuals behaviour is based on his rational strategy for maximising utility. The unique equilibrium, state b, can therefore be thought of as occurring where the individuals strategies intersect. Game theory seeks to model these strategies on which an equilibrium is based. We construct small economies (self-contained worlds in which all other things are equal because there are no other things) and we watch what the players do, what choices they make and what equilibria they arrive at. This may strike the layman as a rather arcane activity, but the results have practical usefulness, in so far as game theory helps us to explain the strategies which underlie real-world outcomes, empirically observed.

A solution concept is a model of strategic interaction. In the real world, strategic interaction always produces a unique equilibrium outcome, a solution. In the world of games, the ideal solution concept is the one which similarly always produces a solution. No such concept exists at present. There are, as we shall see, many games in which we cannot produce any equilibrium outcome. Such situations represent modelling breakdowns. However, as already suggested, the strategies on which an outcome is based are of greater theoretical importance than the outcome itself. And it is always possible to demonstrate (for games with a finite number of players, each having a finite number of possible actions to choose from) unique (Nash) equilibrium in strategies, be they mixed strategies or pure strategies. And by running these strategies off we will always produce an outcome. The fact that we cannot say which outcome, in a given situation, is of secondary importance. We need, finally, to make one refinement of our discussion so far before we proceed. When we speak here about producing an equilibrium outcome, we mean logically deducing the outcome with recourse only to the players strategies, as defined, and the rules and parameters of the game. The actual mechanics of how the equilibrium is arrived at are not touched upon, and may even be regarded as something of a mystery.

Dominant Strategy Equilibrium Concepts

The dominant strategy equilibrium concept is the most basic solution concept, in some sense, the orginal or natural solution concept and therefore a good starting point. I will illustrate the workings of the concept with reference to the following games.[23]

The Prisoners Dilemma

		Player 1	
		Deny	Confess
Player 2	Deny	(-1,-1)	(-10,0)
	Confess	(0,-10)	(-8,-8)

(Confess, Confess) is the unique equilibrium solution in the one-shot Prisoners Dilemma. We speak here of an equilibrium in strictly dominant strategies. For either row or column to play confess is the unique optimal strategy (ie that strategy which maximises game utility) regardless of what the other player will do.

Game 1

		Player 1		
		Left	Middle	Right
Player 2	Top	(4,3)	(2,7)	(0,4)
	Bottom	(5,5)	(5,-1)	(-4,-2)

Game 2

		Player 1		
		Left	Middle	Right
Player 2	Top	(0,2)	(1,0)	(0,1)
	Bottom	(0,3)	(0,1)	(1,0)

Games 1 and 2 represent refinements of the dominant strategy equilibrium concept. In Game 1, there is no equilibrium in strictly dominant strategies. But by the process of iterative deletion of strcitly dominated strategies we can arrive at (bottom, left) as a unique equilibrium. A strictly dominant strategy is each players best reponse to the other players optimal strategies. The iterative process relies on the perfect rationaltiy of all players and each successive iteration adds another layer of assumed rationality: Im watching you watching me watching you and so on. In Game 2, we see the idea of weak dominance. In general, the weakly dominant strategy is player 1s best response (to any strategy played by player 2) in the sense that it is at least as good as any alternative strategy player 1 might employ and better than at least one alternative strategy.

The dominant strategy equipment offers us four solution concepts: strict dominance, iterated strict dominance, weak dominance, and iterated weak dominance. The concept of weak dominance is not entirely satisfactory because it relies upon indifference. Strict dominance, on the other hand, is a very satisfactory, and very strong equilbrium concept. We can regard it, in some sense, as the natural equilibrium concept.

Introducing the Nash Equilibrium Concept

Generally, however strict dominance proves too strong. That is, the rigorous conditions which define the strictly dominant strategy equilibrium do not occur often, as is the case with the following games.

Boxed Pigs

		Small Pig	
		Press	Wait
Large Pig	Press	(5,1)	(4,4)
	Wait	(9,-1)	(0,0)

In Boxed Pigs, there is no dominant strategy equilibrium of any kind. We require a new concept, the Nash equilibrium. A Nash equilibrium involves a set of beliefs (each players beliefs about the others utility-maximising strategy), confirmed in equilbrium, where each players strategy is the best response to the others. Put precisely, in a Nash equilibrium each players strategy is optimal given the others strategy. (We must always remember, however, that the other persons strategy is given only in the sense that it is expected.) In Boxed Pigs, the Nash equilibrium is (press, wait). Each pigs strategy is the best reponse to the others. We immediately see that the concept of Nash equilibrium is similar to the concept of iterated dominance, and, in fact, by theorem, every dominant strategy equilibrium is also a Nash equilibrium. But the converse is not true. Nash equilibrium is a much more powerful concept.[24] It produces equilibria where the dominant strategy concept cannot.

The basic thesis of this essay is that the Nash equilibrium concept, while not being ideal, is nevertheless an intuitively plausible and universally applicable concept. I will now move on to discuss some of the problems with the Nash equilibrium concept, and will look also at some of the ways in which these problems can be tackled.

Problems and Solutions in Static Games

Battle of the Sexes

		Female	
		Opera	Fight
Male	Opera	(4,2)	(1,1)
	Fight	(0,0)	(2,4)

The problem which arises in the Battle of the Sexes is that we have two equilibria and no way of deciding between them. (This is a specific instance of the general problem of multiple equilibria.) Battle of the Sexes is a co-ordination game. To maximise utility, each player must co-ordinate his strategy with the other players strategy. Communication, by definition, is not allowed. Various solutions to this problem are suggested in the literature. If the game is continuously repeated, for exmaple, it may prove possible for the players to evolve some form of mute communication through how they employ their strategies. It is also suggested that significant payoff dominance for one or both of the players may act as a focal point, making that particular equilibrium more attractive than the alternative(s). Another idea is that of correlated strategies. These involve exogenous focal points, which are certainty important in reality, but they have one or two drawbacks from a theoretical point of view unless they can be incorporated into the model. One possible method of doing this would be to give the first (or a pre-game) move to Nature and allowing all players to observe the move. But this alters the model rather than solving the problem, a distinction which may be important.

Welfare Game

		Pauper	
		Try to Work	Be idle
Government	Aid	(3,2)	(-1,3)
	No Aid	(-1,1)	(0,0)

In problems of multiple equilibria where none of the equilibria are compelling and/or where re-specification is either impossible or undesirable, the mixed strategies approach can be employed. We now discuss this method with reference to a different problem: namely, when there is no equilibrium in pure strategies. This is the case in the Welfare Game, which, in a sense, is the inverse of the co-ordination game. The players are trying to out-guess each other. We solve the problem by allowing the players to randomize their actions. Since Nash equilibrium is an equilibrium in beliefs, the beliefs in this case will be probabilities. The government will have a belief about the probability of the welfare-maximising pauper deciding to idle. And the pauper will have a belief about the probability of the welfare-maximising government granting welfare. These beliefs are confirmed in equilibrium. Since the welfare-maximising paupers strategy (which is equal in

equilibrium to the govenments beliefs about the paupers strategy) is his best response to the welfare maximising behaviour of the government, we derive the paupers equilibrium strategy by maximising expected government utility using the von Neumann-Morgenstern method. The governments equilibrium strategy is calculated in the corresponding way. And we arrive at a Nash equilibrium in mixed strategies. This is a remarkably powerful methodology. It guarantees us an equilibrium in strategies in every case (subject to the conditions of the Nash theorem), and is, in that sense, an ideal solution concept.

There are, however, certain difficulties with the mixed strategies approach. To begin with, we emphasised in the introduction how important the assumption is to economics that human behaviour is non-random. In the mixed strategies approach, we have allowed players to randomize their actions. The behaviour, however, is not truly random. It simply appears random. It is sometimes suggested in the literature that we interpret the equilibrium probabilities in the following way. In aggregate, the pauper will choose to be idle four out of five times. He may consciously randomize his decision the first time around (by tossing a coin for example) but thereafter his strategy is no longer random. It depends on all previous moves of which he is assumed to have perfect recall. This can create a problem though. If the pauper looks for work in the first of a series of five games, then the government knows that hes going to idle for the next four and will deviate from its Nash equilibrium strategy and play no aid in each of the subsequent moves. Another way of looking at the randomization is to have the government facing a population of paupers, and believing that 20% will choose the pure strategy of seeking work, and that the other 80% will choose the pure strategy of idling. Alternatively, one individual may have been drawn randomly from a population of the above characteristics. The government, in this case, does not know whether the one pauper it is facing is an idler or a worker. A more serious technical problem than the idea of randomness, is the fact that a player who chooses a mixed strategy is indifferent between the pure strategies on offer. Should this cause the player to deviate, even slightly, from the equilibrium mixed strategy, then the equilibrium collapses. Mixed strategy equilibria are thus not as robust as would be desired.

On the subject of robustness, and of slight deviations from equilibria strategies, we ought to mention the trembling-hand concept. Rasmusen (1989) defines trembling-hand perfection as an equilibrium concept which says that for a strategy to be part of an equilibrium it must continue to be optimal for a player even if there is a small chance that the other player will pick some out-of equilibrium action (that the other players hand will tremble). We sketch the idea here in its simplest and most intuitive form.

Game 3

		Player 1	
		Left	Right
Player 2	Top	(10,0)	(5,2)
	Bottom	(10,1)	(2,0)

Consider the game above. The original equilibrium in this game is (bottom, left). But if there is any chance that row is aware of, that column might make an error and deviate from his equilibrium strategy (even if there is only a small chance of such an error). Bottom is no longer strictly optimal row, depending on his aversion to risk, may play top instead. Column anticipates such behaviour and responds optimally with right. The trembling-hand perfect equilibrium is (top, right). We can note here that (top, right) is a better outcome for column and a worse outcome for row. Column has gained by having a reputation for unsteadiness of hand.

We do not regard the trembling-hand concept as being particularly intuitively plausible, although it clearly does depend on the individual circumstances of the game. People tend to neglect small risks. Life is too difficult as it is for us to worry about getting knocked down everytime we cross the road, especially if we want to get to the other side. There are a multitude of small risks that we could worry about. Which trembles are we to fear ?

Subgame Perfection In Dynamic Games
Entry Deterrence Game

		Incumbent	
		Fight	Don't Fight
Entrant	Enter	(0,0)	(2,1)
	Stay Out	(1,9)	(1,9)

In the above game, the entrant moves first and is observed by the incumbent. This game has one weakly dominant equilibrium (stay out, fight), which is also a Nash equilibrium, and one further Nash equilibrium (enter, dont fight). If we look at the game in its extensive form, it becomes clear that one of the Nash equilibria is not viable.

fight (1,9)

stay out B1 dont fight (1,9)

A enter B2 fight (0,0)

dont fight (2,1)

If the entrant enters, that is a *fait accompli*, and the rational welfare-maximising response of the incumbent is to choose dont fight. However, it is the threat of fight that supports the (stay out, fight) equilibrium. And if that threat is not credible (as is the case when the welfare-maximising incumbent will always play dont fight in response to enter), then the (stay out, fight) equilbrium is unviable. A rational entrant will never play stay out when he knows that he can maximise welfare in the (enter, dont fight) equilibrium. For the threat to be credible, the incumbent would have to commit himself to the fight strategy regardless of what the entrant plays. But it is hard to envisage how such a commitment itself could be credible. Once we arrive at node B2 (the B2 subgame) the rational welfare maximising strategy of the incumbent is to choose not fight regardless of what he has said beforehand. This is the issue of subgame perfection. It arises only in dynamic games and involves the notion of the equilibrium path. As we can see in the Entry Deterrence Game, it is an off-equilibrium response (the lack of credibility of the fight threat) which undermines the (stay out, fight) equilibrium. Or to put it another way, the (stay out, fight) equilibriumm depends on irrational off-equilibrium behaviour. A Nash equilibrium can thus be redefined as subgame perfect if the players strategies constitute a Nash equilibrium in evey subgame (Selten, 1965).

Final Comments

I have confined myself here to an introductory and intuitive discussion of the Nash equilibrium concept. There are many refining (strenghtening) mechanisms which we simply havent mentioned. In dynamic games of incomplete or asymmetric information, for example, players update their equilibrium beliefs by observing what of the hidden or private information has been revealed (or suggested) in the preceeding move(s). We use the Bayes Rule for this updating process, and the refined Nash concept is termed perfect Bayesian equilibrium. There are also other refinements such as the Kohlberg and Mertens (1986) concept of stability, and the Myerson (1978) concept of properness. In general, as we would expect, as games get richer (ie as the models get more complex) the solution concept needs to be refined. But the basic idea of the Nash equilibria remains intact.

Varian (1987) writes as follows: Another problem with the Nash equilibrium of game is that it does not necessarily lead to Pareto efficient outcomes. This misses the point. It may not be optimal for people to behave in a Nash fashion. But if people do behave in a Nash fashion (or, to be more precise, in what approximates to a Nash fashion), as evidence suggests that they often do, then Nash is the optimal way of modelling that behaviour. As Rasmusen says in his introduction : It (game theory) has moved away from maximization by calculus, and inefficient allocations are common. The players act rationally, but the consequences are often bizarre, which makes application to a world of intelligent man and ludicrous outcomes appropriate.

Bibliography

Gibbons, R (1991) *A Primer In Game Theory*

Kreps, DM, (1990) *A Course in Microeconomic Theory*

Rasmusen, E, (1989) *Games and Information*

Varian, HR, (1987) *Intermediate Microeconomics*

ECONOMICS OF AN INDEPENDENT EUROPEAN CENTRAL BANK

Myles Clarke[43]
Senior Freshman

> . . . . the heart of what a Central Bank does is essentially monetary policy and price stability, which includes the stability of the monetary and payment systems. Eddie George

With the establishment of the European Monetary Institute (EMI) earlier this year the existence of a European Central Bank (ECB) comes a step nearer. If a single currency were to be established, this institution would have control of European monetary policy, with the basic objective of price stability. This essay attempts to give the reader some idea of the environment in which an ECB could exist, given the direction in which central banking, in general, is going. I stress the word attempt, because the structure and operations of an ECB are very difficult to predict. The Maastricht Treaty sets out that an ECB would have to be independent; this is consistent with the international tendency for countries to grant their Central Banks independence or at least make them less subservient to their associated governments.

This essay is divided into four sections. The first section discusses the possibility of a causal relationship between independence and price stability and briefly outlines a definition and measurement of independence. The second section discusses the contentious issue of a Central Banks credibility and the accountability of its actions to the public. The third section outlines why an ECB should exist and if its process of establishment has anything to learn from the experience of the Bundesbank and the US Federal Reserve. The final section concludes the essay and comments on the recent changes within the European Monetary System (EMS).

Independence and Price Stability

Initially we do not pose the question why a Central Bank (CB) should be independent, but rather why it ought not be subservient to its government. Firstly, assuming the ultimate aim of the CB is price stability, a subservient CB faces trade-offs between low inflation and other goals like low unemployment, excess creation of seigniorage revenues and balance of payments objectives. Secondly, the high rate of time preferences of governments is exacerbated by the Political Business Cycle and policy vacillations of changing governments. Thirdly, while the relative preference of the government for price stability is subject to political shocks, it is constant for the CB. Lastly, governments may have a penchant for allowing high levels of inflation to persist as they reduce the real value of national debt. There is no incentive for a CB, with the ultimate objective of price stability, to create surprise inflation.

The growing consensus is that a highly independent CB with the overriding objective of price stability results in lower mean levels of inflation and less uncertainty about future inflation. While the former is important the latter is most salient as uncertainty results in sub-optimal investment decisions by shortening effective planning horizons.

Source: *Central Bank Independence and Monetary Stability* Issing, O (1993)

Attempting to prove a causal relationship between low inflation and highly independent CBs has been difficult. Measuring independence is an arbitrary process. Independence does not mean a CB can do anything it pleases. It means the freedom to pursue price stability possibly at the cost of other objectives. It involves the following integral elements:

1. Formalised Independence: the CBs charter must allow it legal authority to do what is necessary to achieve its objective(s) without seeking constant government approval.

2. Material Independence: the CB should be permitted to execute day-to-day operations without obligation to exchange government gilts.

3. Personal Independence: the executive council of the CB must be granted security and optimal duration of tenure.

4. Financial Independence: a revenue pool that is adequate and guaranteed to satisfy daily operational expenses.

Cuckierman (1990) minimises anomalies by including such variables as level of legal independence, turnover rate of CB governors and results from returned questionnaires in his independence index (see Diagram). Swiss and German CBs, observed during the period 1980-90, have successfully combined low inflation with a high degree of independence. Not featured is New Zealand. Had it been, these results, having been formed before its CB was made legally independent (l990), would have shown a mean level of inflation of 14%. In the subsequent years of independence, the New Zealand CB has achieved a mean inflation level of 1.5%. But not all countries fit the trend. It must be noted that this inflation-independence relationship breaks down when applied to less developed countries due to their inadequate macroeconomic and bureaucratic frameworks for accommodating such an institution.

Adam Posen (1993) refutes the suggested relationship between CB independence and price stability as being causal. He claims that low inflation will be a corollary of CB independence only if a consensus exists within the respective financial markets promoting that independence. He goes on to say that open economies, because they are in direct threat of imported inflation, are more aware of its costs and so more averse to it. This implies that the financial markets in an open economy, in realising the accompanying benefits of price stability, would welcome their CB becoming independent. Their confidence in that CB would lessen the risk of imported inflation.

This essay will now focus on the role which formalised independence will play in the future of central banking and especially that of the ECB. For the ECB to promulgate true independence (i.e. an independence in which the public will believe), its charter will have to accommodate a strong sense of formalised independence an independence which is encapsulated in the essence of credibility.

Credibility

One of the main factors which will decide how independent an ECB (or indeed any CB) will be is its charter. It will also determine how democratically accountable the CBs actions will be to parliament, a point which is in direct conflict with the credibility[44] of an essentially apolitical institution. The borders within which a CBs charter defines its role in the economy determines the credibility it deserves. The financial markets will question the credibility of a CB whose projections have been subjected to government interference, fearing a hidden agenda implicit in such forecasts. It is imperative that a CBs credibility is preserved as this will be a major factor in determining the efficient allocation of resources in the money and capital markets. Credibility may also reduce the costs of an anti-inflationary policy. If wage and price-setters recognise that an independent central bank will not accommodate inflationary pressures, they are more likely to show moderation in wage and price rises. This will allow inflation to be reduced with less of a loss in output than otherwise. However, the central banks independence must not be undemocratic. It needs to be answerable to parliament for its anti-inflationary performance.

What charter could consolidate these two conflicting factors? A charter must be decisive in defining the CBs objectives, (the principle objective being price stability) and must restrict the CBs activities to only those that are necessary to achieve these objectives. Such a charter leaves little room for government intervention and keeps the nations best interests in focus. The latter also satisfies the condition of democratic accountability. But because it is impossible for the charter to include appropriate instructions for the panoply of contingencies that may arise, it cannot be followed verbatim in every situation. This infers that in the case of extraordinary circumstances, the charter should allow independence to deviate from that level which is legislated. Therefore, the CBs charter must be flexible enough to allow its incumbents to act intuitively in certain situations, while maintaining optimal credibility.

The issue of central bank independence is a direct illustration of the time inconsistency theory developed by Kydland and Prescott in 1977. A policy may seem optimal in period of announcement t, but may become sub-optimal in the period in which that policy impacts upon the economy, t+i. But at this stage contracts have already been agreed upon given the policy projections of period t. Future policy formulation is considered

incredible in the eyes of the public, thus they heavily discount announced policies believing that actual policy will deviate from its announced counterpart. If government policy announcements are time inconsistent, then the monetary policy of a subservient CB will suffer from their inflationary bias. In this instance should the independent CB be allowed operate with complete discretion? No. This will conflict with the accountability obligation imposed on all CBs. The rules alternative for CB operations is too inflexible as it does not allow for contingency situations like economic shocks. Is there an optimal operational mix within the spectrum of CB regulation? The New Zealand Reserve Bank (granted independence in 1990) has struck an interesting balance between these two regulation extremes. It sets a target every three years for inflation and includes a punishment clause to affirm a commitment to achieve that target if inflation exceeds 2% (subject to re-evaluation in exceptional circumstances), then the governor of this CB is dismissed.

Another factor which will enhance the likelihood of a successful and efficacious independent CB is the reduction of its responsibilities and workload. The too many cooks analogy is drawn by The Economist in this regard. It speculates that a CB with too many things to do is unlikely to do any of them well. The servicing-out of subordinate activities allows a CB to focus on achieving its primary objective(s). This was the intention of the Bundesbank when it created a national treasury agency , thus off-loading the task of debt management to a separate entity. In Europe this could lead to the introduction of a Securities Exchange Committee (SEC) type authority with the specific chore of supervising commercial banking activities. This would also eliminate the risk of imprudent banking which damages the reputation of the associated CB, as occurred with Threadneedle St. and the BCCI controversy.

History Lessons for the ECB

It seems suitable at this stage to introduce an historical perspective to the essay. By considering the operations of two of the most independent CBs in the world, the US Federal Reserve and the Bundesbank, we may attempt a more informed speculation about the possibility of a successful EuroFed.

At the time the US Fed was founded, world trade had developed to an extent that it was sufficient to describe the banks goal as, . . . to furnish an elastic currency. The currency was defined as elastic vis-à-vis seasonal variations in the needs of business. The bank would be facilitated in achieving this goal by the Real Bills doctrine. This proposed that the bank should confine its discount rate manipulations and open market operations to . . . short term and self-liquidating paper. The subsequent increased lending (due to lender of last resort obligation), derived from this doctrine, would finance stock market loans and fuel speculation. This seems far from conducive to the projected charter of a EuroFed. Furthermore, such intervention could seriously impair the reputation of an ECB.

The US Fed and the Reichsbank (predecessor to the Bundesbank) share the similarity that both were formed after political unification of their dependent states/landes, was achieved. Furthermore, even before unification, Germany had a single currency being issued by a single bank the Prussian which was integrated with the Reichsbank on the completion of unity. The European Union (EU) will be attempting to form a similar institution without such political and economic cohesion. Given that the Reichsbank primary functions were to serve as the governments fiscal agent and earn seigniorage to augment its reserves, it seems anachronistic and even irrelevant to hope for inspiration in structuring a CB on the scale of the ECB from the experiences of the German model of independent central banking. I acknowledge that the Maastricht version of an ECB was based on the Bundesbank, but this was only used as an example of a precedence where combining independence with an overriding objective of maintaining the value of money in a European country proved successful. Secondly, the Bundesbank was formed in a background of hyperinflation resulting in a highly inflation averse bureaucracy. All this indicates to the a priori structure and operations of an ECB as being a whole new ball game!

Penultimately, this essay discusses the extent to which fiscal and/or monetary policy will feature in the daily operations of the ECB. The ECB will not be obliged by a Bretton Woods type convention to intervene on international markets to stabilise third currencies. Nevertheless, the Louvre Accord (1987) may have set a precedence, whereby in the case of certain currencies in exceptional circumstances, foreign exchange rate intervention would be deemed as highly desirable. This could open the EU to imported inflation. The degree of independence granted to the ECB will determine how willing it will be to threaten price stability in order to stabilise exchange rates and hence stabilise trade with the associated countries. Thygesen and Gros (1992)

have suggested that pooling of international reserves to finance currency stabilisation activities would imply partial pooling of certain monetary instruments and thus by definition, some pooling of operations. Sufficient pooling of authority will be necessary for the interest rate arbitration function of the ECB and for reserve requirement co-ordination. This calls on EU Members to voluntarily relinquish control over some important fiscal and monetary instruments. In Article 104.1 of the Maastricht Treaty, the EU was adamant about the role which fiscal policy would play in the ECBs operations: Overdraft facilities or any other type of credit facility with the ECB or with the national central banks of the Member States in favour of Community institutions or bodies, central governments, regional, local or public authorities, other bodies governed by public law, or public undertakings by Member States shall be prohibited as shall the purchase directly from them by the ECB or national central banks of debt instruments.

Independence is also a function of the extent to which a government can borrow from its associated CB. The above regulation seems very stringent, but it is vital if the ECB is to be attributed with optimal credibility. With the Belgians tradition of high budget deficits, compulsory independence will restrict its CBs ability to facilitate fiscal policy, a facility the Belgians would claim is crucial to them for protecting their own economic agenda. This is true for other Members including Ireland and Greece. It illustrates the problems associated with lack of cohesion among Member CBs.

Conclusion & Comment

This essay asked the questions why should a CB be independent and why an ECB should exist and be independent? History provides little precedence in developing a charter with which resources can be allocated efficiently within the bank. The ECB is in a unique situation and therefore must be of a unique structure and design. The trust inherent in the Bundesbank could not be extended to a multi-cultural ECB. The corrupt histories of certain European bureaucracies (most notably Italy) might cause the public to consider the activities of the associated CBs with an air of suspicion, thus impairing the credibility of announced ECB policies. Such suspicion would demand greater transparency and accountability with respect to the activities of the ECB. But the area of accountability conflicts with the credibility aspect of central banking which is crucial to the effectiveness of its policies. Confidence in a CB has to be earned as much as bestowed. Such confidence would develop from a charter which was austere in delegating a few key tasks to the ECB combined with decisive detailing of price stability as its overriding objective. While an ECB could not function in a political vacuum, current EU political and economic cohesion criteria will be sufficient to accommodate this relatively apolitical institution and so perfect political unity will not be a necessity.

Finally, I will consider the recent developments within the EMS which could threaten the future prospects of the ECB. The break up of the ERM has been blamed on the Bundesbanks inconsiderate response to the European currency crisis. Yet blaming the Germans only contradicts a belief and perpetuates a disbelief in the efficiency of independent central banking. Even if the Bundesbank had reduced its Lombard rate at the time, the face of the EMS would still appear quite war torn. This experience highlights the problems of converging the economies and financial markets of countries with such diverse historical and cultural backgrounds. Ultimately, it will be the success of stage 2 of the Delors road to EMU which will dictate that of stage 3, a stage in which the largest, most powerful and influential financial institution in the world will be established. It must be said that whether they were aware of it or not, the people of Europe ensured the future of the ECB when they ratified the Maastricht Treaty. So while its structure and operations may be in moot, its future existence is not.

Bibliography

Crawford, M (1993) *One Money For Europe*

Cukierman, A (1992) *Central Bank Strategy, Credibility and Independence*

DeCecco, M (1989) *A European Central Bank?*

Gross, D & Thygesen, N (1992) *European Narrow Money & Too Many Pies for the Old Lady*

Issing, O (1993) *Central Bank Independence and Monetary Stability*

Mankiw, N (1992) *Macroeconomics*

Posen, A (ed. 1993) *Finance and the International Economy*

THE DESTRUCTION OF THE SECOND BANK OF THE UNITED STATES RATIONALE AND EFFECTS

Gareth Davis
Senior Sophister

I have always been afraid of banks Andrew Jackson

Few greater enormities are chargeable to politicians than the destruction of the bank of the United States R Caterall

The motivation behind this paper is to analyse from the perspective of a historian of economic thought and policy the rationale and implications of the destruction of the Second bank of the United States. The account is valuable as an account of the way in which economic thought, political ideology and vested interests can combine to shape policy. The debate also raised issues which are relevant to our modern economic system. The case for state supervision over the banking system is considered by almost all economists as to be so self-evident. But this has not always been the case and the debate over the banks future is a pointed reminder of this fact. Some of the arguments furnished against the bank may challenge some of the complacently held axioms of modern thought.

Other issues of this period relevant today include the benefits or otherwise of inter-regional monetary union. (Frass (1974) has shown how the BUS acted to standardise local regional exchange rates and nominal price levels and the effect which this had on peripheral areas.) Likewise the conflict laid bare some differing, and still widely held, preconceptions regarding both the optimal and the legitimate magnitude of government intervention. This account may not offer exact policy prescriptions for modern economists, since the economy, society and prevailing values have changed so much, but they can offer fresh perspectives to modern thinkers.

The Banks Destruction In Historical Context

The second Bank of the United States (BUS) was founded in 1816 on the basis of a twenty year charter. This charter empowered the bank to act exclusively as the federal governments fiscal agent, holding its deposits, making inter-state transfers of federal funds and dealing with any payments or receipts with which the federal authorities would be involved. Like all other chartered banks, the BUS also had the right to issue bank notes on the basis of a fractional reserve system and to carry out the usual commercial banking activities. In return for these privileges certain conduct of a central bank-like nature was expected of this institution: in the words of the charter the bank will conciliate and lead the state banks in all that is necessary for the restoration of credit, public and private.[45] Despite being 80% privately owned, its operations were subject to supervision by Congress and the President. Pessen gives details which show the banks size and the nature of the activities immediately prior to the assault on it in 1830. It was large relative to other banks, responsible for 15-20% of bank lending in the USA and accounting for 40% of the bank notes then in circulation. It was cautious in its note issuing function, holding a specie reserve of 50% of the value of its notes whilst the norm for the remainder of the banking system was 10-25%.

The 1820s and 1830s in the United States were a time of extremely rapid, but also volatile economic growth. New natural resources were being exploited as the frontier expanded and the new techniques of the industrial revolution were being introduced. The old money supply of gold and silver specie was stretched and found inadequate for the liquidity needs of the growing economy. (Temin indicates that in 1830 the total value of the gold and silver specie in circulation in the economy amounted to only l/30th to l/50th the value of GNP.) The emergence of a number of banks operating fractional reserve note-issuing systems was the result. The notes were underwritten by varying proportions of specie and although not legal tender were widely accepted in payment for debts, although usually discounted below their par value.

The quality of bank notes varied. Fraud was commonplace by unscrupulous bankers who could persuade or bribe the local state legislature to grant them the charter necessary to commence a banking business. For instance Pessen notes that in 1828 the 17 banks chartered in Mississippi circulated notes with a face value of $6 million from a specie base of $303,000. It was in such an environment that the Bank of the United States operated. One of its functions was to discipline and support state banks. As the federal governments fiscal agent it received bank notes in payment for taxes. The BUS would then present these to the issuing state bank in order to redeem them for the gold necessary to pay the taxes it had collected to the federal treasury. In this way state banks were forced to keep a higher stock of specie on reserve than would otherwise be necessary. [46] Conversely the U.S. could also act as a lender of last resort to banks in trouble by not presenting these notes for redemption but rather allowing these banks to run into debt to it.

The political environment of that period was marked by the ascendancy of an ideology termed Jacksonism. Focused around Andrew Jackson, elected president in 1828, this ideology was an uncomfortable, perhaps inconsistent, mixture of agrarianism, nationalism, populism and libertarianism. However the one element which unified this group was a deep hostility to a privileged east-coast based aristocracy. The Philadelphia-based bank of the United States with its obviously patrician president, Nicholas Biddle, could hardly prove to be popular with this new regime.

The Jackson administrations assault on the bank began in 1830. In 1832 Jackson used his presidential veto to thwart the Banks supporters attempt to use Congress to enact a new charter for the Bank. Jackson then used his second presidential election victory later that year as a mandate to order the withdrawal of all federal funds from the bank in 1833. When the Banks original charter expired in 1836 it succeeded in being re-chartered, albeit now only on as a much reduced state bank under the auspices of the Penneslvyania state legislature as the United States bank of Penneslvyania. In 1841 it went bankrupt as a result of speculative dabbling in the cotton market. I shall now consider the motives which inspired the attack on this institution.

The Anatomy Of The Anti-Bank Forces: Vested interests

The role played by vested interests in motivating the anti-bank forces has been given particular emphasis by both Caterall (1902) and Hammond (1947). They point to the substantial personal gains which would accrue to key members of Jacksons administration should the bank be destroyed. Hammond ascribes an important role to the New York financial community which at the time was competing with Philadelphia to be the countrys premier commercial centre. Martin Van Buren, Jacksons 2nd vice-president and eventual successor was particularly identified with the Wall Street element in this Wall Street (New York) versus Chestnut Street (Philadelphia) battle.

Both Pessen and Hammond add an additional group to this coalition; the state banks, who disliked being constrained by the BUSs policy of redeeming their bank notes. This enforced a much higher reserve ratio and hence restricted their lending activities. Hammond also adds to this element the class of nouveau riche entrepreneurs and speculators, a class to which, he maintained, Jackson and many of his associates belonged, and which also disliked the restriction of credit. However, I would argue that the importance of this proposed group in effecting the banks destruction has been over-emphasised by pro-BUS writers such as Hammond and Caterall.

Firstly, the actual existence of such a coalition is questionable. Pessen gives evidence that the New York financial community were divided over the question of the wisdom of the attack on the Bank. Also he shows that at least some of the state banks grudgingly acknowledged one banks role in disciplining the banking system and its activities as a lender of last resort. The homogeneity of Hammonds speculative entrepreneurial class is one for which he offers merely anecdotal evidence and no quantitative evidence. Secondly, to concentrate upon vested interests is to ignore the other influences on political action. Ideologies and economic logic also play a role. Hammond, the primary exponent of the self-interest theory, fails to explain satisfactorily why the measure was extremely popular.[47] Only a tiny proportion of the population would have gained directly and immediately from the destruction of the institution. We must examine the political philosophy and economic logic behind the opposition of the bank. These arguments had much public support which was vital to Jacksons destruction of the bank.

The Political Ideologies

The ideology which underlay the struggle is a highly variegated, and perhaps ultimately inconsistent one. It was a blend of moral judgements, economic argument and populism to attack both the political legitimacy of the bank and its economic rationale.

One branch of the school consisted of states rights advocates, who strongly opposed the substantial power wielded by the federally-chartered bank.[48] Many considered the chartering of the bank an unconstitutional extension of the power of the federal congress. Their position was summarised by Jackson who described the bank as a threat to democratic institutions by the federal authorities. With the destruction of the bank, the power of intervention in the banking and monetary system was left in the hands of individual states until the civil war.

Another stream within the anti-bank framework were the libertarian thinkers. They postulated the illegitimacy (on moral grounds) of any government intervention in the economy or in society beyond a bare minimum. This period was the golden age of Laissez Faire. This group was related to and associated with the Free Banking school which challenged on economic grounds the necessity of government intervention in the monetary system.

The Free Bankers

This group were in favour of a paper currency based on a fractional reserve system. They argued that the banks regulatory function was unnecessary and inefficient because in a completely unregulated financial system free competition would ensure that the public receives whatever security against fraud it so desires.[49] They argued that what was wrong with the banking system was that free competition was obstructed by the monopolistic privileges granted to the BUS in its charter. It is important to place these views in the context of the dominant economic paradigm of the day. Today as I have outlined, the importance of the states role in regulating the money supply is considered self-evident by most economists. (Hayek, Glasner and Greenfield and Yeager being some noteworthy exceptions to this consensus.)

This was not the case in 1832. We have Schumpeters (1954) comment that in the first part of the 19th century most economists believed in the merit of a privately provided and competitively supplied currency. Glasner shows how Smith differed from Hume in advocating state non-intervention in the supply of money. Smith argued that a convertible paper money could not be issued to excess by privately owned banks in a competitive banking environment. Today we see money as a natural public good owing to the externalities caused by variations in its quantity. So the free bankers views were consistent with economic logic of the day.

The Hard Money School

The anti-monopolistic and anti-regulatory free banking school were joined by unlikely bed fellows from the opposite end of the spectrum of economic ideas, agrarian and proletarian mistrust of banks in general and paper money in particular. This mistrust may have been justified in the context of the widespread level of fraud within the system, relative to today. Many saw paper money as a tool used by employers and rich financiers to trick working men and farmers out of what was due to them.

This groups most prominent exponent was Andrew Jackson himself. In his farewell speech he refers to the paper money system and its natural associates monopoly and exclusive privilege. The value of paper, he states, is liable to great and sudden fluctuations and cannot be relied upon to keep the medium of exchange uniform in amount. Jacksons views on this topic may be due to an incident early in his career when he was almost bankrupted after accepting bank notes which turned out to be worthless in return for a debt.

In contrast to the free-banking school this group could be termed conservative, wishing to destroy the system of fractional reserve paper money by removing the kingpin of the banking system which produced it; the Bank of the United States. Even within this group there was a severe division between those advocating gold specie, those advocating a silver specie, and those advocating a bi-metallic medium of exchange.

The Battle Of Ideas And Its Outcome

Thus, was the coalition against the bank of the United States. Both advocates of the free banking (or soft money) school and proponents of a return to a specie economy (or hard money) saw the destruction of the bank as very important, but for both of them, its destruction was a means to divergent and conflicting ends. Against this coalition supporters of the bank such as its president Nicholas Biddle and politicians such as Henry Clay and John Quincy Adams were placed in extreme difficulty. Both anti-federalism and laissez-faire were in the ascendancy at the time.[50] On the economic front the bank was being assaulted from both the left (free-banking advocates) and from the right (anti-paper advocates).

Advocates for the bank did emphasise its moderating role in regulating, informally, the fractional reserve system and hence its publicly-interested central-bank type nature. Such arguments were almost certain to fall on barren ground. Only two major institutions were available for comparison. The first one was John Laws bank from early 18th century France, and the chaotic and inflationary experience of this scheme was hardly one to inspire confidence. The other example was the Bank of England which was at the time subject to scathing attacks during the bullionist controversy, (Hammond (1947) and Glastner (1989)).

Evaluating The Arguments

The final verdict on the validity or otherwise of the differing arguments must wait until the consequences of the banks destruction have been fully considered. However the following points can be made at this stage. Firstly, the arguments of those opposed to the banks existence on ethical grounds, namely the classic libertarians and the states rights advocates, cannot be assailed on empirical grounds given that they are normative judgements.

Secondly, those who attack the bank on the grounds that it was a predatory monopolist within the banking system have had their arguments somewhat refuted by the evidence garnered by Highfield, OHara and Wood, who carried out a systematic econometric survey with regards to the banks decision variables during its existence and found no evidence that its dealings with its competitor banks or with its markets were marked by any of the predatory practices associated with monopolists. However perhaps the anti-bank forces could argue that it was the banks potential as a monopolist rather than its actual behaviour which justified the withdrawal of its charter. The methodology of such evidence and the quality of the data upon which they are based may also be attacked but such studies must be considered none the less.

The Implications Of The Banks Destruction For Output And Employment

Firstly, over the period 1790 to 1860 the general movement in the price level was downwards (with some fluctuations). Output followed a similar fluctuating and variable pattern over the same period, albeit with a strong upward trend. The pattern over the 1830s and 1840s diverges somewhat from this . The period 1830 until end of 1833 was marked by a slight though pronounceable upward trend in prices. 1833 marked the withdrawal of federal deposits from the bank and in early 1834 Nicholas Biddle, trying to convince the government of the need for the bank, massively contracted credit. The result was a short sharp fall in prices and output in what was termed Biddles contraction. However, by late 1834 prices, output and the money supply were strongly rising as boom-like conditions prevailed once again. In 1837 this upward movement was again sharply reversed. It was not until the early 1840s that output began to expand significantly once again. Over this period of the 1840s the money supply also began to grow once again, although more moderately such that the extra output seemed sufficient enough to soak it up, and so the price level resumed its long term downward path.

There are two different interpretations of these events. The first one, expounded by Hammond and Caterall, blames Jacksons actions in destroying the bank for the inflationary boom and resultant recession over the period 1834 to 1837. In their view, dismantling the BUS took a restraint off the fractional reserve system and led, post-1834, to an increase in the money supply which caused the boom. Of course, this was checked in 1837 by a downturn, a downturn made worse by the fact that at this stage the banking system, due to its low reserve ratio, was now very unstable and experienced significant levels of bank collapse. On the face of it the hypothesis has some factual support. Over the period 1833 to 1837 the amount of bank notes in circulation rose from a value of $10.2 million to $149.2 million.

However data from Temin and Engerman show that the banks aggregated reserve ratio did not fall over 1834-1837. It had remained steady from the mid-1820s. Thus, the BUSs demise had not caused the money supply to rise by allowing reserve ratios to fall.

An alternative hypothesis was advanced by Temin. He argued that monetary expansion did not come from a falling reserve ratio but rather from an inflow of silver into the United States in the 1830s. He backs his argument up by showing how this inflow in the 1830s would have resulted from increased silver production in Mexico, from an increase in British investment in America and from the fall in US imports of opium from China, which stopped the outflow of silver. So it is possible to dismiss the relationship between the Banks demise and the panic of 1837 as a coincidence.

The Regional Dimension: Monetary Union For The USA

Frass (1974) notes how the bank acted under a congressional mandate to establish and maintain a uniform national currency by policing the state banks to ensure that convertibility was maintained at a high level. Using data from Ohio, Frass argues that this involved restricting the lending of banks along the western frontier in particular. This, as well as having adverse effects in itself in a capital-scarce region, depressed the general price level in this area relative to the price of unsettled land (which was set arbitrarily by Congress), and discouraged movement to settlement in these areas.

Frass study does not extend beyond 1834 but we can assume that the removal of the bank led to the cessation of these harmful activities. However, we must note that a trade off would have had to be made. A higher level of financial instability may have been the price paid for freer availability of capital and cheaper land in these peripheral regions.

The Long Term Impact On Americas Monetary And Financial Structure

Ultimately the Banks destruction marked a pyhrric victory for the hard money forces. Van Buren, Jacksons successor, was no supporter of a purely metallic currency. Return to a purely metallic currency would have met severe opposition from their former allies, the free banking and libertarian schools (the latter felt that the state had no ethical right to regulate any commercial transactions between consenting individuals including paper currency). The changing economic and social structure would have made it unfeasible to return to a purely specie exchange economy. Cameron posits that a medium of exchange based on bank liabilities and a fractional reserve system and/or government taxable capacity is essential to an industrialising economy. Abolishing paper money could be considered, the modern economic equivalent of attempting to dis-invent the wheel. The true victors then in the struggle were soft money or free banking advocates. Instead of destroying the fractional reserve system the hard money men had removed a force which acted to restrain it.

Similarly, after 1837 the reserve ratio of the banking system was much higher than it had been during the period of the BUSs existence. This reflected public mistrust of banks in the wake of the panic of 1837 when many banks failed. This lack of confidence in the paper money system, could have been ameliorated by a central-bank type institution. Hence one result of the demise of this bank may in fact have been a higher reserve ratio, less availability of credit and a lower money supply during the 1840s and 1850s. The evolution of the American banking system was also probably affected. The BUS was one of the first and last banks chartered by the federal authorities for commercial banking activities nation-wide. Had it survived it is unlikely that Americas retail banking market today would have been so localised and fragmented in a way

which is extremely uncharacteristic of other large industrialised economics. After the banks destruction, banking returned to being a decentralised business in which institutions were chartered by the individual states.

The banks defeat also had profound implications for the role of the state in America in managing monetary policy. Large scale Federal intervention in the supply of money did not take place again until the American Civil War. However Jacksons victory had imbued US political culture with dislike of centralised institutions with large influence over the banking system. The United States did not develop a central banking agency until 1913. This institution was highly decentralised consisting of twelve autonomous components one in each of Americas largest cities. One result of this de-centralisation may have been the incoherent response of the monetary authorities to the 1929 crash and the resultant run on the banking system, possibly one cause of the 1930s great depression. Hence one interpretation might see the destruction of the bank of the United States as leading to the worlds most severe economic recession a century later.

Conclusion: The Case For Further Study

This topic offers an area where rich analytic rewards may be reaped by further studies which employ modern economic techniques. The episode marks a crucially formative event during the nascent period of the monetary system of what is currently the worlds dominant economy. In spite of these facts this subject has been much neglected, the attention given to the English Bullionist controversy. In over twenty-five years not one book has been published dealing specifically with this topic. This paper, I hope, contributes to correcting this deficiency.

Bibliography:

Cameron, R (1967) *Banking In The Early Stages Of Industrialisation*

Dorfman, A (1957) *The Economic Mind In American Civilisation*

Caterall, R (1902) *The 2nd Bank Of The United States*

Engerman, P (1970) Economic Consequences Of The 2nd Bank Of The United States in the *Journal of Political Economy*, 78(4)

Frass, A (1974) The Second Bank Of The United States: An Instrument For Inter-regional Monetary Union in the *Journal Of Economic History*, 34(2)

Glasner, C (1989) *Free Banking And Monetary Reform*

Greenfield & Yeager (1978) A Laissez-Fairez Approach To Monetary Stability in the *Journal Of Money, Credit And Banking*

Hammond, B (1947) *Banks And Politics In America From The Revolution To The Civil War*

Hayek, F (1967) *The Denationalisation Of Money*

Highfield, OHara & Wood (1991) Public Ends, Private Means: Central Banking And The Profit Motive 1823-1832 in *The Journal Of Monetary Economics*, 28(2)

Pessen, E (1985) *Jacksonian America: Society, Personality And Politics*

Schumpeter, J (1954) *A History Of Economic Analysis*

Temin, P (1968) The Economic Consequences Of The Bank War in the *Journal Of Political Economy*, 76(2)

White, L (1986) William Leggett: Jacksonian Columnist As Classical Liberal Political Economist in the *History Of Political Economy*, 18(2)

DEBATES WITHIN MONETARISM[28]

Ian Rowell
Junior Sophister

There is a consensus that monetarism, as applied in the English speaking economies in the early 1980s, was a failure, and criticism of the attempts to apply it has led monetarists to distance themselves from the actions taken in their name. Many opponents of monetarism would argue that the costs incurred by the reduction of inflation in the early 1980s were exceptionally high and that the policy should not have been tried. In this essay, however, I intend to concern myself more with the criticism made by policy-makers, not that monetarism was too costly, but that it proved impracticable because the relationships on which monetarism relied turned out to be unstable. In their turn, monetarists replied that their policies had not been implemented as they had advised. In this essay, I will describe that debate, focussing on both the choice of targets and the instruments used.

Monetarism

Monetarism is the belief that the sole objective of macroeconomic stabilisation policy should be the control of inflation, and that output growth should be achieved through microeconomic (or supply-side) policy. Monetarists rely on the stability of V, the velocity of money, in the definitional equation MV=PY. They claim that money demand (and hence V) is a function of several variables, exogenous to the model, and that the function can be relied upon to evolve more slowly than the other components of the equation. Therefore, for policy purposes V may be regarded as stable. They also argue that M is exogenous, and that changes in M cause changes in PY. Finally, they believe that the economy is on average at the natural rate of unemployment and Y is exogenous. This means that, in the long term, changes in M are the sole cause of changes in P and the relationship is stable. The empirical work of Milton Friedman and Anna Schwarz confirmed just such a stable relationship between M and P. The implication of this theory is that the authorities, being unable to anticipate short-term fluctuations in real income, should set growth in M equal to long-run income growth plus an acceptable rate of inflation. This will be the best automatic stabilizer, and the knowledge of the existence of such a policy rule will force wage- and price-setters to moderate their claims.

In order to implement a monetarist policy, the government must set a target for some measure of the money supply and pursue this target through some policy instrument interest rates, monetary base operations or regulation of the financial sector. It is necessary that there should exist (i) a stable relationship between the measure of money chosen and the price level, and (ii) a predictable relationship between the monetary aggregate in question and the instrument used to influence it. Most of the debate between monetarists over the means of controlling inflation reduces to finding instruments and aggregates which obey these two criteria.

Particular debate has taken place over point (i): which monetary aggregates have a stable relationship with the price level? Instability of this relationship can be attributed to three main causes. Firstly, the relationship is stochastic and when implementing such a policy one must allow for errors in an individual year. Secondly, even in the long term V may be unstable due to institutional changes which affect the money demand function. Thirdly, some monetary aggregates may not cause changes in the price level and may in the past only have exhibited a stable relationship with PY because the causality went the other way. The problem in implementing monetarism is to trade off the instability caused by institutional change, which mainly affects broader measures of money, against the instability caused by the absence of causality, which mainly affects narrower measures. Many of the failures of supposedly monetarist policies in Britain and the US in the 1980s can be attributed to one or other of these problems.

Breakdown

In Britain, the Medium Term Financial Strategy of the Conservative government from 1979 to 1985 consisted of targeting the broad money aggregate sterling M3 ([[sterling]]M3), which contained currency and all bank

accounts except those denominated in foreign currencies. Until 1982 it was the sole target. [[sterling]]M3 was chosen for two reasons. Firstly, by an accounting identity, a large component of it was the PSBR, and thus government policy had a predictable influence upon it. (This was especially important since one of the main means of reducing the money supply was to be fiscal rectitude.) The second reason for choosing [[sterling]]M3 was that it had given advance warning of the Heath inflation of the 1970s, where other aggregates (notably M1) had not.

A broadly stable velocity was observed for [[sterling]]M3 through the 1960s and 1970s, rising slightly over time. It was thought to have a predictable relationship with inflation. However, it turned out not to have when tried for policy. The target was overshot in every year except 1982-3, and yet inflation turned out to have fallen dramatically. In the early 1980s, narrower monetary indicators were showing that policy was far tighter than [[sterling]]M3 indicated. By 1983, it was clear that the velocity of circulation had fallen rapidly as soon as policy began to be applied. Many conservatives, who had hoped for a more gradual approach to the reduction of inflation, blamed the breakdown in stability for the overtightness of policy. Unemployment in Britain rose from 1.3m to 2.2m in 1980. But monetarists had always warned there would be short-term costs and output and employment would recover later. A more serious criticism, from their point of view, was advanced by the monetary authorities: that monetarism was impossible to apply.

In the US, the Federal Reserve attempted to implement monetarism between 1979 and 1982. They targeted the narrow monetary aggregate M1 through a form of Monetary Base Control (MBC). MBC entailed controlling the money supply through the non-borrowed reserve base rather than short-term interest rates. Although the year-on-year M1 targets were met, this policy was unsuccessful, since (i) the volatility of different measures of the money supply actually increased and (ii) as in Britain, the velocity of circulation fell. Inflation was thus reduced more quickly than expected and monetarists were blamed for the resulting recession. The theme of this essay, however, is to focus on the debate within monetarism. The Federal Reserve concluded that monetarism was impracticable and reverted to controlling borrowed reserves.[29] In Britain, Canada, the US and Austrailia previously stable velocities all broke down in the early 1980s. The reason for this breakdown is the source of the controversy between policy-makers and theorists.

The Policy-Makers Criticisms

In both Britain and the US the authorities blamed institutional change for the breakdown in stability. Abandoning the [[sterling]]M3 target in 1985, Nigel Lawson attributed the targeting failures to money demand shifts because of financial innovation. It was monetarisms misfortune to be first tried in Britain at a time of rapid institutional change, and by a government which believed in deregulating the financial markets. A number of developments destabilised the relationship. Firstly, the government abolished the corset, a restriction on bank lending which, significantly, Denis Healy had imposed to help him control broad money growth. Secondly, domestic legislation had increased the competition between banks and building societies for deposits, which led to the development of accounts which became more attractive relative to other kinds of financial assets, increasing [[sterling]]M3. Thirdly, high short-term interest rates increased the use of many accounts as vehicles for savings rather than for transactions demand, and that component of [[sterling]]M3 grew as interest rates rose. So there was a breakdown in (i) the stability of the relationship between the aggregate and inflation, and (ii) the stability of the relationship between the interest rate and the aggregate. Similar explanations were advanced by Chairman Paul Volcker of the Federal Reserve for the Feds abandonment of MBC in 1982. He blamed the unwinding of All Savers Certificates and changes in Individual Retirement Accounts for shifting money demand.

Why did such instability not show up in Friedmans and Schwartzs work; on which the assumption of the stability of V was based? Did money demand functions only begin to be unstable in the early 1980s? An explanation is provided by Laidler. He notes that Friedman and Schwartz used definitions of monetary aggregates which incorporated institutional changes, rather than representing them as a separate term in the model they were testing. So stability of V in the face of money demand changes depends on the definition of monetary aggregates. It should be said, however, in Friedmans defence, that velocity of circulation changed far more slowly over the period he tested than when monetarism was applied.

The Monetarists Reply

Monetarists disclaim responsibility for the breakdown in stable relationships, criticizing both the instruments used and the choice of aggregate. I will deal firstly with their criticisms of the instruments. They claimed that the type of MBC used by the Fed was unlikely to be effective in controlling the money supply. The Fed operated a form of MBC where a targeted non-borrowed reserves. The accounting for reserve requirements was on a lagged basis; in other words, banks were forced to require reserves to meet reserve requirements based on their position two weeks before. They could do this by borrowing reserves from the Fed through its discount window, given the level of non-borrowed reserves that the Fed had. Monetarists opposed the lagged accounting and control only of non-borrowed reserves, arguing that they gave too much freedom to banks. This led to far greater fluctuations in the money supply than under full MBC.

The monetarists also blame the Fed for shifting the relationship between M and P. They claim that intervention by the Federal Reserve, in the form of deposit rate ceilings, shifted the money demand function. In addition, they claim that the instability of the money supply growth, which they attribute to the Fed (see above), caused uncertainty among the public and induced them to demand more of M1. Both these effects on M1 imply increases in speculative demand but not in transactions demand, and so the relationship between inflation and the monetary aggregate in question was shifted.

American monetarists regard the British experiment with incredulity. The choice of instrument was the biggest difference between monetarism as espoused in Chicago and as practised in Britain. Friedman said that when one does not know the interest-rate sensitivity of money demand, MBC is a more precise instrument than the interest rate. The Americans did not recognise as monetarism the use of interest rates to control inflation. The instrument is a very ambiguous one theoretically. If one is using the interest rate, it can be argued that one is controlling the money stock through reducing money demand not money supply, which is not a monetarist policy since it assumes that the money stock is supply-determined.[30]

The second criticism made by the monetarists of policies used in Britain and the US was the choice of aggregate. Friedman claimed that the Feds M1 target was broader than, and thus conceptually different from, the M1 used in his empirical work with Anna Schwartz. As a result it contained much money not used for transactions purposes, and did not have a stable link with inflation. In general, broader aggregates are more subject to shifts in money demand. He says the Feds M2 was conceptually equivalent to his M1, and its M1 to his M0.

Likewise in Britain the choice of aggregate was condemned by US monetarists as being too broad. A narrower measure would have been stabler. Friedman, testifying to a British Treasury Committee, said he had little hope for monetarism as it was being practised in Britain. In his memoirs, Nigel Lawson records the contempt of the Swiss-American monetary expert Karl Brunner for your M3. It was because [[sterling]]M3 was so ill-defined conceptually that its velocity of circulation is unreliable. The uncertain effect of interest rates on [[sterling]]M3 was due not to financial innovations but to the choice of an aggregate which contained a lot of money used for savings and with no clear link to transactions. Many of the effects of high interest rates on [[sterling]]M3 could have been predicted, claimed monetarists.

Is the monetarists defence plausible? Friedman can reasonably argue that what was applied in Britain was nothing like monetarism as he knew it. The aggregate was far different conceptually from anything he had recommended targeting and the use of the interest rate had never been advocated at all. (British monetarists, of course, like Alan Waters and Patrick Minford, have no such defence, since they initially supported the policy). In the United States the targeted aggregate was a lot narrower, and the criticism here looks a bit weak. Whether the American breakdown was the monetarists fault depends on the technicalities of MBC, and a conclusive answer to this question will only come if full MBC is applied. However, we can conclude from this debate that the reformulated monetarist position is as follows: to target a narrow aggregate and to use MBC to do so.

Narrower Aggregates?

However, there are severe problems with the targeting of narrow aggregates. This is best illustrated by what happened in the United Kingdom between 1985 and 1988. British monetarists like Walters and Minford began agitating for the use of MO, cash in circulation, which Lawson had already been looking at since 1982. It was chosen because of its excellent record as a predictor of inflation with a six-month lag. But such a narrow measure of money, as I have pointed out, also runs the risk of producing an unstable V. Causality

cannot possibly run from the amount of cash in circulation to the level of economic activity since cash is such a small fraction of the total money stock. This is a particularly strong form of black box monetarism, relying on a statistical relationship without specifying the underlying theory. Statistical regularities with no causality behind them are apt to break down when used for policy.

Lawson is aware of this criticism, but in his memoirs rejects it in terms which make one doubt whether he believes in monetarism at all. He claims that M0s value was as an indicator of movements in nominal GDP. But such policy rests on the assumption that causality runs from PY to M, and this is not monetarism, but Keynesianism. The essence of monetarism is that changing the money supply changes nominal income and not the other way around. To summarize, either Lawson was using M0 as a target or he was targeting nominal income directly through interest rates. The former would inevitably fail, and targeting of M0, say Pearse and Tysome, may have impaired its efficiency as a predictor of inflation. (Lawson himself records that a ... serious shortcoming was the failure of M0 to give a true picture...in early 1988 the beginning of the boom.) The latter policy conflicts with the basic tenets of monetarism, of which the raison detre is that the authorities cannot know in time what is happening to output.

MO was mistakenly used as a narrow aggregate because British monetarists had followed American monetarists in advocating a narrow target. However, the Americans were advocating that the target should be the monetary base itself (either reserves or reserves plus cash in circulation) which, unlike MO in the British financial system, has a clear causal relationship with nominal income through the multiplier.

No single aggregate?

Friedman has acknowledged that targeting any single aggregate incurs severe risks. Broader aggregates are liable to have their velocity shift, whereas with many narrower measures there may not exist causality. Moreover, Goodharts Law states that any statistical relationship will break down when used for policy purposes. This is because in a sophisticated financial system private agents will be able to substitute between instruments to avoid monetary rules. This suggests that control of all aggregates is required. But Friedman dislikes the Federal Reserves system, practised after 1982 of targeting several aggregates. That is why he has advocated targeting the monetary base directly, rather than any aggregate measure of money. In monetarist eyes, such a policy would have a clear effect on a number of aggregates, would reduce the risk of targeting any one, and yet would eliminate the discretion involved in looking at a number of these aggregates at the same time. This method relies upon the stability of the multiplier. It is true, as monetarists have claimed, that the multiplier is less vulnerable to the type of financial innovation which affects broad aggregates. However, as pointed out by Phillip Cagan, himself a monetarist, other developments like ATM machines could erode the influence of the monetary base. Transactions services, says Cagan, could be then provided by non-deposit-taking institutions, on which the government would have difficulty imposing reserve requirements. As electronic transactions replace currency, the authorities will be excluded more and more from the process of creating money. It was the development of ATM machines in Germany which led the German government to change from targeting the monetary base to targeting broader aggregates, with which ironically they have since had problems.

Rules vs Discretion

An alternative solution is that the authorities should target a broad aggregate, but with qualifications to allow for structural change. One possibility is that the authorities should severely regulate the financial markets. Indeed, countries following broad targets like Japan and the Netherlands were forced to restrict bank lending at certain times. In addition, one of the reasons for the backwardness of Germanys financial sector relative to her manufacturing industry is the governments control of financial innovations to ensure that they do not destabilize the relationships governing the determination of the price level. Another possibility is revision either of the target itself, or of the definitions of the monetary aggregate involved, to allow for institutional change. Chairmen of the Fed Volcker and Greenspan always believed in being allowed some discretion in implementing monetary policy. There are two possible objections. Firstly, changes of target signal to agents that the government is not determined about reducing inflation. A more fundamental criticism is that if you are a monetarist, you must be highly sceptical of the ability of governments to correct properly for shifts in money demand. This is Friedmans objection to the use of discretion in pursuing a target.

In my opinion, the monetarists place far too much importance on excluding discretion from government policy. The argument that discretion reduces agents confidence in the policy is flawed. It seems clear that what changed inflationary expectations in Britain and the United States was the recession, not the perception that the authorities were adhering to rules for monetary growth. Only the financial markets would have the sophistication to make the esoteric distinction between voluntarily committing oneself to rules and changing those rules to allow for changing circumstances. What matters to wage-setters is whether or not the authorities mean business on inflation.

The second monetarist argument in favour of fixed rules is that the authorities do not have the knowledge or the timing to offset shocks. The expectation of income is its trend value and there having been a shock to income in a previous period does not affect its expected value in the next period. Therefore, there is no point in the government reacting one period late. But although this is true of shocks to income, it does not hold for changes in money demand. Institutional change does not fluctuate around a certain value. Once a change is made, it applies to money demand in the future as well as the present. Therefore it may be possible for the government to react after the change has taken place and for this reaction to be of some use. When it comes to government action to correct for institutional change, lags are a nuisance but not at all fatal.

I believe the example of the Swiss and German authorities is the most hopeful for the future prospects of monetarism. They have often made important changes of target once they observed shifts in velocity, but they have successfully pursued monetary growth rules since the 1970s. The Swiss have varied between Monetary Base Control and short term interest rates as instruments; the Germans changed from a narrow to a broad aggregate; both changed the targeted growth rates frequently. Yet because it was done responsibly and credibly, it worked.

Conclusion

I have left out all mention of the exchange rate. Many people believe that it was Britains exchange rate in the early 1980s, which was high due to North Sea oil and the high short-term interest rates, which rought inflation down due to competitive pressures from abroad. I did not deal with the exchange rate as a possible tool to reduce inflation as since the demise of the Echange Rate Mechanism this looks unlikely to be useful for some time. So my analysis is incomplete, but I thihk I can conclude that monetarism is practicable. Whether or not it is desirable is another matter. It may be more appropriate to the United States, where the labour market adjusts more quickly to a monetary contraction than in Europe. However I believe that if policy-makers wish to control inflation then they should not judge monetarism by the experience of the 1980s. Laidler points out that Friedmans model was devised at a time of low inflation to keep it low. The same degree of precision could not be expected in combatting high inflation. Velocity does not remain stable when inflation is high. One of the components in Friedmans original equation describing the effects on velocity was inflation. Brunner and Meltzer are quoted by McCallum as blaming inflation for many of the financial innovations which shifted the demand function, and claiming that instability of V would not reoccur under the monetarist regime. This view is extreme, but it does suggest that the authorities might have found monetarism easier to practice if they had not given up as soon as inflation had fallen. Informal support is given to this view by the German experience. In the early 1980s they did not experience the fall in velocity which the English-speaking economies did. This may be due to their financial system, or to their experience, or to their lower inflation, but it does suggest that monetary relationships are not always subject to breakdown when used forpolicy.

Bibliography

Congdon, T (1989) British and American Monetarism compared in Hill, R *Keynes, Money and Monetarism*

Felderer, B and Homburg, S (1991) *Macroeconomics and New Macroeconomics*

Foot, M (1981) Monetary Targets in Griffiths & Wood *Monetary Targets*

Friedman, M (1984) Comment on the 1979-1982 Policy Experiment in *American*

Goodhart, D *Techniques of Monetary Control*

Economic Review

Gowland, D (1991) *Money, Inflation and Unemployment*

Laidler, D (1989) *Taking Money Seriously*

McCallum (1985) On Consequences & Criticisms of Money Targeting in *Journal of Money Credit and Banking* 1981

Pierce, D and Tysome, P (1985) *Monetary Economics*

STOCKS, HORSES, CHAOS AND EFFICIENT MARKETS

Donagh Lynch
Junior Sophister

It is a difference of opinion that make horses race. Mark Twain

Since the early 1960s a considerable amount of empirical work and not a small amount of computer time has been spent determining whether or not financial markets are efficient. While the evidence in support of the efficient markets model is extensive and, somewhat unique in economics, contradiction is sparse[51], the robustness of any model depends on the span of its applicability. Too often it is assumed that if a hypothesis is proven in one econometric study, that the results hold for all possible scenarios. My objective in this essay, therefore, is to augment the traditional analysis of securities markets with an examination of the efficiency of on-course betting markets.

If securities markets and horse racing seem worlds apart, the difference is more apparent than real and can be largely explained by the stigma society attaches to the latter. Both have common characteristics: a large number of participants; a wide range of publicly available information; and ease of entry. Like securities markets, horse racing provides an opportunity to study the rationale of decision making under conditions of risk and uncertainty. Furthermore, since betting schemes give rise to wager markets with equilibrating functions similar to commodity markets, they afford opportunities for the study of market mechanisms under a widened class of contingency and institutional conditions.[52]

In section one I examine the various empirical tests applied to the securities and horse racing markets. I intend to show that on the whole both markets are efficient, although some inefficiencies persist. In the second section I explore some of the new mathematical and computer techniques used to exploit these inefficiencies to earn super normal profits.

First it is necessary to define what exactly is meant by efficiency. When we speak of market efficiency, we refer not to the efficiency of the services industry, but rather that of information markets. The Efficient Market Hypothesis (EMH) states that at any one point in time, prices reflect all available information. This implies that no amount of data mining can predict future prices. Furthermore an analysis of past or current data cannot identify undervalued stocks. Applying this to the securities markets, the EMH implies that no trading mechanism can consistently beat the market. Hence, for a given level of risk, speculators cannot earn supernormal returns. Similarly, no betting system can consistently earn super normal returns: professional punters, therefore, cannot earn a greater return on betting markets than you or I.

There are varying degrees of market efficiency, with Fama (1965) providing the traditional framework through which the EMH is examined. The weak form simply states that all past information is reflected in current prices. The semi-strong form states that all publicly available information is incorporated in prices, while the strong form, an extension of the first two, states that all information, including insider information, is included in share prices. Each will be examined in turn.

The Weak Form

To understand, we search for sequence and order. It is of little surprise, then, that fund managers look to past prices as indicators of the future. However, the weak form of EMH states that all available past information is rapidly incorporated into current prices. I emphasise rapidly because many portfolio managers believe that they can identify under-priced stocks through elaborate techniques such as charting and filters, collectively known as technical analysis, and use these to extract market beating returns. A weak efficient market implies that this analysis is a waste of both time and resources.

To test for weak form efficiency in security markets one tests the serial correlation of prices through time. If there is a relationship between prices in one period and those in the next, then *Corr(P*t-1, *Pt)* will be either positive or negative. However, repeated studies [53] have shown that there is a near zero correlation between stock prices through time. This flies in the face of technical analysis, and suggests that although in the short run there may be short lived patterns in price movements, they are unstable and inconsistent. In the long run, therefore, the best strategy to adopt is a buy and hold policy.

In considering betting markets, the construction of a weak form test is more difficult. Unlike stocks, there is no history of betting prices to correlate. Instead a more intuitive approach is needed. Snyder (1978) examines the subjective winning probabilities punters assign to horses, at eight different levels of expected return. His findings show that, unlike securities markets, there is a negative relationship between returns and odds. This is accounted for by the utility derived from a bet: punters derive enjoyment from reading the form, riding their luck, and in particular picking long shots to win, thus proving their ability to out-perform the market. [54] Hence, unlike their speculator counterpart, gamblers have different risk preferences: they prefer low probability-high return bets to high probability-low return combinations. This results in what is known as an overlay, whereby smaller proportions are bet on lower-odds horses than their actual chances of winning justifies.

Snyders study, which was compiled from data on over 300,000 horses in more than 30,000 races, shows that at lower odds levels there is a strong bias: punters bet less on favourites than the horses ability merits. This suggests that favourites are good bets and long shots are bad bets. However, when the bookies or tote take is added back in, the bias is not large enough to consistently beat the market. Hence, betting markets, like securities markets, are weak form efficient.

The Semi-Strong Form

Semi-strong tests deal with whether or not markets fully reflect all publicly available information. Fundamental analysts forecast earnings, dividends and other factors which affect share prices, in search of undervalued stocks. However, the semi-strong form states that publicly available information is quickly assimilated, if not anticipated, by the market, so that as more players enter the market, inefficiencies are quickly eliminated.

Tests of semi-strong efficiency in securities markets have centred around the study of events. These events include public announcements, such as a company reporting a substantial new holding in a subsidiary, and dividend disclosures. A more interesting example involves testing analyst predictions to determine whether these can be used to earn excess returns. The Heard on the Street column in the *Wall Street Journal* publishes analysts tips a week or two after the opinions were circulated to the analysts clients. The *Journal*, however, is the first wide scale dissemination of such information. Davies and Canes (1978) examined the differential between expected and normal returns in the twenty days prior to and subsequent to the Heard on the Street recommendations. The results show that prices rise in the final days prior to publication and decline thereafter. As more investors with knowledge of the analysts predictions enter the market, so demand for the stocks rise. In response to this prices rise so that the markets anticipate new information. This suggests that while excess returns cannot be earned after the announcement, they may be earned prior to publication. However the residual returns prior to the announcement were between 1% and 2%. Hence, only if transaction costs were lower than this residual could excess returns be earned. Securities markets are therefore semi-strong efficient: public information is rapidly incorporated into, and sometimes even anticipated by, prices.

An interesting extension of these tests has been to test for market volatility previous to and after events. Roll (1984) looked at the production of orange juice in Florida and the effects of the weather and public announcements. His research shows that there is excess volatility in security markets: prices over react to new information. This highlights an inefficiency in information markets, to which I will return later.

Turning to betting markets, Snyder (1978) states that no satisfactory semi-strong test of horse racing has yet been developed. Like the securities market example above, what is needed is a model which incorporates previous performance data, and allows for the incorporation of new information, be it new weights, jockeys or workout information. This is a point of contention, to which I will return after examining the most stringent form of EMH.

The Strong Form

The strong form of EMH states that all information, be it public or private, is incorporated so quickly into prices that no investor can make excess returns. Tests of the strong form have focused on insiders, those who occupy a unique position due to their monopoly of information. As insider trading is illegal, data is somewhat scarce and so empirical studies have focused on those with most information: portfolio managers.

Several studies have analysed mutual funds in the United States, the classic example being Jensen (1968). In this study, the performance of 105 funds were tested over a 19 year period and compared to the market return. Of the 105 funds, only one performed significantly better than the market, while 14 performed significantly worse than the market over the same period. Jensen concluded that securities markets are strong form efficient, and suggested that mutual funds spend less resources conducting technical and fundamental research.

A number of criticisms have been made of Jensens method. First of all he analysed mutual funds and not the individuals who make the arbitrage decisions. Successful analysts manoeuvre regularly through the labour market, and so it is the individual who should be assessed and not the fund. Secondly, and more importantly, Jensen assumes that the market return is the return on the whole market. While portfolio managers can diversify risk, they cannot ensure a market return. Hence the benchmark by which these funds are analysed is arbitrary. Later studies[55], have responded to these criticisms, and while they dont produce the same startling results, they nevertheless support Jensens original hypothesis.

Horse racing is notoriously full of insiders, be they owners, trainers, jockeys or stable lads. Like the securities markets, data on the returns these insiders earn goes largely undocumented. Tests have therefore focused on those who publish their tips. Snyder (1978) tests the returns from the tipsters predictions in the Daily Racing Form and four daily newspapers. His findings show that not one of the tipsters earned a positive return after the bookies take is subtracted, from which Snyder deduces that betting markets are strong form efficient.

This analysis raises a number of difficulties. Firstly, newspaper predictions are made up to 30 hours before the off, so new information is not fully accounted for. Secondly, newspaper tipsters are employed to tip a horse in each race. Professional punters, on the other hand, will often sit a whole race meeting out without placing a single bet. A third deficiency in Snyders methodology is the assumption that newspaper tipsters are examples of insiders on betting markets. In effect many are employed on a part-time basis, and have access only to information that anyone in the general public can obtain. Hence, Snyders strong form test actually tests whether public information only, and not inside information, is incorporated into betting prices. Hence his strong form test should be relegated to a test of the semi strong form.

Similarly, mutual funds are not a satisfactory proxy for those in security markets with inside information. That analysts better utilise the public information available to them is not inconsistent with the semi strong form of EMH. To test for strong market efficiency we are compelled to analyse the returns of *real* insiders. In the US, shareholders with a substantial interest in a company, as well as management who own shares, are required to list their purchases and sales with the Securities Exchange Commission (SEC). An insider trading on specialised information can be expected to purchase shares in months prior to price rises. Jaffe (1974) tested for this and found that these insiders earned greater than normal returns. There is no reason to assume that substantial shareholders hold superior analytical skills to other shareholders. It can be deduced, then, that insiders can earn excess market returns. Hence, securities markets are strong form inefficient.

To recapitulate it has been shown that financial markets are largely efficient. Inefficiencies exist nonetheless. I have, however, highlighted a special subset of players, insiders, who due to their monopoly on information can earn excess returns. It has also been shown that information sometimes is temporarily misreflected in prices, suggesting that those who receive better information quicker are at an advantage. Finally, it has also been shown that there are short lived patterns in share prices. While in the past these were thought to be unstable and inconsistent, a new breed of analyst, armed with superior data, are utilising new statistical, mathematical, and computer techniques to anticipate market reactions to exploit these inefficiencies.

In a study on our attitudes to risk, Tversky (1988) demonstrates that we are non-linear in our thinking: we are risk averse when expecting a gain and risk seeking when facing a loss.[56] This is consistent with Gruens (1976) analysis of betting markets. He concludes that gamblers, when facing a loss, are especially risk

seeking. Recent analyses of securities markets have now shifted from looking at stock behaviour to analysing the behaviour patterns of people. Market theorists are now looking for slight non-linearitys in markets, as a means of anticipating future prices. One theory, GARCH (Generalised Auto-Regressive Conditional Heteroskedacity) assumes volatility in stock markets is clustered. In times of high volatility markets are contrarian, a point Keynes first highlighted in the thirties. GARCH theory, however, concentrates on clusters of low volatility, where it anticipates that trends last longer than an examination of market fundamentals suggest. Investors often wait until they see a price rising before they purchase. Hence a rising price is a bandwagon, with herds of investors entering the market, second guessing each other. If we are non-linear in thinking, then technical analysis is justified.

If statistics are helping us understand market movements, then mathematics will help us anticipate them. The new breed of market analyst is now looking to chaos theory as a means of predicting financial markets. If markets trend in one direction longer, then prices do not follow in a random walk as stated by EMH. Instead of following the normal Gaussian distribution, returns on financial markets are leptokurtotic. A leptokurtotic distribution is one in which the occurrence of event, *E*t, is a function of *E*t-1. A study by Peters (1991), which analyses S&P 500 price changes in the period 1928-1989, shows that securities markets are highly leptokurtotic. This implies that information is not immediately reflected in stock prices, a direct attack on the semi-strong form of the EMH.

If securities markets are leptokurtotic in distribution, then fractal theory can be used to predict future prices. A fractal is an object with a non integer number of dimensions. An example is a piece of crumpled paper which is estimated to have 2.5 fractal dimensions: in parts it looks two dimensional and in others it appears three dimensional. Any one point in a fractal is a function of those surrounding it. Similarly, in a leptokurtotic financial market, any one price is a function of previous prices. Peters estimates that the S&P 500 index has 2.3 fractal dimensions. This implies that as few as three variables can accurately predict market movements. However, markets are in a perpetual state of change and so these three variables are constantly changing. Chaos theory seems better at explaining financial markets than predicting them. Nevertheless, in times of low volatility, pockets of predictability arise, permitting the potentially successful exploitation of chaos theory. This is where neural networks enter the frame.

A neural network is a system of computers which learn by trial and error. Developed to mimic natural evolution, these super computers breed: those networks which are most productive survive, and generate new sibling networks. Those computers which are least productive are shut down, their power being used to develop the new siblings. Neural networks are the perfect complement to chaos theory, given their ability to learn quickly and develop. However, they are not without fault. Firstly, neural networks tend to overfit data: leave them processing long enough, and theyll find a relationship between planetary movements and stock prices; a relationship which can only be coincidental.[57] Secondly, neural networks are a black box: they offer predictions without rationale. Now while most traders act on little more than a gut feeling, investors of the future are unlikely to place their hard earned cash in the hands of a breeding computer.

Despite these criticisms, Fidelity, the largest Mutual Fund company in the world, have seven funds utilising neural networks. One of the funds has consistently beaten the S&P index by between 2% and 7% per quarter for the last three years. Yet the market has been rising consistently during those three years. Neural networks are not good at revising their predictions and so when market conditions change, successful neural networks of the past will have to be replaced with newer models. However, the question remains as to when the optimal time to stop using a neural network is.

Conclusion

This paper began with an analysis of the efficiency of securities and betting markets. I concluded that both markets were, on the whole, efficient, but highlighted that some inefficiencies exist. In section two, I outlined some of the statistical, mathematical and computer techniques being used to explot these inefficiences. Yet the overriding question remains: can analysts earn a greater return than the ordinary investor? Research of the past suggests they cant. However, the analysts of today are perhaps keeping their findings to themselves.

Efficient markets imply analysts who adopt a passive strategy of picking stocks are as likely to be as successful as those who pursue an active policy of research and analysis. Yet, as the Grossman-Stiglitz paradox outlines, if everyone adopted a passive strategy, some securities would be undervalued and

opportunities for successful technical and fundamental analysis would arise. One thing is clear, however. Computer trading software is as heterogeneous as traders. Hence, the use of computers in stock analysis extends to the public (Primitive Neural Networks can now be purchased for PCs to be run at home), entry to the securities markets will rise. This will increase liquidity and, in the long run, reduce costs thus making markets more efficient.

Bibliography:

Ali, M (1977) Probability and Utility estimates for Racetrack Bettors in the *Journal of Political Economy*, Vol.85, No.4

Asch, P & Quandt, R (1987) Efficiency and Probability in Exotic Bets in *Economica*, Vol.54

Asch, P, Malkiel, B & Quandt, R (1982) Racetrack Betting and Informed Behaviour, in the *Journal of Financial Economics*

Craft, N (1985) Some Evidence of Insider Knowledge in Horse Racing Betting in Britain in *Economica*, Vol.52

Davis, P & Canes, M (1978) Stock Prices and the Publication of Second-Hand Information in the *Journal of Business*, Vol.51, No.1

Dowie, J (1976) On the Efficiency and Equity of Betting Markets in *Economica*, Vol.43

Dimson, E and Marsh, P (1984) An Analysis of Brokers and Analysts Unpublished Forecasts of UK Stock Returns in *The Journal of Finance*, Vol.39, No.5

Economist, The (1993) Frontiers of Finance, edition dated October 9th.

Elton, E & Gruber, M (1991) *Modern Portfolio Theory and Investment Analysis*

Elton, E., Gruber, M., and Grossman, S (1986) Discreet Expectational Data and Portfolio Performance in the *Journal of Finance* Vol.XXXXI, No.3

Fama, E (1965) The Behaviour of Stock Market Prices in the *Journal of Business*, Vol.38

Fama, E (1969) The Adjustment of Stock Prices to New Information in the *International Economic Review*, Vol.10, No.1

Gruen, A (1976) An Enquiry into the Economics of Race-Track Gambling in the *Journal of Political Economy*, Vol.84, No.1

Jaffe, J (1974) Special Information and Insider Trading in the *Journal of Business*, Vol.47, No.3

Jensen, (1968) The Performance of Mutual Funds in the Period 1945-64 in the *Journal of Finance*, Vol.23

Jordan, R & Fisher, D (1990) *Security Analysis and Portfolio Management*

Losey, R & Talbott, J (1980) Back on the Track with the Efficient Markets Hypothesis in the *Journal of Finance*, Vol.XXV, No.4

Peters, E (1991) *Chaos and Order in the Capital Markets*

Roll, R (1984) Orange Juice and Weather in the *American Economic Review*

Snyder,W (1978) Horse Racing: Testing the Efficient Markets Model in the *Journal of Finance*, Vol.XXXIII, No.4

Stewart, Ian (1990) *Does God Play Dice? The New Mathematics of Chaos*

Tversky, A (1988) *Decision Making: Descriptive, Normative and Prescriptive Interaction*

DOES MONEY MATTER ?

Alan Dunne
Junior Sophister

Money matters if variations in the money stock exert a systematic effect upon macrovariables that economists feel are important. Two categories of macrovariables exist : real and monetary. Real variables comprise the level of output, employment, real wages, and real interest rates. In general the criterion used for evaluating the importance of money is whether or not it influences the real equilibrium profile of the economy. If variations in the money supply have no effect on the real system then money is neutral or money does not matter. While some economists also recognise the influence of money on some nominal measures, notably the rate of inflation, the importance of this is seen in the context by which variations in inflation (caused by variations in the rate of growth of the money supply) destabilise the economy and knock it from its equilibrium growth path.

The Quantity Theory, the Classical Dichotomy and the Neutrality of Money

> .....We may conclude that it is of no manner of consequence, with regard to the domestic happiness of a state, whether money be in greater or less quantity.[1]

The debate regarding the role of money in the economy finds its origins in the quantity theory of money, an identity developed to illustrate the classical dichotomy - the idea that the real variables in the economy, such as real interest rates, relative prices and real income, are determined by real forces and that monetary forces only affected nominal quantities. Thus, in the classical model money was said to be neutral or money is a veil.[2]

The income version of the quantity theory states that

$MV = PY$

where M is the nominal stock of money in the economy, Y is real income, P the price level and V the velocity of money in circulation defined as the average number of times per unit time that the money stock is used in making income transactions i.e. $V = PY/M$

Thus, the quantity theory is a mere tautology i.e. the total value of output equals the stock of money in the economy multiplied by the number of times it is used in transactions. For the equation to become a theory of the price level, restrictions must be placed on V and Y and an assumption regarding the determination of M must be made. Y, real output, was seen as being determined by real factors such as the size of the capital stock and the labour force, the state of technology, etc and was assumed to be at its full employment level. V was seen as determined by the payment and expenditure practices in the economy and therefore was assumed to be independent of variations in M. M is assumed to be determined independently of PY. With V and Y predetermined and M exogenous, P is the only endogenous variable in the system. In the long run variations in M are reflected in equiproportionate changes in P. Hence in the long-run money does not matter.

Short-Run Non-Neutrality

It should be noted that the neutrality of money is dependent on a number of conditions- price/wage flexibility, an absence of money illusion, an absence of distribution effects and price and interest rate expectations of unitary elasticity. If the assumptions are violated money will not be neutral. These assumptions are most likely to be violated in the short run. This was recognised by the quantity theorists. Hume recognised that prices do not rise proportionately to the increased quantity of money and that in the intervening period this stimulates production.[3] This transmission mechanism can best be seen by reference to the Cambridge Cash Balance version of the quantity theory of money.

M = kPY

where k is the reciprocal of the velocity of circulation of money.

Given the initial situation where peoples portfolio holdings are in equilibrium i.e. peoples desired holdings of money balances equal their actual holdings[4]. Assume the money supply increases[5] so that peoples holdings of money balances exceeds their desired holdings. As a result, people will try to off-load their excess money balances and bring their portfolios back into equilibrium by means of increased expenditure. However, as an aggregate net excess money holdings can not be reduced as one mans spending is another mans receipts. One man can reduce his nominal money balances only by persuading someone else to increase his.[6] If price and income are free to adjust, the increase in total expenditure will initially have an effect on output because of slow adjustment of prices (i.e. short-run non-neutrality), but ultimately prices will be bid up to the level where the real stock of money in the economy is restored to its original level.

A more sophisticated analysis of the effects of monetary policy was later developed by Marshall and Wicksell,. They pointed out that by engaging in an open market operation to buy bonds, the monetary authority will normally be obliged to pay a higher price for bonds. This will bid up the price of bonds and depress interest rates. Thus in the short run they recognised that money did matter in that it determined a real variable, namely the rate of interest. In the short run, the rate of interest will be below the natural rate of interest. [7] This will increase investment and reduce saving and produce what Keynesians regard as an inflationary-gap, i.e. the aggregate demand for output as a whole exceeds the maximum amount of output that the economy is capable of producing. In consequence the level of prices will rise, increasing money demand as consumers require more money to make a given level of transactions. The increase in money demand will restore the rate of interest to its natural rate (assuming that the increase in money supply was a once off). The net effect will be a higher price level. In the long run money will be neutral.

the monetarst - keynesian debate

In the thirty years following the *General Theory* Keynesian economics became known first and foremost as a theory of effective demand emphasising the role of fiscal policy in stabilising output. Monetary policy was subordinated to the role of maintaining interest rates at a low level in order to stimulate investment. Indeed, in most instances (e.g. UK pre-1968) the government targeted the rate of interest and allowed the money supply to react passively. Despite the almost universal acceptance of Keynesian principles, a small number of economists, particularly in the Chicago School, kept faith in the significance of the quantity theory of money. Friedmans restatement of the quantity theory as a theory of money demand combined with the statistical evidence indicating a strong correlation between monetary fluctuations and changes in the level of business activity in the short-run, and in the price level in the long-run, were the building blocks for monetarism. While classical analysis had shown that money supply fluctuations may have real effects in the short run, in the 1960s Keynesians and Monetarists were debating their relative positions of money does not matter at all versus money is all that matters.[8] The debate focussed on two interrelated issues:the stability of the money demand function and the transmission mechanism

The Money Demand function

The Cambridge Cash Balance equation states that M=kPY. The central difference between Keynesians and Monetarists is that for Keynesians a change in M will in certain circumstances, most notably in the case of the liquidity trap, lead to increased hoarding[9]of money reducing the velocity of money, V (increasing k). This is the extreme position and in general most moderate Keynesians while subordinating the role of monetary policy would acknowledge some effect of a change in M on Y (however, changes in the velocity will reduce the impact of such effects). Monetarists, arguing along a similar line to the classical economists, see increases in the monetary supply impinging mainly on Y in the short-run but on P in the long-run (V and k are assumed to be stable functions). To see these effects we must analyse their respective theories of money demand.

Keynes Liquidity Preference Theory

One of Keynes critical departures from classical economics was his theory of the rate of interest which incorporated his liquidity preference theory of money demand. This theory is an extension of the Cambridge Cash Balance equation to include a speculative demand for money.

$M=M1 + M2 = kY + f(r-r^*, r^*)$

Where M equals total money demand, M1 equals money demand for transactions purposes, M2 equals money demand for speculative purposes, r equals actual rate of interest and r* equals the normal rate of interest (i.e. what investors perceive as being the long-run rate of interest).

The theory of liquidity preference is based on the idea that individual wealth holders have a certain opinion regarding what constitutes the normal rate of interest r*and hold either money or bonds in order to reap capital gains from movements in r. If an individual believes that the present rate of interest exceeds the normal rate of interests he expects r to fall and bond prices to rise therefore he will move entirely out of money and into bonds. Conversely, for r<r* he holds only money.[10] This implies that an individuals demand for money is discontinuous at the normal rate of interest. Since bond traders opinions regarding the nominal rate of interest are heterogeneous, summing individual liquidity preference schedules yields a downward sloping speculative demand for money curve.

For a given r Keynes believed that the speculative demand for money would be highly elastic at r= r*.The elasticity of money demand at an observed value of r depended on how homogenous the expectations of different holders of money are and how firmly they are held. At more extreme levels people will tend to converge in expectations. Thus at very low rates of interest, absolute liquidity preference may exist. Under such circumstances monetary authorities will effectively have lost control of monetary policy. As Keynes (1936) pointed out circumstances can develop in which a large increase in the quantity of money may exert a comparatively small influence on the rate of interest. The reason for the loss of control of monetary policy is as follows. If the monetary authorities sought to increase the supply of money by buying bonds this would tend to raise bond prices and lowering rates of return. However,only an infinitesimally small reduction in interest rates is required to entice holders of bonds to substitute into money because of a consensus among bondholders that interest rates will rise in the future.

Keynesian Transmission Mechanism

The implication of the theory was that people would not hold a fixed proportion of their wealth in cash. People will tend to hoard based on their evaluation of the costs and benefits of liquidity.[11] In this situation an increase in the money supply will mean people will hold more money maintaining interest rates at their present value and reducing the velocity of money.[12] Thus, in this case the traditional Keynesian cost of capital effect is not operative.[13] However, Keynes acknowledged that while this limiting case might become practically important in the future, I know no example of it hitherto. Nevertheless, while the possibility of a liquidity trap can be questioned, it seems Keynes remained convinced of only a minor role for monetary policy.[14] He states that it seems unlikely that the influence of banking policy on the rate of interest will be by itself sufficient by itself to determine an optimal rate of investment. I conceive, therefore, that a somewhat comprehensive socialisation of investment will prove the only means of securing an approximation to full employment.[15] However, it would be false to suggest that the Keynesian position is encapsulated by money does not matter just because of the theoretical possibility of the existence of a liquidity trap. Regardless of whether the liquidity trap exists[16]the interest inelasticity of investment is a more important barrier to an influential monetary policy. Money may matter in its effect through alternative transmission mechanisms such as wealth effects and credit availability.

The monetarist position was based on Friedmans restatement of the quantity theory of money. Monetarists analysis is highly similar to the classical analysis and reaches similar conclusions. Friedman introduced the idea of Permanent Income into his demand for money function. He regarded this as a crucial difference between his model and that of the Keynesians (where current income was the relevant variable), in that money demand will be less volatile because it will respond less to changes in transitory income. Friedmans restatement was that money demand was a function of the current price level, the rates of return on bonds and equities, the rate of inflation and the stock of wealth (calculated from permanent income).Thus, money demand (and also velocity) is a stable function of these variables.

This version of money demand is obviously very similar to the quantity theory of money the crucial difference being that the velocity of money in the monetarist model is a stable function i.e. it is determined by a limited number of variables which evolve slowly overtime, whereas in the classical model it is taken as a numerical constant. However, like the classical model, the monetarist money demand function is assumed to be relatively stable as individuals are again assumed to hold a relatively stable amount of money in cash.[17] Furthermore, it assumes that certain factors that influence the money supply do not effect money demand. The impact of this is that variations in money supply will have important effects.

The Monetarist Transmission Mechanism

As a result of the stable demand function monetarists believe in a more direct transmission mechanism of M on Y, unlike the Keynesians. They argue that the crucial difference between themselves and the Keynesians lies in the range of assets considered as substitutes for money. They believe that every asset, including physical assets, has a rate of return (yielded in the form of a flow of services in the case of physical goods) and that every asset is substitutable for money whereas Keynesians believed that only financial assets were close substitutes. [18] The crucial issue ..... is not whether changes in the stock of money operate through interest rates but rather the range of interest rates considered.[19] Thus, since monetarists believe that all assets are substitutes for money an increase in the money supply can lead agents to offload excess money holdings by purchasing capital goods. Thus a direct link is said to exist. However, this is only one possible portfolio adjustment[20] and the specifics of the transmission mechanism have proven difficult to model. Hence monetarists have often been labelled as believing in a black box transmission mechanism.

Black-box monetarism

However, to many people the direct link is implausible because of the mechanism by which increases in the money supply occur. Unlike the classical system increases in the money supply occur through banking operations, therefore using the analogy of dropping money from a helicopter and analysing the effects on spending as Friedman (1969) did is not the same as analysing the effects of an open market operation . The helicopter example did much to provoke claims of a black box transmission mechanism in monetarism. Furthermore, Friedmans riposte to such criticism that if anybody asks what is the mechanism by which an increase in the money supply brings about an increase in prices, what David Hume has to say answers that question about as well as anybody else I know[21]did little to dispel such criticism. Trevithik points out that it is ridiculous to view variations in the total volume of bank deposits as having the same effects to those of increases in the supply of gold in 18th century England.[22] However, monetarists now interpret the transmission mechanism along similar lines as Keynasians such as James Tobin.

The Empirical Evidence

The basic differences among economists are empirical not theoretical[23]

It is often said that the main differences between the monetarists and Keynesians surround empirical magnitudes. If this is the case, an examination of the evidence may give some insight into what school may offer the best theory.

In the US a number of empirical studies found a statistical relationship between money and the level of economic activity. Firstly, the pioneering work of Friedman and Schwartz (1963) was one of the building blocks of modern monetarism. Using US data from 1867 to 1960 they investigated the length and variability of the time lag involved in the influence of money. Their most crucial finding was that peaks in the rate of change of the money supply preceded peaks in the level of economic activity by an average of 16 months and the corresponding figure for troughs was 12 months. The implications drawn from the results were not only that variations in the stock of money caused fluctuations in the business cycle but also that since the lags were long and variable a constant rate of money growth was the optimal monetary policy.

However, a number of problems exist regarding the study. Most importantly, Friedman and Schwartz compared the *rate of change* of the money supply with the *level* of business activity. However, for any time series of data the rate of change will peak before the level does, so that even if, in terms of levels both cycles

were exactly contemporaneous the peak in the rate of change of the money supply would precede the peak in the level of business activity. Thus the study can be criticised on the basis of invalid statistical techniques.

Even allowing for this statistical error temporal precedence does not necessarily imply causality. Firstly, changes in the money suppply may be reacting to changes in money demand caused by increased income. This will be true in the case where the monetary authority targets the rate of interest and allows the money supply to react passively. Tobin (1970) derived the Keynesian type model in which the money supply was demand determined and yet the turning point of the rate of change of the money supply actually led the peaks and troughs in nominal income. Also in a Friedman type model where the money supply was exogenously determined and the demand for money was a function of permanent income the turning points in the rate of change in the money supply actually lagged those of nominal income.

Secondly, both variables may have been influenced by another variable (e.g. the budget deficit) with money reacting more rapidly than income. An increase in the budget deficit must be financed by printing money or issuing bonds, and *ceteris paribus* this will lead to an increase in the money supply also having an impact on income. In this instance the transmission mechanism by which the budget deficit affects income may run through money but the exact transmission mechanism will be subject to the same disputes already mentioned.

Thirdly, if planned expenditures fall people will demand less money and the money supply will fall. In this instance the subsequent fall in expenditure is not caused by the fall in the money supply. Instead the causation runs from expenditure to the money supply. This point also illustrates the general Keynesian proposition of variations in M do not cause variations in Y but that the reverse holds.

Further studies in the US[24] attempted to model the individual influence of monetary and fiscal variables on income. Both studies came out strongly in favour of the importance of monetary policy but both were subject to criticism on methodological grounds and often the results changed dramatically when a different model was used.[25]

Consensus ?

While the major issues have not been resolved by empirical evidence, both schools positions are now more modified. Both Neanderthal Keynesians, denying that monetary authorities have any control over the money supply ,and Neanderthal Monetarists denying any efficacy of fiscal policy even in the short-run, have been left far behind.[26] Blaug (1980) argues that in one sense the Monetarists have won - governments now take greater account of fluctuations in the money supply. Keynesianism has developed to take account of the several channels through which monetary disturbances influence the economy. In a more profound sense, however, monetarism has lost. Blaug argues that monetarism never succeeded in clarifying the causal mechanism that produced their empirical results, sometimes even denying that the results required interpretation in the light of supporting causal theory.[27] Monetarists now recognise that the money supply can only be regarded as exogenous in the long-run.

In the 1960s debate centred on one particular transmission mechanism - portfolio adjustment. In the context of portfolio adjustment, Keynesians were portrayed as regarding monetary policy as impotent because of the existence of the liquidity trap, whereas monetarists were portrayed as believing in a black box transmission mechanism. Both positions are an exaggeration. Keynesians recognize that changes in the money supply may have effects through wealth effects and increased credit availibility. Monetarist theories of portfolio adjustment are now very similar to Keynesian theories (such as Tobins). Monetarists also emphasise the important influence of monetary policy on wage and price expectations Thus, monetarists and Keynesians have tended to converge to some degree, although differences of emphasis still exist.

The empirical evidence has proven that the money supply not only effects the level of business activity but is also influenced by it. This means that it is difficult to test empirically for the monetary effects on activity because allowance must be made for the feedback effect of economic activity on the money supply. The extent to which variations in the money supply are demand-determined is strongly influenced by the reactions of the monetary authorities. Thus, to model the feedback effect it is necessary to model Central Bank behaviour. Attempts to do this have been relatively unsuccessful. This presents a problem in assessing empirical evidence regarding the role of money in the economy. However, Cagan (1987) points out that while the concurrent mutual interaction between money and economic activity remains difficult to disentangle, the

longer the lag in monetary effects the less likely that the feedback from activity to money can account for the observed association.

CONCLUSION

While monetarists and Keynesians have modified their positions as a result of the theoretical debate of the 1960s, the exact impact of money on the level of activity is still disputed. The instability of money demand and the unpredictability of the velocity of money in the early 1980s, and the abandonment of the monetarist experiment in the UK and US in the 1980s were the basis of the criticisms of monetarism.

A new challenge to the money matters school came from the body of new classical economics. Based on their theory of rational expectations new classical economists, most notably Sargent and Wallace, believed in the idea of policy irrelevance. With regard to monetary policy since agents make rational decisions, only unanticipated changes in the money supply will effect output. Since individuals are rational, anticipated increases in the money supply will be reflected in higher price expectations, and compensatory wage bargains will ensure that real wages remain constant. Furthermore, the assumption of perfect price and wage flexibility implies that any increases in the money supply will be reflected in higher prices in the short run (i.e money is neutral even in the short run).

New classical views are difficult to test empirically. For example, it is very difficult to measure what part of a monetary policy can be regarded as anticipated and unanticipated. However, since their predictions are based on the restrictive assumptions of flexible prices and wages, in the short run, this tends to suggest that monetary policy will have some impact. Apart from the new classicists, most modern economists would acknowledge that the basic postulates of the quantity theory still holds - in the long run prices respond and moneyis neutral, but in the short run prices are slow to respond to demand fluctuations and money does matter.

Bibliography

Blaug, M (1980) *The Methodology of Economics*

Cagan, P (1987) Monetarism in Eatwell et.al. (ed) *The New Palgrave*

Dennis, EJ (1981) *Monetary Economics*

Friedman, M (1956) The Quantity Theory of Money : A Restatement in M. Friedman (ed) *Studies in the Quantity Theory of Money*

Friedman, M (1969) *The Optimum Quantity of Money and Other Essays*

Friedman,M (1987) The Quantity Theory of Money in Eatwell et. al (ed) *The New Palgrave*

Gowland, D (1991) *Money, Inflation and Unemployment*

Keynes, JM (1936) *The General Theory of Employment, Interest and Money*

Morgan, B (1977) *Monetarists and Keynesians*

Patinkin, D (1987) The neutrality of money in Eatwell et. al. (ed) *The New Palgrave*

Pierce, D and Tysome, P (1985) *Monetary Economics*

Pigou, AC (1917) The Value of Money in *Quarterly Journal of Economics* 32, November

Tobin, J (1970) Money and income: post hoc ergo propter hoc *Quaterly Journal of Economics*

Trevithick, J (1992) *Involuntary Unemployment*

An Inquiry into the Growth of Nations

Cloda Lane

Junior Sophister

The single most important force leading to long-run increases in living standards is economic growth. The benefits of economic growth include increases in income levels (or 'welfare'), greater room for redistributive policies and favourable changes in lifestyle. The costs include forgone consumption today and environmental damage due to increased production. In this analysis I am going to assume that growth is a 'good' thing and it is the desire to know why growth has been achieved by some countries and has not been achieved by others that motivates this inquiry. The investigation is also being carried out to empirically test whether Keynesian demand management or Classical supply-side policies should be used to stimulate economic growth. Finally, this phenomena is being explored because it is not understood. As yet there are no firm theoretical conclusions about the causes of growth or about the reasons why growth rates differ. My investigation is therefore aiming to shed some light on a currently uncertain area.

STAGE 1 : SPECIFICATION

Y-Variable

My dependent variable in this investigation is the average annual rate of economic growth between 1970 and 1985. This is measured as the average rate of change in Real GDP per capita (RGDPCH). I chose to average growth over this time-span rather than investigate it in any one year because there could be too many unique factors influencing it in any one year that would defer it from its normal pattern. Rates are used instead of absolute growth levels because they show how growth is related to the country's initial level of income, they are easier to use in a cross-sectional analysis and they are the most commonly used measure of growth in economic analysis.

Economic growth is a measure of the change in national income over time. In my analysis the change in RGDPCH is used as the most preferred measure for national income because it eliminates the influence of inflation or deflation on changes in GDP by taking a real level, and the per capita figure is used because it is a relevant measure of welfare change.

X-Variables

In order to understand why economic growth occurs independent X-variables must be introduced into the investigation. Economic theory has proposed numerous variables which may determine the rate of economic growth. These range from population growth to consumption expenditure to confidence levels to political stability. It must also be emphasised that different factors may influence the country's growth rate depending on the country's level of development. Thus, it can be seen that by limiting my exploration to three independent variables my analysis is going to be imperfect due to the absence of many other possible determining factors.

X1 : My first X-variable is the average annual rate of investment expenditure (as a % of RGDP) between 1970 and 1985 (**I**). This variable was chosen because it is often assumed by theorists and policy makers to have a large influence on growth. The promotion of investment to increase RGDPCH is a central demand management policy that is supported by Keynesians. Thus, the relationship between **I** and economic growth is being investigated to test the Keynesian viewpoint. The relationship is also being explored to decipher whether it is worth our while taking money, which could be used for consumption today, to put into investment so that consumption could be higher in the future. That is, will economic growth (a benefit of investment) occur to outweigh forgone consumption today (a cost of investment).

X2 : My second X-variable is the average rate of government expenditure (as a % of RGDP) between 1970 and 1985 (**G**). Like investment, this variable was chosen because it is often cited as having a large influence on growth. Specifically, Keynesians argue that government expenditure is needed to stabilise an economy and that it will cause RGDPCH to increase. This variable is also being used in my analysis because I want to see if the growing levels of public sector expenditure (PSE) worldwide have been the cause of economic growth. If this is the case then the hypothesis that increased **G** will lead to a better future would provide a justification for large levels of PSE.

X3 : My third, and final X-variable is the average number of school years in 1970 (**HK**). This variable was chosen as a 'proxy' measure of human capital in a country. I chose to use the level of human capital in one year rather than over the 15 year period because it is the educational level of the workforce (assumed to be determined by the length of time they spent in school before entering the labour market) over the 15 years which influences the rate of change of income and not the educational level of the students in school over the period. It is generally hypothesised that an economy will grow if its workforce is well - educated. The purpose of the investigation into the relationship between growth and **HK** is to test this classical supply-side argument.

The Model

In this investigation I am using the estimation technique of ordinary least squares. This allows me to use regression analysis to get a line of 'best fit' for my data. The model takes the form of

$$Y = 0 + 1X1 + 2X2 + 3X3 + e$$

for the multiple regression case. The purpose of this exploration, then, is to estimate the size and sign of the unknown parameters (0 , 1 , 2 , 3).

STAGE 2 : ESTIMATION

The Data

In my exploration of economic growth I have 40 observations. I left myself with this many countries because I feel that the growth rate is too diverse on a global scale to narrow my analysis of it down to a particular region of the world. Instead, I wanted to get a general overview of world economic growth rates. This is done by breaking my observations down into four subsections (10 countries in each) with each section representing a national income category as defined by the World Development Report (1993).

The data for RGDPCH (from which the growth rates were derived), **I** and **G** were all taken from the Penn World Tables (Mark 5). These data are consistent because the compilers have taken figures from the benchmark studies of the UN International Comparison Program for 138 countries and have adjusted them all for different purchasing power parities and they have all been indexed linked to 1985 international prices. It should be noted, however, that the data are only as reliable as the UN figures, which in most cases, are only as reliable as the individual country's National Accounts. Thus, for some countries, most notably the Centrally Planned Economies, the RGDPCH figures and thus, the growth rates may be overestimated.

The data for the average number of school years in 1970 was taken from an NBER working paper by Barro and Lee. The authors got the figures from national censuses so the data are only as reliable as these censuses. The measure used should be fairly consistent as interpretations of this straight forward measure can not differ too much from country to country.

The Estimation

The results of the multiple regression analysis are shown in the table below. A priori, the problem of multicollinearity is recognised by the author although it shall not be dealt with in detail. The line of best fit has been estimated as :

Y = 0.58 + 0.06 X1 - 0.05 X2 - 0.05 X3 + e

where 'e' is the error term or residual of the regression. The correlation coefficient , R^2, indicates that 32% of the variation in Y can be explained by the linear influence of all the Xs.

Dependent Variable: Rate of Economic Growth

Independent Variable	Parameter Estimate	t - statistic H0 : i = 0
Constant	0.5786	0.93612
X1 = I	0.0583	2.75994
X2 = G	-0.0450	-2.25457
X3 = HK	-0.0513	-0.73746

$\mathbf{R^2}$ = O.3246

STAGE 3 : EVALUATION

Economic Theory

To evaluate the regression results I am going to compare my estimations with those predicted by theory. In all, the three X-variables have, as theory claims, a relatively significant impact on economic growth (32%). This result is especially interesting when we consider that consumption expenditure, the largest influence on RGDPCH, has not been included in the analysis. Looking more closely at the results, however, we see that the outcome is not so positive. The Keynesian theory that increased **G** means increased RGDP and the Classical Theory that increased **HK** would lead to economic growth are both contradicted by my regression results (i.e. the parameter estimates have a negative sign).

In the case of human capital, no result can be taken as absolute because of the 'proxy' nature of my variable. It should be noted, however, that the number of years spent in school may not have an influence on growth. It may be the quality of the education and not the length of it that matters for increasing productivity and thus, growth. In the case of government expenditure there is probably some truth in the result that increased **G** will not automatically stimulate economic growth. This is particularly true if the expenditure was aimed at sectors that had no growth potential or if it was mismanaged. Thus, badly spent PSE, no matter how large it is, may not lead to an increase in RGDPCH.

The case for investment is far more hopeful. Although the parameter estimate is small (0.06), indicating a not so great influence of **I** on growth, the sign of the coefficient corresponds to the one proposed by economic theory. That is, the positive nature of 1 coincides with the Keynesian theory that an increase in investment will stimulate economic growth.

Statistical Evaluation

To make a simple statistical evaluation of my results I am going to examine the t-statistics given in the table. Considering the hypothesis that there is no relationship between X and Y (H0 : i = 0) we can look at the t-statistic which is the ratio of the estimate to the standard error. An estimate of a parameter is statistically significant if the t-statistic associated with it causes us to reject, at a particular significance level, the hypothesis that i is equal to zero. At both the 5% and 10% significance levels the estimate for 1 is statistically significant while, at the same significance levels, the estimates for 2 and 3 are statistically insignificant.

Policy Relevance

In today's world, economic strength is the key to political power. Thus countries that are striving to become powerful and to increase living standards all look for economic growth. A final means of evaluating my results, therefore, is to see if they are relevant to policy makers who are trying to achieve this primary policy objective.

My research would indicate to policy makers that the best method of getting growth is to encourage investment. It should be noted, however, that we can not assume that increased investment will automatically lead to large scale growth. Only if investment is directed to the right areas will RGDPCH be increased. Secondly, my results will, perhaps, encourage policy makers to reduce PSE. They could not justify large scale expenditure on the grounds that it will lead to an improved future using my analysis. Instead, it would be suggested that lower but better managed PSE is what is needed to stimulate economic growth. Finally, I would discourage policy makers from taking my results vis à vis human capital too seriously. They should be aware of the very slight possibility that the results are right and could use this as a guideline to improve the quality rather than the quantity of education. They must not, however, take the results so seriously that they would actively discourage people from staying in education. This would have disastrous results for society and for the economy.

It is hoped that my research provides an insight for policy makers and other economists into the workings of economic growth across the globe. The results can be taken as possible indicators of how growth may be generated but they should not be taken as being infallible. Like all econometric work, my exploration into economic growth should only become completely acceptable if and when it is verified by the research of others.

Bibliography

Barro, R & Lee, J (1993) International Comparisons of Educational Attainment Mankiw, G (1992) Macroeconomics Summers, R & Heston, A (1991) "The Penn World Table (Mark 5): An Expected Set of International Comparisons, 1950 - 1988" in the Quarterly Journal of Economics

World Bank (1993) *World Development Report*

Customs Unions and Trade Conflicts

GATT and the EC: The case of Agriculture

Deborah Heaney

Senior Sophister

The concluding declaration of a GATT ministerial meeting in 1982 stated the need "*to bring agriculture more fully into the multilateral trading system*".[9] The seven years taken to bring about a successful conclusion of the Uruguay Round of negotiations indicate the complexity of such a task, and the painstaking edging of the EC and the US towards compromise from the diametrically opposed positions set out in the Punta del Este declaration is a result of this complexity.

This paper attempts to demonstrate how the principal areas of dispute have their roots to some extent in the particular legal status of agriculture within the GATT but how their deepest roots are in a contrasting attitude towards the best means by which agricultural support can be disciplined. The paper is divided into four sections. Section I deals with the background to the negotiations - the economic context and the legal preferences. Section II examines the main issues of dispute using the main groupings of proposals as a basis. Section III analyses the roots of these differences, showing how the GATT negotiations became a table for domestic policy and outlines the differing philosophies regarding the discipline of agricultural support. Section IV draws some conclusions.

SECTION I

Background

"*Agricultural trade has not been effectively governed by the institutional framework for international trade relations that has been provided by the GATT since its inception in 1957*" [10]

Before moving to an assessment of the issues under dispute in the Uruguay Round, it is, I believe, vital to understand the situation in which countries found themselves in the 1980s. As a result of price, production and consumption developments the USA and the EC began to accumulate stocks at increasingly rapid rates. [11] Two main effects were notable. Firstly, to get rid of surpluses, the Europeans subsidized exports, which led to trade disputes with the US and the Third World. Secondly, budgetary pressures became intolerable and set the US and the EC on the path of reform, albeit in different directions.

As suggested by Rayner et al (1993) the primary problem lay with the operation of agricultural trade policies within a preferential GATT framework, which had not only the effect of increasing agricultural distortions, but also undermined the credibility of the GATT as an organisation. Both GATT rules designed to conform to policies of the previous three decades, and domestic policies of the 1970s became totally inadequate in the 1980s environment.

The special treatment given to agriculture under the GATT is, I would argue, a major source of dispute in the Uruguay Round. Had agriculture been subject to the same rules as other sectors then there would be little leeway for separate negotiations. Of crucial importance to the disputes are the special provisions regarding export subsidies; prohibited by Article XVI, except for primary products. Article XVI.3 says "such a subsidy shall not be applied in a manner which results in that contracting party having *more than an equitable share of world trade* in that product, account being taken of the share of the contracting parties in such trade in the product *during a representative period*." (my italics). It is clear from this quote that a minefield of definitional issues were latent, ready for such a time when agriculture was fully placed on the negotiating table.

Normally protection from imports is allowed in the form of tariffs only, however, the use of import restrictions in agriculture was permitted as long as domestic supply was restrained concurrently (Article XI). It was the US who in 1955 instigated a special waiver allowing it the right to impose import restrictions without corresponding supply restrictions - a particular case of US insistence that GATT rules fit US farm policy.[12] The position of agriculture in the GATT was also influenced by the EC who insisted that the CAP (and in particular the variable levy system) was integral to the success of the Common Market (which the US supported), and as such was non-negotiable, and never effectively challenged in the GATT.

SECTION II

The Issues of Dispute

The initial US proposal, the double zero option, indicated the absolute extremities of the gap between the US and the EC. It proposed the complete elimination of all forms of agricultural support over a ten year period, allowing only decoupled payments and verified food aid programmes. Import barriers to be eliminated included phytosanitary barriers. Policy changes were to be commodity specific, using producer subsidy equivalents to measure initial levels of protection and monitor changes. This option would not only have removed agriculture from its privileged status within the GATT, but would have made it the most liberalised sector under the GATT.

The EC favoured no such dramatic changes in agriculture's status within the GATT nor an ultimate goal of support elimination. The CAP was seen as integral to the success of the EC and its flagship achievement. Proposals for short term market sharing and long run reductions in support measured at the aggregate level were put forward. Such diametrically opposed proposals allowed no agreement and the impasse of Montreal was the inevitable result. By April 1989 an agreed agenda was produced with a compromise position agreeing on a *`progressive reduction in support'* - signalling a US abandonment of the zero option and an EC concession that reduction should take place. This statement allowed negotiations on the key areas of dispute under three primary headings:

· *export subsidies*

· *import access*

· *internal supports*

This structure in itself represented an EC concession in allowing negotiations on types of support and not only aggregate levels.

Export subsidies

Hathaway (1990) presents the disputes as an interpolation of reform and restraint and with this analysis he comments "It is in the area of export competition that the difference between reform and restraint is most pronounced and the political confrontation most stark." It is in this area, as noted in section I, that agriculture is particularly favourably treated. Specific issues relating to export subsidies centred on how they should be defined and how then they could be disciplined. Disagreement surrounded the definition of export subsidies. Should any policy that stimulated exports be regarded as an export subsidy? More specifically, should US deficiency payments be regarded as export subsidies? The US argued that domestic production subsidies were a separate dispute and should be disciplined as such.

How export subsidies should be disciplined was the second major issue - what level of reduction would be allowed and at what level of aggregation and volume. Confrontation was between reform and restraint, the former being favoured by the US, as total phase-out and prohibition, while the EC favoured the latter, involving the definition of an equitable market share and an export subsidy, i.e. clarification of what already existed. Definition of an equitable market share was seen as feasible for many products, but for others such as sugar it was more complex because major producers such as Cuba did not subscribe to GATT.

Market Access

"*This is almost certainly a case in which either everybody reforms or no one does*".[13]

The areas of dispute under market access included the EC variable levy which had resulted in total insulation of the market, the US section 22 waiver on quotas and voluntary export restraints (VERS) such as those between the EC and New Zealand for butter.

The increasing importance of non tariff barriers in world trade, especially in agriculture relative to manufacturing[14] meant that agreement on tariffication of non-tariff barriers was an important issue. The EC favoured tariffication with a two part tariff - a fixed element plus supplementary additions and a variable element. The US favoured a formula approach for tariffication (to increase transparency) based on the difference between the domestic price and an external reference price. These tariff equivalents were to be used for a transition period, over which reduction was to take place.

Internal Supports

The US attitude changed over the negotiation period from an overall emphasis on discipline in all areas, to an emphasis on border measures as the most distortionary and the instigation of a system of classification. "This shift was predictable once it became clear that only very limited progress would be made in the Uruguay Round in reducing domestic subsidies."[15] Overall the US defined trade distorting strategies according to red, green and amber, where decoupled payments were given the green light.

Internal subsidies may in reality be more self limiting than other protective measures due to their prohibitive cost. The choice of a measure and its level of aggregation was a continuous source of disagreement. Originally the US wished the elimination of all supports over ten years, this was later reduced to the red supports only over ten years with amber supports to be monitored. The EC desired aggregation for support reduction to allow rebalancing and the retention of domestic control.

This brief outline of some of the main issues in dispute is naturally only the tip of the iceberg but provides a launching pad for the analysis of the roots of these main disputes.

SECTION III

The roots of the issues

Philosophies about how best to discipline agricultural support are to some extent the roots of many of the disputed issues on agriculture outlined above. However it seems that an awareness of political realities is also important as is the potential for protectionism in agriculture. If a group of countries has taken advantage of the more liberal legal regime for agriculture, then others will do so, if not for national gain, just to level the playing field once more.

US concerns as expressed within the Uruguay round, focus on the type of support in existence (the rules base approach) and wish to take account of differing distortions imposed by different types of protectionism encompassing supports in a scheme of tariffication with income supports unrelated to current production levels. Such a philosophy on the discipline of agricultural support is based on the view held that US farmers are so efficient that they would thrive in an unregulated agricultural market.[16] Furthermore the US sees itself to a large extent as having been forced into export subsides as a retaliatory measure and thus sees the best way to reduce costs as being through enforcing protection reductions elsewhere.

For the EC the CAP is a symbol of European unity, and it will seek maximum discretion and control over agricultural policy. This being the area in which the EC has most control over its member states control will not be easily ceded. Hence CAP reforms were being worked out even while GATT negotiations were taking place. The reform package was then presented as *fait accompli* as a negotiation package thus retaining the image of the control of the EC on agriculture.[17] In a sense then the GATT was used as a negotiation tool to enable tough domestic policy decisions to be taken in a multilateral context and domestic policies were confused with those on trade. The EC's philosophy is that support can best be disciplined by controlling its amount not its nature (allowing maximum flexibility).

These philosophies were played off in a strategic way, but I would argue that equally crucial were the political realities. If we set up a game[18] between the US and the EC using payoffs calculated by Tyers and Anderson then we can see a strategic interaction based on the net gains and the producer gains. The Nash equilibrium for net welfare gains is to liberalise, but where only producers are taken into account the Nash equilibrium is to not liberalise. Hence the striving towards discipline of agricultural support largely depends on how the governments weigh interests within the economy, and not just on an overall self interested philosophy.

Payoffs (EC, US) Net Economic Gains

	US (do not liberalise)	US (liberalise)
EC (do not liberalise)	(0.0, 0.0)	(-1.9, 3.3)
EC (liberalise)	(21.4, 1.7)	(17.6, 3.1)

Payoffs (EC, US) Producer Gains

	US (do not liberalise)	US (liberalise)
EC (do not liberalise)	(0.0, 0.0)	(5.0, -21.5)
EC (liberalise)	(-88.6, 7.6)	(-73.7, 3.1)

SECTION IV

Conclusion

The legal status of agriculture within the GATT has been shown as the main culprit of the painstaking negotiation process from dramatically opposed positions by the EC and the US. For too long, agriculture was accorded special status within GATT, natural interests superseding those of international trade. Once agriculture was included under the GATT umbrella, a minefield of issues was unearthed. Productive negotiations focused on three areas: export subsidies, market access & internal supports. Within these areas disputes focused on the time periods for phase-out supports, the base period from which to calculate reductions, levels of aggregation and definitional issues. For the EC, maintaining a semblance of autonomy and allowing CAP structures to remain was crucial to negotiations -- thus a domestic reform package predated the final GATT outcome. Linked with this, I have argued that positions were based not so much on deep held philosophies, but on political realities, and the interaction of interest groups in this context -- and that these were crucial to the final outcome and the process of negotiation towards it.

Appendix

Table 1

Level of Public Stocks : EC : 1000t year end

	1979	1983	1986	1987	1988
Cereals	2677	9542	14717	8147	8312
Beef	293	410	576	776	425
Butter	310	686	1297	860	120

Source: European Commission (1989) "A Common Agricultural Policy for the 1990s"

Table 2

% Imports subject to NTBs

	Manufactures	Agriculture
EC (from LDCs)	29.9	26.9
EC (from ind. countries)	15.2	47.7
US (from LDCs0	18.6	25.1
US (from ind. countries)	16.5	23.5

Source: Zietz and Valdes 1988.

Bibliography

Australian Bureau of Agricultural and Resource Economics (1990) "Proposed strategies for reducing agricultural protection in the GATT Uruguay Round"

European Commission (1989) "A Common Agricultural Policy for the 1990s" *European Documentation Periodical*

Fitchett, D (1989) "Agriculture" in Finger, J & Olechovski, A eds. *The Uruguay Round, A handbook for the multilateral negotiations*

Greenaway, D (1991) "The Uruguay Round of Multilateral Negotiations : last chance for GATT?" in *Journal of Agricultural Economics* Vol.42 No.3

Hathaway, D (1987) *Agriculture and the GATT, Rewriting the Rules*

Hathaway, D (1990) in Schott, J ed. *Completing the Uruguay Round: A results orientated approach to the GATT trade negotiations*

Matthews, A (1993) *Gatt Impact: Agriculture and Food*

Rayner, A, Ingersent K, Hine, R C (1989) "Agriculture in the Uruguay Round: from the Punta del Este declaration to the Geneva Accord" in *The Journal of Agricultural Economics* Vol.40, No.3

Rayner et al., (1993) "Agricultural Trade and the GATT" in Rayner A and Colman D *Current Issues in Agricultural Economics*

Sumner, D (1993) "The economic underpinnings of the Uruguay Round Proposals", in Becker, T ed. *Improving Agricultural Trade Performance under GATT*

The Economist September 22nd 1990

Zietz, J, Valdes, A (1988) *Agriculture in the GATT, An Analysis of Alternative Approaches to reform*

The Japanese Trade Surplus with the European Union: Extent, Causes and Implications.

Olaf Gersemann

One Year ERASMUS

This essay examines the European Union's trade deficit with Japan. It argues that this deficit, a cause of concern to the EU, could be even larger if trade barriers were removed. Firstly, I discuss, using data from the EU's Statistical Office, the magnitude of the deficit, then the reasons for it on the Japanese side, and finally the welfare implications for the EU. Examples are taken from the car market, as cars are the most visible product of Japan's economic success.

EXTENT AND COMPOSITION OF THE JAPANESE TRADE SURPLUS

Extent

Trade with Japan accounts for a large proportion (almost three fifths in 1992) of the EU's overall trade deficit. The Japanese surplus with the EU rose from 0.7 billion ECU in 1970 to 31.0 billion ECU in 1992.

Even when the yen was strong, the import-export ratio increased from 1.3 in 1960 to 2.9 in 1986. In 1992, the value of EU imports from Japan was 2.5 times higher than the value of the Union's exports to Japan. The penetration of Japanese markets by European products, on the other hand, was even lower in 1992 (0.72 %) than in 1960 (0.74 %).

Japanese products, however, still account for not more than 10.6% of EU total imports, less than the imports from Austria and Switzerland together. Hence, the attention focused on the Japanese trade surplus is "scarcely proportionate to its relative significance"[1] for the EU.

Composition

Japan has few raw materials or energy sources of its own. As a consequence, it has an enormous demand for these primary products and is forced to produce goods and services for export. Japan's economy has specialised in sophisticated products like machinery and transport equipment. In 1992, three such sectors (road vehicles, office machinery, telecommunications and sound equipment) contributed

* 51.7% of total imports from Japan and

* 74.2% to the EU's total trade deficit with Japan.

Thus, it is obvious that Japanese pressure is felt particularly in specific sectors and in the countries where these sectors are located.

Voluntary Export Restraints and Direct Investments

Given free trade, the Japanese exports to the European Union would be likely to be even bigger, since this would imply that tariffs and non-tariff barriers (NTBs) would be removed.

In addition, for example, to the Union's common external tariff of 10.3%, five EU member states restrict car imports from Japan by Voluntary Export Restraints (VERs). From the exporters' point of view, a VER is the lesser of two evils, since it not only raises the price to consumers, in the same way as a tariff, but also "allows

the exporters to obtain the rents associated with the artificially high prices"[2] by letting them do the restricting.

However, Japanese car manufacturers should look forward to the removal, by the year 2000, of the VERs on car imports from Japan, as agreed by the EU in 1991. Smith and Venables (1991) showed in a model based on the world car market in 1988 that a removal of Europe's VERs would push the Japanese market share from 8.8% to 16.5%. This increase in volume will, according to Smith and Venables, outweigh the reduction in price (which in general cannot be taken as given, as Greenaway and Hindley (1985) point out). The profits of Japanese car manufacturers are expected to rise by one billion ECUs per year.

In 1977, Japanese firms began significantly to export not only goods but also industries. Japanese-owned car factories in the Union, for instance, produced some 500,000 cars in 1993 and are expected to make about 2 million cars annually by the year 2000.

Japan's Foreign Direct Investments (FDIs) in Europe can be considered "an effective counter to the accusation that Japanese imports are `destroying' jobs"[3]. Circumventing barriers might not be the only reason for the FDIs in the EU, but the investments tend to be concentrated in industries which face protectionism. However, the FDIs, at least, reduce the EU's trade deficit if the Japanese factories use, as do the car manufacturers, a high proportion of locally made parts. Toyota, for example, expect the European content of its cars to reach 80% by value this year.

CAUSES OF THE JAPANESE TRADE SURPLUS

Europe's manufacturers are clearly failing to penetrate Japan's markets: Thurow (1992) says that Japanese imports from Europe are 25% to 45% less than would be expected given Japan's circumstances. This might be due partly to the preferences of Japan's consumers. However, given roughly similar preferences of consumers in Japan and Europe, there are two possible reasons for the European lack of success: firstly, Japan might be more protectionist than Europe and secondly, Europe's economy, when compared with its counterpart, might suffer from a lack of competitiveness.

The Openness of Japanese Markets

Undoubtedly, the Japanese built a protective wall in post-war years in order to develop certain industries. However, in the last few decades Japan has cut her tariffs and quotas down significantly. Hence, according to Hanabusa (1979) the discriminatory effects on imports to Japan today have to be considered dubious. Nevertheless, Laura D'Andrea Tyson (1992), Chief Economic Adviser to President Clinton, considers that despite the tariff reductions, the country "remains significantly more closed to foreign trade and investment" than other advanced industrial nations. The closure, argues Tyson, is a consequence of structural market barriers created by the integration of large firms with one another. Many Japanese companies are members of *keiretsu*, groups consisting of companies with cross-shareholdings in one another. The members of these business groups give preferential treatment to each other as preferred suppliers and customers, i.e. by co-ordination and economies of scale. Further co-ordination is provided by the Ministry for International Trade and Industry (MITI). However, to argue that this is protectionist is subjective and difficult to prove.

The Competitiveness of Japan's Economy

A country's competitiveness is reflected by its ability to increase its share of export markets, given free trade (or an equal level of protectionism) and depends upon the productivity and costs of its factors of production.

In the 1960s Japanese productivity was significantly below that of the USA and Germany. The gap had been nearly closed by 1970 and had turned into an advantage by the end of the 1970s. Today Japanese manufacturers can, for instance, assemble a luxury car with only one quarter of the labour that is required in Europe[4].

Japan's competitiveness can also be seen by comparing the performance in third markets. In the US car market, for instance, Japanese producers proved to be far more competitive than the Europeans. While Volkswagen, the last European mass manufacturer left in the US market, "is in the process of being driven out", Japan's market share declined in recent years but is still higher than one fifth. (Although one must be

cautious about comparing a single sector since it is possible that Japan has competitive disadvantages in other areas.)

The Japanese cost advantage is due to several characteristics of Japan's economy and society. Three of them will be presented in more detail in the following section.

System of Values

"The Japanese secret is to be found in the fact that they have tapped a universal human desire... to belong to an empire". The goal of Japanese firms is not profit maximisation, but market share maximisation. Accordingly, "a big part of the Japanese success in coping with the rising value of the yen"[5] is found in their willingness to accept lower rates of return than their competitors.

Moreover, the identification of Japanese employees with their companies is far stronger than elsewhere. This is because firms offer life-time job security (which would in the view of European managers undercut motivation). Even in recessions Japanese companies tend to accept `in-house unemployment' rather than dismissals).

Savings, Investments, and Consumption

In the Japanese economy, savings and investments are systematically promoted. Compared to European countries

* the work force gets a low share of national income,

* rates of personal income and commodity taxation are low

* dividends are low relative to after-tax profits.

The outcome is a traditionally high savings ratio which in turn leads to a low price of capital. Hence, the required rates of return are far below the corresponding rates in the EU. As a result, Japan invests relatively more money which in turn improves productivity and thereby supports economic growth. High savings ratios, on the other hand, are synonymous with low consumption ratios. This means that Japan's domestic demand is traditionally weak. Therefore market penetration by imports is impeded.

Research and Development

Japan traditionally has a high level of research and development (R&D), encouraged by MITI, the existence of *keiretsu*, and the low price of capital. Another advantage is that it has concentrated its R&D since the 1950s not on new products but on new production processes. Therefore Japanese firms are often, as is the case of the CD player, not the inventors of a new product but the cheapest producers.

IMPLICATIONS OF THE JAPANESE TRADE SURPLUS

Consumers

European consumers gained enormously from the increase of Japanese imports. Access to Japanese goods offers additional choice through product differentiation. It provides, as Winters (1992) expresses, `life-enhancing variety'.

Furthermore, the penetration of EU markets by Japanese products stimulated competition. This led, for example, to "vast improvements in the quality of automobiles which have been pioneered by Japan and then imitated and developed"[6] in the Union.

However, the gains for consumers would be far bigger if the EU stopped protecting its industries from Japanese competitors. According to Smith and Venables (1991), a removal of the VERs in the European car market would imply a consumer surplus in the order of 6 billion ECU per year, an amount that clearly outweighs the predicted losses to European producers of approximately 3 billion ECUs per annum.

Employees, Producers and Taxpayers

If exports and imports increase in a balanced way, it is not likely that the expansion of trade between the European Union and Japan would create major (net) employment problems. In the case of trade between developed countries (or trading blocks, respectively) it can under certain conditions be assumed that job losses which emerge from additional imports are roughly compensated by job gains through additional exports.

If, however, imports grow faster than exports the contraction of one sector is not accompanied by the expansion of other sectors; the losses of profits and employment in the penetrated industries are not compensated. Moreover, the speed at which the penetration by Japanese imports boomed and still booms creates serious problems by making adjustments in the affected sectors more difficult and, thus, costly. Producers therefore usually plead for time. If they are, like the European car makers, successful, competitive pressure is taken off by protective measures such as subsidies and VERs at the expense of European taxpayers and consumers.

One might consider that the speed of contraction of domestic sectors justifies protection. On the other hand, protective measures are according to the theory of public choice easy to introduce but difficult to remove. Moreover, protection leads to consumer losses that "in total will almost always outweigh producers gains"[7] at least in the long-run.

Growth Potential and Future Wealth

A trade deficit allows a society to consume more than it produces at the expense of its heirs, since a trade deficit requires that a country "sells off the assets that determine its future standard of living"[8] the possible future growth of the domestic economy declines.

The European long-run growth potential is further weakened if the key for future growth lies in the high-technology industries. Particularly in these so-called strategic industries the Europeans seem to be unable to cope with the Japanese, at least at present. While the EU typically sustains the largest trade deficits with Japan in sophisticated products, the EU recorded the biggest surplus with Japan in 1992 for clothing and accessories.

CONCLUSION

It is not the extent of imbalance which implies economic problems for the European Union but the highly concentrated composition of imports from Japan. Therefore, the EU member states which are most affected are those in which sectors which face pressure from Japanese import penetration are located (several countries, for example, do not have a car industry). A removal of the Union's protectionist measures would strengthen the pressure.

Japan is possibly more protectionist than the EU. But this is clearly not sufficient explanation for the trade imbalance. There are reasons to believe that Europe, compared with Japan, suffers from a significant lack of competitiveness. The EU has tried to remove pressure from the penetrated sectors with protectionist measures as well as, the introduction of the Internal Market. On economic grounds, however, import penetration is clearly an inferior solution to an increase in exports to Japan.

Bibliography

Commission of the European Communities (1993) *European Economy* No. 54.

Conrad, K (1989) "Productivity and Cost Gaps in Manufacturing Industries in US, Japan and Germany" in the *European Economic Review,* Vol. 33

Der Spiegel, (Feb 14th, 1994) *Billiger und Besser*

Eurostat, (1993) *External Trade and Balance of Payments,* Rapid Report, No.6

Eurostat, (1993) *External Trade and Balance of Payments,* Statistical Yearbook.

Frey, B (1984) *International Political Economics*.

Greenaway, D & Hindley, B (1985) *What Britain Pays for Voluntary Exports Restraints*

Hanabusa, M (1979) *Trade Problems between Japan and Western Europe*

McAleese, D (1990) "External Trade Policy" in El-Agraa, A. *The Economics of the European Community.*

OECD, (1993) *Obstacles to Trade and Competition*

Smith, A & Venables, A (1991) "Counting the Cost of Voluntary Export Restraints in the European Car Market" in Helpman, Elhanan & Razin, *International Trade & Trade Policy*

Suzuki, Y (1994) "Bright Spots in the Gloom," in the Financial Times, January 5th 1994.

The Economist,

* June 12th 1993

* January 30th 1993

* July 18th 1981.

Thomsen, S (1993) *Japanese Direct Investment in the European Community: The Product Cycle Revisited*

Thurow, L (1992) *Head to Head. The Coming Economic Battle among Japan, Europe, and America.*

Tyson & D'Andrea (1992) *Who's Bashing Whom? Trade Conflicts in High-Technology Industries*

Winters, A (1992) *Goals and Own Goals in European Trade Policy*

Customs Unions

Alexandra Berndt

Junior Sophister

A customs union is an association of two or more countries to encourage trade. The countries making such an arrangement agree to eliminate tariffs and other restrictive regulations on trade among them. It is a discriminating trade arrangement since the liberalisation only includes the countries that are members of the customs union and they formulate and administer a common foreign trade policy in regard to tariffs and other trade restrictions against third countries.

The best-known customs unions have included the Zollverein, Benelux and the EEC, now called the EU. The Zollverein was formed by German states in the 1830's. These states became the German nation in 1871. Belgium, the Netherlands and Luxembourg established Benelux in the 1940's. Belgium, France, Italy, Luxembourg, the Netherlands and West Germany set up the EU in 1957.

Customs union theory builds on relatively strict assumptions such as perfect competition in commodity and factor markets and hence it is often referred to as orthodox customs union theory. It also only deals with the static welfare effects of a customs union. It has both positive and negative welfare effects, compared to a situation in which every member state practices protectionism. Therefore no conclusion can be drawn in advance as to the net welfare result of a customs union. It must be mentioned that some alternative theories have appeared which try to give an economic explanation of protectionism including customs unions but these shall not be discussed here.

In this essay I will start by looking at a small country in a small customs union. I will assess its effects on trade creation, trade diversion and trade expansion. This will be followed by the incidence of positive and negative effects of a customs union. I will then proceed to look at a large customs union in the world economy and link it to the terms of trade. Long term restructuring effects will be discussed, then the effects of 1992, and finally I will assess customs unions in relation to the EU and GATT.

A Small Country in a Small Customs Union

As already mentioned, the term `orthodox customs union theory' is due to the relatively strict assumptions of this theory, i.e. perfect competition in the commodity market and factor markets, perfect factor mobility within individual countries but not among the countries, foreign trade equilibrium and full employment. The opportunity cost in production is reflected in the relative commodity prices in each country and transport costs are not included since tariffs are assumed to be the only kind of international trade barrier.

Here I will look at a partial equilibrium model for a good X. It is a situation in which the small home country (H) forms a customs union with a small partner country (P). dH and sH indicate the home country's demand and supply of good X. sP and sW indicate the perfectly elastic supply by the partner country and the world.

Initially H imposes a non-discriminating specific tariff (T) on its imports, implying that the import supply curve is sW + T. Country H now consumes quantity O(1), where quantity O(2) is supplied by the home producers and quantity (2)(1) imported from the world producers. The creation of the customs union between country H and P will mean that the supply curve of country H will be sP. The price in the home market then decreases to pP. From this it can be seen that domestic production decreases from O(2) to O(4) while domestic consumption increases from O(1) to O(3). The imports of country H now also increase to (4)(3) and are no longer bought from the world producer but from the partner country at a price of pP.

Effects on Trade

To discover whether this new situation of a customs union has led to trade creation, trade diversion, trade expansion or a combination of the three, it is necessary to define them. Trade creation will occur when there

is a shift from a higher cost to a lower cost producer, i.e. in country H demand will shift from the expensive protected domestic product to the cheaper product from the partner country, implying a shift from a less efficient to a more efficient producer. This is the case in the diagram and is shown by the area (4)(2). Trade diversion will occur when imports from the efficient or cheap world producer are replaced with imports from a less efficient and higher cost partner country. That country's product can be sold more cheaply in the home country than the world's products because the customs union imposes a protective tariff on the imports from the world, while leaving the imports of the partner country tariff free. This is also the case in our diagram and can be seen by the change from (2)(1) to (4)(3). Finally, trade expansion will occur if the lower market price in country H stimulates total domestic demand which will be satisfied by foreign trade. In our diagram the increase in total consumption can be seen by the area (1)(3).

Welfare Effects

The total welfare effects to H are also illustrated in this diagram. The increase in consumer surplus is equal to the areas (1)+(2)+(3)+(4). The decrease in producer surplus is area (1). The tariff revenue equal to areas (3)+(5) disappears. Areas (1) and (3) are internal redistribution from domestic producers and the government to the consumers. Hence the trade creation effect of the customs union is the sum of (2) and (4), and the trade diversion effect is equal to area (5). The question of whether the customs union leads to a welfare gain or loss cannot be answered unless we have further information about the relative sizes of the areas (2), (4) and (5). It can be said with accuracy that the larger the price elasticity and the difference between pH and pP, the larger the gain in welfare resulting from the trade creation effect and the larger the difference between pP and pW, the larger will be the size of the trade diversion effect of a customs union.

Determinants of gains from Customs Unions

Various factors exist which influence the occurrence of negative and positive effects of a customs union. I will mention five of the most important, the first being the production structure. Two countries can be complementary or competitive. If either country is a potential competitor of the other, specialisation in the products which either country can make best and cheapest is probable, and the advantages of a customs union are likely to be relatively important. The opposite is true if the production structures are complementary. The second factor is the size of the union. The more and the larger the countries participating in the customs union, the larger is its share in the total world trade and the smaller the risk of trade diversion. The third factor has to do with the level of the tariffs. If the initial tariffs of the trade partners are higher, the inefficiencies will probably be worse and the welfare effects of the abolition of the tariffs will be greater. Also the introduction of high tariffs against the world producers will reduce the positive effects. The fourth factor is transportation and transaction costs. For increased trade we will need efficient transport systems, the lack of which will replace tariffs as an obstacle for further specialisation. Clerical procedures at the frontier and linguistic differences in Europe also tend to make transaction costs higher. Finally, the advantages of forming a customs union are greater if member countries can respond more flexibly to new prospects.

A large Customs Union in the World Economy and Terms of Trade

Assuming the customs union is large there are two considerable implications for the total welfare of the customs union. The larger the customs union, the larger the possibility that the most efficient producers of various goods will be inside the customs union and hence, the smaller the potential for trade diverting effects. Secondly, when the large customs union fixes it's common external tariff rate then the possibility of it affecting its external terms of trade is increased and thus it can obtain an additional welfare gain.

So far I have assumed that the terms of trade relative to world producers will not be affected by the creation of a customs union but for a large customs union like the EU this assumption is not appropriate. For a large customs union a general tariff imposed on imports can lead to gains in the terms of trade which exceed the negative welfare effects resulting from a decrease in imports to the benefit of the domestic production. This is the rationale for the optimal tariff rates. I will illustrate this in a diagram.

In the initial situation the member states practice free trade on their own. Initial equilibrium is where the demand curve for imports of the potential customs union cuts the world export supply curve; imports being O(1) and price being pWo. With the formation of the customs union the optimal volume of imports from the world producers will be O(2) where the marginal cost of imports (sW) equal the marginal import utility measured along the demand curve for imports dCU. The optimal tariff rate for the union is therefore tCU.

The creation of this customs union leads to an increase in the home market price to pCU and a decrease in the world price of good X from pWo to pW1.

There is a reduction in the import consumer surplus inside the customs union (area (1)+(2)) and the export producer surplus of the world is reduced by area (3)+(4). The total tariff revenue (1)+(3) accrues to the customs union. In this case the customs union has a trade diverting effect and there is a global efficiency and welfare loss of areas (2)+(4). Since the tariff rate tCU is optimal the customs union achieves a net welfare gain since efficiency loss (2) is less than the welfare gain resulting from the improvement in the terms of trade on the quantity of imports O(2), i.e. area (3).

Long Term Restructuring Effects

Restructuring or dynamic effects occur with the creation of a customs union because firms, workers and governments react to new situations and adapt the structure of production and the economy. Firms faced with increased competition will try to lower their costs to stay in the market and increased technical efficiency due to increased competition can have a welfare effect, exceeding many times the limited static effect. An establishment which can produce larger quantities cheaper than smaller ones and is constrained in its outlets by a market of limited size, would profit from the extension of the market, for example, by a customs union. The justification for the creation of a customs union on the point of economies of scale depends on the net effect for the entire customs union and their division between the partners.

Advantages of a customs union internal to the company depend on the size of the company, its growth rate and its learning curve. The larger the company the more efficient is its production and the stronger is its negotiating position. They are also more able to build up stable market positions in export countries. The growth rate of companies tends to have a positive effect on efficiency. Fast-growing firms have the most up to date machinery etc. but they tend to be less flexible in their response to entirely new markets. The learning curve indicates that companies learn to produce more efficiently by the production of greater numbers. Expansion permits producers to offer products of higher quality that are better adapted to specific consumer needs and demand will increase. Also, when a customs union puts a company in a better position, the positive influence is not confined to that company but extends to all related suppliers and buyers. That effect will be greater the better the various parts of the economy are equipped to respond to the impulse.

As barriers such as tariffs, quotas etc. are eliminated, domestic producers have to reduce their price to the level of the partner country. Excess profits will disappear and inefficiencies like overstaffing will have to be reduced. Consumers gain from these price reductions as they obtain more goods at lower prices and producers offset the loss by price reductions.

High Stakes for Europe: The 1992 Challenge

In the integrated Community market post-1992 a dramatically new environment awaits consumers and producers alike. The removal of a whole range of non-tariff barriers, i.e. government protection in procurement markets and a plethora of differing product standards leads to an immediate downward impact on costs. More substantial gains will be generated by completion of the EU internal market. There will be a new and pervasive competitive climate and firms can exploit new opportunities and make better use of available resources.

There are four major consequences which are expected from the combined impact of the removal of barriers and the subsequent boost to competition:

* a significant reduction in costs through the reorganisation of business and economies of scale

* improved efficiency within companies due to the downward pressure on costs due to more competitive markets

* new patterns of competition since real comparative advantages will play a determining role in market success

* increased innovation because new business products will be generated by the dynamics of the internal market.

These effects will be spread over differing time spans but the overall effect will be an increase in the competitiveness of business and the general economic welfare of the consumer.

The consumer will no longer be confronted with enormous price differences depending on their country of residence, as is the case in today's Community. Due to the reduction in costs, the level of this price will be on the downward journey. The consumer will also be faced with a wider choice as a result of market integration and increased competition leading to differentiating products as well as economies of scale.

1992 has led to the end of firms relying on the national soft option. Those who are able to scale up their performance to the demands of increased competition, will have an outlook for sales and profits which is dynamic but for others profits will be clearly squeezed by Europe's competitive renewal.

Strengthening European competitivity leads to the reconquest of the European market, but failure to do so will not mean that the challenges of the European market will not be mastered. They will, but not by the Europeans.

EU, GATT and Customs Unions

The basis of the EU according to Article IX of the Treaty of Rome is a customs union. Articles XII to XXIX give a detailed description of phasing out the internal tariff rates and establishing common external tariffs. In July 1968 a customs union for industrial goods had been realised. The idea of the customs union was to establish a totally free internal commodity market in the EU. This goal has still not been fully realised due to the use of non-tariff trade barriers such as technical trade barriers, government subsidies, etc. which became increasingly important in the 1970s.

Customs unions are discriminating trade arrangements and hence violate the rules of GATT. Under GATT's `principle of most favoured nation' member countries have to give each other the same favourable treatment that they give to any other country. Customs unions between a limited number of countries is a clear violation of the principle, but Article XXIV of the GATT treaty gives the right to form regional customs unions if certain conditions are satisfied because the overall aim of GATT is to promote international trade. When GATT was created in 1947, it was the widespread belief that customs unions were a step closer to free trade and it was not until later that it became clear that customs unions could in fact be a form of protectionism.

Conclusion

The creation of a customs union has some positive and some negative welfare effects. It can only be well founded in economic terms if the former exceeds the latter. The welfare effects of the customs union as a whole are uncertain. Only if it is possible for a customs union to affect the external terms of trade through the optimal tariff is it possible for the union to achieve a gain in net welfare.

In relation to the short term effects which affect consumers, producers and governments, customs unions tend to have more positive effects as production structures are more competitive, initial tariffs are higher and also, as customs unions are larger transaction costs are lower. Competition and economies of scale are long term effects and are better reasons for creating customs unions.

Bibliography

Cecchini, P (1988) *The European Challenge 1992*

El Agraa, A (1981) *Theory of Customs Unions*

Henderson, W (1959) *The Zollverein*

Molle, W (1990) *The Economics of European Integration*

Nielson, H (1992) *An Economic Analysis of the EC*

Palmer, J (1989) *1992 and Beyond*

Yannopoulos, G (1988) *Customs Unions and Trade Conflicts***GATT and the EC: The case of Agriculture**

The Economic Legacy of Ceausescu

Nicolas Holman

Senior Sophister

Romania is an important country to study for many reasons. The fact that it is an example of a command economy which took Stalinism to its absolute extremes without any previous attempts at reform makes it an ideal case study for transition economists and theorists. Furthermore, it is also a country which had no national debt immediately after the revolution yet seems to have received no economic benefits from this fact. Finally, it is adopting the gradualist approach to transition as opposed to the ÒBig BangÓ approaches of Bulgaria and Poland and thus the effectiveness of this approach remains to be seen.

However, to be able to understand the current Romanian economy fully, it is important to understand the huge personal impact that Ceausescu[[Otilde]]s ideals had on his country. For example, the desire that he had for Romania to be completely independent of the West caused him to adopt a policy of depressing consumption to pay off the foreign debt which resulted in a huge decline in living standards and eventually sowed the seeds for revolution. This paper, therefore, seeks to outline the key features of the Ceausescu regime[[Otilde]]s policies before discussing the relevance of these to current economic decision making in Romania.

The Ceausescu Regime

Although originally regarded by the West as a sort of maverick Communist, due to his opposition to the 1968 Soviet invasion of Czechoslovakia, it soon became apparent that he was one of the most brutal leaders that Europe has ever seen through his attempts to personally dominate the Romanian people.

Nevertheless, Ceausescu was actually very consistent in his policies. His unyielding Stalinist ideology to create a `homogenous Socialist population out of the traditionally peasant based Romanian people' was what caused Romania to be in the situation that it is today. To achieve his aims, the population would have to be subjected to his control - most easily achieved if they were contained in large urban centres.

The economic consequences of such a policy was an all out drive to create a heavy industrial base in Romania and a determination to make Romania self sufficient through the elimination of its foreign debt. A policy of `systemisation' was also proposed in which the rural population was to be moved to larger urban centres - but this was later abandoned.

Initially, the development strategy was very successful, as vast pools of underutilised labour in agriculture was mobilised for industry with the proportion of the non-agricultural labour force increasing from 30.3% in 1956 to 63.5% in 1977. However, this growth was not sustainable - being based on structural shifts - and soon the labour force was faced with inadequate employment and income opportunities with a reduced supply of food and other consumer goods. However, Ceausescu[[Otilde]]s ideological inflexibility allowed for no changes in his policy and the regime resorted to coercion to achieve the production targets which enterprise managers were then forced to fabricate.

The effects of this flawed system soon became apparent as the benefits from the labour force shift were reduced. Economic growth fell from 10% in the early 1970[[Otilde]]s to 3% in 1980 with food and other consumer goods becoming very short in supply.

The situation was worsened by the energy crisis of the early 1980[[Otilde]]s. Despite the fact that Romania had one of the best endowments of natural resources in Europe and in 1985 produced more electricity per capita than Spain, Italy or Portugal (excluding imports of electricity) the streets were dark and people died from the cold in their homes. In 1989, household consumption of electricity accounted for only 5.1% of the total.

This crisis was a consequence of the continuation of the drive towards heavy industry, making Romania a net importer of electricity at a time when World energy prices were at a peak. When the crisis became apparent, instead of curbing this drive, it was, in fact speeded up - hitting the population more.

Apart from the obvious social consequences of this policy, the demand for many household appliances fell sharply while light industry was often forced to run at less than full capacity. Agriculture was also affected as a high proportion of irrigation couldn[[Otilde]]t be used.

Yet Ceausescu[[Otilde]]s vision was not swayed and he turned to even more ambitious projects such as the Danube-Black Sea canal and the Bucharest civic centre - to create a capital worthy of the New Socialist Man. All this while his people starved and sat in the dark at night.

Ceausescu's policies soon became seen for what they were as the economy became unable to meet the basic needs of the population such as food and health care. In 1981, rationing was reintroduced as agricultural output fell sharply. This was mainly due to unrealistic food prices and agricultural wages that were too low. As a consequence, the private sector became more important for providing food and by 1985 it accounted for over half the total production of milk, wool, eggs and potatoes despite only holding 13% of arable land.

Ceausescu[[Otilde]]s response? The `new agrarian revolution'. A plan which aimed to incorporate the agricultural sector properly into the plan. Economically it made no sense as farmers were now forced to deliver quotas of food to the state at a third of the market price - with these quotas being even higher for private farms. University graduates were forced to spend several years in the countryside and the introduction of the `closed town' scheme made it very difficult to get back to the cities.

To worsen the situation, in 1988 Ceausescu announced that systemisation was being reintroduced. The idea behind this was to structure all villages into a well-defined hierarchy with some 300-500 villages selected for promotion into modern towns while another 6, 500 villages were to be phased out entirely.

It can be seen, therefore, that by this time Ceausescu had completely lost control of himself. He was now insisting on visiting factories himself and choosing their production targets according to his own personal whims. In other words, by 1989 he had succeeded in subordinating the entire Romanian economy to his own personal `off the wall' ideologies. Happily, Ceausescu was deposed before either the New Agrarian revolution or systemisation could do damage but it is clear what their effects would have been. Firstly, the agricultural production base would have been eroded further while removing any form of production incentives from the farmers. Furthermore, the systemisation imposed rural-urban migration would have reduced the supply of labour further while alienating the Romanian peasantry even more. Ironically the net result of this would have been to increase the population[[Otilde]]s dependence further on private production - something which Ceausescu abhorred.

Eventually, in 1989 as GNP fell by 10%, shortages became more pronounced and unemployment began to rise, Ceausescu was deposed. The economy was in a shambles. The entire information base for planning had been eroded and physical infrastructure was severely run down - despite high government investment rates. The actual physical well-being of the population had deteriorated through malnutrition, pollution and the decline in the health sector. Furthermore, the fall in investment in education had been so severe that the numbers enrolled in University education had actually fallen despite the increases in the size of the 18-22 age bracket.

Thus, Ceausescu left Romania with an ineffective, underproductive industrial base, a highly weakened, unmotivated agricultural sector and an unhealthy, unproductive population. It is this legacy which provides the foundation for the `basic needs' approach to the Romanian economic programme. Many economic resources have been diverted back to the population with the net effect of providing confidence in the economy and through the increased supply of food, reducing the need to hoard.

Nevertheless, it would be wrong to think that all of economic problems have been resolved. Although the population has a sense of real improvement, the macroeconomic indicators show a bleak situation. Output has plummeted, labour productivity has fallen and hard currency imports have risen significantly.

Much of this decline is merely due to the adjustment process that is necessary in economic transition. However, it would be wrong to underestimate the extent to which structural factors such as oversized

industries, price controls and highly inefficient managers have contributed to the decline. It is these factors for which the Ceausescu regime is responsible and it is these distortions which will prove the greatest challenge to reformers.

The dire structural damage that Ceausescu had done to the Romanian economy, in addition to the massive decline in living standards that had occurred, posed a dilemma for the post revolution government. It was clear that something had to be done quickly if the economy was to have any success in recovery. However, due to the hardships that had already been suffered the social costs of any reform programme had to be minimal. Thus, a policy of `gradual reform at a rapid pace' was adopted. The key features of this policy were as follows: firstly, introduce market forces into economic decision making; secondly, privatise State Owned Enterprises, housing and land; and thirdly, reduce the role of the government in economic decision making.

The crucial need to raise living standards, however, has severely curtailed the government in achieving these objectives. It is politically difficult to impose positive real interest rates on the population, and any attempts so far to do this have resulted in debts being reneged upon. Clearly, this has made monetary policy very difficult and, so far, ineffective. On the fiscal policy side, despite plummeting revenues it has been difficult to curtail expenditure on such items as food subsidies and, of course, social welfare benefits. Finally, on the privatisation front, many of the state owned enterprises are so hugely inefficient with little prospect of profits that nobody wishes to buy them. Nevertheless, it is difficult to shut them down as often whole towns depend on a single firm for employment. This of course raises many questions about the whole privatisation process. Up to now it has been an elaborate charade with a series of mutual funds being established to manage the newly privatised enterprises. As expected, few firms have actually been sold.

It is clear, therefore, that although the Romanian economy is slowly reforming, the structural distortions are deeply embedded in the economy. Ceausescu[[Otilde]]s attempts to run Romania as a firm failed because he neglected his own people for the purposes of gaining international prestige. Furthermore, the `export at all costs' policy resulted in Romania[[Otilde]]s natural resource base being eroded significantly. Thus, where Romania was once self sufficient in energy, it is now a net importer; something which could prove to be disastrous if World energy prices rise.

In conclusion, therefore, Romania under Ceausescu will go down in the history books as a sort of economic curio; a failed social experiment. However to the Romanian people his policies were very real and his legacy will undoubtdedly affect many generations to come.

Bibliography

Ronnas, P (1991) "The Economic Legacy Of CeausescuÓ in *Economic Change in the Balkan States*

Demekas, D, Borensztei, E & Ostry, J (1993) "An Empirical Analysis of the Output Declines in Three Eastern European CountriesÓ in *IMF Staff Papers,* March 1993.

Earle, J, Sapatoru, D (1992) "Privatisation in a Hypercentralized Economy: the case of RomaniaÓ in *Privatisation in the Transition to A Market Economy*

Teodorescu, A (1991) "The Future of a Failure: the Romanian EconomyÓ in *Economic Change in The Balkan States*

OECD, (1993) *Romania: An Economic Assessment*

REGULATION IN IRELAND IN A EUROPEAN CONTEXT

Martha OHagan
Junior Sophister

In an era of high budget deficits and high tax rates, regulation as a cost-efficient means of government intervention is both advisable and necessary. In many ways as effective as other means of intervention, such as government expenditure, its advantages are being praised at a time when there are many calls for cuts in public spending. òGovernments need not and should not be idle. Choosing the right policies is a matter of cost-effectiveness. Do what costs little and is sure to work well, avoid what costs a lot and probably will not work at all [1]. Le Grand et al describe regulation as follows; òGovernment regulation is an alternative to provision to the problems caused by the market of social care by imperfect information and irrationality. Instead of replacing the market providers by government ones, the market providers could be strictly controlled by a system of monitoring undertaken by field social workers and by inspectors[2]. Irelandõs integration into Europe affects the role of regulators. With free movement of goods, services, capital and people between member states of the EC, foreign ownership and pollution on an international scale, European co-operation is needed, to fulfil the role that previously the Irish government regulators fulfilled on a national scale.

This essay firstly analyses why the need for regulation arises. Secondly, regulation is analysed in relation to market failures and primary policy objectives of the government; economic growth and equality. In conclusion I will evaluate regulation as a means of intervention in Ireland in the 1990s, in light especially of European integration.

Why intervention?

The market price mechanism operates on two principles. (i) A consumer reveals preferences for a good or service by buying it. (ii) The benefits of a good or service are incurred by its consumer or producer only. The existence of externalities, public goods and imperfect information leads to market failures. The market also assumes perfectly competitive behaviour which operates on two assumptions. Firms are profit maximisers and consumers are utility maximisers and both have perfect information. This should ensure an efficient outcome and is referred to as Adam Smithõs ôinvisible handõ. Non-competitive behaviour such as monopolies or oligopolies cause the assumption of an efficient outcome to fail and government intervention is needed. Where market failures or an inefficient outcome occurs, consumer welfare may not be maximised. The ideal aim of a government is to maximise its peopleõs welfare, therefore intervention is justified. Apart from efficiency, income redistribution is also an objective. The government can regulate the price or quantity of a good or service and entry into markets. Regulation in Ireland is used to regulate monopolies (to ensure competitive behaviour for allocative efficiency), to ensure equality (a primary policy objective), to correct externalities, to provide information and to stop ôexcessiveõ competition.

Correcting market failures

Examples of market failures are the existence of externalities and imperfect information. The costs and benefits of goods and services are often not restricted to the consumer or producer. In an unregulated market negative externalities such as pollution tend to be produced. Although some pollution cannot be avoided if we want economic growth, regulation is necessary to control the amount. The optimal level of pollution is difficult to determine. Usually it is reached by setting uniform standards across firms, which may not be the cheapest way of reaching the given level. Firms who find it cheaper to abate pollution will have no incentive to reduce pollution below the level set by the standard. The two main problems concerning the regulation of pollution are lack of information and regulator ôcaptureõ (bias in favour of other objectives). Regulator capture expresses the idea that, if companies are publicly owned, politicians are also responsible for the performance of the firm and may therefore be more concerned with keeping down costs rather than

increasing them to maintain environmental standards. If the firm is privatised, regulators will only be concerned with quality control. However, the information regarding the firmõs activities may be imperfect, posing problems for regulators.

Information is like a public good: supplying an extra person costs next to nothing, yet its acquisition by a single person would cost a huge amount. Regulations for health and safety are needed due to imperfect information about the possible consequences of consuming a good. Health warnings on cigarettes and safety belts in cars are due to regulation. Regulation also ensures that drugs are stringently tested before being marketed. However, this may push up the price of drugs (increasing costs of the producing firm), making them inaccessible to those on lower incomes. Overregulation may discourage firms from research which could potentially be very beneficial to society. Also, regulating what people consume impinges on their personal freedom to make decisions concerning their welfare. Some argue that information concerning the potential consequences of certain drugs be provided but the drugs be made available. òThe urgency of AIDS has changed the ways in which drugs are tried and approved....For those with fatal illnesses, a succession of inconclusive trials and the associated delays in approval may mean that drugs which could help are not available. Moreover, tight controls on those enrolled for trials means that many who want to try the drug often cannot[3]. Because of illnesses such as AIDS, the role of regulators must be revised. Those who contract the disease are not likely to thank the government for regulating a potentially beneficial drug because it has not been tested over a five-year period. Its urgency changes the argument for government regulation.

Policy Objectives :

(i) Growth

Being a small open economy with few natural resources to speak of, we rely heavily on trade. Without imported inputs no output can be produced and without output to export no foreign exchange can be earned in order to pay for imports for domestic consumption. Therefore Irelandõs economic activity relies heavily on the traded sector. The output of an economy depends on productivity, human capital, natural resources and technology. In order to grow, an economy must experience an increase in any of these. The key to this is gross investment, which is therefore also the key to growth. Thus in order to evaluate the determinants of growth, it is necessary to investigate what determines investment in a small open economy. Whereas in a closed economy, domestic savings determine domestic investment, there is no such relationship in Ireland given the integration of our capital markets into the world markets. Investment in Ireland is a function of world investment and of the k-variable, our attractiveness to foreign investors relative to other countries. World investment is exogenously determined, therefore the k-variable is the means by which we can influence our growth rate. Because the k-variable depends largely on factors in the non-traded sector, inefficiencies there are of great relevance to the Irish economy.

The importance of the non-traded sector, rests on the determination of the k-variable (infrastructure, legal costs, telephone costs etc.), input costs for the traded sector (with given prices and wages, lower costs increase the profits and the competitive position of the traded sector) and the percentage of GNP spent in the non-traded sector. As prices of traded goods are given exogenously, lowering the prices of non-traded goods can directly affect living standards. In 1992, 60% of GNP was spent on non-traded goods and services. Because the non-traded sector is not open to international competition, monopoly power causes inefficiency and high prices. Therefore regulating the non-traded sector will be of great relevance to the growth of the economy and to consumer welfare.

Two types of monopoly exist in the non-traded sector. Firstly, economies of scale and natural monopolies have led to the provision of public utilities such as water, gas, telephones to be undertaken by monopolies. Secondly, in order to protect the public from inadequate services, for which information would be difficult to obtain, barriers such as length of training or qualification licenses have restricted competition in services, banking and insurance. In order to ensure competitive quantity and price these monopolies must be regulated. òIt has been argued that the medical professionõs power over the length of training has been used to limit the

supply of doctors (and hence raise their income) rather than to protect the public from poor quality treatment[4]. Thus, further regulation is needed to reduce doctorsõ fees and to increase the supply of doctors.

Furthermore, other non-competitive behaviour may include ôexcessiveõ competition, where large firms operate below cost price in order to push smaller firms out of the market, creating their own monopoly. Hence, improving efficiency of the non-traded sector is essential for the competitive position of the traded sector, for living standards, for the attractiveness of Ireland as a location for investment and thus for the growth of the economy. òThose who call for an industrial policy to help manufacturing are missing the point. If services account for over half the sales price of the good, then improving efficiency there will help to trim the cost of the final product and thus become more competitive, rather than tinkering with the production process [5] However, regulation has costs and can cause inefficiencies which I will discuss later.

(ii) Equality

There are two definitions of equality, equality of opportunity, where everyone has equal access to all markets, and equality of outcome, where everyone has an equal level of utility. Using income as a crude measure of utility, equality of outcome will be measured by the extent to which income is evenly distributed among the population. A combination of the two is desirable. Where one exists without the other, inequality still occurs, e.g. equality of opportunity may result in gross inequality of outcome and vice versa. Equality, being a social issue, often conflicts with economic issues.

Equality of opportunity ensures equal access for everyone to all markets. In the labour market regulation ensures equal pay for equal work and makes discrimination illegal. Firms may be required to hire a certain quota from specified minorities. However an article in *The Economist* entitled Discrimination and the market[6] points out the costs and problems of regulating market access. In the case of banks, insurance companies and other financial institutions, problems arise. In England a bank charging unmarried couples higher mortgage rates than married couples, because the former were likelier to default, provoked outrage. Similarly, banks are reluctant to lend to inner city areas because of the risk involved. Should health insurance companies offer the same cover to those whose families have a history of a hereditary disease or to homosexuals (because of the risk of AIDS) as to everyone else? Apparently so. However covering the costs of these risks cripples some financial institutions. Often the government must support financially troubled banks who offer equality in lending. Therefore equality of opportunity imposes a cost on the taxpayer, a cost that they should know about. Other costs of regulation include the expenses of the regulatory body and its staff.

The two main regulatory means for improving equality of outcome are the minimum wage law and rent control. In the labour market, supply and demand equate at the equilibrium wage. Imposing a minimum wage may have distortionary effects. If, however, the minimum wage exceeds the equilibrium wage, workers may be priced out of a job. Costs accruing to the employer arise. These costs may be passed on to the consumer by increasing prices but unless there is inelastic demand for the good, quantity demanded will fall. Producers may cut back on output or employment or replace labour with machinery. Those who keep their jobs gain at the expense of those who lose jobs. If wage differentials within employment are maintained, then equality within employment will not increase. Therefore minimum wage laws are an imperfect method of achieving equality. Some argue, however, that those on lowest wages are teenagers on a part-time temporary basis. Therefore unskilled, manual workers may gain at their expense (assuming they are not priced out of the labour market).

Rent control, or the imposition of a rent ceiling, transfers money from landlords to tenants. If demand for housing increases, rent may rise to what the government may consider an unacceptable level. Imposing a rent ceiling will reduce the number of houses rented, so that those lucky enough to obtain accommodation will gain at the expense of those who are not. Lower rents, however will make investment in housing a less profitable venture and supply may decrease, reducing the supply of housing at the rent ceiling even further. Therefore, more tenants will be worse off than before. In addition, it is not always the case that landlords are among the wealthy in society. Surveys have shown that often low earners or manual workers own property, and reducing their income does not reduce inequality.

Regulation on grounds of equality is sometimes ineffective and costly. Uninformed regulation such as imposing an excessive minimum wage or imposing laws which cripple insurance companies can have serious

consequences.

The effect of Europe on regulation

Many believe that a large service sector means slower growth in an economy, because of its flagging productivity. European integration has opened up trade possibilities, making many goods and services tradable which were previously non-tradable (i.e. opening them to international competition). This is of great relevance in Ireland as services will be forced to become more efficient, in particular banks and insurance companies. òCocooned for years by restrictive practices, liberalisation and deregulation have opened services to competition. Privatisation is having a much bigger effect on services than on manufacturing, forcing airlines, banks and telecom firms to become more efficient[7]. By deregulating some services, they will be forced to act more competitively. These services are not only vital for the k-variable and for the traded sector, but they themselves are becoming tradable. Therefore there is less of a role for regulation, as monopolies will be forced to act as competitive firms Indeed some deregulation may be beneficial.

Externalities, especially pollution, are no longer simply on a national scale. The English nuclear plant, Sellafield, pollutes the Irish sea and environment, yet we are powerless to regulate its emissions. A European central regulatory board is however needed to regulate on a European level. Undertaking measures concerning the ozone layer would be useless without international co-operation. However, problems will be faced: European authorities will not have adequate information or be perfectly impartial in regulating pollution levels. Also, setting international standards will mean inefficiencies on a large scale. However, with international externalities, since we are useless in regulation alone, Europe must take over the role to some extent.

Since January 1993, customs barriers have been removed. Free movement of goods and services means that safety and health standards are an international issue. The rule at the moment is, however, that if an international standard has not been set for a good, the standard set by the country in which the good is produced is the one used throughout Europe. Ultimately standards for all goods will be set at a European level. Imported cars will have to be concurrent with Irish safety regulations. Similarly, laws against drugs are useless if they can be purchased abroad and brought here. Therefore regulations which are set to protect individuals from the potentially harmful consequences of goods due to imperfect information must be set through government co-operation and agreement.

Finally, the European social charter will influence equality measures. Minimum wages and rent ceilings will be surveyed by European boards. The European Commission accepts complaints of unfair discrimination, ensuring equality of opportunity in all countries. Therefore equality regulation will be very much affected by European set standards. òInternational trade barriers have been pushed forward by three main things: falling regulatory barriers to overseas investment, tumbling telecommunication and transport costs and increased mobility in domestic and international capital markets[8].

Deregulation and closer ties with Europe, which in many cases will assume the role of regulator, given the international nature of many factors, is beneficial to Ireland for growth, control of pollution, safety and health standards and for equality. Regulation exclusively on a national level is in some cases no longer effective.

Conclusion

In the 1980s deregulation was all the rage. In the 1990s re-regulators are getting to work throughout the world. However, because of our closer ties with Europe the role of regulator will be increasingly centralised. Regulation may impose costs such as regulation-caused inefficiencies (x-inefficiency). This is particularly evident in Ireland in the professions. Restrictive practices have given professions the position of monopolist, increasing the price and decreasing the supply of their services. Opening services to competition may overcome some of the regulation-caused inefficiencies. General laws, for example concerning nuclear plants throughout Europe will cause large scale inefficiencies, but will also provide a means for countries affected by pollution from other countries to control that pollution.

With changing times and changing situations, the role of regulation changes. Although a potentially very useful and effective means of intervention, its adverse effects can be very serious and costly. When needed,

deregulation must be as readily implemented as regulation. European integration will also change the role of regulators in Ireland, to which they must be quick and willing to respond, in order to maximise the welfare of the individuals in the economy. Regulation, however, is particularly relevant in Ireland because of the constraints placed on government spending because of the national debt. One of the preconditions set by the Maastricht Treaty is the reduction of our debt and no monetary financing of budget deficits. Therefore, to comply with this we must reduce the debt, which reduces expenditure available for government intervention. Although there is still a role for government intervention in Ireland, in the context of Europe it is changing.

Bibliography

Di Cagno, D (1990) *Regulation and Banks behaviour towards risk.*

Le Grand, Propper & Robinson (1992) *The Economics of Social Problems*

OHagan, J (ed. 1991) *The Economy of Ireland*

Spulber, D (1989) *Regulation and Markets*

The Economist - relevant articles as quoted in footnotes

SOCIAL PARTNERSHIP

Muireann Kelliher
Senior Freshman

Discussion of the area of planning and social partnership is particularly relevant at the present time, given the recent Programme for Competitiveness and Work. After the last eight years under corporatism, I feel that the area needs a more central analysis. In doing so, I will discuss the rationale for planning itself and the role for government in the process. Next, I examine the Irish example and the consequences it has had for the economy, in particular the impact on unemployment and our ability to deal with the consequences of international fluctuations. I conclude that wage agreement has had a higher cost than it is worth.

The Role for Planning

It is commonly believed that a private market economy, if left to its own devices, fails to achieve the social optimal rate of economic growth. One contributing factor to the inefficient intertemporal allocation of resources is the great uncertainty concerning future movements in aggregate demand and prospects for different sectors in the economy given the small size of the Irish economy, the above is the most applicable. Hence in theory consensualist policies should be pursued to maximise social welfare. However theory doesnt always hold true. In the face of uncertainty, risky projects will be too heavily discounted and individuals will be reluctant to sacrifice current consumption for the prospect of future growth.

The problem is that the information necessary to make more measured decisions is costly for individuals to obtain in a decentralised economy. Large economies of scale and the existence of public goods suggest a role for government intervention in acquiring and publishing information. The improvement of the performance of the economy through better information is thus one objective of indicative planning. Allied to the notion of planning, is the complementary social partnership concept. This is required because the success or otherwise of a governments intervention strategy depends on the response of the private sector. For example, it is regarded as critical that since wage-setting behaviour is a major determinant of an economys performance, the government should be a part of the process. One mechanism for intervening is an institutional arrangement with employers and trade unions. The consequence of such an agreement is that the requirements of the overall economy become an issue in the labour market. Examples of corporate-style policies like those used in Germany, where centralised bargaining seemed to work well, were often cited. It supposedly delivered social stability and moderate pay rises because unions were forced to take account of the impact of pay rises on employment.

The Recent Irish Experience

Corporatism was revived in Ireland in the 1987-1990 period with the agreement on the Programme for National Recovery (PNR), which was followed between 1990 and 1993 by the Programme for Economic and Social Progress (PESP). Its successor is the newly agreed Programme for Competitiveness and Work. In theory, these agreements were meant to ensure wage moderation and hence make job creation easier. However, increases have been substantially above the rate of inflation, and wage moderation and jobs have been sacrificed for the sake of industrial peace.

There is one very strong objection to wage agreements *per se* irrespective of whether wage moderation is achieved. To the extent that they impose a uniform rate of wage increase on different sectors of the economy, they impair the efficient functioning of the labour market. Wage differentials are compressed irrespective of changes in economic activity. This may be bad for unemployment, for two reasons. Firstly, compression of wage differentials means a reduction in the efficiency of the economy because labour will be suboptimally allocated since workers wages will not be made equal to their marginal product. Workers will not be given the incentive to move to firms where they would be more productive. This welfare loss is likely to lead to job losses somewhere although it is hard to predict exactly how. In addition, there are likely to be direct job losses under the Irish tax system.

Firms where productivity increases less than the average will be paying too much for workers and may lay them off. With the disincentives to hiring extra workers, these jobs are unlikely to be made back by an increase in employment among the firms whose wage bills have grown less than productivity.

Apart from these general problems with wage agreements, there are even more specific problems with them as practised in Ireland. Over the six years covered by the PNR and PESP, unemployment has risen from 215,000 to 290,000. Meanwhile, insiders in the labour market have done very well. The pay rises agreed in the PESP were based on a forecast of an average growth of 2.5% in GNP which did not materialise. This has meant that all gains in the economy have gone into pay rises for those already working and nothing into creating new jobs. Most of the wage increases came in the public sector. During the six years of the PNR and the PESP, public sector pay rose by 46% , while inflation over this period was just 15%. The poor wretches in the private sector, whose taxes paid for this largesse, received 23% over the same period, or 20% if they did not benefit from the local bargaining clause introduced in the PESP. Looking at trade union participation in centralised wage bargaining, one is led to the conclusion that the numerical strength of the public sector unions within the ICTU places them in an advantageous position compared with other groups. They have used this advantage to override the interests of the more heterogeneous private sector membership of the Congress. However the public sector unions have proven themselves adept at gaining not only huge basic increases but also special increases. Interestingly the unemployed are not represented at the discussions.

The success of the public sector unions in getting large wage increases from the government has had severe budgetary implications, since public sector pay is the largest single component of government spending. With the ink barely dry on this years budget statement, the Minister for Finance is already reaching for the Tippex. For the numbers in the 1994 budget do not add up and government spending, the CBD and the EBR will all have to be adjusted upwards. The new public sector pay agreement, it was announced, was an attempt to stop rises in the pay bill due to the special increases which were common under the PESP. So an 8% basic pay rise was granted as well as a further 3% over the duration of the agreement. Apparently productivity increases would be required in return for this 3%, but already the government has given the first 1% automatically. In the light of this performance, further capitulations must be on the cards. The money given in pay increases to the public sector could have been used to begin rectifying our tax distortions in the labour market and so the excessive pay awards must be considered to have had a direct effect on unemployment.

The government has acceded all budgetary power to the trade unions in return for industrial peace. Many of the measures in this years budget were designed specifically to get the trade unions to sign up to an agreement. The tax wedge, for example, was not comprehensively reformed at all and most changes in taxation were given specifically to the union side, such as the abolition of the health and training levies. Meanwhile the employers PRSI contribution on higher-paid workers was increased. The tax wedge was merely rearranged rather than being tackled properly.

Moreover we discovered during the currency crisis that the PESP was much too rigid a framework for the economy during periods of exchange rate volatility. At the time the economy needed a lower real exchange rate, which required flexibility either in the nominal exchange rate or in wage claims. The unions would agree to neither. Devaluation would undermine the low-inflation forecasts on which agreement was based. Ironically we are now making another agreement at a time of exchange-rate uncertainty. The punt is now tracking sterling, not the deutschmark. The British clearly have no objections to a loose monetary policy in search of recovery. Whether the new agreement will be able to cope with this situation is highly debatable.

The main risk of abandoning the consensus approach to wage bargaining is the chance that lower pay rises, especially for the public sector, will provoke industrial unrest. But the risk is probably much smaller now than it was ten or twenty years ago due to changes in the size and structure of the labour force. The fall in strikes during the PNR and PESP was only following an already declining trend and there is no evidence that it was due to these agreements.

In conclusion, the whole Irish approach to wage bargaining borders on the ridiculous. The notion that the government should intervene at all in private sector wage-setting, rather than letting the market sort it out itself, is questionable. Even more dangerous is the idea that the public sector should dictate enormous wage increases of which the private sector bears the brunt, since public sector workers are less vulnerable than anyone in the country to unemployment. In theory partnership is an excellent idea but this idea has been misconstrued in Ireland with a relationship more like that between victim and bully. Given the climate and

expectations that the government have now created, wage agreements have little left to offer except an increasingly wide definition of the victim category.

EMIGRATION FROM IRELAND - VICE OR VIRTUE?

Patricia Broderick
Junior Sophister
This country is too small to support all her people[9]

For too long emigration has been dismissed as a necessary evil; a stop-gap solution to our unemployment problem. Ireland has witnessed the departure of entire generations and yet the pattern of net migration has never been so pronounced as in the last thirty years. In this period, emigration reached levels comparable to those at the time of the Great Famine and also, for the first time, Ireland experienced massive immigration, a phenomenon that has never been repeated. More recently however, the return to high levels of emigration has denuded many rural areas of their young people, and hence their future. I feel it is unrealistic and naive to believe that such dramatic population swings have had no effect on our economy. As such, the purpose of this paper is to examine some of the factors which caused such large migration swings, to assess empirical evidence pertaining to net migration and finally to consider related policy implications and the possible effects on the economy as a whole.

Migration - causes of mass movement?

The causes of migration can be divided into two categories: push factors and pull factors, which cause involuntary and voluntary migration respectively. I have chosen to concentrate predominantly on push factors, as I believe them to be most harmful.

Nonetheless, the determinants of voluntary migration are not to be forgotten. It can, to a large extent, be explained by high wage differentials, better experience and job opportunities, family links abroad, or simply *wanderlust*. These variables are however, hard to quantify and as such, data collection is difficult. I have thus turned my attention to unemployment statistics which, although they have some definitional difficulties, are quite readily available. The chosen push statistics are the unemployment rate in Ireland, the United Kingdom and the rate in the ten remaining EC member countries. It is obvious why the unemployment rate in Ireland should affect migration levels. As the search for work is one of the primary reasons for migration, high levels of unemployment in Ireland will be a disincentive to stay here. By the same token, the United Kingdom and Europe are migration destinations and thus, low unemployment rates in these regions should encourage emigration from Ireland. It could be argued that America and Australia should also be included as large numbers of Irish people have emigrated there in the past, but visa restrictions, and in the case of America, high labour market flexibility, make these figures less sensitive, as statistics show permitted, rather than desired migration. Other possible spurs towards emigration include the structure of our tax system which heavily penalises young, single workers, but again, such influences are hard to measure.

The Model: Individual Regressions

To test the validity of the hypothesis, I regressed the dependent variable (net Irish migration) on the independent variables of Irish, British and EC unemployment rates (where the X variables are the percentages of unemployment in the respective labour forces). Y represents net migration, which is usually negative in Ireland, thus representing emigration as a negative figure. I expected Irish unemployment to have a negative relationship with the Y variable, i.e. as unemployment rose, net migration would decrease, indicating a rise in emigration. I expected the opposite to be true of the other two variables, as high unemployment in emigration destinations would encourage immigration to increase, thus increasing the Y variable.

When net migration was regressed on each of the independent variables, the results were quite surprising. The R-squared values in each case were lower than I expected, although the F-statistics indicated that some

relationship (however slight) existed. Relatively high t-statistics (arbitrarily choosing the absolute value of 1.8 as my reference point) indicate that the models are quite a good fit, but although the nature of the relationship between Y and unemployment in Ireland was as expected (i.e. negative), I found, contrary to expectations, a negative relationship between Y and my other two variables. One possible explanation for this is that most emigrants to Britain and Europe are highly qualified and thus, in times of recession, it is the indigenous unskilled workers who suffer the brunt of the unemployment rises, making migration less sensitive to changes in the unemployment rate in these countries. It is also true that rising unemployment in the rest of the world will be reflected in the Irish economy. Thus the situation abroad may still be better than in Ireland, which implies that in spite of rising unemployment abroad, migration outflows from may still persist.

Multiple Regressions

The multiple regressions had a slightly higher R-squared value:

R-squared = 0.45312

Y = 19656 - 3061(X1) + 6495(X2) - 7123(X3)

but low F-statistics suggest that there is probably only a weak relationship between the variables and net migration. I re-ran the multiple regression and included a dummy variable. In doing so, I hoped to achieve a weighted regression to show that there has been a shift in emphasis in the importance of the United Kingdom, relative to the other EC member states, as an emigration destination. I arbitrarily selected 1984 as the first year this change occurred. Since Irelands entry into the EC in 1973, the importance placed on learning foreign languages has increased. Indeed, further integration within Europe through the Single European Act and the Maastricht Treaty will augment this. Although my results show no conclusive evidence, I believe that in the coming years, this trend will become more apparent, although effects will be considerably lagged, as the necessity to learn a new language and essentially, a new culture, slows down the process.

Another problem I encountered was that of multicollinearity between my supposedly independent and exogenous variables. Huge R-squared, F- and t-statistics indicate very close relationships between unemployment in Ireland, the UK and in the rest of Europe. This is, of course, only to be expected. Ireland is a small, open economy, and as the countries of the EC make up a substantial part of our trading partners, and as our interest rates (and hence our monetary policy) are determined in Europe, it is not surprising that we are inextricably locked into the fluctuations of the European business cycle. This effect has naturally become more pronounced as we move towards closer ties with Europe.

Some of the problems of multicollinearity may be overcome by lagging variables. Another method I tried, was to examine the differences in the rate of unemployment between Ireland and the UK, and Europe. I hoped that the relationship between unemployment rate differentials would be more sensitive to net migration than absolute unemployment rates. The results were however, only slightly improved.

Policy Implications - Does Migration Really Matter?

Although my results are inconclusive, it does not mean that issues concerning migration are trivial or unimportant. Large swings in migration levels have serious consequences, whether it is at a regional or nation level. The decimation of rural communities by the return of mass emigration, places a strain on the very fabric of society and the mass exodus of recent years may have severe effects on the economy in both the short and long run.

In the basest of terms, mass emigration represents a fall in the number of consumers which, in Keynesian analysis, represents a fall in aggregate demand and hence, output. Evidence of the feedback of emigration into growth in the Irish economy is mixed and conflicting[10] but it is indisputable that growth in the economy has not been matched by growth in the employment level. It is not unreasonable to assume that emigration could have played some part in this.

With regard to my own limited analysis, it is possible that push factors may account for a higher proportion of migration than suggested by the results. Inadequate data, oversimplification, poor selection of arbitrary values or even misspecification of the model, may have distorted the true values. If this is the case, and we are serious about halting forced emigration, the government must stimulate demand for labour and growth, so as to provide sufficient employment, and no longer rely on emigration to provide the answers.

However, if my findings are correct, the implications are even more far-reaching. This would imply that the majority of emigration is voluntary. The impact of this on growth is immeasurable. Firstly, it means that there is some innate factor in the Irish economy which encourages emigration. Whether it is the legacy of tradition or the product of our current tax and employment systems, the net effect is the same. Already, distortions of the demographic structure have created a worryingly high dependency ratio, which hinders the much needed reformation of the taxation system, thus fuelling the vicious circle of loss.

As has been the case in recent years, the brain drain is bleeding the country dry of her most precious resource. This creates an anti-entrepreneurial environment and undermines the countrys business and financial structure. The very people who are highly educated by the state, and should be creating new employment and promoting growth, are heading for foreign shores. If this is the case, the costs to society are twofold; the state loses the return on its investment in human capital and secondly, society loses jobs which would have been created and the development of an enterprise culture which is sure to follow.

The only possible response to such findings is to discover and correct the imperfections in the economy which drive these people away. This however is a long run goal. In the meantime, it is important to tap into the vast experience of our emigrants. If there truly is a shift away from the UK towards Europe as an emigration destination, it is important to know now, so that we can develop industry similar to that abroad and attract Irish nationals home to run it. In this way, it is possible to compete within an international market and still nurture an enterprise culture.

Conclusions

Emigration is undoubtedly caused by both push and pull factors, although the importance of each cannot be conclusively determined from the empirical work in this paper. However, it is not only important to discover the exact causal relationships of migration, but also to take a normative view of this phenomenon and assess the effects of migration, and emigration from Ireland in particular. It is also essential that we consider the effects of emigration as a whole on the Irish economy and examine some of the policy implications of these effects.

If we believe that future growth in the Irish economy depends on our ability to harness the experience and knowledge of our people, recognition and implementation of policy options which will stem emigration becomes even more urgent. In the past, we have actively encouraged emigration in the hope of shortening the dole queues. Yet, it is a false economy. Mistakes have been made in the past, but we can still use the wealth of knowledge that should be Irelands, but is, instead, distributed around the world.

Bibliography

European Council, (1993) *Statistics Bulletins.*

Dowling, B & Durkan J (eds.) *Irish Economic Policy: A review of Major Issues*

Fell, J P C (1989) Growth, Migration and Causality in *The Economic Social Review.*

Keenen, J G (1978) Unemployment, Emigration and the Labour Force in

Lucey, B (1989) Tests for Macroeconomics Effects of large scale migration based on the Irish Experience, 1948 - 1987. A Comment. in *The Economic and Social Review.*

NESC, (1991) *The Economic and Social implications of Emigration*.

Walsh, B (1989a) Tests for Macroeconomic Effects of large Scale Migration based on the Irish Experience 1948 - 1987. in *The Economic and Social Review.*

Walsh, B (1989b) Tests for Macroeconomics Effects of Large Scale Migration based on the Irish Experience 1948 - 1987. A Rejoinder. in *The Economic and Social Review.*

WELFARE ECONOMICS: FOR LIGHT OR FOR FRUIT

Justin Morton
Junior Sophister

Welfare economics is the branch of economics which applies theory laid down by the rest of the subject, mainly microeconomics. Basic theory is applied to the area of social choice in a bid to assist policy decisions in this area. Hence, the objective of welfare economics is a fruitful one. However, a certain ambiguity remains over the question of whether it has been successful in its aim. In this essay, I will show how welfare economics has succeeded in its original task, and has not failed into academic obscurity that is, it is not for light.

The literature in welfare economics can be divided into two areas: efficiency and distribution. The marriage of these two concepts is welfare economics and is what directs policy choice. Section I will deal with efficiency criteria, Section II with distribution criteria and Section III with the marriage of the two and what results.

Section I

To measure welfare, and hence welfare changes, we appeal to the microeconomic notion of utility. A societys utility is dependent on the utilities of all its INdividuals

W = f(U1,U2,.....Un).

As a proxy for utility, we employ the notion of willingness to pay.[1] The higher the willingness to pay, the greater utility he or she must obtain.

With a proxy of welfare in mind, we now turn our attention to maximising the economic cake. Vilfredo Pareto introduced the efficiency criteria which now form the basis of welfare economics. Pareto defines an improvement in social welfare to have taken place when at least one individuals utility has risen, and nobody elses utility has fallen. This is known as the Pareto criterion and is based on certain assumptions. According to Andrew John[2] these are:

(i) social welfare depends positively on the welfare of individuals

(ii) welfare of individuals depends on the goods and services they consume

(iii) individuals are the best judge of their own welfare, and act in their own self interest.

Although these are accepted, they are open to criticism. Granted they may not hold in certain instances but these cases are the exception rather than the rule. The key criticism of Paretos criterion for social improvement is that it does not allow for losers. (Responses to this will be discussed later.)

Writing after Pareto, the Irish economist Edgeworth furthered this work by constructing his boxes of consumption and production based on a two-person, two-good, two-firm economy. The first order conditions for Pareto optimality, a case where no-one can be made better off without someone losing, are:

(i) marginal rate of substitution (MRS), the rate at which a person would exchange one good for another while keeping utility constant, must be equal for both people in society. If this condition holds, efficiency in exchange is guaranteed and utility is maximised.

(ii) marginal rate of technical substitution (MRTS), the rate at which a firm can exchange the factors of production between two goods, while at the same time keeping quantity constant, must be equal for both firms. If this condition holds, efficiency in production is guaranteed and profits are maximised.

(iii)MRS is equal to marginal rate of transformation, the slope of the production possibility frontier. This ensures that the rate at which firms can reallocate to produce good 1 instead of good 2 is equal to the rate at which consumers want to exchange good 1 for good 2. The outcome is harmonious across all markets.

This analysis leads to the utility possibility frontier (UPF) which maps all combinations of utility which result from this general equilibrium (i.e. the locus of all Pareto optimal points). Consider the point A inside the UPF. This is not a Pareto optimum, since the movement to point D means that both individuals gain and there are no losers.

Diagram 1

As mentioned earlier Pareto gives a non-complete ordering of possible allocations. The point D has the property that there is no feasible Pareto superior point. It is therefore Pareto non-comparable to C. How then do we choose between two points that lie on the UPF?[3]

In an attempt to overcome this, Kaldor[4]developed his ingenious compensation tests. This test deems a project desirable if the gainers can hypothetically compensate the losers. That is, the policy should be implemented if there is a net monetary gain to society. In this case, the point K is more desirable than A. Notably the actual redistribution is not required, it is merely hypothetical.

However, Kaldor introduces another value judgement in order to validate his test. He asserts that there must be equi-marginal utility of money. Thus ten pounds to a millionaire yields the equivalent utility to a less well off person. While this may not necessarily hold, in the main, comparisons will be away from the extremities of the UPF, validating the Kaldor criterion.

Unfortunately for Kaldor, Scitovsky noted a paradox in the test. This arises when UPFs cross, with the present allocation on the first UPF, and the allocation after a potential policy implementation on the other. The paradox, that given the policy is not implemented it is preferred, and given the policy is implemented the former state is preferred! Hence Scitovsky introduced his reversal criterion to overcome this paradox.

The first order conditions for Pareto optimality will only be fulfilled in an entirely perfectly competitive market structure. Due to market failures such as public goods, externalities and monopolies, reality does not result in such a structure. Although this does not nullify the analysis it does have serious implications which are addressed in Section III.

Section II

Although both Pareto and Kaldor remained neutral about actual redistribution of welfare it is a key issue in modern life. According to Pigou:

The misery and squalor that surround us, the injurious luxury of some wealthy families..., these are evils too plain to be ignored.[5]

Hence, we attempt to compile a social welfare function (swf), a map of different levels of utility for each individual that gives rise to a given level of social welfare. The swf allows us to reveal the bliss point where it is tangental to the UPF. To find this point, we need to find the form of the swf. Appealing once again to the concrete foundations of microeconomics, we can think of an indifference curve as showing combinations of two goods which leaves the consumer at the same level of utility. A swf is an aggregate of n individual indifference curves and thus shows the combinations of n peoples utility which leaves society at the same level of social welfare.

However, a major difficulty arises when attempting to reveal social preference. Inconsistencies arise and paradoxes occur when using majority voting. In fact Arrow, in his Impossibility Theorem, claims that there is no way of deriving preferences consistent with social preferences.[6]

As an immediate result of this problem, there is debate over the shape of the swf. John Rawls posits that societys welfare only increases when the utility of the poorest person increases. A swf based on this viewpoint is L-shaped. Conversely, Jeremy Bentham believes that an increase in utility is equivalent and

desirable regardless of the wealth of the individual. This leads to a straight-line downward sloping swf, a third swf was outlined by Bergson and Samuelson. In this case, extra negative weights are given to cases where the distributions of utility are highly skewed. The shape of this curve is convex to the origin. The very existence of at least three alternatives highlights the political nature of the swf.

Section III

Combining the efficiency and distribution criteria laid out in Sections I and II, in a perfect market where social preferences could easily be revealed, leads to a complete ranking of all social states, and hence we end up at the bliss point.

However, noted above are certain practical problems relating to both the efficiency and distribution areas. In the former we come across market failures and in the latter we encounter preference revelation difficulties.

The market failures imply a role for the government. However, due to complex market interrelationships, it appears implausible for the government to come up with a set of rules to apply uniformly across the economy to lead to optimum efficiency. Hence the government is limited to individual proposals to change welfare, such as a new public park or a new bypass. Thus their role cannot be to rank all social states but rather to rank certain proposals open to them at a given time to move the economy towards efficiency. This means that the governments role is discretionary. To rank these individual changes the government appeals to welfare economics, changing the actual role of the subject from ranking all social states to ranking a few. Kaldor-Scitiovsky criteria fulfil the task of choosing between proposed projects by estimating if there is a net monetary gain to society. This work is the foundation of Cost Benefit Analysis (CBA), welfare economics most powerful tool. CBA assesses all the costs and all the benefits of a given project in quantitative money terms.

So, with an (imperfect) proxy for the efficiency criteria, how then do we get over the problem of preference revelation? This is done, again imperfectly, by the voting system. The elected government give weights to different projects depending on their effects on distribution. Whereas a labour government may value a distribution friendly project highly, a conservative government may value an efficiency friendly project highly. These weights reflect public opinion (through the voting process) and are cleverly encompass in CBA. Depending on the regime, the Benefits/Costs ratio for a given project will vary as the relative benefits of the distribution or efficiency are calculated. Hence, CBA combines both the efficiency and distribution elements in one tool.

Is the notion of the swf redundant and the search for the bliss point vacuous? Yes would be the answer from certain naive economic killjoys. What these authors fail to realise is how absurd the alternative to the economic approach to social choice theory is. Whereas economics provides firm systematic analysis, the alternative is a rag-bag of ad hoc techniques and value judgements.[7] In fact while economics is able to quantify notions as abstract as social welfare in actual monetary terms, the alternative approach is engulfed in qualitative nonsense. Granted, welfare economics is an incomplete subject. For example massive problems arise in social preference revelation and the imperfect world in which we live does not result in Pareto optimality. As Culyer (1973) puts it

> The economic approach to social policy is, in general, more comprehensive than any other, and though it has many half filled boxes, it has no empty ones.

Conclusion

Welfare economics is fruitful in two ways. Firstly it provides us with CBA, a vital policy instrument. Secondly, and more essentially, it has brought us a long way down the road to being able to rank all social states and globally say whether or not a policy is desirable. Considering the importance of this, I propose we accept welfare economics in its entirety, and strive to further ourselves along this road to being able to rank all social states successfully.

Bibliography

Culyer, A (1973) *The Economics of Social Policy*

Johannson, P (1991) *Introduction to Welfare Economics*

Ng, Yew-Kwang (1979) *Welfare Economics*

OHagan, J (ed. 1991) *The Economy of Ireland* 6th ed.

Price, C (1977) *Welfare Economics in Theory and Practice*

Rowley & Peacock (1975) *Welfare Economics - a liberal restatement*

Varian, H (1992) *Microeconomics*

THE IMPORTANCE OF EXTERNALITIES IN RESOURCE ALLOCATION.

Mark Kehoe
Senior Freshman

This essay is divided into three parts. The first part deals with how externalities arise and how they theoretically affect resource allocation. The second deals with the importance of externalities in actual resource allocation and the third deals briefly with ways of correcting for externalities and for better resource allocation.

Externalities and Resource Allocation

Externalities arise when the consumption and/or the production of one or more individuals unintentionally alters the utility and/or the production functions of one or more individuals without those persons being compensated or forced to compensate others for that economic activity (Arrow, 1971). Externalities can either be positive or negative.

Externalities affect resource allocation because the market fails to fully price the external effects generated by some economic activities. This is because market prices tend to reflect the cost sellers charge buyers of a commodity, a price based on the personal utility derived, while ignoring the costs/benefits imposed on third parties. Thus the pricing mechanism fails to reflect the true or social costs of economic activity so private costs may diverge from social costs. Resources will be allocated on the basis of private consumption and/or production decisions and not on social welfare maximising ones and for this reason resources will be allocated inefficiently.

The failure of all relevant effects to make their impact on the pricing system will result in a sub-Pareto optimum allocation of resources as the social marginal cost (MSC) of an activity will not equal its marginal private cost (MPC) which equals its price, The real price of the commodity does not fully determine its allocation so the function of the market to efficiently allocate resources based on their true prices breaks down resulting in a misallocation. The existence of externalities will thus lead to a sub-optimal allocation as either too many resources are used in processes conferring uncompensated social costs or too few are used in processes conferring uncompensated social benefits as the profit maximising output is less than the socially optimal output. This misallocation of resources is best seen by an example.

Let a firm be in perfect competition with a given market price, p, and a profit maximising output of xl and a marginal cost curve as in figure 1. Suppose now that the production of x creates air pollution which imposes a cost on local residents of [[sterling]]1 per unit of x produced. To obtain the MSC of x [[sterling]]1 must be added to the MPC of x. As a consequence of the negative externality the profit maximising output xl, exceeds the socially optimal level x* where the MSC=MPC=P.

If the firm is permitted to pollute, the firm produces too much of x, the reason being that part of the real cost of production, [[sterling]]1 per unit, is not recognised as a cost by the firm (Johansson, 1991). Therefore the existence of uncorrected externalities implies that resource allocation is inefficient as a Pareto improvement is possible. Thus externalities, which tend to be mostly negative, result in an inefficient resource allocation as commodities are not allocated on the basis of their true economic price.

Externalities and Cost Benefit Analysis

The importance of externalities in actual resource allocation is best seen by studying cost- benefit analyses. Such a cost-benefit analysis is Barretts and Mooneys (1982) cost benefit study of the Naas Motorway Bypass. This study showed how the construction of the bypass resulted in significant positive externalities and a few negative ones.

There were three main positive externalities. The first was a time savings of 10.28 minutes from reduced traffic congestion in Naas centre at peak hours accounting for 90.5 per cent of the total benefits from the bypass, making the social feasibility of the project was very dependent on this positive externality. The second was a reduction in road accidents in Naas centre due to the transfer of traffic to the safer motor way, while the third was a fuel saving accounting for 2.6 per cent of the benefits.

There were other positive externalities such as reduced lead and smoke pollution in Naas centre and especially reduced noise pollution which constituted a serious negative externality. There were also negative externalities on the environment associated with the construction of the motorway. The Letich committee (1977) detailed some of these costs on non-road users such as the demolition of property, visual intrusion and the impact of farm severance. There exist substantial problems in pricing these externalities and for this reason they were excluded from the study as no accurate price could be put on them. This pricing problem will have affected exceeded the negative ones so yielding a net positive externality suggesting that the real cost of the project was lower than its private cost. This would imply that similar projects should be undertaken for congested towns on the national primary routes (Barrett, 1984), thus leading to a more optimal resource allocation.

Resource allocation will not be optimal unless all costs and benefits associated with the project are calculated. This is the major difficulty with cost-benefit analysis as we do not know how to accurately measure externalities. Some economists such as Roth have suggested that it is impossible to price them, so much so that he ignores them in his road pricing study. This is also echoed in the Smeed Report (1964).

By ignoring to price externalities resource allocation will suffer as projects which would be socially profitable when including all externalities may not be so if only private costs are calculated. This point is addressed by Mishans horse and rabbit stew analogy (1990). He says that economists can easily ignore externalities as they are quantitatively difficult to measure, but doing so could result in a sub optimal resource allocation as such analysis would favour mostly commercially viable projects. There is another school of thought which says that social expenditure cannot be justified largely on the grounds of correcting for externalities. Lees makes the point that only 5 per cent of health expenditure can be justified in terms of correcting for externalities as most medical expenditure centres on non-contagious diseases where the benefits are quite private. Peacock and Wiseman make a similar point saying that the positive externalities of education of the individual may be exaggerated. They suspect that the recipient may appropriate most of the benefits in the form of higher wages and salaries (Allan, 1971).

Therefore externalities, although important in causing resource allocation to be sub-optimal, have varying effects. One thing that is sure, however, is that they must be included in cost-benefit analysis even if they are estimated very roughly.

The existence of externalities implies that unless special arrangements are made resource allocation may not be Pareto optimal. One way to Pareto optimality is by modifying the pricing system to reflect the true price of the resource. Through this process of internalisation/shadow pricing resources are allocated on the basis of their true prices. It may not be possible to internalise all externalities (Mishan, 1990) so government intervention in the market may be needed.

An area where externalities may be tolerated even though they affect resource allocation is where the correction of them may have regressive social welfare distribution effects and clash with other government objectives (OHagan, 1991). This arises because the Pareto criterion takes no account of welfare distribution. The other case is where the benefit of internalisation exceeds its cost. But society is always worse off with existence of externalities even when they are corrected than without them. By internalising them we are doing no more that making the best of a bad job. We are certainly not as well off as we should be if they had not appeared on the economic scene.(Mishan, 1990). Thus we are forced to the theory of the second-best.

Externalities effect resource allocation by distorting the pricing mechanism, resulting in an allocation of resources that is not optimal. Therefore the importance of externalities in resource allocation is crucial if it is to be optimal and it is observation that gives cost- benefit analysis some of its justification as it is necessary to measure those created by activities and to intervene to correct them.

Bibliography

Allan, C (1971) *The Theory of Taxation*

Arrow, K (1971) *Public Expenditure and Policy Analysis*

Barrett, S (1982) *Transport Policy in Ireland*

Barrett, S & Mooney, D (1984) The Naas Motorway Bypass - A Cost Benefit Analysis in *Quarterly Economic Commentary* 1/1984

Johansson, P (1991) *An Introduction To Modern Welfare Economics*

OHagan, J (ed. 1991) *The Economy of Ireland*

Mishan, E (1990) *Cost Benefit Analysis*

Mishan, E (1988) *Elements of Cost Benefit Analysis. Road Pricing: the Technical and Economic Possibilities* [The Smeed Report], (1964).

THE COASE THEOREM

Brendan Fitzpatrick
Senior Sophister

The Coase Theorem is false on purely economic analytic grounds. Furthermore, the conventional interpretation of Coases message, and the policy prescriptions which follow from this do not observe the strict qualification which Coase placed on his result. There is, therefore, a divergence of meaning between the orthodox statement of the Coase Theorem and what Coase intends (as revealed in his 1989 retrospectives: Notes on the problem of Social Cost.). This may also be dissonant with what he meant in his original 1960 article.

What is the Coase Theorem: A clarification

For a working definition of the theorem consider that professed by Cooter in the New Palgrave:

> The initial allocation of legal entitlements does not matter from an efficiency perspective so long as they can be exchanged in a perfectly competitive market.

To illustrate, consider Coases famous example whereby locomotives emit sparks which set fire to farmers fields. Suppose the legal intervention is by an initial governmental legislative allocation of spark emission permits, which allows the railway to emit only a restricted amount of sparks. Assume these can be traded with the farmers. While the law apparently controls the extent of the damage, the theorem dissents, stating that it is the market which determines the final efficient allocation of permits between farmers and the railway, and hence the extent of the damage.

This is because the farmers and the railway will face the above aggregate marginal willingness to pay and marginal cost curves for the trade in permits. Starting at, say, an initial allocation Q1, the individuals most preferred level of emission reduction, firms will buy permits until the efficient outcome Q* results. Indeed Q* will result regardless of the initial allocation of legal entitlements.

The rationale for Stiglers (1966) definition of the Coase Theorem is now evident: Under perfect competition, private and social costs will be equal. Coase (1988) argues that zero transaction costs are a suitable proxy for perfect competition and notes that zero transaction costs were assumed in his original (1960) paper. An alternate definition by Cooter, then, is the best indicator of what Coase means: The initial allocation of legal entitlement does not matter from an efficiency perspective so long as the transaction costs of exchange are nil.

Criticism

We may first note three initial reasons why the Coase Theorem, as specifically formulated above, is wrong. These are based on the inadequacy of zero transaction costs as a proxy for perfect competition. Firstly, monopolies, such as a typical railway company, are unlikely to act like competition despite Coases claim to the contrary. Secondly, Arrow (1969) demonstrated *a priori* that the characteristics of many externalities make unfeasible a competitive market in them. Consider a market in rights to produce a negative externality which affects the whole public. Each person has a permit, which he may sell to a firm, allowing the production of a certain amount of the externality. But his selling of that permit creates a negative externality for other members of the public. Many externalities are impossible to internalise. Thirdly, imperfect information will thwart the bargaining process for legal entitlements, leading to Prisoners Dilemma type situations.

Regardless of the inadequacies of the zero transaction costs approximation, the initial allocation of entitlement always has distribution consequences for wealth. As a result, demand for other goods and services will be dependent on the initial allocation of property rights. The shapes of our aggregate marginal

willingness to pay and marginal cost curves will then change, being dependent on the demand for other goods and services. Consequently, the efficient outcome Q* is dependent on the initial legal decision. This amounts to outright theoretical falsification of the theorem.

Varian (1987) has noted, however, that in the very special circumstances, where preferences are quasilinear, the strict version of the theorem may hold. Quasilinear preferences suggest the absence of income effects, implying demands for the emission permits are independent of the income distribution. This is seen in the form of a horizontal contract curve yielded in an Edgeworth Box analysis of the situation.

Varian doesnt acknowledge Arrows aforementioned objection, however, which would deny the theorems veracity even in this special case.

A final objection to the theorem is that the notable externalities of our age, such as ozone depletion and nuclear pollution, affect a large number of people, implying large transaction costs. This further consolidates the political irrelevance of the theorem.

The Meaning and Relevance of the Theorem

One should note Coases recent rebuke to the allegation that the world of zero transactions costs is a Coasian World. He states (1988) that he merely wished to bring to light some of the properties: I argued that in such a world the allocation of resources would be independent of the legal position, a result which Stigler dubbed the Coase Theorem. Coases argument to this effect was, as we have shown, wrong. But this belated emphasis on the qualifications of the result suggest that Coase now means his theorem to be a mere theoretical nicety.

It also indicates the removal of Coases present position from the, at times, rash and barely qualified espousal of its result, by his disciples. Consider Stiglers (1966) description of the original statement as a profound article and its result as a remarkable proposition to us older economists who have believed the opposite.

It also suggests a divergence of Coases present position from what he once meant. Circumscribing criteria seem comparatively de-emphasised in the original article, in which he states that the usual courses of action (on externalities) are inappropriate, in that they lead to results which are not necessarily, or even usually desirable.

Conclusion

The Coase Theorem, in its traditional exposition as a policy prescription is false and of no political or legal relevance whatever. Moreover, the restrictive set of postulates whereupon it is questionably derived tend to be de-emphasised in contemporary espousals of the result. This inappropriately bestows unwarranted practical relevance unto it which Coase (belatedly) seems to recognise in recent statements.

Bibliography

Arrow, K (1969) *Readings in Welfare Economics*

Coase, R (1988) *The Firm, the Market and the Law*

Cooter (1987) The Coase Theorem in *The New Palgrave*

Stigler, G (1966) *The Theory of Price*

Varian, H 1987 *Intermediate Microeconomics*

ROAD PRICING: THE CASE FOR AND AGAINST

Carsten Rolle
One Year ERASMUS

It has always been a major task of economists to set incentives so that resources are allocated in the most efficient way. Nevertheless, the Transportation Sector in general was characterised for a long time by government intervention. Radical changes in transport policies - especially in the UK and later also in other European countries came along at the beginning of the 1980s and were largely supported on a theoretical basis by the new concept of contestable markets[13]. Accordingly the policy of deregulation and privatisation of, for example, express coaches, buses and aviation, has resulted in significant product innovation and rising efficiency.

As road space is a valuable and scarce resource[14] we would argue that it ought to be rationed by a price mechanism. Road users should pay for using the road network to make correct allocative decisions between transport and other activities.

In the past, the technical possibilities of road pricing were very limited[15] but with the advent of electronic road pricing, cars no longer have to stop to be charged.

Both aspects, the more liberal political climate of deregulation and privatisation in the last years and the new technology, are essential prerequisites for the actual discussion of road pricing. On a more practical political basis we could also identify the interest of the state to create a new source of revenues although road pricing does not necessarily imply a higher burden for car users.

After a brief description of electronic road pricing this paper will outline the costs of using a road and give a definition of the price to be charged. Then the effects of road pricing shall be identified and critically evaluated before the basic results will be summarised.

Economic description

Road Pricing - the technique

The new technology of electronic tolls no longer requires motorists to halt at tollbooths. Therefore, it prevents additional congestion. Drivers would be given an electronic number plate which signals to the recording computer the presence of a vehicle. This would be the most direct way to charge the amount specific to the road and the time of the day. The devise could charge users via bank account or monthly bill. Because there have been many objections against the individuals location by electronic detectors[16] the use of smart cards appears to be more preferable. The electronic licence plate would be loaded with the smart card and would debit payments. Only if the card were exhausted would the central computer monitor and bill for road use[17].

The complete installation of such an electronic system would take some time. In the meantime area licences could be sold for very congested zones, such as city centres. (This solution is used in Singapore for the rush hour traffic with considerable success. The impact of its introduction was an immediate reduction of 24,700 cars during the peak time and a rise of traffic speed by 22%.)[18]

Definition of the good

Before actually defining the price of a good it is essential to characterise the good in terms of the Public Goods Theory to deduce the optimal rule of providing it. Road use is rival in consumption[19] and also excludable with adequate costs of the pricing technique. For that reason road infrastructure is no longer a

pure public good but a private good with some degree of externalities. This implies that private provision of roads may be favourable.

Definition of the costs

The ideal road price would only allow cost justified trips to be undertaken. Economic theory therefore postulates marginal social costs (MSC) to be equal to marginal social benefits (MSB)[20]. But what are the marginal social costs that allow a vehicle to make a particular trip? Besides the private costs of road use (like fuel, drivers time etc.) that are directly paid by the driver, four main costs can be identified[21]:

a) road damage costs

b) accident externalities

c) congestion costs

d) environmental costs.

a) road damage costs

Road damaging is basically caused by heavy vehicles as the damage to the road pavement increases to the fourth power of the axle load. Therefore road damage costs should be proportional to the damaging power (measured in terms of Equivalent Standard Axles). Thus almost all costs should be paid by heavy trucks.[22]

b) accident externalities

Accident externalities arise when extra vehicles on the road increase the probability that the other road users will be involved in an accident. Accident probability depends to a large extend on distance, driving time and particularly the other traffic. This is why accident costs will be treated like congestion costs.[23]

c) Congestion costs

Congestion costs arise due to the fact that additional vehicles reduce the speed of the other vehicles and hence increase their journey time. Economic analysis shows that the traffic flow will be optimal at Q* if the costs of additional traffic (MSC) and the demand are equated. However, the individual user entering the road will typically consider only the costs he personally bears (MPC), i.e. marginal private cost and will thus operate at Q Therefore he takes the marginal private cost curve into consideration rather then the optimal marginal social cost curve (MSC) for the new trip-maker and the existing road users. The difference between the MPC and MSC curves reflects the dead-weight loss of excessive traffic congestion.

In a market system without transaction costs the other road users would be willing to pay the additional car the amount of their opportunity costs of time and additional fuel for not entering the road. As transaction costs have been obviously immense (if a perfect bargaining process would have been possible at all) so far, only an electronic pricing system can overcome the huge existing transaction costs between the road users.

d) environmental costs

The road use of vehicles has various spillover effects on the environment[24] like:

local: emission of CO, NC, NO2

global: emission of CO2, CFC

water pollution

noise and vibrations

land use effects (destruction of wildlife habitats and the landscape)

Definition of the price

As shown in the previous analysis the marginal social costs of road use exceed the marginal private costs. The optimal road price, p* must therefore reflect the differences between MSC and MPC and will generate a welfare gain. Because of the traffic reduction (Q - Q*) consumer surplus will fall but as long as the social gain is greater than this loss the total welfare effect will be positive.

A basic problem still remains, however, : road users have to get the information about the changing road prices immediately to optimise their individual transport decisions. (This could only be guaranteed if the motorist would have access to the prices via a board computer.)

Evaluation and Effects

Advantages of Road Prices:

* In the second section, it was shown that road pricing is a good instrum -ent to use to internalise most of the external effects numerated above.Especial-ly in the case of congestion costs, it appears to be the optimal method of internalisation because a price mechanism would replace the present queuing mechanism which is allocatively inefficient.

* Furthermore, it reveals the true economic costs of the road use (including replacement costs) so that intermodal competition[25] would become fairer. Because road prices would be primarily connected with congestion costs, some distributional and locational effects could arise. Costs of driving in non-urban areas would probably fall whereas urban driving costs would increase so that in the medium run, the quality of the public urban transport system would improve.[26]

* In the case of pricing highways on the continent, road pricing is a good instrument to overcome the free rider problem of foreign carriers using "home country" highways. This is especially interesting against the background that current ways of financing highways are very different. For that reason actual competition between international carriers is not neutral.

* As shown in the previous section, road damage costs of cars are almost zero whereas heavy trucks cause most of the damage. Therefore, a vehicle specific tax depending on the damaging power would be a simple and effective wayof charging efficiently. Road pricing systems could improve this instrument a little by taking the quality of roads that were actually used into account.

Disadvantages of Road Pricing:

* The costs of implementing an electric toll system are very high. The German government estimates that the implementation of the system will cost 2bn pounds for its Autobahn network plus individual costs for every vehicle of 40 pounds each, not including additional costs of controlling the system. On the other hand, controlling the toll system would enable a privatisation of the roadnetwork which could lead to additional revenue for the state.

* In terms of negative environmental externalities, road pricing is (with the exception of noise) probably not the optimal instument for internalisation. Taxes on fuel or emission fees, for instance, charge vehicle emissions in a moredirect way and they are very simple to design.

Furthermore it must be mentioned that the effect of road pricing depends to a large extent on the authority[27] that receives the revenues and its way of using the money. Economists would argue that the profits made should be reinvested into the transportation system to generate an efficient outcome rather than cross-subsidising other traffic modes or other state activities.

Conclusion

To put it in a nutshell, this paper advocates that road pricing is the best instrument to internalise the costs of congestion and road damage. Although the initial costs of installation are high, these costs would probably quickly be exceeded by the efficiency gains of corrected prices. Nevertheless, road pricing cannot perfectly internalise external environmental costs. That is why instuments like "fuel taxation" or "emission fees" will still be necessary to design an optimal price mechanism in the transportation sector that sets the correct incentives.

Bibliography

Baumol/Panzer/Willig, (1982) *Contestable Markets and The Theory of Industry Structure*

Boris,S., (1988) "Electronic Road Pricing: An Idea Whose Time May Never Come", *Transportation Research*, 22A(1), pp. 37-44

Braeutigam, R (???) *Optimal Policies for Natural Monoplies*

Buchanen, J. (1965) " An economic theory of clubs" in *Economica* 32

Button, K (1990) "Environmental externalities and Transport Policy" in the *Oxford Review of Economic Policy*, Vol. 6, No. 2

Glaister, Starkie and Thompson (1990) "The Assessment: Economic Policy for Transport", in the *Oxford Review of Economic Policy*, Vol. 6, No. 2

Keating, G (1993) "For whose the road use", in *Economic Affairs*, June

Newbery, D (1988b) "Road Damage Externalities and Road User Changes" in *Econometrica* ol. 56(2)

Newbery, D (1990) "Pricing and Congestion: Economic Principles relevant to Pricing Roads", in the *Oxford Review of Economic Policy* Vol. 6, No.2

Sandler, T and Tschirnart, J. (1980) "The Economic Theory of Clubs: An Evaluative Survey", in the *Journal of Economic Literature*, Vol. 18

Schmalensee/Willig,(1989) *Handbook of Industrial Organisation*, Vol. 2

WELFARE ECONOMICS : FOR LIGHT OR FOR FRUIT A REPLY

Ian Cafferky
Junior Sophister

The purpose of the reply to the essay Welfare Economics : For light or for fruit is to inject a sense of realism into the study of welfare economics as I understand it. This is not intended as a slight on the author, to whom I address this reply, but it is a personally held criticism of the tired standard approach to the subject. Implicit in my reply is the commonly held assumption that welfare economics should be a policy prescriptive subject and thus its aim is for fruit rather than for light.

The approach to welfare economics to which I refer can be found in any standard textbook on welfare economics and is cogently proposed in Mortons essay. It begins with a discussion regarding the measurement of welfare and then enters into abstract theoretical efficiency conditions for a Pareto Optimum. That aside, the question of distribution is discussed and overall the framework would appear to be plausible. Next market failures are introduced as a necessary step to make the theory acceptable for real world situations. Unfortunately, this step into the real world requires a great mental leap as the apparent consequence of theoretical welfare economics is the applied welfare economics of cost- benefit analysis.

My proposal is straightforward. Theoretically, many aspects of welfare economics are not entirely sound, and furthermore, when the theory is faced with real world dilemmas it collapses. I also believe that the applied branch, cost benefit analysis, is not necessarily a product of welfare economics. Thus, it can be concluded that the most notable achievement of welfare economics is its own redundancy.

Criticism of the standard textbook approach

Standard welfare economics begins with the proposal that the measurement of welfare can be closely proxied by the measurement of utility. Any observer would duly ask: When does utility deviate from welfare? Welfare is the well being of an individual whereas utility is the satisfaction an individual derives from the consumption of goods and services. The justification for the synonymity is the claim that if an individual chooses X over Y, then he or she can be assumed to be happier in X than in Y. This is not always true. Consider the irrationality of preferences which can be due to many factors. Ng (1979) notes that habits/customs, advertising and ignorance are clear examples of possible causes of irrationality of preferences. None the less it is generally assumed that the individual is the best judge of his or her own welfare. This, I would agree is an acceptable assumption. There is one final criticism of the claim to equality of welfare and utility. That is, that welfare is an ex-post measure and utility is an ex-ante measure. Thus, utility as a measure of welfare, is in fact an expectation of welfare. This logically makes the point that ex-ante and ex-post measures are not necessarily equal. It is pedantic to labour these points as there is no absolute measure of welfare and nor is there ever likely to be. It is, as Ng (1979), proposes, convenient to accept these divergences and to continue the study. This criticism of welfare economics is in fact minor, in that it will never be solved and must be accepted. The ultimate criticism lies in the inapplicability of the theory that follows.

Accepting utility as a proxy for welfare the theory of welfare advances taking a two plank approach. On the efficiency side there is the apparent positive economics of Pareto, and on the distribution side there enters the explicitly normative side of welfare in the quest for a social welfare function. I shall consider each of these in turn.

The efficiency side of the theory is especially open to criticism due to its positive approach. Using Paretian analysis, Morton introduces the first order conditions necessary for a social optimum. These conditions, however, are completely and absolutely inapplicable to the real world. The form of competition necessary to achieve this optimum is a theoretical abstraction, perfect competition. In practice, the world suffers from

market failures such as monopolies, public goods and externalities that cause a deviation from this perfect competition theory.

Then, given this reality constraint, what does the theory tell us to do? Perhaps we should try to achieve the Paretian first order conditions wherever possible? To answer this Lipsey and Lancaster (1956) attempted to derive similar first order conditions for an optimum when market distortions exist. Their results, in the form of the theory of second best, were conclusive. The resulting conclusions were far too complex to utilise, involving second order cross partials for substitutes and complements. Furthermore, the analysis showed that fulfilling the Paretian first order conditions for some sectors and not for other sectors could in fact lead to a situation where we are worse off in efficiency terms than before. Thus we must analyse the entire economy and take every problem into account to guarantee an improvement. The dilemma is stated succinctly by Ng, We must leap right to the summit to be sure of an improvement. But it is clear that this task is epistemologically, administratively and politically impossible.[8] The summit being referred to is the fulfilment of each and every first and second order condition necessary for an optimum. It is apparent that the entirety of theoretical efficiency analysis is inapplicable. The purpose may have been to produce conditions for the entire economy to achieve an optimum, yet the result is the inadequacy of the theory to adapt to real world scenarios. Perhaps the discussion of market failures, under the theory of welfare, should be changed to the discussion of welfare theory failure when applied to the market.[9]

The other plank of the theoretical welfare economics approach is the distribution issue. Welfare economics aims to rank all possibilities and thus it must not only consider beyond efficiency but also whether a distribution is good or bad. The crucial problem in this side of the analysis arises in the aggregation of the necessary ordinal utility functions. Arrows Impossibility Theorem shows that from individual orderings of social states, a social ordering consistent with some reasonable condition can not be found in general.[10]Without digressing, it simply eliminates the possibility of a social welfare function based on ordinal utility functions. Furthermore the theorem has similar implications for the voting process which is often the mechanism by which our individual preferences may be revealed.[11]

Finally, even if there were an acceptable method of aggregation we must introduce the reality constraint once more by reminding ourselves that utility is not directly observable and that at best, we will end up using a proxy for utility which is a proxy for welfare. This is clearly an inexact theory, but we are attempting to measure such qualitative concepts as happiness and so it is an inevitable problem.

I believe it is clear that the standard textbook teachings are of little use to anyone. The efficiency conditions are redundant and to put it mildly, the distribution issue suffers from aggregation and measurability problems. It appears that the bulk of welfare economics analysis is fruitless. This brings us back to the title of the essay. There can be little doubt that the intention was to use the abstract efficiency theory as a base from which assumptions may be relaxed and reality introduced. This approach has failed. The theory has merely illustrated the complexity of the problem. Perhaps welfare economics has been for light rather than for fruit?