Econometrics and the Scientific Status of Economics: A Reply

Allan Kearns (JS)

'It is unreasonable to expect the economist to forecast correctly what will actually happen as it would be to expect a doctor to prognosticate when his patient will be the victim of a railroad accident and how this will affect his state of health' Schumpeter
In this paper, Allan Kearns replies to the arguement that econometrics does not add to the scientific status of economics, noting that, without econometrics, economics would be less scientific.

INTRODUCTION

As Schumpeter (1978,p.464) notes, '[t]here is ... one sense in which economics is the most quantitative, not only of 'social' or 'moral' sciences, but of all sciences, physics not excluded.....some of the most fundamental economic facts, on the contrary [to those of natural science], already present themselves to our observation as quantities made numerical by life itself. They carry meaning only by virtue of their numerical character". It is the numerical basis of economics that yields the discipline its standing as a science and implies a role for econometrics as the tool for the process of those 'quantities made numerical by life itself.' This paper will seek to argue that, regardless of the behaviour of some of those who have chosen to practice the subject, econometrics is scientific and therefore adds to the status of economics as a quantitative science. It is not the purpose of this paper to argue that econometrics has the ability to contribute so much, that economics is elevated in popular opinion, to a scientific standing akin to the hard sciences, but merely that without econometrics, economics would be less scientific. One will offer a definition of science, a definition of econometrics and show that despite well documented limitations, the definition of econometrics fulfils the definition of science. One will note in conclusion that it is econometrics contribution to the scientific status of economics which underlined Schumpeter's initial view of econometrics as 'the colossal step forward on the road to the development of our discipline [economics]".

Is Economics a Science?

It would be inconsistent to argue that econometrics contributes to the scientific status of economics if one failed to note that indeed economics had a scientific nature pre econometrics. Thus as Schumpeter notes, 'we have no common credo beyond holding: first, that economics is a science'.

When is a discipline Scientific?

In defining what is scientific, one has chosen two definitions, each of which highlight what one feels are important characteristics of items scientific, important not only in the context of this papers argument, but also in any definition of science. At a broad level, the encyclopaedia Britannica defines science as 'any mode of investigation by which impartial and systematic knowledge is acquired' Mises (1956). However one should observe the following two narrower definitions concerning elements necessary for a process to be scientific. Firstly, Hendry (1980) notes that Keynes in a review on econometrics came close to asserting that 'no economic theory is ever testable' to which Hendry adds, 'in which case of course, economics itself ceases to be scientific' Hendry (1980). One takes from this that if one could provide a method for testing economic theory, that economics would 'begin' to be scientific. Mises comments that 'scientific method is supposed to mean statistical method or the demand that every proposition be 'verified' by numerous sets of statistical data relating to sufficiently comparable situations'. Secondly, Schumpeter notes the old dictum that 'correct prediction is the best or only test of whether a science has achieved its purposes' (a href=#schumpeter>Schumpeter(1978). Therefore, correct prediction within the bounds of what one can reasonably expect of an uncertain future, is a requisite for scientific status. In the following description of

econometrics it will be seen that the two of the characteristics of a scientific discipline noted previously, are to be found in econometrics.

What is Econometrics?

Koutsoyiannis notes that 'applied econometric research is concerned with the measurement of the parameters of economic relationships and with the prediction (by means of these parameters) of the values of economic variables' (Koutsoyiannis, 1978) and that there are five desirable properties of any econometric model:

1. theoretical plausibility - the model must describe adequately the economic phenomena to which it relates,
2. explanatory ability - the model should be able to explain the observations of the actual world,
3. accuracy of the estimates of the models parameters - the parameters should approximate as best as possible the true parameters of the structural model, that is, they should be efficient, consistent and unbiased,
4. forecasting ability - the model should provide satisfactory predictions of future values of dependent variables,
5. simplicity - the model should represent the economic relationships with maximum simplicity.

It is noticeable that these desirable qualities of any econometric model reflect the characteristics noted earlier, as requisites for a scientific discipline; the ability to verify results and to produce accurate forecasts. Therefore, by showing that econometrics methodology is essentially scientific, one hopes to argue that this methodology then adds to the scientific status of economics.

How does Econometrics contribute to the scientific standing of Economics?

Repeated Verification of the Model's Explanatory Ability

Hendry (1980) notes that 'the three golden rules in econometrics are test, test and test' yet one must take account of Machlup when he states that 'it is said that verification is not easy to come by in the social sciences, while it is the chief business of the investigator in the natural sciences' (Machlup, 1978). Therefore, if economics had access to a procedure of verification, then this social science could then engage in the chief business of the 'hard'sciences. As described earlier in the definition of econometrics, a process of verification is allowed for, for instance it is possible for one to construct a measure of closeness of ones estimate to the true parameter. One concludes that if economics could not refer to econometrics methodology for examination of its theories, then economics would be less scientific. As Hendry (1980) concludes about the ability to consistently verify results, 'rigorously tested models, which adequately described the available data, encompassed previous findings and were derived from well based theories would greatly enhance any claim to be scientific'

The Forecasting Abilities of Econometric Models

As noted in the definition of what is scientific and of the desirable properties of any econometric model, one encountered a consensus between the two in the importance of any scientific methodology to provide 'satisfactory' forecasts. Econometrics models provide the ability to make predictions, the accuracy of which is constrained only by the practicing econometrician and the economic theory upon which the model is based. Remembering Schumpeter's view that a science cannot achieve its objectives without correct prediction , this would imply that without the forecasting abilities of econometrics, economics would be less scientific.

Yes, but is it a constrained contribution to Economics scientific status?

Two lines of thought that one is subjected to in arguments that the contribution of Econometrics to the scientific nature of economics is constrained and possibly non-existent, is firstly that the unsatisfactory nature of the data generating process and secondly the standing of economics itself as a social science, questioning the ability to have a scientific study in the field of human behaviour.

In response to this first approach, the faults of the data generating process and hence of econometric models have been well documented; omission of variables, faulty measurement, misspecification and inadequate sampling. However one feels that these do not undermine what one has shown to be the scientific basis of econometrics, the provisions in its methodology for verification and prediction.

Hendry (1980) concludes that 'the ease with which spurious results could be created suggested alchemy, but the scientific status of econometrics was illustrated by showing such deceptions are testable.' Knowing that such errors exist allows one to account for them. As Machlup (1980) reports, 'everyone of us could name dozens of propositions that have been disconfirmed, and this means that the verification process has done what it is supposed to do.'

With regard to the standard of forecasting held against econometrics, one feels that this is not attributable to the methodology of econometrics but to what Machlup (1978) notes as the fundamental difference between the soft and hard sciences, 'experts in the natural sciences usually do not try to do what they know they cannot do and nobody expects them to do it' whereas 'social scientists, for some strange reason, are expected to foretell the future and they feel badly if they fail.'

With regard to the second line of reasoning noted earlier, questioning the ability to have a scientific study in the field of human behaviour, one refers to what Mises (1956) calls the inferiority complex of the social sciences. It has been offered that economics can never produce consistently accurate results. In essence, the argument is that a scientific status akin to that of the natural sciences can never be achieved. Mises (1956) notes that such individuals 'are apparently ashamed of the one thing that really distinguishes social sciences from natural sciences, namely, that fact that the student of human action is himself an acting human being and therefore has at his command a source of knowledge unavailable to the student of the phenomena of nature.' Similarly for hard sciences it must be noted that 'there exists no method-oriented definition of science under which all parts and sections of physics, chemistry, biology, geology and other generally recognised natural sciences could qualify as sciences' (Mises,1953).

In conclusion one must note that when comparing the hard and the social sciences, it is common practice to refer to the controlled laboratory experiments of the natural sciences in which predictions have proved so eminently successful and then to look at the social scientists inability to apply this methodology with the resultant level of accuracy. however, one would support Machlup (1978) when he notes that 'if a comparison is made it must be between predictions of events in the real natural world and in the real social world.'

Conclusion

As stated in the introduction, one has sought to argue not that econometrics is capable of raising economics to a scientific status akin to that of physics, but that without econometrics, economics would be less scientific. Given ones definitions of what constitutes a science and of what constitutes econometrics, one concludes that even taking account of its limitations, econometrics contributes to economics scientific standing. One wholeheartedly agrees with Schumpeter's (1978) conclusion about this union of economic theory and statistical science, 'I expect much from it; even the most modest result, no matter how much it will be ridiculed and criticised - as it surely will be, since nothing is easier than to criticise first attempts - will be a colossal step forward on the road to the development of our discipline [economics]'.

Bibliography

Hendry, T. 'Econometrics-Alchemy or Science?' in Economica November 1980

Koutsoyiannis, A. 'Theory of Econometrics-An introductory Exposition of Econometric Methods.' 2nd ed. 1977 Macmillan London

Machlup, F.'Methodology of Economics And Other Social Sciences' 1978 Academic Press

Machlup, F.'Are the Social Sciences really inferior?' as edited by Machlup, F. 'Methodology of Economics And Other Social Sciences' 1978 Academic Press

Mises, L. Von'The inferiority complex of the social sciences' 1956 as edited by Machlup, F. 'Methodology of Economics And Other Social Sciences' 1978 Academic Press

Nagle, E. 'If matter could Talk' 1969 as edited by Machlup, F. 'Methodology of Economics And Other Social Sciences' 1978 Academic Press

Schumpeter, J. 'Economic Methodology' as edited by Machlup, F. 'Methodology of Economics And Other Social Sciences' 1978 Academic Press

The Pearsonian Stock Market: Testing the Fama-French Model of Share Returns

Peter Nolan

Ordinary least squares is the most commonly used technique in any practical application of econometric methodology. In this paper, Peter Nolan using this technique, seeks to explain a cross-section of weekly returns quoted on the international stock exchange.

I can calculate the motion of heavenly bodies, but not the madness of men. Sir Isaac Newton

INTRODUCTION

Apart from love and war, nothing arouses violent human passion like the stock market. Ever since gentleman speculators studied hand-drawn charts of stock price movements in the time of J.P. Morgan, investors have looked to scientific and economic methods in their quest for models to predict the future more accurately One of the greatest advances in finance has been the Capital Asset Pricing Model (CAPM) which postulates that the return on a share was positively related to the beta of this share is a measure of the volatility or dispersion of returns (which, in finance, is defined as risk) of that asset relative to the volatility of the entire set of assets investors can choose as investments. The expected return on an asset can be predicted using estimates of the risk-free rate of interest (Rf), of the expected return on the market (Rm) and of beta.

Ri = Rf + Bi(Rm - Rf), where

Bi = (Si,m)/(Sm^2)

Paul Samuelson has said that this is an excellent model because it explains between thirty and ninety percent of the cross-sectional variation in stock returns using a single factor Bernstein (1992). An article by Fama and French used regression to show that did not explain the observed cross-section of returns observed on US equities. The authors noticed that two other variables, the ratio of market equity (ME) to book equity (BE), and the size of the firm (which equals ME) had more explanatory power. Indeed is likely to be a proxy for ME, as it is almost perfectly negatively correlated with firm size (r^2 = 0.98). In this paper I shall use Ordinary Least Squares (OLS) regression to test whether the ratio of ME to BE, together with ME, alone explains *ex ante* the cross-section of weekly returns on holding each one of a sample of shares quoted on the International Stock Exchange over a particular trading week in mid-December 1994. Firstly I shall explain how these two variables could have an effect on the return of a stock. Then I shall describe how I gathered my data and the behaviour of the variables over the period. Next I report the results of my regression analysis. Finally I explain some of the possible implications of my results.

Model Specification

My objective is to use regression to capture the influence of both company-specific risk factors and investors' expectations in determining returns, unifying the two conflicting schools of financial thought, the fundamental and the technical. Fundamental analysts concentrate on finding the 'real' value of the firm using economic and financial analysis; technical analysts aim to forecast the changing sentiment, often using charts that show the past movement of prices. In an efficient stock market, the expected value of discounted future cashflows from purchasing an equity will equal the price at which it may be purchased Fama (1966). Book equity is the accounting figure for the acquisition value of the firm's assets. Hence the ratio of ME to BE will represent the expectations held by investors about the firm. It seems, although this is not stated in the Fama & French article, that this is an estimator for Tobin's Q. The higher this ratio is, the higher investors expect the firm's return on the value of its assets will be, as shown by the premium over the acquisition cost of the firm's assets they pay for a share in the firm's dividends and capital growth. Therefore I expect that there will be a positive relationship between this ratio and the observed return.

Variables

The size of the firm will also affect the return. I expect a negative relationship. Large firms are typically more diversified and would tend to exist in more mature or low-growth industries, which is confirmed by a casual examination of the larger companies in my sample - mainly made up of multinational giants (Shell, BAT) or utilities (BT, the regional electricity companies), or diversified industrial holding companies (BTR, Hanson). Hence the riskiness of their returns would most likely be low, leading to a lower return under capital market efficiency. Since seems to be a proxy for size, this conclusion is simply restating the conclusions of the CAPM, which were shown formally in the introduction. I could have used an estimate of the for each stock as a third independent variable, but I decided not to because it would introduce multicollinearity: has a high correlation with ME. Return over that predicted by the CAPM could have been used as my independent variable. I was aware of the difficulties in estimating the figures needed for such a computation, as the predictions of the CAPM are very sensitive to the market index chosen as a proxy of the theoretical most efficient portfolio out of all investment assets Malkiel (1992).

I ran regressions using the weekly return on the stocks as my dependent variable, **Y**. The first independent variable, **X1**, was the market equity at the opening of business on Monday. The second independent variable, **X2**, was the ratio of book to market equity at the opening of business on Monday. To minimise the effect of shifting parameters, the equity data was taken in a week in which no potentially significant new information, such as new macroeconomic data, was released. All the shares are quoted in sterling, and BE and ME are both measured in sterling, so that currency rates will have no direct effect on the variables.

Data collection was not difficult. Every transaction is recorded by the exchange clearing house. The cheap and easy method for obtaining price data is to use *Datastream*. The drawback of this data is that the price is based on averages of bid and offer prices prevailing at any particular time, so there may be many small errors present. It would be too expensive for me to obtain data for the actual first observed transactions. The individual stocks were chosen at random from the Financial Times (18/12/94) . Only non-financial companies were used, which seems to be conventional in financial studies; this avoids the high leverage of banks and other types of financial holding companies and the assosiated effects on risk and return of equity. The property and insurance companies I have included had low leverage ratios - less than thirty percent according to *DataStream*.

My dependent variable, **Y**, was the weekly holding return on each of a sample of fifty-nine shares expressed as a percentage rate of return. This was made up of capital gains only and excluded dividends paid during the period. I feel confident doing this because the assumption was used by Fama & French article also. Typical weekly returns varied between plus eleven and minus four percent, while the dividend yield for the week would be insignificant. At an average for the market as a whole of three percent per annum, the weekly yield was calculated as

(1.03)^1/52 = 0.00059%

I gathered figures for the independent variables using *Datastreami* programmes numbers 101S and 900B. The BE figures were taken from the 'total share capital and reserves' entries in the *Datastream* balance sheet database. I found the data for the ME/BE ratio using the analysis program number 101S. I was surprised that the actual figures were in the database of technical data, rather than my having to calculate them. This could indicate that the ME /BE ratio has already been commonly adopted as a tool by technical analysts.

From the Financial Times that week, I found out that none of these companies had their financial year-end during the week so this data would be as up to date as possible.

Econometric Analysis

In this investigation, I will reject the model as unworthy of further investigation if it explains less than thirty percent ,using Samuelson's criteria for a good financial model Bernstein (1992). Anything less than that would probably be too risky to use as a basis for investing my own money. I am trying to estimate the parameters of the equation

Yi = 0 + B1Xi + B2X2i + i

Results

I ran the cross-sectional regressions on the HUMMER econometric package. The coefficient of correlation (r^2) was very low, less than 2%. This seems to indicate that the significance of my chosen X-variables in determining the observed return was negligible. The relationships observed seemed not to comply with my model. The coefficient for ME was approximately zero while that of the ME/BE ratio was -0.3. I then carried out a regression of one X-variable against the other. Again the r^2 was tiny, about 0.0001%. This indicates a low degree of collinearity between the independent variables. A final set of regressions showed that either ME alone or the ME-to-BE ratio alone is insufficient to explain returns, with r^2 of 0.2% and 1.13% respectively. The coefficients were approximately equal for the independent variables in the multiple regression and the single-variable regressions.

Conclusion

It is possible that my results are so flawed that I cannot inductively draw conclusions inductively from them. This could occur if the variable RETURN, as empirical studies suggest, cannot be approximated by a Gaussian distribution Peters (1991). This could be verified by a Q-Q test, which involves regressing the distribution of observed variables against those expected in a Gaussian distribution JP Morgan (1994). Another plausible explanation is that my model is misspecified, so that the two independent variables I chose do not apparently determine cross-sectional return. Consequently, it is possible that the result obtained by Fama & French (1992) is spurious. It is possible that this could be due to data-mining. If the CAPM is disproved, several econometric studies that contradict the Efficient Markets Hypothesis (EMH), of which Fama is a leading proponent, can be dismissed as being based on bad data. An example would be Basu's (1981) work on the 'P/E effect'. He states that stocks with low price-to-earinigs (P/E) ratios will return more than the level predicted by the CAPM.

From the fact that ME was shown to be uncorrelated with returns in this sample, one could infer that the variable, and hence the CAPM, does not produce an accurate estimate of risk or expected return.

The CAPM is used for determining the level of profits a firm should earn, given the risk of its business Carlton & Perloff (1994). The degree of market power held by the firm can then be estimated by comparing actual profits with the profits the CAPM says the firm should earn for the level of risk undertaken. Hence, if CAPM is incorrect, as my results suggest, anti-trust action based on risk-adjusted profits as a measure of market power could be either too zealous or too lax.

The function of capital markets is to provide funds for investment by firms and increase the aggregate level of risk-taking by lowering the risk to individuals. CAPM can play an important role in allowing us to tell whether this function is being carried out properly already or whether structural reform or new financial techniques are needed. CAPM is used to measure how much risk fund-managers have taken over a period, compared to the risk of an index-tracking fund, to achieve the given return. This provides a means of evaluating the skill of the fund managers in stock-picking or market-timing , in other words, how good they are at fundamental and technical analysis respectively. If this is done incorrectly, investment funds may be placed in the hands of incompetents who will then misallocate the money between good and bad investment projects, lowering the return from investment and lowering future national income.

As I stated above, CAPM is the foundation for many tests of stock market efficiency. Although the disproving of CAPM will discredit some tests that attempt to disprove the EMH, the hypothesis is threatened more by results involving the application of new techniques of investment analysis. On the other hand, perhaps the market is efficient, and some investors have taken note of the result published by Fama & French (1992), with the result that any potential profits have already been arbitraged away, thus undermining the original econometric relationship. This was the reason Fama (1966) gave for the lack of observable linear relationships in security price time-series data.

In summary, the evidence I have produced is of poor quality and conflicting conclusions can be drawn from it.

Biliography

Books

Bernstein,P (1992) : *Capital Ideas*, Macmillan, London,

Carlton & Perloff (1994): *Modern Industrial Organisation*, Scott, Foresman and Little.

Malkiel, B. (1992)*A Random Walk Down Wall Street*, Norton, London.

Peters, E. (1991)*Chaos and Order in the Capital Markets*, Wiley Finance Editions, New York.

Articles

Basu, S. (1981) : 'An Explanation of the Small Firm Effect', *Journal of Finance*

Fama, E. F.(1966) : 'The Behaviour of Common Stock Prices', *Journal of Business*.

Fama, E. F. & French, K.(1993) : 'The Cross-Section of Expected Stock Returns', *Journal of Finance*.

Financial Times, Dec. 12th to 18th., 1994.

Reports

Riskmetrics Technical Document, J. P. Morgan Inc., 1994.

An Econometric Model of Non-Stationary Data

Peter Lambert

The Classical Linear Regression Model is unfortunately not universally valid in econometrics. In this essay,

Peter Lambert examines one of these cases, that of non-stationary data.

INTRODUCTION

In the following report, I intend to set out the results of an attempt to model a non-stationary data series in terms of (mainly) stationary explanatory variables. In the first section, I will outline the results of a conventional analysis of the data ignoring the problem of non-stationarity. The second section will describe the details of the non-stationarity problem. In light of this I will review the conventional results, commenting on their validity. Finally I intend to reach a conclusion as to the applicability of standard procedures to non stationary data in this specific case and in general.

Data

I hoped to model the demand for electricity in Ireland in terms of its main uses. In particular, I looked at industrial use, heating and lighting. Each of these requirements varies over time and I intended to represent the effects of these changes on the demand for electricity. Industrial demand for electricity was represented by an index of industrial production ('OUTPUT'). There is an implicit assumption here of a stable relationship between industrial production and the amount of electricity used to produce it. While not true for longer periods as the ratio of capital to labour would change, over a five-year period I hoped that this effect would not be very significant. Meteorological reports of the average temperature ('TEMP') and the average sunlight ('SUN') for each month were used to represent the demand for heat and light respectively. Other uses of electricity such as home entertainment and cooking were not represented due to lack of data. It was assumed therefore, that this 'other domestic' element was constant over the time period in question. Although data was available over quite along period, a sample size of sixty-six months was considered sufficiently large for this analysis without being too cumbersome.

Software

Access to computer resources was limited. Econometric software such as PCGive and Microfit were available while I was testing preliminary models, but the actual model selection procedure and testing were carried out without the benefit of powerful econometric packages. Wallace and Silver's 'HUMMER' programme allowed basic regression to be carried out and provided a limited amount of statistical information. Microsoft Excel spreadsheets provided even less statistical information than HUMMER through their regression command, but it allowed easier calculation of further statistics. Excel also provided greater accuracy, with up to 15 significant figures. For these reasons most of the estimation work was carried out using Excel. Combining the output of its regression command with its calculation abilities allowed most tests to be conducted. I also used Excel's matrix multiplication facilities to solve (XTX)-1XTY manually in some cases (this allowed me to find the t-statistics for the parameter estimates.) HUMMER was used for the less critical and more regression-intensive Dickey-Fuller tests.

Conventional Analysis

Model selection

The procedure used for model selection was primarily a 'general-to-specific' one. A very general model was first estimated and irrelevant explanatory variables were then excluded. Suspicions that there may be a lag of some time between changes in an explanatory variable and adjustments in the explained variable led to a distributed lag model. A further suspicion of an autoregressive component led to the use of an autoregressive distributed lag model. In keeping with the general-to-specific approach, it was intended that these hypotheses could be dropped should those suspicions prove unfounded.

As the data series were in monthly time series form, it seemed most appropriate to include lags of up to twelve periods on each of the Y variable and the main X variable. With only sixty-six observations it was not feasible to include such lags on the other X variables, though this may have been desirable.

Starting with this general model of 29 explanatory variables, variables were excluded one by one starting with those with the largest t-statistic. Eventually the null hypotheses of a variable's insignificance was rejected for all explanatory variables at the 5% level. Interestingly, SUN was one of the variables excluded in this way. The following model resulted. The reduction from 29 to 8 explanatory variables obviously involved a cost in terms of explanatory power, but this loss was quite small: the R2 figure fell by just 0.014 to 0.9749. The Rbar-squared figure remained constant.1

Table 1: Regression Results

Variable	Coefficient	(XTX)ii	SE	T-Stat	~t49

Constant	81.4394	35.5259	4.2389	19.2124	0.0000
elec-3	-0.3879	0.0030	0.0388	-10.0045	0.0000
elec -7	-0.4326	0.0095	0.0692	-6.2469	0.0000
output	0.0317	0.0002	0.0104	3.0626	0.0036
output-6	-0.0714	0.0002	0.0107	-6.6704	0.0000
output-9	-0.0344	0.0002	0.0108	-3.1844	0.0025
temp	-0.5512	0.0072	0.0603	-9.1462	0.0000
t	0.2992	0.0011	0.0234	12.7946	0.0000

Theoretical note

The autoregressive component of this model ensures that we cannot assume independence between the disturbance term and the explanatory variables:

E(ytt-) 0, for 0

This is very serious for small samples, but for the purposes of this analysis I will invoke a combination of the Mann-Wald theorem and Cramérs theorem which show that for large samples the OLS distribution theory is still asymptotically valid.

Misspecification tests

Functional form

A number of tests were carried out to check the validity of omitting certain variables and for the validity of the model as a whole. The results of these tests are shown below.

Table 2: F-Tests

F-Tests

Null Hypothesis (H0)	DoF 1	DoF 2	URSS	RSS	F-Stat	Prob.>F
Zero Slopes	7	49	24.7831	986.833	271.7319	0.0000
No Lags	4	49	24.7831	80.9449	27.7602	0.0000
General ardl	21	25	11.3367	24.7831	1.4120	0.2034
reset (Y2)	1	48	1.7977	24.7831	613.7357	0.0000

In the first test, the null hypothesis that all of the variables were irrelevant (have coefficients of zero) was tested against the alternative that they do not. This hypothesis is rejected as there is only a tiny probability of its validity. Similarly in the second test, the null hypothesis that all of the lagged variables are irrelevant is rejected. For the third test the simplified model is tested against the alternative of the original general ARDL model with all of the lags. This test does not reject the simple model at the 5% level of significance. The RESET test tests for the validity of including Y2 as an explanatory variable. This hypothesis is definitely not rejected at the 5% significance level. This suggests that there exists an alternative model which would better explain the Y variable. Unfortunately I was unable to find such a model in the time available.

Normality

Based on asymptotic theory, we can test the for non-normality by looking at the standard third and fourth moments of the residuals. Looking individually at the statistics for skewness

Table 3: Tests for Normality

Normality

Test	Statistic	Distribution	Value	Prob. >
Skewness	Root b1	~AN(0,9.5)	0.10	0.83
Kurtosis	b2	~AN(3,0.42)	2.47	0.10
Normality	N	~2	0.78	0.68

---

and kurtosis we find that in neither case is the null hypothesis of normality rejected. Similarly, a test based on a standardised combination of the two statistics ('N') does not reject it either.

Heteroscedasticity

Graphical techniques show few signs of heteroscedasticity. There were no noticeable trends in the error terms graphed against time or any of the other explanatory variables, nor against the explained variables. A Goldfeld-Quandt test gives a more subjective answer. This showed only a small difference between the R2 statistics for separate regressions for the first 19 observations and the last 19. The F-statistic of 1.860911 is less than the 5% critical value for (11, 11) degrees of freedom (2.817927), so the null hypothesis of no heteroscedasticity is not rejected. Both the graphical techniques and the Goldfeld-Quandt test used here test only for time-related

heteroscedasticity. The Breusch-Pagan test is more general, testing for possible heteroscedasticity involving the original regressors. Again this does not reject the null hypothesis of no heteroscedasticity.

Autocorrelation

Finally, the Durbin Watson-statistic for this regression is 2.30, well above the upper bound for this test when k=8 and n=54. This indicates rejection of the hypothesis of positive serial correlation. This is supported by the first-order serial correlation coefficient of only -0.17.

Conclusion

The model outlined above has good explanatory power (high R2) and passes most of the usual diagnostic tests (except of course the RESET test). This would appear to be a successful attempt at modelling explaining the demand for electricity in terms of an index of industrial production temperature and time. The RESET result indicates that a better model could be found which would be far more powerful, but in its absence this model appears to be a good alternative.

Non-Stationarity

Implicit in the above analysis are certain assumptions which I have neglected until now to prove. Here I will show that these assumptions are not in fact valid and examine the implications of this result for the conclusions of the preceding analysis.

The Problem

Looking at a correlogram for electricity demand in figure 1, it is quite clear that the sample autocorrelations do not 'decline relatively rapidly' as the order of autocorrelation is increased. In fact it follows a clear trend, rising close to unity for orders which are multiples of twelve and then falling again in between.

Figure 1: Correlogram

Failure of the autocorrelations to collapse is an indicator of non-stationarity in the data. It is not proof. To prove the existence of this problem Dickey-Fuller tests were carried out on the data.

Table 4: Dickey-Fuller Tests

Variable	Lag	Std.	t-stat	5% Critical
		coefficient	Error	value2
elec		-0.20913	0.0795	-2.6315 -3.4500
output		-0.69748	0.1251	-5.5772 -3.4500
temp		-0.20633	0.0780	-2.6464 -3.4500

As we can see the null hypothesis of a unit root is only rejected for the output variable. Both ELEC and TEMP may have unit roots.

Implications

The implications of this problem are devastating. Asymptotic theory has been shown to be completely different in cases of non-stationarity from the usual textbook asymptotic theory. R2 has a non-degenerating limiting distribution while the Durbin-Watson statistic converges to zero (one would expect therefore that the R2 values which I found might have been accompanied by a lower Durbin-Watson statistic.) More specifically, a non-stationary regressor cannot fulfil the condition

plim (n-1X ij 2) = constant

This assumption is crucial to all of the statistical inference made above. It embodies the 'standard regularity conditions'. These allow us to show that the OLS estimator is asymptotically normal. Without this condition we can say nothing about the distribution of the OLS estimator or that of the OLS residuals. All of the tests reported in section one are therefore invalid. Each of them made implicit assumptions about the distribution of the OLS estimator or its residuals. These assumptions cannot be shown to be valid and in the absence of any proven valid theory we cannot proceed with the tests.

Conclusion

Non-stationarity of the explained variable means that no valid statistical inference can be made about any autoregressive model, as these include at least one non-stationary regressor. A number of alternatives exist. It is possible that TEMP and ELEC are co-integrated, this would allow us to model some linear function of them which was stationary. Alternatively, as both TEMP and ELEC appear to be seasonal it is possible that the seasonal difference version of them would be stationary.3 Unfortunately I do not have the resources to re-estimate the model taking into account the problem of non-stationarity. I must conclude therefore that the non-stationarity of the data used in this model means that it is not possible to estimate except in special cases such as co-integration.

Notes

1 An alternative model was also estimated which maximised the level of Rbar-squared by only excluding a variable if its inclusion would increase the Rbar-squared figure. This model had seventeen explanatory variables, though, and achieved a gain of only 0.01 in the adjusted R2

2 The 5% critical value here are in fact those for a sample size of 100. Those for smaller samples are smaller in absolute value.

3 I have checked this possibility and found that it does give a stationary Y variable, but TEMP still has a possible unit root.

Diamonds Are Forever: An Econometric Investigation

Muireann A. Kelliher (Junior Sophister) [1]

The diamond cartel now operated by De Beers was set up by the South African mining magnate, Sir Ernest Oppenheimer, in 1934. It is now the vehicle through which over 80% of world sales of rough diamonds is administered and marketed. In this paper, Muireann Kelliher explores how one company has come to dominate an industry.

'People buy diamonds out of vanity and they buy gold because they are too stupid to think of any other monetary system which will work - and I think vanity is probably a more attractive motive than stupidity.'

The above statement was made by Harry Oppenheimer, current chairman of De Beers and a man more aware than most of the amazing illusion surrounding diamonds that De Beers has managed to create and sustain. Diamonds are in fact cheap to produce and would be lower in price but for the global cartel operated by De Beers. Its aim is to maintain a strong monopoly position, an objective which it has successfully achieved for several decades, notwithstanding the pundits' predictions to the contrary. In controlling the diamond market, De Beers have exploited a relatively simple idea: put an armlock on production and keep prices high. What makes De Beers so special is its execution of the idea. Over the past 60 years the cartel has done for diamonds something that eluded the oil producers of OPEC and even the cocaine barons of the Medellin cartel. It had the muscle and the nerve to impose its own order on the market and it built a syndicate not for weeks or months but for decades.

This paper will attempt to examine these diamond prices as set by De Beers. To this end both the market structure and control system as operated by the cartel is introduced, before the choice of variables and model is specified. The third section presents the results of the regression analysis, and in the final section these results are evaluated.

Market Structure and Control System

Gem diamonds, as distinct from industrial diamonds, are different from most other minerals in that they are non-homogenous. Unlike gold for instance, they do not have a standardised unit price. Each diamond has to be individually classified and valued and different diamonds of the same price are not necessarily tradable substitutes. At the micro-level, the diamond market is not a single market, but embraces many sub-markets, with different prices, supply and demand characteristics. Price differentials exist of course, in most commodity markets but the lack of homogeneity in diamonds is such that market segmentation is much more of an issue than in otherwise comparable situations. Owing to the wide variation in diamond qualities, the volume of a producer's output is not a reliable indicator of the value of his production. As a result the interests of producers obviously vary greatly. Due to the considerable differences in the scale of production, producers are not solely concerned with market share. Marketability is a major consideration and co-operation between producers can be a strategy for maximising revenues. This is because different productions are often complementary and different types of stones are easier to sell at different times. A mix from different sources may be more marketable and realise a higher long-term average price than could individual marketing. What drives the major rough diamond producers to co-operate is that there is no guarantee that independent marketing will realise higher sales revenues than centralised selling over the life of a De Beers' sales contract.

Certain key characteristics of the market as described above render it suitable to cartelisation. Firstly, there are a small number of significant suppliers and rigorous barriers to entry. Secondly, the product is durable, has a high value to volume ratio and is easy to store. Thirdly, the demand for jewellery-quality gems, which make up 90% of the $5 billion rough diamond market is relatively price inelastic. The rewards to a cartel

from controlling supply are therefore likely to be considerable, provided that the difficult tasks of establishing such a cartel and managing it can be accomplished.

Operation of the Cartel System

De Beers and its associate companies buy of rough diamonds from the mines, value them and then sell them to 'sightholders'. The policy of the cartel is to maintain stability in both of price and supply, flattening out short-term fluctuations in the demand-supply equation. The central thrust of its strategy is to exercise control over the market and thus maintain rough prices at the highest sustainable level, in essence aiming for long-run revenue maximisation over the whole of the demand cycle rather than short-run market clearing at spot prices. Unlike other commodity cartels, the diamond cartel it controls both supply and influences demand, combining the roles of major distributor, marketing agency and buffer stock manager. It has developed an expertise in matching supply to demand and the financial strength to operate an extensive buffer stock, capable of holding rough diamonds temporarily off the market.

The cartel has six key characteristics. The first of these is the existence of producer quotas. Most significant producers have a long-term and exclusive contract to supply a certain proportion of De Beers annual sales. This proportionate contract ensures that the burden of weak markets is passed directly back to the mines. This commitment by the cartel to purchase regardless of the state of demand, provides guaranteed cashflow and price stability, which induces the manufacturing trade to continue buying confident that the market has long-term support. Thus, in a rising market the cartel benefits from both higher prices and stock appreciation as goods are sold from the buffer stock. Yet in a weak market, it bears the full brunt of financing the buffer stock and its commitment to purchase from producers.

Secondly the cartel has created a strong antidote to any producer's incentive to cheat, with the ability to release from its own stocks a supply of any type of diamond. This means that the stockpile-supported price can drop dramatically as the market is flooded with a similar type of stone, released from the buffer stock by De Beers. This punishment was enacted against Zaire in 1981 when it chose to market its production independently of the cartel to traders in Antwerp. Market clearing of similar type stones lead to substantial price discounting in that segment which took a considerable length of time to work itself out, even though Zaire had been considered the dominant producer in its market segment. Thirdly, De Beers acts as a swing producer. It can play this role credibly since its own mines are among the cheapest sources of fine diamonds in the world.

A fourth element of market control is operated at wholesale level. The cartel's role as buffer stock manager is reinforced through its external buying offices, competing with independent traders for diamonds mined outside its own production network. When markets are weak it mops up excess supply by outbidding the independent traders and conversely allowing them to bid as they wish when markets are strong.

An important element of the system is the participation of rough diamond cutters, the middlemen. On every fifth Monday of the year, a select group of diamond dealers is invited by De Beers to the London Headquarters of the cartel, to collect a sack tied with a ribbon. Inside is a cardboard box containing a selection of uncut diamonds worth on average between $2-$5 million, which more or less correspond to the dealer's prior request, but ultimately the choice of diamond supplied in the box is at the discretion of De Beers. A box must be either accepted or rejected in full and price haggling is not permitted. This ritual is called a 'sight'. The invitation to a sightholder to attend is considered a privilege within the industry, with a corresponding loss of status if permission to purchase is withdrawn.

Finally, the cartel pays careful attention to demand management, operating a, highly successful worldwide advertising campaign. This campaign is aimed at attracting the consumer's attention to the particular type of stone De Beers needs to sell, corresponding to the composition of the buffer stock at a point in time.

Specification

Dependent Y-Variable

Given the context of the earlier discussion, my dependent variable in this investigation is the wholesale price of diamonds between 1976 and 1990. More explicitly this corresponds to the price paid by De Beers to the

mines when it makes the level of purchases agreed in the contract. During the 1970s commodity prices in general boomed. Gold, silver, platinum, art objets and diamonds all soared in price. Diamond prices rose sixfold, with the investment motive strong and the trade suffered from bouts of trade speculation. Market conditions changed rapidly in the 1980s. The early 1980s brought Reagonomics, a stronger dollar and a preference for financial rather than physical assets. A global recession ensued and the diamond market spent the best part of the next decade recovering and adjusting to a new set of circumstances.

Independent X-Variables

As already detailed diamond prices are artificial in the sense that they are set by a controlling interest and are supported by a stockpile. As such numerous variables may be employed in the price setting decision. Thus, it can be seen that by limiting my investigation to two independent variables, my analysis is going to be imperfect due to the absence of many other possible determining factors. Nevertheless, I have chosen two variables to examine their respective influences on diamond prices.

My first X variable, X1, is the annual volume of diamonds produced at mine level. The variable was chosen as all major producers operate under a long-term, quasi-exclusive contract to De Beers providing a certain proportion of the cartel's annual sales. Thus, weak market conditions directly affect production at mine level. The production figures used here are introduced with a one-year time-lag to diamond prices. World rough supplies more than doubled during the 1980s, so it is particularly interesting to examine how De Beers coped with this huge increase, coupled with a slowdown in demand. One would a priori expect a negative relationship between price and supply. Investigation should detail to what extent the cartel has managed to control this.

My second variable, X2, is the US/Yen exchange rate. This variable was chosen because of the growing importance of Asian demand in the diamond market. In the mid-1980s, high US real interest rates and the overvalued dollar were beginning to depress world demand. In 1985, following the Plaza Accord, the major economies embarked on a policy of economic co-operation and the dollar started to depreciate. The market for larger diamonds benefited and the growing Asian markets, notably Japan, Taiwan, South Korea and Thailand, started buying. Prices started rising as the dollar fell against the yen and other hard currencies. Thus one might posit a negative relationship between the exchange rate and prices.

The Model

I am employing the ordinary least squares estimation technique in this investigation, which will yield a line of 'best fit' corresponding to the data. The model takes the following form:

Yi = beta0 + beta1X1i + beta2X2i+ epsiloni

for the multiple regression case, where i represents the error term of the regression or the residual. The investigation will arrive at an estimate of the sign, size and significance of the unknown parameters beta0, beta1 and beta2.

Estimation

The Data

The estimation of the regression line was achieved using the HUMMER econometric package, using 15 observations of annual data, dating from 1975 to 1990 with a one-year lag for production. The diamond market is highly secretive and prices at wholesale level are virtually impossible for the outsider to acquire. Thus, I have constructed a price index (base year 1976=100), using the price increases for rough diamonds as supplied to me by De Beers. It must be reiterated that these refer to wholesale level where an amazing feature of the system is that De Beers has never announced a decrease in nominal terms. Prices are in dollar terms and the index has been deflated using World Bank deflators. The production figures are measured in millions of carats and were obtained from an industry journal, Mining Annual Review. These should be accurate as they are updated on an annual basis where necessary to ensure precision. The exchange rate series was obtained from the Financial Times.

The Estimation

The results of the multiple regression are detailed below. The line of Best Fit has been estimated as:

Y = 2.19391 - 0.005276X1 - 0.004593X2

The correlation coefficient, R^2, indicates that 75 percent of the variation in Y can be explained by the linear influence of all the independent variables.

Independent Parameter Estimate t-Statistic
Variable Ho: i = 0
constant 2.919391 9.13874
X1 0.005276 -2.40822
X2 -0.004593 -5.21805

R2 = 0.7598

This is quite a satisfactory result but as with any econometric result it cannot be taken in isolation. By regressing Y on X1 alone, an R^2 of 0.21537 was yielded, showing that 21 percent of the variations in diamond prices were explained by variations in production figures.

Independent Parameter Estimate t-Statistic
Variable Ho: i = 0
constant 1.322931 8.28737
X1 0.004107 1.88899

R2 = 0.21537

By the same procedure on X2, an R^2 of 0.64398 was yielded, showing that 64% of variations in Y are explained by variations in X^2.

Independent Parameter Estimate t-Statistic
Variable
constant 2.194303 17.56708
X2 -0.002853 -4.84918

R^2 = 0.64398

Evaluation

This evaluation shall deal first with the results of the multiple regression, followed by an examination of the individual regression results and ending with a statistical evaluation of these results. In evaluating the results yielded by the regression, I intend to compare the estimates with the a priori assumptions of the author. As expected, in all, the two X variables, have a considerable impact on diamond prices (75 percent). The relationship with X1 (production) is negative as expected, as is that with the exchange rate. Examine the results more closely and the picture is not as clear-cut. A separate regression of the independent variable X1 on the independent variable X2 reveals an R^2 = 0.67365, indicating a high degree of muticollinearity, the problems engendered by the presence of which are a priori recognised by the author though they shall not be elucidated on here. Indeed, this is evident to such an extent that an individual analysis of Y on X1 yields a differing nature of relationship between the two than was revealed in the multiple regression case - this time indicating a positive relationship.

Statistical Evaluation

In order to make a relatively simple statistical evaluation of my results, I am going to examine the accompanying t-statistics in the previous tables. Considering the hypothesis that there is no relationship between the X and Y variables (Ho: i:= 0), one can examine the t-statistic which measures 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. From the multiple regression, the estimates for both i and i, at both 5 percent and 10 percent significance levels are statistically significant. The estimate provided by regressing Y on X1 is statistically insignificant at a 5 percent level, but significant at a 10 percent level with a t-statistic of 1.88899. Examining the regression of Y on X2, the t-statistic is -4.84919, showing a high degree of statistical significance. This is a strong indication of a causal relationship. It suggests that the dollar/yen exchange rate plays an important role of diamond prices.

Conclusion

In the markets of today, anti-competitive rules are outlawing commercial agreements like cartels. But the legal competence of the nation state can only extend so far. Cartels, historically speaking, given the stringent conditions necessary for success, do not have a high survival rate. De Beers success in this environment is testimony to its ingenious domination of the trade. Unlike other commodity cartels, it both controls supply and influences demand, combining the roles of major distributor, marketing agency and buffer-stock manager. It has developed an expertise in matching supply to demand and the financial strength to hold diamonds temporarily off the market. But the cartel can never be infallible. Several factors in the present environment could pose a potential threat to its commanding position Economist Intelligence Unit (1994). De Beers sway over customers for rough diamonds is not matched by its influence on customers for polished diamonds, namely, jewellers and their suppliers. They are increasingly well organised, good at hard bargaining and make it difficult for De Beers set prices to pass smoothly through the system to the end buyer. Russia always adds uncertainty to the diamond market. Rumours of a Russian stockpile persist and increasing numbers of Russian stones have found their way to the market outside of the 'single channel' as operated by De Beers. Assuaging this threat will require all the political and diplomatic clout of the corporation. Yet, there is a degree of consensus within the industry that centralised selling and buffer stock management are in its collective and long-term interest. The diamond is the strongest material known and it is therefore fitting that the diamond cartel has been so resilient slightly ironic that the diamond cartel is not only extremely strong but also extremely old. Only time will tell whether De Beers will continue to control this market and whether diamonds really are forever.

Notes

[1] The author wishes to gratefully acknowledge the contribution of Professor Dermot McAleese to an earlier research project on the same subject.

Bibliography

Economist Intelligence Unit (1994) 'Diamonds: A Cartel and Its Future'

How does Econometrics contribute, if at all, to the scientific status of economics?

Gavin Falk

In this paper Gavin Falk discusses the role for econometrics, if any, in adding to the scientific nature of economics.

'The justification sometimes advanced that a multiple regression analysis on observational data can be relied upon if there is an adequate theoretical background is utterly specious and disregards the unlimited capacity of the human intellect for producing explanations by the carload lot.'

- K.A. Brownlee

Introduction.

Econometrics was developed as a means of testing economic theory empirically. It was thought that by developing an empirical means of analysis, the scientific credentials of economics would be firmly established forever. On the contrary, the disputation as to whether economics is a science or not has become more contentious than ever before. This essay begins by presenting a definition of econometrics and discusses some of its unignorable limitations. Part two discusses the scientific status of economics and the different attitudes that exist towards econometrics. Finally, part three presents the usefulness of econometrics as a tool of policy makers and as a means of deepening our understanding of economics today.

What is Econometrics?

'A science brings together a great number of similar facts and finds that they are special cases of some great uniformity which exists in nature. It describes this uniformity in a simple and definite statement or law,'Thiel (1971).

Econometrics can be defined as the study in which the tools of economic theory, statistical inference and mathematics are systematically applied, using observed data, to the analysis of economic laws. It is therefore concerned with the *'empirical determination of economic laws'*, Brown (1991). Economic theories are written in mathematical form and are then analyzed using statistical methods. If the observed data are found to be incompatible with the predictions of the theory, it is rejected. Theories are accepted if the data are found to fit the theory.

Although this process may appear to be essentially simple and straightforward, there are many problems intrinsically related to every econometric study. The biggest problem econometricians must overcome involves the 'data generating process' (DGP). The data generating process is *'the underlying behavioral function that actually produces the data that are being analyzed'*. In economics, our data is provided to us by nature. Often, there is insufficient data available, which means the theory cannot be tested in the first place. Other times, the available data is inaccurate which leads to flawed results, which have little or no value. Even assuming that accurate data is readily available, we still cannot be certain that we will be able to perform an analysis. Many economic theories may not be testable with econometrics. In the real world, many variables are inter-related, a consequence of this is the problem of multicollinearity. Multicollinearity is almost always present in multiple regression models, especially those using time-series data. This can also lead to flawed results as it is difficult, sometimes impossible to solve the problem by adjusting the data. Unlike multicollinearity, specification errors can lead to biased parameters, potentially a very serious problem. Other problems which econometricians encounter include those of identification and aggregation.

Econometrics therefore provides us with a quantitative basis to aid the formulation and simulation of economic policies and the provision of economic forecasts. Its limitations however must not be ignored.

The Scientific Status of Economics.

The traditional method of pure economics has been deductive, theories being deduced by logical reasoning based on the behavior of representative consumers making rational decisions over time. As an alternative to this deductive method, some economic problems are approached on inductive line. In this instance, a mass of data is obtained from actual experience, and then used as a basis for generalisations.

The question has to be raised, during any discussion about the scientific status of economics, as to whether or not it is possible to have a scientific study in a field of human behavior. Before the development of econometrics, the natural scientists rejected claims by economists that economics was a scientific study, on the grounds that the ability to test empirically is a precondition for scientific status. What is at issue here is the whole inductive method of economics, of finding in Marshall's words the *'great uniformity'* when facts are collected and arranged. Of course all modern economic analysis proceeds not only from the obvious facts of human behavior, but also from the mass of statistics which are available to us today. So, in theory, economics now has the necessary credentials to be truly 'scientific'. It is based on logic, theories can be empirically tested, and accurate predictions about future economic developments can be made. In practice, econometrics is not perfect. Economists are divided as to the usefulness of econometrics. Two schools of thought have emerged.

One contends that economic theory can never achieve true scientific status in the 'natural' or 'physical' sense of the word. Their argument is that all economic models will have omitted variables, because *'everything depends on everything else'* in economics. The actual variables used will depend on what was believed beforehand, the result being that all parameter estimates are potentially biased. This in turn leads to biased predictions. In this light, economics can never overcome its inherent subjectivity and can never produce consistently accurate results. Scientific status, akin to the natural sciences can never be achieved.

The other argument claims that as econometrics is merely a process of 'number-crunching', it is an objective method for carrying out empirical testing. As long as there is no manipulation of the data to suit a particular theory or point of view, then consistently accurate results are obtainable and economics deserves full scientific status.

Uses of Econometrics.

In my opinion, econometrics has not contributed to the scientific status of economics. While natural scientists do not claim to be able to empirically test all their theories, they are generally able to acquire more consistent and accurate results than those acquired from the research of econometricians. This accuracy and consistency also leads to more accurate predictions. Economics however is based on psychological grounds. There is enormous uncertainty regarding future events and expectations which are largely unquantifiable.

The ongoing development of econometrics has not solved these inherent problems. In fact, the converse has occurred. Econometricians are faced with a multitude of difficulties with every analysis. One of the consequences of these difficulties is that econometricians are wont to try different functional forms, lag structures, and so on until the desired results are obtained or at least results that are interesting enough to warrant publication. This practice, termed 'data mining' is common, although considered unethical. Data mining certainly does not contribute to the validity of econometrics as a science. Every result obtained is imperfect and widely open to subjective interpretation. It is logical that accurate predictions are unobtainable without accurate results. This analysis may seem to present a gloomy picture of econometrics but that the subject is exceedingly complicated does not entail that it is hopeless. Even with all its weaknesses econometrics is still a very useful component of the field of economics.

In my opinion, the fundamental use of econometrics is as an aid in the policy process. However it is essential to remember that econometric analysis is prone to error. Any predictions deduced from the results of an analysis must be seen only as guidelines and not definitive reflections of the future. In economics we are only certain of what has gone in the past. Econometrics is not a crystal ball that allows us to see every fluctuation in the economy before it happens. Policy makers have tended to forget this and in so doing contributed to economic crises such as the massive inflation of the early 80's or the stock market crash and recession of the late 80's. When used logically and correctly econometrics is an invaluable tool available to policy makers in formatting and evaluating economic policy.

Conclusion.

Economics is a social science which studies the behaviour of consumers and firms. Expectations and uncertainties play key roles in the translation of economic theory into real world economics. These uncertainties mean that econometrics cannot supply us with a perfectly accurate model of the workings of the economy. If econometrics is used with these limitations kept in mind, it can be extremely useful in increasing our depth of knowledge and understanding of the economy. If however, we attempt to use econometrics in the same manner as the 'natural sciences' use controlled experiments, without respecting its inherent limitations of methodology and data collection, we are negating its value and running the risk of making totally inaccurate and useless predictions.

Bibliography

Brown, W.S., (1971) :*Introducing Econometrics*, West.

Thiel, H., (1971) :*Principles of Econometrics*.

Does Europe Need a Monetary Union?

Daniel Gallen (Junior Sophister)

'A single currency is the cement that binds our economies together' - Sir Leon Brittan'

'The only guarantee there will be a parallel between economic and monetary policy is a political union.' - Jacques Delors

---

Introduction

EMU, denoting *'Economic and Monetary Union'* means the creation of a single market in commodities, capital and labour (economic union, now constituted) and a single currency (monetary union). No historical precedent exists for the current EMU process. There are cases where separate political entities formed a political union before adopting a common currency and single central bank, such as Italy following unification (1861). There are examples like the monetary union of the 1870s between Belgium, France, Italy and Switzerland, where special sovereign countries standardised their coinage but without establishing a common central bank. Yet there are no precedents for Europe's current course, in which countries with histories of monetary sovereignty and well established central banks form a common central bank accountable to them jointly with the control of their national monetary policies, including the power to issue a common currency. As Eichengreen (1993) elucidates, *'It would seem paradoxical that the European Community, the one part of the world that has succeeded in largely insulating itself from exchange rate variability, is where the call for monetary union was taken up.'* The objective of this paper is to examine whether a single currency is a technically necessary concomitant of a single market in capital, labour and goods. The essay is divided into three parts. The first section lays out the historical backdrop to EMU. Next, the advantages and drawbacks of proceeding with a monetary union will be weighed. The final section suggests some urgent measures that need to be taken in the transition to inevitable monetary union and concludes that the rationale for a single currency as a corollary of a single market derives to a large extent from political economy considerations.

Developments to Date

Since the 1969 Hague Summit, EMU has been an official objective of the European Community. In a period when economists were hopelessly divided over the advantages of a system of free-floating exchange rates, on the one hand, and the best road to monetary stability on the other, a decision was taken by Giscard, Schmidt and Jenkins to get out of the deadlock. The European Monetary System (EMS) with a European currency unit, the Ecu was proposed, came into force in 1979, and has been successfully implemented. Within the system, member states' exchange rates could fluctuate only slightly in relation to one another (by no more than 2.25% on either side of a fixed rate). This was the first step towards the ultimate goal of fixed exchange rates in the final stage of EMU. The first decisive step to economic and monetary union came in June 1988 when the Hanover European Council appointed a committee of experts (chaired by Jacques Delors) to examine ways and means of gradually bringing about EMU. The decision to set up this committee was promoted by the implementation one year earlier of the Single European Act (SEA), which brought about the completion of a frontier-free market at the beginning of 1993 (economic union). In June 1989, the Madrid European Council set the date of July 1990 for the beginning of Stage I (convergence) of economic and monetary union, a target which was met. The first stage saw the removal, with just a few exceptions, of all restrictions on capital movements between member states co-ordination and multilateral surveillance of the economic policies of the member states was intensified and cooperation between the central banks was stepped up through the Committee of Governors of the Central Banks.

The EMS had allowed long periods of exchange rate stability due to fixed rates, with periodic realignments redressing serious competitive problems. This was possible because capital controls protected central banks' reserves against speculative attacks motivated by anticipations of realignment. The Maastricht Treaty ruled

out their use from the beginning of Stage II. After undermining the viability of the 'Old EMS', monetary unification followed inevitably. This second stage (institutional) commenced at the beginning of 1994 when member states had to start even more intensive preparations for the final stage. The European Monetary Institute (EMI) has been created as a prefiguration of the European Central Bank. The ECB will be solely responsible for defining the Community's monetary policy This centralisation removes a priori any possibility of conflict between different monetary policies within the zone. This institution would also guarantee unlimited convertability of the different national currencies and the fixity of their parities. These currencies would thus be perfectly substitutable and by definition there would be a single monetary policy.

Guidelines in the Maastricht Treaty make transition to EMU conditional on several convergence criteria of national economic performance. Observed over a period of one year before the examination of the final stage, a member states average rate of inflation and nominal long term interest rate must not exceed by more than 1.5. and 2 percentage points respectively, that of the three member states who have the best results in terms of price stability. The deficit and total government debt for any member must not exceed 3% and 60% of GDP respectively. Finally, member states must have respected the normal fluctuation margins (2.25%) provided for by the exchange rate mechanism of the EMS without severe tensions for at least two years before examination. Before the end of 1996, the heads of state or government will decide, by a qualified majority, whether it is appropriate for a majority of the member states to move to the final stage of EMU. Even if by the end of 1997 no agreement has been reached for the beginning of the third stage, it will commence on January 1, 1999. The number of member states launching this final stage of economic and monetary union would not then have to constitute a majority. A mere four years before the 1999 deadline, however, public interest is focused on mass unemployment and the persistent recession rather than on feverish preparations for monetary union. Turning fifteen segmented European markets into an integrated economy whose constituents can specialise fully in producing goods and services in which factors of production can flow freely to wherever they gain highest returns is hard to challenge on efficiency grounds (economic union). Yet is there an economic logic for a single currency and a European Central Bank accompanying this process.

Benefits and Costs

Only a single currency completely eliminates the costs of exchanging currencies, a saving of as much as 19 billion ecu (or 0.5% of Community GDP). These gains are larger for small open economies like Ireland (0.9% of GDP) than for larger economies (0.1% to 0.2% of GDP). Even if exchange rates are irrevocably fixed, doubts about this fixity will encourage residents to use their home currency instead of a foreign one. There will be a complete elimination of exchange rate uncertainty inside the Community which varied from 0.7% in monthly variations for the original members of the exchange rate mechanism of the EMS to 1.9% for the other Community countries. The addition of single money to a single market will brighten the business climate considerably, given recent surveys provide strong evidence that foreign exchange risk is still considered a major obstacle to trade. Europe's business sector has declared an overwhelming preference for a single currency as distinct from a common thirteenth currency. While economists argue that compared to a floating exchange rate regime, EMU improves greatly on the stability of inflation and real economic activity. It also improves on the EMS; especially as regards the stability of real economic activity. Reductions in uncertainty will permit a reduction in the rate of return of investment demanded by shareholders, which in turn boosts investment and growth. Only a modest reduction in this rate of return of 0.5 % is already sufficient for a gain in output accumulating to 5% of GDP. The dynamic gains, occurring through a widened increase in investment, will strongly amplify the static efficiency gains and will be instrumental in leading the community to a higher sustainable growth path.

However, de Grauwe (1988) found exchange rate uncertainty not to be a substantial barrier to international trade. The IMF (1984) reached a similar conclusion. Since intra-EMS rates are already stable, the gains are unlikely to be vast. A lower risk due to less exchange rate variability can have a double negative effect. It reduces the real interest rate and it reduces the expected value of future profits of firms. Similarly de Grauwe (1992) indicates that consumer surplus is greater when prices are more variable. When prices are low, the consumer increase demand to profit from lower prices. When prices rise, he does the opposite, thereby limiting the negative effect the price increase has on his welfare. From this, we may deduce that if gains from a common currency and an ensuing drop in risk are to be expected, they will be found elsewhere than in these static welfare gains. Empirical evidence also suggests a weak link between monetary union and economic growth. The greater exchange rate stability that the EMS countries experienced in the 1980s did not provide a

great boost to the growth rates of output and investment. In fact, these growth rates have been larger in non-EMS countries that experienced relatively large movements in their exchange rates.

The rationale for retaining flexibility in rates of exchange rests on the assumption that governments aim to achieve both internal and external balance, and as Tinbergen (1952) has shown, to achieve these simultaneously at least an equal number of instruments is needed. Internal equilibrium is tackled via financial instruments (monetary and fiscal policies) while external equilibrium is sought after using the exchange rate. If countries deprive themselves of rates of exchange as policy instruments they impose on themselves losses that are essentially losses emanating from enforced departure from internal balance (Corden, 1972a). The central element of the solution to the problem of loss of autonomy is policy coordination on stabilisation and budgetary policies. This approach finds its theoretical underpinning in game theory (e.g. Hamada, 1985). In all types of games in which the policy of one country affects the variable making up another country's welfare function, better results are possible with a cooperative rather than with a non-cooperative attitude. Steinherr (1984) studies in detail the reasons why cooperation is the exception rather than the rule in a Monetary Union. The efficiency of the coordination process (consistent movement of partners) and the effectiveness of its outcome (credibility in markets) depend critically on the gradual reinforcement of the regulatory and institutional set-up (Molle, 1994). At any rate, in most cases, there is an alternative to using exchange rates. For example, when confronted with a loss of domestic competitiveness, countries can use contractionary demand policies aimed at regaining competitiveness, though such a measure can be more difficult than simply changing the exchange rate.

Moreover, Allen and Kenen (1980) and Allen(1983) have demonstrated that although monetary policy has severe drawbacks as an instrument for adjusting cyclical imbalances within a monetary union, it may be able to influence the demand for goods produced by member countries within the short term, provided the markets of the member nations are not too closely integrated. Their model indicates that economic integration in this sense, can be created as a consequence of the substitutability between nations' commodities, especially their financial assets. The moral of this is that the central bank of a monetary union can operate disparate monetary policies in the different partner countries without compromising their internal and external equilibria, a severe blow to those who stress the costs from monetary integration (El Agraa, 1990).

Another area to be examined is that of price stability. The cost of bringing down inflation to a low level is minimised if there is a credible commitment to stable prices. A monetary union cannot be sustained if inflation rates diverge. In the long run, higher than average inflation rates of EU countries inevitably lead to a corresponding depreciation of their currencies. Thus, inflation rate convergence and at a lower rate is necessary for the group of countries aiming at EMU. Low inflation is associated with low variability of inflation, and therefore of relative prices. The economic literature suggests that a 1% decrease in relative price variance could increase real output by 0.3% of GDP, for instance. The costs incurred in a disinflation process arise because of wage rigidity and a lack of credibility of the national authorities. A credible exchange rate commitment will minimise output losses by reducing nominal interest rates and accelerating the adjustment on wage and price setting, as exemplified by the French, Danish and Irish disinflation experiencesin the EMS during the 1980s. In addition, as I have outlined using Allen's analysis, changes in real exchange rates remain possible and desirable within EMU. This is why wage and price flexibility is a necessary condition of success. National exchequers will lose revenues from the inflation tax (seignorage) and gain from the lower interest rates to be paid on public debt. These benefits, which though of a transitional nature, could lie between 2% and 5% of GDP, will far outweigh the loss of seignorage revenue (0.5% in Spain/Italy, to 1% in Greece/Portugal) to be experienced by some countries. When a monetary union establishes a central fiscal authority with its own budget, then the larger the size of this budget, the higher the degree of fiscal harmonisation (MacDougall Report, 1977). This has many advantages; regional deviations from the internal balance can be financed from the centre, and the centralisation of social security payments financed by contributions or taxes on a progressive basis would have some stabilising and compensation effects, modifying the harmful effects of monetary integration (Corden, 1972a).

Furthermore, if a member country of a monetary union is in deficit, it can borrow directly on the union market, or raise its rate of interest to attract capital inflow and therefore ease the situation. However, the integration of economic policies within the union ensures that this help will occur automatically under the auspices of the common central bank. Since no single area is likely to be in deficit permanently, such help can be envisaged for all the members. Hence, there is no basis for the assertion that one country can borrow indefinitely to sustain real wages and consumption levels that are out of line with the nation's productivity

and the demand for its product (Corden, 1972a). The introduction of ecu will mean that equivalent securities become subsitutes and the 'law of one price' will be obtained in securities market throughout the European Unoin. This implies that debt can be raised at lower interest rates, which will reduce the cost of financing the National Debt.

EMU could tighten constraints on national fiscal policy, a benefit or a cost, depending on one's viewpoint. If capital and labour are freely mobile within the economic union, however, borrowing today which implies higher taxes tomorrow may induce mobile factors of production to move to lower tax jurisdictions, reducing the tax base. Investors know a government's ability to borrow now is limited by its ability to tax in the future, which in turn is limited by factor mobility. Therefore, they may refuse to lend to governments threatening to exceed their borrowing capacity. Yet the fact that certain factors will remain relatively immobile in Europe compared to say the United States, will allow greater variation in tax rates. Moreover, European governments require their central and commercial banks to hold substantial amounts of their debts~ so some fiscal autonomy will remain. The principal cause of fiscal restraints will be the clauses of Maastricht, (summarised earlier) which advise member states to *'regard their economic policies as a matter of common concern and (to)... coordinate them within the Council.'* EMU would also lower the costs of economic shocks. For example, simulation of the shocks from 1970-1990 (a turbulent period) under alternative exchange rate regimes suggests that EMU with a single currency might, compared to the EMS of the mid-1980s, witness a decrease in inflation fluctuations of one-fourth (0.6% - 1.5% less variation) and of output growth fluctuations by one fifth (0.3 to 0.7% less variation) (the ECU report, 1991). A symmetric shock, one which impacted proportionately on product markets throughout the EU, such as a commodity market shock, would require no policy response, merely resource allocation best achieved through decentralised market mechanisms. On the other hand, asymmetric shocks would be brought to a fast resolution with an exchange rate adjustment, which could quickly bring a change in prices that would otherwise only come from a prolonged and expensive deflation. Furthermore, the Commission professes that EMU will deepen the process initiated by the 1992 programme, that is intra rather than inter-industry trade, lessening the possibility of product specific shocks, eliminating the shocks that occur when domestic interest rates surge in response to speculative selling of the Irish pound associated with a sudden depreciation of Sterling.

The benefits of EMU will be particularly significant for smaller countries like Ireland whose economies are much more exposed to international trade. The foreign trade ratio (exports plus imports as % of GDP) is 115% for the Republic compared to 50% for the UK. Therefore, anything that reduces or removes the impediments to trade is of greater significance to small open economies. In addition, the current less well-off regions have a real opportunity for rapid catch up if they maximise potential synergies between EC policies and national development efforts. Economic and monetary union, like 1992, is a positive sum game (Commission, 1990). Whilst sovereignty may be lost at the national level, it is regained at the centre.

Finally, the importance of the Community in the world economy will be strengthened with ecu as a single currency, assuming the role of an international currency as powerful as the Yen or the U.S. Dollar, a larger weight of the Community in macroeconomic policy coordination, possibly evolving into a more balanced multipolar international monetary system. As an international currency, the ecu would allow a saving in the Community's exchange reserves of about 160 Billion ecu (allowing this financial capital to be allocated more optimally), increase the share of ecu - dominated assets in the world financial portfolio by 5% and increase the invoicing in ecus (with 10% of EC trade). This amplified use of the ecu would provide for a decrease in the transaction costs with third countries (up to 0.05% of EC GDP) and bring international seignorage revenues from a possible long-term accumulation of some 28 billion in ecu notes outstanding abroad (Ecu Report, 1991). However, the use of an integrated area's currency as a major reserve currency doubtless imposes certain burdens on the areas but in the particular case of the EU, it would create an oligopolistic market situation which could either lead to collusion, resulting in a permanent reform of the international monetary system, or intensify the reserve currency crisis and lead to a complete collapse of the international monetary order. The latter possibility is, of course likely to result in the former outcome; it is difficult to imagine that the leading nations in the world economy would allow monetary chaos to be the order of the day.

Since to a large degree, labour immobility and fiscal restraints will be prevalent in a European monetary union, the domestic economies must try to be flexible to cope with adverse economic shocks. To avoid large-scale emigration and rises in unemployment, and to benefit from monetary unification, we must end

restrictive practices, increase deregulation and encourage a stringent competition policy in the non-traded sector.

The Future

Some believe in light of a recent deterioration in Europe's economic circumstances (20 million unemployed), that the date for EMU should be postponed. After all, most countries have moved away from the Maastricht criteria thresholds. For example, Molle indicates that if 1993 had been chosen as the reference year, the inflation threshold would be 3.4% not 1.5%. Only six currencies would have met the exchange rate criteria at the end of 1993 (Germany, Denmark, France, Benelux, and Ireland) and three more did not participate in the ERM (Itlay, Greece and the United Kingdom). The 3% deficit-to-GDP ratio was achieved by just Ireland and Luxembourg in 1993 and in Belgium, Italy and Greece, the government debt-to-GDP ratio stands at a very high level(above 100%). Despite these shortcomings, a long period of transition to EMU is also undesirable. EMU cannot include all EC countries from the start. This is recognised by the provision that a majority of countries can go ahead on their own. Should those countries which are ready now go ahead and reap the benefits of *a 'small monetary union?'*

There could be a strong temptation to go down the path of a clear 'two speed ' EMU. For some countries it would be sensible from an economic point of view to tie their monetary fate to others, for example Benelux with Germany. Other EC countries may find it favourable to wait and see, until the benefits outweigh the costs. Britain is opting for the slow lane. Whilst there would be political difficulties in taking such a decision, although we must evaluate the cost to the individual member states, and to the collective endeavours, of doing nothing. Such a process may encourage Mediterranean countries to speed up convergence to enjoy the benefits of a monetary union. Doubts about Ireland's ability to enter EMU focus on our dependence on Britain, which is now in fact less than commonly perceived. The reliance is not 32% as often quoted, but 15% to 20%, as trade in agriculture and energy products is largely unaffected by sterling exchange rate movements. Moreover, entry by Ireland into EMU will further accelerate diversification from the UK market.

If speedy transition to EMU is impossible, as I believe it is, there are numerous temporary remedies. Larger and more frequent exchange rate adjustments during the transition to monetary union would be one option. This would require allowing exchange rates to flow more freely and fluctuate more widely. A more secure interim measure would be to request all institutions to take open positions in foreign exchange to make non-interest bearing deposits with their central bank to slow down speculation (Eichengreen and Wyplosz, 1993). This would provide time to organise realignments, ensuring the survival of the EMS over the remainder of the transition, Corden (1972b) suggests a *'pseudo union'*, where exchange rates are fixed but monetary policies are not fully integrated and there is no monetary authority. Yet fixed exchange rates, no matter how earnestly enforced, always raise the spectre of devaluation. To quote Portes in Eichengreen (1993), *'Permanently fixed exchange rates are an oxymoron'*. The single currency is an essential feature of EMU. In the same way as it is not acceptable that a single monetary policy should be determined by one of the existing national central banks, it is not desirable that the single currency should be one of the existing national currencies. Furthermore, it should not be a new thirteenth currency. As we move closer to Stage III, the basket definition of ecu will have less and less practical relevance. If all currencies perform well, the average is as good as the best and the ecu becomes de facto a hard ecu. This is already more and more the case.

Factor and producer-market integration could proceed under floating exchange rates as well as under a common currency. Yet with floating exchange rates and greater integration, we will move to a point where only the most efficient firms will survive. This is the basis of a single market and an economic union. Yet if national industries under pressure to remove trade barriers found their competitive position eroded further by unforeseen exchange rate swings, the opposition to the single market would rise sharply. Therefore, it seems to be for reasons of political economy, not economic efficiency, that monetary union is sought. Monetary unification serves as a stepping stone to deeper political integration. German insistence on strengthening the powers of the European Parliament in conjunction with progress on EMU supports this hypothesis. As Jacques Delors once pronounced *'Our community is the fruit not only of history and necessity but of political will.'*

Bibliography

BOOKS

Eichengreen, B., (1993) *'European Monetary Unification'* (1993) Journal of Economic Literature, Vol 31.

Eichengreen, B. and Wyplosz, C., (1993) *'The Unstable EMS'*

El-Agraa, A.M., (1990) *'Economics of the European Community'*

4. Bayoomi, T. and Eichengreen, B. *'Shocking Aspects of European Monetary unification'* (1992)

de Grauwe, P.(1988) *'Exchange Rate Variability and the Slow-down in the Growth of World Trade'*

de Grauwe, P.(1992) *'The Economics of Monetary Integration'*

Molle, W. (1994) *'The Economics of European Integration'*

Corden (1972a) *'Economies of Scale and Customs Union Theory'* , Journal of Political Economy, Vol. 80

Corden (1972b)*'Monetary Integration'*, Essays in Internatinal Finance, No. 93

Allen and Kenen (1980)*'Asset Market, Exchange Rates and Economic Integration'*

Allen (1983)*'Cyclical Imbalance in a Monetary Union'* Journal of Common Market Studies, Vol 21 No. 2

Tinbergen, J.(1952) *'On The Theory of Economic Policy'*

Agriculture and GATT: How the Compromise was Reached

Joanna O'Riordan

Agreement on free trade in agriculture was always going to be difficult, with the existence of two mutually exclusive systems on either side of the Atlantic. The GATT accord thus stands as a remarkable, if tardy, triumph for international negotiators. In this essay, Joanna O'Riordan examines the initial strategies of both parties and details how agreement was reached.

'The deal [GATT] will create a better, more prosperous and politically safer world, as it will mitigate the tensions that would otherwise exist between peoples.'

Peter Sutherland, Director General of GATT, quoted in The Irish Times (3/12/94) after the US Senate approved the World Trade Treaty.

Introduction

The GATT treaty currently being ratified by the governments of the contracting parties is the product of a lengthy series of trade negotiations which began in Punta del Este, Uruguay in 1986, Gardner (1993). The aim of those who initiated the current round of talks was to liberalise trade and decrease protective barriers throughout the world. It has been estimated by GATT that the gains from trade due to the Uruguay round deal will be $500 billion. This net welfare gain will accrue from increased competition and efficiency, economies of scale and higher world market prices. It is an understatement to point out that the deal was hard fought and no where is this more true than in the case of agriculture. In this paper I intend discussing the process by which agreement on the reducing of protective policies in agriculture was reached. The narrative is confined to the United States and European Union. This is for two reasons. Firstly because these two trading blocs represented the polar extremes in the debate on agriculture during the round and secondly because the other two major trading blocs with respect to agriculture, Japan and the Cairns Group (CG) largely hold the same views as the EU and the US respectively.

The first section of this paper will deal with the background to the Uruguay round, the treatment of agriculture in the previous GATT rounds and the reasons why it was essential that agriculture should be more fully incorporated into the GATT framework. In the section and third sections I will discuss the background to both the US and EU's involvement in the Uruguay round, the positions they adopted in Punta Del Este and their ambitions as to what the round would achieve. In section four I will link up the two viewpoints and outline the process towards ultimate agreement in Washington in December 1993. Finally I will conclude by assessing the impact of the round on world agricultural trade.

Background to the Uruguay Round

In previous GATT rounds agricultural trade was virtually ignored, despite the fact that trade in agricultural products represented 13% of total trade in goods and that extremely high protection levels were present throughout the world. While export subsidies and quantitative restrictions on imports were outlawed in other sectors, manufacturing for example, these regulations were waived with regard to agriculture. The absence of effective international rules and disciplines in evident in Producer Subsidy Equivalent (PSE)a percentages for agricultural produce for the US and EU throughout the early 1980's. The trend was clearly a rising one.

PSE percentages measure the total financial transfer to the farm sector. This figure is calculated by dividing the total PPS by the net value of production. In monetary terms the 1994 Agra Europe report estimated that the cost of supporting agriculture in the developed world amounted to $300 billion per annum during the 1980's.

The principal problem with such a high level of agricultural support was the fact that 82% was in the form of market price support. This had a number of critical consequences. Its adoption had the effect of both depressing and detribalising world market prices. The former conclusion is based on the simple logic that increased production, which leads in turn to excess supply. In the face of static demand prices will inevitably fall, thereby necessitating even greater levels of protection. Both the US and the EU are large net exporters of agricultural products. As a result of their protectionist policies, the rest of the world must bear substantial negative terms of trade. Protectionist policies to stabilise a country's agricultural prices also means the burden of adjusting to world market conditions is thrown onto the result of the rest of the world. This is shown diagrammatically below.

Due to the inelastic import demand curve of the protected domestic market the world import demand curve also becomes more inelastic. This has the effect of increasing the volatility of world market prices with particularly adverse effects for developing countries.

By the mid-eighties the budgetary and political costs of agricultural subsidisation and the extent to which consumers and taxpayers were bearing this burden were becoming increasingly well documented. It was estimated by Roningen and Dixit (quoted in Ingersentet al (1994) that for every $1 transferred to producers that consumers and tax payers had to give up $1.5. Put differently, half of the total value of agricultural transfers represented a deadweight loss to the world economy. In addition to anxieties from an efficiency perspective, the general malaise in world trade and the increasing tensions between the four main trading blocs (the US, EU, CG and Japan) resulted in agriculture not only being on the agenda in Uruguay but given top priority.

The US Perspective; Background and Aims

Modern agricultural protection in the US dates back to the thirties and the legalising of import quotas by the 1933 Agriculture Adjustment Act. For over forty years American agricultural policy was predominantly inward looking and protectionist. The farm lobby was very strong, but by the seventies there were indications that things were changing. Increased competition, lower farm prices and rural depression resulted in the US losing its position as leader in world agricultural exports worth $44 billion. Five years later that figure had slumped to $26 billion.

The US's firm line in the Uruguay round was as a direct result of its desire to reclaim its share of world trade. In order to achieve this the US proposed the phasing out of all trade distorting measures over a ten year period. It is important to note that the US government saw GATT as an excellent opportunity to pursue their domestic objective of reforming agricultural policy. Agricultural groups were still strong in Washington and it was felt that change, which the Reagan administration was determined to introduce, could be more readily implemented through an international and wide-ranging trade treaty.

In the years immediately prior to 1986, US agricultural policy was one of qualified ambivalence. While the 1985 farm bill had succeeded in decreasing subsidies, Section22 of the 1933 Adjustment Act which allowed for the use of import quotas even when domestic constraints didn't operate, was widely still in use. It is estimated that US transfers to producers in 1986 were still worth a huge $88 billion (Ingersent et al., 1994). In the light of the above the 'zero option' which the US adopted at the start of the Uruguay round and which would have necessitated the sweeping elimination of all trade distorting agricultural subsidies by the year 2000 is all the more surprising. It is debatable whether the US was fully committed to the 'zero option'. Some commentators have suggested that the US was calling the EU's bluff and foresaw that their strong opening position would inevitably have to give way to something like the final outcome. However, others point to the length of time the US took to fully back down from the 'zero option' and argue that the US was fully determined to liberalise agricultural commodity markets and the only reason they backed down was because they had miscalculated the strength of the resistance from the EU.

Two final points with regard to the US opening position must be made. Firstly support for the notion of 'decoupled payments' was mooted by the EU at this time. These payments would be in the form of a direct income transfer and would be divorced from production incentives. They would allow agricultural support to continue but would avoid the adverse effects of market price supports. Secondly the US proposed the tariffication of all non-tariff barriers (NTB), which could then be subject to a phased reduction. Tariffs were more desirable than other support mechanisms because the deadweight loss is less.

By the end of 1988 at the Montreal mid-term review.the US stance had changed little. It demanded the prior commitment of all parties to the eventual elimination of support before negotiations could begin on merely reducing it in the short-term. Furthermore the US wanted the EU in particular to quantify the extent to which it was prepared to reduce certain specific support mechanisms. This approach was to become known as 'rules-based'.

The EU perspective: Background and Aims

By the start of the Uruguay round in 1986, the EU had become the world's largest exporter and second largest importer of agricultural produce. Clearly it would play a crucial role in the GATT talks. The EU however, found itself in a difficult position. Mounting budgetary pressures and surpluses within the Community had triggered off a series of demands for support levels to be reduced. The need for reform was therefore widely accepted. At the same time the EU was not yet prepared to yield on the fundamental principles of the Common Agriculture Policy (CAP), which served to insulate and protect EU producers from world prices. The CAP had been the first truly common policy within the EU and in a sense was held sacred . The power of the farm lobby in Europe was a further contributing factor. The initial EU position was therefore an extremely vague one. In the short term the EU proposed action to decrease instability in agricultural markets and to deal with the problem of excess supply. Only in the longer term would any move towards a redistribution of support be contemplated, and indeed, it wasn't till after 1990 that the EU was prepared to quantify the size of the reductions it was prepared to concede.

The one early concession which the EU did make was in its willingness to contemplate an Aggregate Measure of Support (AMS) to measure and monitor all trade distorting mechanisms with a view to reducing them in the future. The adoption of such a system would be conditional on the EU being permitted to rebalance external protection by increasing it on some commodities and decreasing it on others. While such an approach certainly amounted to a compromise as far as the EU was concerned, it also had its advantages. If support levels were to be reduced on an aggregate level the appropriate measures to be used would be left entirely up to each country. The European Commission therefore believed that it might be possible to avoid having to cut export subsidies, which politically were especially significant.

The determination and intransigence of the EU is evidenced by the fact that it was prepared to jeopardise the whole Uruguay round if the US did not make major concessions on the 'zero option'. Despite statements by the Commission to the contrary, the importance of CAP reform to the eventual GATT agricultural agreement can not be over emphasized. Progress on agriculture was only really achieved after the MacSharry plan was adopted.

How the Compromise was Achieved

It was only after the Montreal deadlock that the US finally signalled a willingness to move away from the 'zero option'. The Geneva Accord of 1989 salvaged the negotiations from complete breakdown, with the US accepting that instead of the elimination of all trade distortions, 'substantial and progressive reduction of agricultural support and protection' would suffice. Agreement on agriculture was due to be finalised in Brussels in 1990. Two months prior to that the US issued its final proposal. This consisted of a renewed emphasis on tariffication, coupled with the reduction of all trade-distorting agricultural subsidies over a ten year period. Included in this would be a 90% reduction in export subsidies. This was in stark contrast to the stance of the EU who proposed that export subsidies should be reduced only in the context of overall levels of support. It was inevitable that the Brussels meeting would breakdown. The crises was heightened by US threats of a trade war.

However in early 1991 three independent events helped to diminish tension and enable talks to be restarted. Firstly, the departure of US trade representative, Yeutter, so long identified with the US 'zero option' heralded in a softening of the US line. Secondly, aware of the Fact that there were significant pressures on both sides to enact reform, GATT secretary general Dunkel began a series of reductions aimed at reaching a compromise. Simultaneously with the 'Dunkel Shuttle' the EU was preparing proposals for CAP reform. The MacSharry Plan, named after the then Commissioner for Agriculture, Ray MacSharry, recognised the fundamental imbalances inherent in the CAP system and proposed the guaranteeing of farm income rather than market prices.

The Uruguay round was formally revived in February 1991 with all parties agreeing to reach 'specific binding commitments to decrease farm supports in each of the three areas'(Ingersent et al 1994 ch. 4), namely internal protection, border protection and export competition. This represented a significant EU concession, in that it implied an accession to America's demands for specific support-reducing commitments, in other words rules-based approach. In December 1991 Dunkel submitted his draft agreement which the US was prepared to accept. The two most critical features of the agreement were, firstly, the provision for improved market access and, secondly, the twin requirements of a 36% decrease in budgetary outlays on export subsidies and a 24% cut in the subsidised export volume. While the EU was beginning to accept that concessions on export subsidies were going to have to be made, it still held back for a number of reasons. Dunkel contained no concession towards allowing Europe to rebalance its overall commitment to decrease protection. More importantly there was doubt as to the compatibility if the draft agreement and MacSharry's proposals for CAP reform. In particular it was found that the compensatory payments at the centre of the MacSharry plan would not meet the criteria laid down by Dunkel for direct income support to be classified as 'decoupled', and therefore exempt from reduction.

Following on from the Dunkel Draft agreement the US and EU eventually reached an agreement during bilateral talks in Washington at the end of 1992. This compromise, which was essentially a slightly watered down version of Dunkel, became known as the Blair House Agreement (Nov. 1992). The essential provisions of the agreement were as follows:

- Domestic Support - A total aggregate measure of support (AMS) for agriculture was to be calculated as the sum of such measures for all agriculture commodities. The AMS was to be decreased by 20% over six years.
- Market Access - The tariffication of all import restrictions, which were to be bound and decreased by 36% over six years using a base period of 1986-88. Secondly, a minimum access requirement for imports of domestic consumption, increasing to 5% by the agreement, was included.
- Export Subsidies - volume of subsidised exports to be decreased by 21% over six years and budgetary expenditure on export subsidies to be simultaneously decreased by 36%.

Following on from the compromise achieved in Washington the agricultural Final Act was agreed on in Brussels in December 1993. The principles incorporated were largely the same as those agreed on between the US and EU a year previously. One difference however, was that the Final Act allowed for 'frontloading' of subsidised exports to enable existing EU stocks to be reduced.

Three years behind schedule agreement on agriculture within the Uruguay round had been reached. Writing this at the end of 1994 all that remains is for the governments of the EU and Canada to formally ratify the treaty.

The Final Position Analysed

The bilateral agreement between the US and EU that clinched the GATT deal represented a considerable compromise by both parties. As far as the US was concerned the levels of support reduction agreed on were a far cry from the 'zero option'. However, it is debatable whether the US ever seriously expected to secure the total liberalisation of agricultural trade. Meanwhile the EU finally gave up insisting that the CAP couldn't be the subject of international negotiations. In particular the issue of qualified export subsidy reductions was conceded.

The rules-based approach of the US, was adopted in so far as the final Blair House Agreement does hone in on specific forms of support. At the same time the agreement also shows the influence of the AMS approach with all measures of domestic support being aggregated. Furthermore, AMS, as defined by GATT, excludes all 'green box' policies, thereby exempting all decoupled payments. This has advantages for both the EU and US. The formers direct payments under the MacSharry reforms and American deficiency payments will both be exempt from reductions.

Ingersent, Raynor and Hide (1994) estimate that the net economic welfare of both the US and EU will increase with unilateral agricultural trade liberalisation by $3.3 and $21.4 billion respectively. Clearly it was essential that the Uruguay round incorporated agriculture, and that moves were made to reduce protection. However, while much work has been done , two issues remain at large. Firstly to what extent is the Uruguay

round agreement compatible with the MacSharry plan? Some commentators have suggested that a further series of CAP reforms may be necessary. Secondly the next round of GATT talks will need to consolidate and extend progress on agricultural policy reform. Whether either of these can be easily achieved is in moot.

Bibliography

Agra Europe Report, No.74, 1993.

Gardner, B., (1993) Urugual Round Agreement, Effects on European Agriculture.

Ingersent, Raynor and Hide, (1994) Agriculture in the Uruguay Round.

The Irish Times, 3/12/1994.

Fiscal Reform in Eastern Europe

Michelle Phillips (Senior Sophister)

The current desire among Eastern Europeans to correct their formerly planned economies to market ones of limited intervention is one of the most significant features of the present phase of international economic developement. Such a process provides fascinating study for economists. In this essay, Michelle Phillips focuses attention on fiscal reform, addressing such associated issues as public expenditure, tax reform and the financing of emergent fiscal deficits.

One of the most prominent debates among different schools of economic thought has concerned the appropriate role of the state in the economy. The divergence of opinion on this issue can be seen by conkasting the Stalinist-Marxist with the Smithian conceptions. According to Stalin (in his interpretation of Marxist economics), the state should determine, inter alia, *'allocation of resources, distribution of income and consumption, levels of saving and investment, and relative prices of goods and services'* (Tanzi, 1992). On the other hand the Smithian view was that the state is *'a spectator that steps in to correct the actions of the private sector when it commits errors of commission or omission'*(Tanzi, 1992).

The contrast between these two viewpoints epitomises the difference of approach of the state between the former centrally planned countries of Eastern Europe and the Western democratic states. However, since the fall of communism, it has been widely recognised in the former that a reduction of the role of the state is appropriate. Since the most direct form of government intervention in the economy is fiscal policy, this requires, first of all, a redefinition of the role of fiscal policy, and then total reform of the fiscal system. This is possibly the greatest challenge facing these Eastern European administrations, as fiscal policy is both complicated and multi-faceted, and the required fiscal reform involves 'opting out' of the economy to an extent.

In this essay, I shall examine the issues associated with such fiscal reform in these countries. I will begin by focussing on public expenditure, and the need to reduce it drastically. Although there do exist taxation systems of sorts in these economies, there is urgent need for tax reform, as the revenue requirements of the transition process are substantial. I shall therfore discuss the issue of taxation at some length in this essay. It must be recognised that fiscal policy cannot be discussed or examined in isolation of the rest of the economy, so I will focus on how fiscal policy can relate to the stabilisation process. The importance of separating fiscal policy from monetary institutions and policy must be pointed out at this conjecture. Finally, the revamped nature of the fiscal system means that tax revenue is reduced and government expenditure may need to be increased; this leads to fiscal deficits. In the last section of the essay, I shall focus on the methods available to those financing fiscal deficits in Eastern Europe.

The Transition from 'Big' Government to 'Small' Government

In a nutshell, what Eastern European governments do as part of fiscal reform is to decrease the level of, and change the structure of, public expenditure. On first examination of data, one would be surprised to find how low the ratio of public expenditure to GDP actually is. In the late 1980s this ratio for Eastern European countries averaged at around 55%, a level very similar to many Western Economies with large social welfare systems, for example Ireland and Luxembourg, both of which had ratios exceeding 50% in 1988 (Chu and Holzmann, 1992). However we must approach these data with some cynicism. The actual level of government intervention via expenditure is understated by such figures, due to the existence of much implicit expenditure, for example enterprise subsidisation in the form of tax incentives. Also the quality of data comes under question because of the existence of price distortions, unrealistically low interest rates and a multitude of exchange rates in former centrally planned economies (Chu and Holzmann, 1992). But there can be no doubt that the role of the public sector in these economies was, and still is, too large, and a substantial downsizing of government is required.

Before transition began, the level of public expenditure was determined by the plan. If state enterprises needed subsidies of £100m to produce what the plan dictated, then this was accepted without question by governments and automatically provided. Government subsidisation was quantity-based. The high level of expenditure was also as a result of the governments' perceptions of themselves as employment guarantors to the people, which warranted further subsidies to enterprises to provide employment for a surplus labour force. But it is clear that an abandonment of 'the plan' in the transition to market orientation means that the level of public expenditure will have to be determined by different criteria. A distinction will have to be drawn between the public and private sectors, and the role of the government in the public sector will have to be made explicit. The central challenge here is to decide on the appropriate sizes of the two sectors, and therefore the role that the government will have.

Possibly of greater significance than the level of public expenditure is the structure it has had in the past, and what its composition should be. The major focus of expenditure in the past was on subsidies of state enterprises, the subsidisation of housing, the financial support of the state monobank, and military expenditure. Much expenditure also went on cushioning the effects of interest rate fluctuations on enterprises and households, and a very broad generalised social benefit system of a universal nature.

Reform of the composition of public expenditure should include some very obvious measures, such as the removal of (or at least downscaling of) military expenditure, and a social benefit system that is less universal and more focussed on poverty alleviation, targeting the poorest sectors of society. The widespread subsidisation of state enterprises (most of which will become private enterprises as transition progresses) should be stopped, and if this proves too difficult, it should at least be based on criteria of profitability or efficiency, rather than output.

Essentially, public expenditure should have a growth-orientation (implying contributions to national savings, expenditure on infrastructure, incentives to invest and cleaning up of the remaining public debt between enterprises and banks). It should also be more market-orientated, undertaken like private investment with cost-benefit analyses done prior to investment (Chu and Holzmann, 1992). For the public good of these middle-income countries, there should also be more expenditure devoted to health and education.

Of course, there are large difficulties associated with this downsizing of government. The cessation of subsidisation of enterprises means that the previously unheard of problem of unemployment will become an issue. Expenditure outlays will have to be increased to pay for increased unemployment benefits. The restriction of expenditure means that greater attention will have to be paid to forecasting of economic conditions when appropriations are being made for the budget, a difficult task when faced with an unstable economy. The implicit nature of much government expenditure makes this difficult to remove. Finally, during all of this, the government will have to be strong, as many of these measures will be (and have been in the past) fiercely opposed by certain political activists and interest groups. The need for reform will have to be re-emphasised many times in certain cases (Tanzi, 1993).

In conclusion then, although it seems like a very difficult task at first, the importance of firm restructuring and downscaling of public expenditure cannot be over-emphasized in the transition to market economy. Stabilisation of the economy is impossible unless the budget deficit is at least stabilised. Even with tax reform, this is impossible without reform of the public finances.

In the next section, I shall highlight the importance of the tax side of the budget, and the urgent for reform in this area, often considered much more important than expenditure reform.

The Transformation of the Tax System

In the days of the command economy in Eastern Europe, the methods employed to extract revenue from the people were totally unlike anything ever seen in Western, market-based economies. In a market-based economy, the government only has the power to define the tax base, set rates, and try to penalise evaders. However, in a centrally-planned economy, the collection of tax is very simple, as it really only represents the transfer of funds from one area of 'the state' to another.

Tax revenue accruing to the government in centrally-planned economies came through four channels, namely the enterprise profits tax, the tumover tax, payroll taxes, and foreign trade taxes. The latter are mainly used to

protect the planning process from the effects of intemational trade. Their importance from a fiscal revenue point of view, however, is often negligible. The remaining three channels of tax extraction are much more significant in revenue terms. The payroll tax is a sort of implicit personal income tax, except it is not taken from individual incomes, but rather as a lump-sum from the payroll of the state enterprises which employ the individuals. Therefore, the individuals do not see the tax being taken from them, and are not obliged to file tax retums (although there are exceptions to this). The amount of the tax is just automatically debited from the enterprise's account with the state-owned monobank. The tumover tax is collected in the same way, except it is levied as the difference between the administratively-set producer price and the administratively-set consumer price; as the name suggests, it is a tax on the tumover of enterprises. Any remaining surplus accruing to the enterprise, a somewhat arbitrary figure, is then extracted via its bank account as the enterprise profits tax.

The system as it has stood in the past has perfommed its functions extremely well, and the revenues accruing to the state in these centrally-planned economies were very impressive. However, in recent years, as these economies have embarked on the transition to market-orientation, their tax systems have failed. For example, in the years 1989, 1990 and 1991, the level of tax revenue received by the state in Bulgaria fell from 60% of GDP to 53% in 1990, and then to 39% in 1991, an absolute decline of 20% (Bristow lectures, Nov. 1994). Other European Countries have experienced similar declines as the central planning system collapsed. The implications of this for fiscal deficits are startling, as the ability of these govemments to cut spending by similar proportions is very limited. It is clear that the reason for this revenue collapse lies in the incompatibility of the old taxation system with a market-based economy.There are many reasons why the old system should fail utterly when faced with a market regime. Firstly, whereas previously the state had to extract taxes from hundreds, or perhaps thousands, of enterprises over which they had great control, with privatisation, the state was faced now with millions of individuals and privately-owned enterprises, which meant a reduction of infommation available to tax collectors, and increased difficulties associated with collection (Tanzi, 1993). Secondly, the liberalisation of prices and wages to market-detemmined levels, meant that neither payroll nor tumover taxes could be relied on as much for revenue, as the state could no longer administratively set the payrolls or the tumovers of enterprises. The excessive levels of inflation which accompanied liberalisation also spelled a fall in the remaining profits of enterprises, which meant that there was less taxable income in the fomm of company profits, and less potential tax revenue. Thirdly, the onset of a market-based economy led to the privatisation of what were state-owned enterprises, making the private sector larger, and more fimms difficult to tax. Finally, the financial reform which had to accompany transition meant that the state had to sell off the banking sector to private interests. However, this meant that withholding taxes at the source was no longer possible; rather than just debiting enterprises' bank accounts, the state now had to rely on the enterprises to pay the tax.

There can be no doubt, then, of the need for tax reform in Eastem Europe. Or rather, what has been suggested instead of refomm of the existing system, is the creation of a totally new system, which many agree should broadly mirror the systems in place in Westem economies, without, of course, the inherent flaws in these systems. It goes without saying that some form of value-added tax and personal income tax should be simplemented. However, it is unreasonable to expect that these can be just introduced ovemight with no effects on revenue. There will have to be some sort of transitional arrangements made to ensure that the effects on the fiscal deficit of the tax reform are minimal. With highly unstable prices, the introduction of a VAT may take some time, as , with a view to prospective EU membership, there has to be a change from a multi-rate system to a system with one or two rates of VAT and a very broad base of taxable goods. The introduction of a personal income tax shouldn't pose too much of a problem if it is levied at a low rate, on a PAYE basis, with a tax allowance system that is initially quite simplified (Gandhi and Mihaljek, 1992).

In addition, two extra transitional taxes have been suggested by Holzmann (1992) to ease the effects of economic stabilisation on the public finances, namely a transitional import surcharge (which would have revenue benefits and has the advantage of protecting domestic industry from extemal competition) and a capital gains tax (which would ensure an equitable distribution of income). But it must be emphasised that these would have to be transitional measures and specified as such from the outset, as their imposition in the long run would be detrimental for many reasons.

The orchestration of tax reform should be done as speedily as possible, but various experts in the field have laid down some golden rules on this issue. There should be discontinuation of negative discrimination against the private sector via the profit tax. It should be recognised that the sole purpose of taxes are to collect

revenue for the state; they should not be used to achieve the objectives of the plan. There have been many calls by public sector economists to experiment with new and untried taxes in Eastem Europe. However, according to Vito Tanzi (1992), these countries should 'refrain from excessive experimentation, that is, from the temptation to go for untried taxes, which may look good in theory, but may be difficult to administer'. A tightening of tax accounting regulations has also been suggested because a reduction in the profit tax base will result from a market-oriented accounting system (Chu and Holzmann 1992).Finally, in conclusion, tax reform will have to be focussed on economic efficiency (minimising the deadweight loss of taxes), administrative efficiency, equity, and, above all, simplicity.

Fiscal Policy in a Broader Context

There can be no doubt that elements of fiscal policy have a role to play in the stabilisation process, or at the very least can help to ease the hardships associated with the transition to a market economy. In an ideal situation, fiscal reform should reduce expenditure sufficiently so that budgetary surpluses result, and with these surpluses the state is in a position to mobilise savings in the private sector, facilitating ease of privatisation, supporting the banking sector and fostering investment. However, no situation is ever ideal, and we can only hope for some positive effects of fiscal policy in the transition period, while ensuring a balance between private and public savings. Since alternative institutions tend to be too weak in the early stages of transition, it is often up to the state (via public expenditure) to subsidise key industries, improve infrastructure, ensure minimum subsistence wages for all in the face of spiralling prices, foster goals of equitable income distribution and encourage foreign investment. But even though fiscal policy should foster private sector development, it is, however, imperative that it *'contribute to stabilisation without delaying the necessary fiscal restructuring'* (Chand and Lorie, 1992), as the latter should take priority.

But, as mentioned previously, the stabilisation process usually spells disaster for fiscal coffers. Instability in the economy as well as terms of trade deteriorations lead to a fall in enterprise profits (be they public or private sector), and this means a fall in revenue. The emergence of current account deficits for the first time in these former centrally-planned economies is daunting enough, but even more so when we consider the possibility of this leading to hyperinflation, causing depression on the economy and hindering investment. I shall deal with the problem of deficits in the next section.

One of the more alarming aspects of the legacy of centrally-planned economies is the degree to which fiscal policy is fused with monetary policy. These fiscal-monetary links are numerous. The banking systems in Eastern Europe have been giving out credit at will for many years to struggling enterprises, and then relying on the state to finance them when they turn into bad loans. Thus, in effect, fiscal policy (i.e. the subsidising of state enterprises) has been exercised via the monetary institutions. The banks have also been keeping interest rates low to reduce the burden of the internal public debt, traditionally a fiscal responsibility. In more recent years, as government revenue has collapsed, and budget deficits have emerged, there has been a reliance on monetary expansion to finance any extra public spending (Tanzi, 1992). As these economies approach market status, it is becoming increasingly important that many of these fiscal-monetary links be severed, so that each policy is directed at the area in which it is relatively more efficient.

Dealing with Deficits

It may seem an oversimplification to assert that under the old command economy systems in Eastern Europe, budgets always balanced, but essentially it is the truth. The planning system largely ensured that government expenditures were matched by government revenues. If there was ever any shortfall, due to the administration mechanism of the budget, or, more significantly due to exogenous shocks to the economy, the deficit was usually camouflaged by *'creative accounting'* (Cheasty, 1992), or prices were adjusted so that more money could be extracted from enterprises. However, with the onset of the market-based system, such methods were no longer available to the state, and all deficits had to be announced and publicly financed (i.e. they became 'below the line' deficits). This became more of a problem considering the aformentioned revenue collapse experienced by these countries, and the difficulties associated with reducing public expenditure. As they began their journey into market-orientation, the governments of these formerly centrally-planned economies found themselves, for the first time, exploring Western methods of deficit financing.

In her paper *'Financing Fiscal Deficits'* (1992), Adrienne Cheasty has discussed various options open to these countries for the financing of deficits. In doing this, she has orchestrated various simulations analysing the various effects on the economies in question of the different options, assuming an initially loose policy that gets progressively tighter. Although some specific recommendations are made for specific countries, these are beyond the scope of this paper. Some general conclusions did emerge, however. Foreign financing is the least costly option, but is only appropriate for those countries that do not already have a large external debt, and have a good trade balance. However, it should really only be a last resort, to relieve inflationary pressures. Theability to use this method also depends on the domestic-world interest differential. As a temporary option, bank financing is often suggested for those countries with an existing large external debt. However, this invariably leads to inflation, and the danger of hyperinflation. The third option examined by Cheasty is domestic non-bank financing, the main disadvantages of which are that the interest burden is often quite substantial (relying on the domestic interest rate), in that it tends to depress the economy. It also defeats the purpose of reducing public sector interests in the economy. The broad conclusion which can be drawn from this is that different financing options are more appropriate in different countries, and no sweeping policy recommendations can be made for Eastern European countries in general.

Conclusion

In conclusion, then, complete fiscal reform remains one of the greatest challenges facing the governments of economies in transition. They must try to *'step back'* from the workings of the economy in order to allow the market to do its job, while simultaneously trying to construct modern Western-style tax systems, all to be done as quickly as possible. This is made all the more difficult by the fact that they must try to minimise the size of, or at least the effects of, deficits resulting from the precariousness of the system. It can only be hoped that valuable lessons can be learned from the experiences of the well-established fiscal systems of the West.

Bibliography

Bristow, John <1994) *Senior Sophistor European Transition*, Course Notes, Economics Department, Trinity College, Dublin.

Chand and Lorie, (1992): *Fiscal Policy*, Ch 1 of Tanzi V. (ed.), Fiscal Policies in Economies in Transition, IMF.

Cheasty A. *'Financing Fiscal Deficits'*, Ch 2 of Tanzi V. (ed.), ibid

Chu and Holzmann *'Public Expenditure: Policy Aspects'*, Ch 12 of Tanzi V. (ed.), ibid

Gandhi and Mihaljek *'Scope for reform of Socialist Tax Systems'* , Ch 7 of Tanzi V. (ed.), ibid

Holzmann R., (1992) *'Tax Reform in Countries in Transition: Central Policy Issues'*, Public Finance Supplement.

Tanzi V., (1992) *Fiscal Policies in Economies in TransitionIMF.*

Tanzi V.,(1993) 'Financial Markets and Public Finance inthe Transition Process', in Transition to Market : Studies in Fiscal Reform, IMF.

Stabilisation, Privatisation, and the Development of Capital Markets in Poland.

Neil MacDermott

The path from the command economy to free-market system has been a difficult one for the countries of central and eastern Europe. The initial euphoria of freedom has been replaced by a distrust of all things western. Neil MacDermott in this essay examines the case of Poland and analyses its success in making the transition.

'There is nothing more difficult to carry out, nor more doubtful of success, nor more dangerous to handle, than to initiate a new order of things. For the reformer has enemies in all who profit by the old order , and only lukewarm defenders in all those who would profit from the new order. The lukewarmness arises partly from fear of their adversaries who have law in their favor; and partly from the incredulity of mankind, who do not truly believe in anything new until they have had actual experience of it.'

Niccolo Machiavelli, *The Prince* (1513)

Introduction

In the emerging transitional economies of Central and Eastern Europe political and economic reform are interdependent. If 'adjustment fatigue' sets in at the present early stage of the process the political consequences will spillover to the detriment of economic reform. This investigation into the progress of these new democracies will focus on the stabilisation of the Polish economy, the development of their privatisation programme and the advance of capital markets. Poland has distinguished itself as a leader in economic reform. It was the first Soviet satellite to reject communism with semi-free elections in June 1989, followed in August by the removal of the old Communist establishment. However this path has not been without difficulties, the Poles still suffer high inflation and unemployment and have recently voted the old ruling communist party back in to power, the effects of which have yet to be gauged.

The importance of a well functioning capital market for an emerging economy, the enterprising actions of the Polish government with respect to stabilisation and the uncertainty inherent in investing in these countries forms the rationale for this piece. The second section will outline the macroeconomic stabilisation process and privatisation policy undertaken by the authorities. Progress in the crucial area of banking reform and the development of an efficient capital market in Poland will be highlighted in the third section. Finally, the conclusion deals concisely with the achievements of the Polish authorities struggling down the road to economic reform.

The Stabilisation Process

Macro-Stabilisation

An unprecedented challenge confronts the centrally planned economies of Central and Eastern Europe in their attempts to reform and stabilise their respective economies. Most commentators argue that the development of a market-oriented structure and institutional framework through a 'simultaneous assault on macroeconomic stabilization, prices and property rights is necessary from the start'.

Price liberalisation is the first step. Prices act as indicators to enterprises of demand conditions for their products. When prices are market-determined they equate supply and demand. 'Big Bang' price liberalisation as occurred on the first of January, 1990 in Poland causes an immediate price explosion for both supply and demand reasons. It was thought that the monetary overhang particularly would encourage inflation as goods became available, but some studies have shown that in the longer run, consumers refrained from price bidding because of the overall aura of uncertainty at the time. Liberalisation also brought about a significant

decline in output. Real incomes fell as did consumption levels. The CMEA collapse depressed export demand in conjunction with the increase in cheaper and superior imports from foreign competition in the West.

Poland attempted to curtail its rampant inflation and maintain internal convertibility by using four stabilising anchors:

1. a fixed exchange rate,

2. incomes policy,

3. real money supply,and

4. real interest rates.

The authorities hoped to restrain inflation and reduce an escalating budget deficit. The achievement of these goals was expected through *'monetary policy means alone(no fiscal intervention), while long term structural changes were to be enterprise-driven....[the development of an efficient financial intermediation system]....They were to take place in the absence of any industrial policy, with weak trade policy and with an underdeveloped financial sector'* . The idea that the necessary restructuring could be achieved without recourse to fiscal activism was a throwback to the liberalism most associated with the first post communist solidarity-led coalition government.

Some Economic Statistics

	1989	1990	1991	1992	1993	1994e
Real GDP Growth %	0.2	(11.6)	(7.6)	2.6	4.0	4.5
CPI%	258.1	558.3	75.9	45.3	36.8	30.0
Current a/c Bal % GDP		(2.3)	1.1	(1.7)	(0.3)	(2.7)	(2.7)

This policy changed quickly when the authorities realised that the economy needed some continuity: markets were simply not developed enough to carry out an effective allocative role. The deflationary measures introduced in 1990 included a 35% cut in the wage bill, a reduction in state sector subsidies from 11% to 7% of GDP and a scaling-down of the state investment programme. Despite these efforts, inflation almost doubled in 1990 while the country experienced severe negative real GDP growth. Private sector growth, although affected by the slump in the state sector, still recorded positive growth. However, these poor figures were worse than most had predicted. Unemployment, which theoretically had been nonexistent, grew to 11% in 1991. Poznanski (1993) informs us that state owned-enterprises have, on the whole, maintained employment. Growth in unemployment is therefore mainly among new entrants to the labour market. As a result Poland's transition programme came under attack from certain commentators. Inflation was curbed significantly and the zloty, the polish currency strengthened against the dollar but the drop in output 'wasn't just a temporary structural adjustment but a deep recession if not depression' in all sectors of the economy. I, however, believe this attack is flawed [Poznanski, R. 1993]. Bruno (1992) suggests that external shocks exacerbated an already depressed situation. The collapse of the CMEA , the recession experienced in most European and North American markets, the slower than expected rate of structural adjustment and privatisation were all contributory factors. GDP growth did resume in 1992-1993 (1.5% and 4% respectively), but the GDP remains 13% smaller than in 1989 all the same.

The drop in output notwithstanding, Poland has sizeble economic potential. Production assets in industry and construction alone are estimated at about US$70 billion. In addition, it has almost 39 million consumers, a skilled labour force and a well-developed transport infrastructure.

Privatisation

Privatisation has proved to be the basic means of dismantling Poland's command economy while raising revenue, improving tax collection and boosting capital markets. Apart from the need for economic efficiency, widespread privatisation is important because the concept of private ownership is consistent with a democratic state. The literature outlines the various routes towards privatisation:

(a) public share issues,

(b) mass privatisation through a voucher scheme, and

(c) employee buy-outs.

However, it is difficult to grasp the sheer scale of the transformation programme that these countries face. Poland had over 8,000 state-owned enterprises that were in line for restructuring, representing more than 80% of GDP9 . Western experience is often alluded to in policy descriptions, yet in Thatcherite Britain the government sold off enterprises representing only 5% of GDP. Therefore no realistic precedent exists. Experiences in Poland will form a testing ground for many of the other ex-Soviet satellites policy options.

The issue of corporate governance must also be considered when choosing a method of privatisation. To achieve the desired dispersal of ownership the Polish government decided upon a multitrack approach. State-owned enterprises had three options: Liquidation, commercialisation, or remaining in the public sector.

The authorities decided to include 'insiders'(workers' councils and managers) in the decision process, granting them an effective veto over privatisation.

State Enterprises Transformed as of 30/06/9411

Commercialised	579	7.0%
Liquidated	2,132	25.9%
Other	5,117	67.1%
Total	8,228	100%

The privatisation process has been particularly successful in the small retail and service oriented industries. In those sectors restructuring is less complex, not only because organisational size is small, but more importantly because their impressive performance is 'linked with the de facto resolution of ownership rights'. The lack of such clarity with respect to large state-owned enterprises, coupled with the insiders' veto, and the large scale restructuring that is necessary, has stymied the process. In response, the government has established a voucher-based mass privatisation programme for medium and large enterprises, Those participating will be turned over to fifteen-to-seventeen National Investment Funds(in an attempt to overcome the corporate governance issue), which in turn will be managed by supervisory boards elected by the Treasury. Delays have dogged the programme, but assuming the problems can be overcome, the Polish Ministry of Privatisation has ambitious plans. The ministry is seeking approval for a 1995 plan to commercialise 1,500 state-owned enterprises and a further 3,800 by 1998.

A need for a set ofre-privatisation rules has been identified by the Ministry. The lack of such rules has slowed down the process, especially in rural areas. A new owner is never sure whether the property he has purchased, be it enterprise or farmland, is burdened with any undisclosed reprivatisation claims. Up to mid-1994 over 100,000 reprivatisation claims had been registered, the majority coming from inhabitants of Polish territories conceded to the USSR who resettled after the second world war. As of right these people should be compensated for the lost property. However, this figure represents only a fraction of potential claimants, if the laws concerning the recovery of lost property are simplified, is planned in the Sejm(the lower legislative chamber). Originally the legislation provided material compensation, albeit incomplete, or compensation in terms of government securities hwere this wasn't feasible. Substitute property could be chosen from the so called undesirable property of enterprises earmarked by the state. However, because of the strain on the state finances, the new Government announced its own version, which stated that the state could only afford symbolic re-privatisation.(From a statement by Wieslaw Kaczmarek, the Minister of Privatisation). The Ministry reported 1993 revenue from privatisation at $225m for the state budget and Minister Kaczmarek expected the 1994 figure to reach $275m.(From a statement by Wieslaw Kaczmarek, the Minister of Privatisation).

Banking Reform and the Capital Markets

Pescetto (1992) suggests the following three reasons for developing an efficient banking system.

(a) Firstly, the banking system facilitates the efficient allocation of capital. Through credit risk analysis it ensures that only the most effective firms receive the necessary finance while imposing a strict budget constraint. Thus banking reform and privatisation are interdependent.

(b) Secondly, the control of money supply and all its aggregates is a crucial function for a central bank. For the government to effect stabilising policy, control must be placed on money supply.

(c) Thirdly, the supervision of the commercial banking system by the central bank is also of great importance, as is the continuing effort to bring informal markets into the legitimate economy.

It is with these in mind that I continue, firstly with a brief discussion of the banking system in Poland and the reform process, and secondly with an outline of the development of capital markets in Poland.

Banking Reform

Until 1988, the shape of the Polish banking system was typical of any centrally planned economy.The National Bank of Poland combined the duties of a central bank (although it had little autonomy with regard to monetary and credit policies) and of a commercial one (handling accounts of state enterprises and natural persons) as outlined in the monobank model. The Banking Act and the National Bank of Poland(NBP) Act were passed on January 31, 1989. The NBP is the current bank of issue, a lending institution for other banks and the central foreign exchange banking institution. Nine commercial banks were separated from the NBP structure. This legislation made it possible to create a two-tier system modelled on that of Germany.

Banks are free to set rates,both deposit and lending, and to use the NBP refinancing rate as a common reference. The NBP has a policy of maintaining a positive differential over the rate of inflation. Most banks have a significant portfolio of non-performing creditors. Chopra (1994) believes that banks will resort to 'charging burden sharing premiums on performing loans and widening of interest rate spreads i.e. imposing a tax on financial intermediation to recapitalise.' As an alternative the state could encourage banks to use debt-equity swaps, turning banks from creditors into owners of the company. Any other alternative involves placing the burden of debt on the state budget.

Finally, it is expected that a new banking act and an act pertaining to the central bank , conforming to the European Communities standards, will be passed in 1994.

Polish Capital Markets

Developing capital markets is crucial to the advancement of a market economy. Stiglitz (1992) places great emphasis on their importance 'if capital is at the heart of capitalism, then well functioning capital markets are at the heart of a well functioning capitalist economy'17 . Chopra (1994) suggests in the recent IMF paper on Poland that the success of transformation programmes is dependent on the development of an efficient financial intermediation system as well as credit and capital markets. Capital markets facilitate the reconciliation of saving and investment flows and therefore play an important role in the real sector of the economy.

The regulated capital market has been in existence in Poland since 1991, operating under the Trading in Securities and Trust Funds Act of March 1991, which was amended in January 1994. It includes provisions governing the public issue of securities, and secondary trading in securities on the stock exchange and the over-the-counter markets. The Warsaw Stock Exchange(WSE) is the principal institution of Poland's capital market. It is run as a joint-stock company with US$4 million in share capital. Only brokerage firms and banks may acquire interest in the Stock Exchange. The Polish system is fashioned after the order-driven market model and all transactions are settled at the fixed price of the day. The Securities Commission has been established to regulate and supervise the public capital market. In addition, this institution monitors the issue of new securities and also licences brokerage houses and brokers. Banks' brokerage divisions and private brokerage houses act as agency brokers(until now 40 permits have been issued so far). Brokerage houses also launch new issues, manage investment portfolios and render the services of investment advisers.

Shares of 23 companies were listed on the Stock Exchange in early 1994, up from 16 at the end of 1992. ' No restrictions are placed on foreign participation in the WSE and non-residents are free to repatriate all profits

earned in Poland. Non-resident investors are subject to the same levels of taxation as Polish residents. Between 1992 and 1995 all capital gains are exempt from tax, with dividend income charged at a flat rate of 20%'.18 Following a period of sluggish development in 1991-92, the Stock Exchange grew dynamically in 1993. The WSE Index(WIG) increased during this period by over 1,100% and its turnover expanded over 100 times. In early 1994, trading exceeded US$500 million a week. Trading is dominated by private investors although the activity of institutional investors - both national and foreign - is on the rise. As of yet, however, the stock market has not been an important source of raising new capital. Low price/earnings ratios have discouraged firms, although they improved dramatically improved in 1994. Chopra (1994) suggests that the cost of capital now secured in the equity market is around 3% of net earnings, in contrast to bank lending rates of closer to 40%. This would lead us to conclude that the boom experienced in Poland will continue as the stock exchange sees greater issuance and greater use of the market is made by newly commercialised firms.

Conclusion

The success of Poland's transformation continues to depend on the rate of structural adjustment, the growth performance of its major trading partners, and the confidence of Western companies and the major financial markets in Poland. Progress has been made to stabilise inflation, minimise the drop in output and encourage private ownership. In addition, a new framework for the banking system has been developed and capital markets are beginning to emerge both of which will prove crucially important to the financing of newly privatised companies and in acting as an important source of capital to the Polish government. They will also promote investment in the economy through their provision of credit analysis. Recent experiences in Mexico have highlighted the fickle nature of market investment. Such experiences only strengthen the notion of the interdependence of all emerging economies in achieving a successful transition.

Bibliography

Bruno, M., (December 1992) 'Stabilization and reform inEastern Europe', *IMF Staff Papers*.

Chopra, A., (1994) *Poland: Path to a Market Economy* IMF Occasional Paper No 113.

Pescetto, G., (1992) 'Eastern Europe and financial m', Ch5 of G. Bird(ed), *Economic Reform in Eastern Europe* .

Poznanski, R., (1993) 'Poland's Transition to Capitalism:Shock and Therapy' in *Stabilisation and Privatisation in Poland*

Stiglitz, J. E., (1992) 'The design of financial systems for the newly emerging democracies of Eastern Europe' Ch 9 of C Clague and G Rausser (eds), *The Emergence of Market Economies in Eastern Europe*.

Stephen Moore

Keynes noted in his summation of the Treaty of Versailles that the representatives at the Paris conference committed a grand error *: 'By excessive concentration on political objects and on the attainment of an illusory security , they overlooked the economic unity of Europe - illusory because security is to be found least of all in the occupation on extended frontiers and also because the political contrivances of the moment will be largely irrelevant to the problems of a later decade.' In this paper, Thomas Stephen Moore writes on the 'Economic Consequences of the Peace'* in Northern Ireland, focussing on the possible benefits from the greater economic unity of the island of Ireland.

Introduction

The dynamic nature of recent political developments in Northern Ireland has caused much debate as to the economic benefits which would result from a permanent cessation of violence. This essay shall in this context be partly concerned with the peace dividends of cross border trade and will be divided into three broad sections. The first entails an analysis of some of the underlying trends relating to cross border trade. Inspection of the data reveals two important issues which must be addressed and these shall be my focus for section two. In my final section I shall look at the future opportunities and benefits which may result from a prolonged period of peace and stability in the province.

Cross Border Trade - The Statistics

There has been a general upward trend in cross border trade over the last twenty years; however one can easily contrast the balance of payments surplus in the Republic with the large deficit in Northern Ireland. Pre 1960 the balance had been in favour of the North, but a small trade deficit emerged in the late 1960s and early 70s, just before the first oil price shock and the emergence of widespread civil unrest. Subsequently there was a spiralling deterioration which has persisted to the present day. It is estimated that the trade deficit now amounts to between £IR300 - 400 million, which is approximately between 1 and 2% of southern GNP. (Business and Finance, Dec. 1994). It is interesting to note that the North - South trade deficit can be directly compared with the global trade pattern. The Republic's current account surplus is officially estimated to be IR£2.4 billion or 8% of GNP, whereas the Northern Ireland overall deficit is roughly equivalent to the annual subvention to Northern Ireland - around £3 billion or over 20% of GNP.(NIERC,1994).

The current market orientation of Northern Ireland remains firmly rooted in the home market and Great Britain, with these two markets alone accounting for over two thirds of manufacturing output by destination. The remainder of the output is sold in export markets with the Republic accounting for around 10% of the subtotal. In the Republic the domestic market is likewise equally important for Irish industry, accounting for 38% of industrial output (B&F, Dec. 1994). Despite this comparison it is vital to note one important difference between the Republic of Ireland and Northern Ireland in respect to external sales. Industry in the Republic has a much lower reliance on the British market, corresponding to 14% compared with 38% for Northern Ireland (NIERC,1994).

The overall level of trade between the two economies in 1992 amounted to £1.25 billion, of which just over £1billion was in manufactured goods with the remainder in agricultural products. It is interesting to note that last year Northern Ireland ranked in the top ten export markets for the Republic; however this must be put in context as Japan ranks higher in terms of Irish exports than does Northern Ireland! The above analysis of Cross Border Trade highlights two fundamental questions which must be addressed. Firstly, why has Northern Ireland had such a large trade deficit with the Republic over the last two decades? Secondly, why is the overall level of trade so low between two adjoining small open economies?

Cross Border Trade - The Answers

In response to the first question it is impossible to ignore the effects of twenty five years of violence on the Northern economy. If one examines the size of employment in the manufacturing sectors in 1960 and 1990,

one observes a dramatic decline in the North (from 184,000 in 1960 to 110,000 in 1990) which can be contrasted with a large stimulation inmanufacturing employment in the South (from 175,000 to 232,000 over the same period). (B&F, Dec. 1994). Why should this be so? Some would argue that the decline in the North was largely due to the decline in traditional industries; however there is no doubt that the Troubles acted as a catalyst in the closure of manufacturing plants in the North. For example the C.B.I. in a recent report *'Peace - A Challenging New Era'* highlighted some of the key areas where the Troubles may have had a negative impact on industry e.g. increased security costs, image problems, employment problems such as sectarian harassment and finally a *'brain drain'* with few talented individuals returning.

Even if one accepts the argument that the decline in manufacturing employment was in indigenous companies, one cannot possibly ignore the adverse effects the Troubles have had on foreign direct investment, with the period 1973 to 1990 witnessing a huge decline in the externally owned sector. Despite offering a wide range of incentives the Industrial Development Board (I.D.B.) has had arelatively poor record of inward investment compared to countries such as the Republic of Irelandand Scotland. For example between 1980 to 1993 just over 13,000 jobs were promoted by the I.D.B. in new inward investment projects, compared to an equivalent figure of almost 116,000 for the I.D.A. over the same period. (NIEC, 1994). In addition some economists such as Teague and Harrison have noted that the I.D.B. has had to accept more relatively risky projects because Northern Ireland is seen as an unattractive location for investment. Also some would argue that much externally owned manufacturing employment in Northern Ireland is in relatively low technology sectors such as clothing and textiles. This is in direct contrast to the South, so the quality as well as quantity of investment is called into question. Based on these figures I feel that it is impossible to ignore the negative impact which the violence has had on foreign direct investment and hence on the size of the manufacturing base in Northern Ireland. In this context, though one must also note that the relative success of the South compared to the North may in part be due to the greater policy flexibility enjoyed in the Republic compared to the UK with regard to foreign direct investment.

One last issue which must briefly be mentioned in relation to Northern Ireland is the rapid expansion of public sector employment over the last two decades. Employment in health, education and public administration accounted for 25% of total employment in 1971, though this share had risen to 36% last year. (DKM, 1994). Though a very rough calculation this means that public expenditure in Northern Ireland accounts for over 60p of every £1 earned. While public expenditure can have positive effects such as the stability of incomes and employment, it can crowd out private sector employment by raising wage levels and attracting talented individuals into unproductive employment. Also high levels of public expenditure can direct companies towards safe strategies in home markets rather than growth strategies in international markets.

Some of the listed reasons may go part way in explaining why Northern Ireland has such a large trade deficit with the Republic. Over the last two decades, the economies have diverged quite considerably, largely I feel as a result of the Troubles. The Republic's manufacturing base has expanded rapidly and the economy is today an export oriented small open economy. The North by contrast has seen a large decline in manufacturing, and what manufacturing output there is tends to be focused predominantly on the domestic market. Also the rapid expansion of the public sector means that the North is more an *'island of Keynesianism in a UK free market area'* than a divergent export oriented economy. (Tomlinson, 1994).

This discussion goes only a small way in answering the question of why the overall level of trade between the two economies is so small. Numerous other factors also apply.Firstly, a limiting factor on the level of cross border trade may be the large degree of similarity between industry North and South. Approximately, one half of gross output in manufacturing in the North and South is concentrated in four sectors - namely food/drink/tobacco, timber/furniture, paper/printing and mineral products in which both countries produce similar products. If we momentarily return to the basic Heckschler Ohlin theorem we can note the result that the potential for trade is greater the larger the diversity between factor endowments (in comparing two countries). If there exists a large degree of similarity between North/South industry, then it may be possible to assume roughly equivalent capital/labour ratios and hence similar comparative costs. In this situation, we can use the Heckschler Ohlin theorem to predict that there may exist little basis for mutual North-South trade, which interestingly appears to hold whereas for example Linder's alternative hypothesis does not. In addition if one examines other sectors, which are largely externally owned, they tend to operate on global/international scales or produce goods such as aircraft/ships in the North or pharmaceuticals/computers in the South which have limited demand in the home markets.

Secondly, one must mention the effects of barriers to entry, both physical and mental, which may have inhibited cross border trade over recent years. A strong and efficient infrastructure is vital to promote strong trade flows. The Troubles have adversely affected the transport infrastructure in many ways. Selected road closures in towns and cities as a direct result of bomb threats and delays and damage to the Belfast-Dublin rail line has no doubt had a damaging effect on economic activity, increasing costs especially for the agricultural and business sectors in local communities. The politics of the Troubles may also come into question as a mental barrier to entry for cross border trade. For example, some authors have noted that cross-border trade during the Troubles appeared to thrive where the buyer/seller were dealing in speciality products, for instance one engineer talking to another. In such cases, politics does not enter as a factor in the buying decision.

Lastly, although the overall level of trade appears small in absolute terms, statistics show that the level is largely in line with other small European economies, such as Denmark and their nearest neighbour. If one looks at sales per head of population rather than absolute levels, one notes that in 1991 sales of Northern Ireland goods was £110 per head of population in the Republic, compared with £37 per head in Great Britain. (NIERC,1994). Nonetheless, the lions share of sales in both economies is retained in the home market at £1,222 per capita for Northern Ireland and £1,600 per capita for the Republic of Ireland.

Cross Border Trade - The Future

Over recent months there has been much debate about the so-called 'peace dividend'. In this section I intend to investigate what may result from a prolonged period of stability in the Province.

From a manufacturing industry viewpoint a cessation of violence may lead to a reversal of some of the negative externalities imposed on the Northern economy by the Troubles. Increased inward investment, a strengthening of the manufacturing base, improved transport infrastructure, a reduction in the size of the public sector and a general *'feel good'* factor may also lead to increased trade between the two economies. On a realistic level, one must recognise that the overall level of trade between the North and South is not out of line with other similarly sized neighbouring countries in Europe, so we should not in any way expect a Trade explosion. In this light some of the CBI/IBEC claims for example trebling current trade levels with a commitment to increase employment by 75,000 may be optimistic to say the least (NIERC,1994). If one took a more modest figure for example, the doubling of cross border trade in manufactured goods from £1 billion to £2 billion over a five year period, then allowing for some inevitable displacement (say 50% of Northern Ireland and Republic of Ireland sales) producers would still generate a net increase in output of £0.5 billion. It is obviously difficult to put a figure on the increase in employment resulting from this expansion of output, but perhaps a total increase in employment of 7,500 over the next five years would not be unreasonable. This simple calculation would suggest that there exists benefits in the form of increased incomes and jobs which could result from a stimulation of cross border trade in a peaceful environment, but on a more modest level than that espoused by CBI/IBEC.

Other factors which may boost this figure even further is firstly the expansion of operations through cross border acquisitions. Note already the example set by the food and financial sectors, for instance Dunnes Stores(which is now the largest company in Northern Ireland), Golden Vale and Kerry Group. As for the financial sector, the acquisition of Northern TSB by AIB consolidated AIB's position in Northern Ireland with First Trust. The Ulster Bank has also shown interest in expanding in the Republic with bids for TSB and NCB stockbrokers. Secondly, it is important to note the increasing role played by small firms in the Northern economy. Generous grant assistance offered by LEDU (the N.I. small business agency) has stimulated the emergence of the small firm as a major source of employment in the Northern economy. In 1991, employment in small firms accounted for 54% of total employment (absolute number 156,000) in the private sectors of the economy . (B&F, Dec. 1994). Stimulating the small firm sector on both sides of the border to a more export orientd outlook could satisfy niche markets on both sides of the border and thereby exploit possible mutual gains from trade.

The European Union Act of 1993 has provided a wealth of opportunities for firms North and South to work together and carve out niche markets in Europe. Within Europe itself, Northern Ireland and the Republic are small players; however by working together there are obvious beneficial economies of scale. A peaceful environment would make the transition to integrated projects much easier. The IBEC/CBI Northern Ireland

Joint Development Programme highlights many areas where joint ventures between the Republic and northern Ireland would bring obvious benefits. Some of the main issues involved include:

(i) A co-ordination approach to the development of North-South transport links including road/rail/port and other related areas.

(ii) The creation of an integrated energy policy between the Republic and Northern Ireland. For example the interconnector between the electricity grids of the Republic and Northern Ireland was damaged by bombs on six occasions before it was closed down in 1975. It is estimated that the absence of the inter connector costs the ESB at least IR£10 million per year. Akin to this, the failure to pipe Southern natural gas to Northern Ireland has no economic rationale.

(iii) The development of an *'Economic Corridor'* betweed the Belfast and Dublin regions. The Joint Council issued a report last year entitled A Corridor of Opportunity which concluded that not only was the corridor feasible but that it was essential to North-South economic linkages.

This new Business Development Programme is aimed primarily at the manufacturing sector, but agribusiness and services are also targeted. The main emphases on these sectors is to encourage more companies to plan their businesses on an island basis, more specifically to businesses within the island economy and also increase opportunities to integrate projects within the European market. Finally, whereas for the majority of this essay, I have focused on issues relating to the manufacturing sector in cross border trade, I do not intend to ignore the services sector. Globally, a new trend has emerged where services are becoming an increasingly important component of world trade. In 1987, global exports of services accounted for 20% of world trade which is equal to the exports of fuel/services.

Trade in services could be crudely divided into three broad areas:

(i) Transport of goods/people

(ii) Tourist expenditure at destination

(iii) Private services and income (e.g. telecommunications, financial services)

One has already briefly discussed (i) and (ii) above and shall now focus on Tourism. While tourism in both the North and the South has been adversely affected by the Troubles, Northern Ireland has undoubtedly suffered more. Between the period 1967 and 1972 real tourist revenue fell by 63% and visitor numbers fell by 60% over the same period. During years of escalating violence on the North one can also notice a negative correlation of visitors to the South.To give some idea of the potential size of the industry, one can look at the 1993 figures compiled from the Northern Ireland Tourist Board (NITB). Of the 1.26 million visitors 373,000 were from the Republic and they are estimated to have spent in the region of £30 million of the £17 3 million total revenue generated by staying visitors. This figure is approximately 1.5% of GDP and is just a third of the level enjoyed on the Republic. It is a fair conclusion to note that the Troubles have had a greater adverse effect on tourism in the North when compared with that experienced in the South.

A permanent cessation of violence would I believe have a dramatic influence on the tourism industry on both sides of the border. One would expect increased flows of visitors Northern Ireland and the Republic, and also increased interest from visitors in Europe and Americas. Of relevance here is the recent announcement of a joint venture marketing initiative between Bord Failte and the NITB which will market the whole island of Ireland as a tourist destination. It is expected that this will lead to an extra 92,000 visitors in 1995.

Conclusion

A statistical analysis of Cross Border Trade over the last two decades highlights two important trends. Firstly the increasing balance of payments deficit Northern Ireland has with the Republic of Ireland and secondly the low level of overall trade between the two neighbouring economies. In this essay I have firstly looked at the relevant statistics, secondly attempted to offer an explanation for these trends and lastly discussed the potential benefits to North-South trade which may arise through a permanent cessation of violence.

Bibliography

DKM Economc Consultants, (1994) *'The Economic Impact of the Northern Ireland Conflict'*

Business and Fianance, December 1994 *'Northern Ireland: Open for Business'*

CBI NI, (1994) *'Peace - A Challenging New Era'*

Northern Ireland Economic Research Centre, (October 1994) *'The New Economy'*

Northern Ireland Economic Council, (October 1994) Autumn Economic Review.

Mike Tomlinson, (August 1994)*'25 Years on The Costs ofWar and the Dividends of Peace'*

Oxley, Ch. 4. *'The Challenge of Free Trade.'*

Sustainable Development and Environmental Economics

Myles Hubert Clarke

For ardent ecologists, only a minimalist development strategy is morally supportable, while other analysts contend that the concept of sustainability contributes little new to economic policy. Sustainable development has two two components: economic growth and environmental protection. In this paper, Myles Clarke balances these objectives and contends that a considerable degree of both is obtainable.

"*Definitions of sustainable development abound. There is some truth in the criticism that it has come to mean whatever suits the particular advocacy of the individual concerned.*"

Introduction

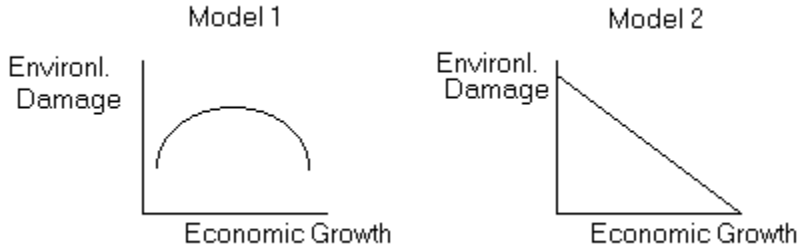
Development can be interpreted as a set of goals or objectives which society aims to achieve. Development is sustainable if it meets the needs of the present generation without compromising the ability of future generations to meet their own needs. It requires that each generation be endowed with sufficient resources to generate its own wealth. However, this agenda has been distorted by a number of factors which this essay will discuss. For instance, current measures of wealth neglect to take account of the value of natural wealth, and hence mislead the policy maker. Despite its inaccuracies, some attempt must be made to estimate the value society places on the environment and incorporate it into economic policy. The endogenising of environmental variables into economic models would reveal the interdependence between the economy and the environment; the healthy maintenance of one being a function of the other. This essay begins by setting the scene in which this discussion has evolved in a macroeconomic context and goes on to illustrate some of the solutions which microeconomic theory suggests.

Simple models have been conjected as possible explanations for the trade-off relationship between economic growth and environmental damage. They ask whether economic growth is compatable with preserving our environment. If not, is it possible to adjust current production processes so as to advance capital wealth alongside a balanced ecology?

Economic Growth vs. Environmental Damage

This model suggests that initially society is unaware of the damage its growth is causing the environment, or at least underestimates the consequences. At some stage, it becomes obvious that unless adjustments are made and contingency plans derived, the damage to the environment may be irreparable. This cycle seems plausible enough, but it is weak in many aspects. While it was derived to reflect US awareness in the 1970s of the damage caused by its growth in the 1950's, it provides little prognosis for developing countries. Although these countries now find themselves in the early stages of revolutionary growth, it is incredulous to suggest that they might be unaware of the consequences for the environment of such change. Furthermore, it takes no account of the environment's assimilative power to absorb a certain amount of growth debris and regenerate its diminishing resources. Nor does this model reflect the time lag between the period in which environmental damage is acknowledged, and the period in which institutional changes actually take impact to stop the damage.

Figure 1.



This second model possesses a stronger policy prescriptive power. It accepts that at the moment, trade-off decisions are biased in favour of growth and against the environment. Accepting this , it also provides a framework for policy makers to shift this bias more in favour of the environment. In relation to the first model, the extreme point on the Y-axis of the above diagram corresponds to the point where policy changes begin to improve the environment. This point would be preceded by a jump off and under the first curve. We start by asking ourselves two questions. Firstly, could any given level of GNP be achieved at a lower level of damage, in other words could the curve be shifted inwards? Secondly, could subsequent growth be less damaging so that the slope of the function begins to fall? Examples of how this might be done are given in the micro section. This function will reach some asymptotic limit representing a level of environmental damage inevitably caused by economic growth. This serves as the policy maker's ideal, and as an objective points policy in the right direction. Yet, we still find ourselves in the early stages of the first model where growth and environmental conservation are incompatible.

Policy makers' attitudes towards environmental damage may be *reactive* or *anticipatory*. It can be argued that each is optimal in different circumstances. However, further research reveals that there are few if any situations in which a reactive mode is optimal, and yet in practice environmental policy is rarely anticipatory. For the sake of further analysis, the introduction of time preference adds a dynamic quality to the argument and helps to determine why this is so. This entails representing the interests of generations which as yet do not exist. Nevertheless, Vilfredo Pareto provides another ideal for policy makers to set their sights upon: the intertemporal allocation of natural wealth will be optimal if no generation can be made better off without another being made worse off.

In the case of scarce resources whose exhaustion is irrevocable, an anticipatory policy should certainly be favoured. For instance, there is an unstable critical level of equilibrium fish stock, below which it is unable to sustain itself. In this instance a reactive policy will simply be too late. Indeed, such a policy can justify current levels of consumption only if expected improvements in technology are sufficient to offset future scarcity. In the case of depletable resources like oil and gas, this belief would only be tenable if substitutes which are currently uneconomical were expected to become financially viable as scarcity raised price and innovation reduced costs. This outlook can be reversed to favour an anticipatory policy if we consider that the myopia of the current generation is negated by the their risk averseness, causing them to discount heavily such technological possibilities. Furthermore, an anticipatory environmental policy is more compatable with sustainable development. Together they ensure successive generations are endowed with as much natural wealth as the current generation inherited.

Anecdotal evidence such as the depletion of the ozone layer, dead Swedish

lakes and diseased woodland in the Black Forest suggest that a reactive attitude has been adopted. But even this reactive approach has been ineffective. Authorities may be drawing up the correct legislation but it may lack the credibility necessary to make firms adhere to its regulations, or provide the incentives which should lead to research development in less "dirty" technology. The more genuine and earnest the motivations for creating the legislation are, the more credible it will be in the eyes of managers, thus leading them to search for the best least cost abatement technology available. It is more productive for a firm to utilise its resources in this fashion, rather than wasting them on rent seeking activities like circumventing legislation by finding loopholes in its stipulations.

If authorities invested more funds into research which identifies earlier how the environment would react to current trends of growth, it could prepare preemptive legislation which restricts the damage this growth might cause. Furthermore, the sooner firms can be informed of the new conditions on production, the less costly the necessary adjustments will be. An anticipatory policy would lend more credibility to legislation as firms could endogenise the imposed constraints into decisions about expanding operations in their own market or entering a new market. *Ex ante* preparation is more efficient than *ex post* alteration. It costs more for extant firms to adapt their processes to new conditions than for new firms to accommodate the same conditions before investing heavily into projects requiring long term commitment. However, if enough firms have been allowed to extend their time horizons due to effective anticipatory policy, then subsidising the alteration costs of the residual firms will cause a negligible distortion to competition. The following anecdote illustrates the significant distortions which "incredible" environmental legislation may cause in the market place. If firms are explicitly aware of a reactive environmental policy, they have the opportunity to create unfair competitive advantages. Large firms in the U.S. do this by investing in trying to predict how the environment will react in the future, given current industrial practice, and hence the legislative steps that a reactive policy would take. They then prepare for the inevitable changes which the legislation will impose but which will be difficult for smaller firms to predict. Even more resources are wasted by engaging in lobbying activities which try to speed up the legislative process and increase the benefits a firm derives from being prepared. If the market acknowledges the environmental damage of production before the relevant authorities, or at least before the authorities have time to prepare the necessary legislation, larger firms will find it rewarding, to set an asymmetric information barrier between themselves and their competitors. It is obviously worth policy makers' time and effort to invest more resources in anticipating the legislation that will be needed to handle future environmental damages sooner and accelerate the legislative process.

Microeconomic conventions have been used to establish markets for resources where they did not exist or where their existence was not evident. Secondly, they identify the role of price as the medium through which resources are allocated efficiently over time. I will proffer that in the case of certain zero priced open access resources, a market will inevitably fail. Most of the literature in this area finds provenance in the U.S. where a more informed public were aware of the damage their consumption habits caused the environment, primarily because those habits were so excessive. Moreover, the size and ecological variation one finds in the U.S. required a single federal body to monitor and regulate firms behaviour. The Environmental Protection Agency (EPA) was established with this very task in mind. It is noteworthy that while the EU has legislated its own directives, it lacks the federal power to enforce them effectively. This point demands more attention if there is the intent to create a more integrated and federal Europe.

In the case of pollution, the challenge before micro analysis is to provide incentives for firms to employ least cost structures which internalise environmental considerations. Because pollution is a negative externality, the costs of production to the firm are below the social cost, resulting in overproduction. The two most popular methods of achieving this are taxation and permit programmes. The EPA adopted the latter method. The criteria which such an institution would use when choosing any method is defined along the following lines. An important requirement it should fulfil is economic efficiency. This subsumes three divisions of efficiency: it is allocative efficient if the marginal social cost of production equals the marginal social benefits, it is cost efficient if the marginal abatement costs to each firm of meeting the given standard is the same, and it is technologically efficient if firms are given the incentive to invest resources into the R&D of cleaner, cheaper technology. These concepts will become clear when we extend them to the example at hand. There are two further important criteria. Firstly, low information requirements impose potential restrictive measures which are crucial to hit the target level of pollution accurately. Secondly, some sense of equity between consumer and producer is preserved, in other words, that the burden of the cost of the firm in adhering to these restrictions is not easily transferred to the consumer. Ultimately, it is the credibility, adaptability and dependability of the chosen policy which determines its success.

The U.S. adopted a command and control approach to its pollution problem beginning with the 1970 Clean Air Act. The EPA primarily set emission standards at a 'suitable' level of pollution, and distributed the corresponding number of permits to firms, each permit allowing a certain amount of pollution to be discharged. The permits were administered to firms with respect to the amount of pollution they historically discharged (known as "grandfathering") as long as they employed the best available control technology (BACT), and kept emissions down to their lowest achievable rate (LAER). The firms could bank, transfer, merge, or, depending on relative cost structures and technological levels, trade permits with other firms. A market had been created for pollution and by the Coase Theorem , identifying the right to pollute as being initially with firms would ensure a significant degree of efficiency. It was technologically efficient in that firms who reduced their costs would be able to produce more without needing to purchase extra permits or could maintain the same level of output and sell the excess permits. It was equitable in that consumers could buy permits themselves, reflecting their willingness to prevent pollution.

The permit programme was a success to the extent that it did result in an overall reduction in industrial pollution. But some vital considerations were overlooked. These must be highlighted if EU policy is to extrapolate any relevant lessons from the U.S. experience. First of all, it was necessary in 1977 to introduce a non-compliance penalty to consolidate the Clean Air Act. Up to this point there were no sanctions in place to enforce firms to fully comply with the BACT and LAER conditions which permit ownership stipulated. The penalty was estimated as the extra profits that a firm earned which could be attributed to non-compliance. Secondly, the permit programme violated the cost efficiency criteria discussed above. While the EPA insisted that firms used the BACT, it completely ignored the costs involved in switching technology. This 'at all cost' approach was crudely inefficient and meant the same or greater improvements in pollution levels could have been achieved at less cost. It furthermore distorted the equalising of the marginal abatement costs to firms on which the restrictions were imposed. It is debatable whether the Europeans have the free market attitude to accept such a *laissez-faire* approach to what is considered an issue which is beyond the realm of economic value. Taxation has been the preferred option. But in a dynamic context, and if the above anomalies are rectified, then

permits are the clear winners. The chosen tax levy is a function of the static conditions in the period of its inception. Over time, as more firms enter the industry, and as inflation erodes the real value of firms' tax burden, the total amount of pollution created will exceed the associated emission standard. The tax has to be adjusted continuously to accommodate the changing conditions of which it is a function. Permits, on the other hand, automatically absorb such vacillations through their pricing mechanism, thus reflecting the forces at work in the market.

While the Coase Theorem was briefly mentioned above, the significance of property rights and their correct identification if a market is to operate efficiently can be revealed by the study of an open access common property case, and explains why the free market offers no solution to such complex problems. The smog problem in Dublin in the late 1980s caused by domestic fires provides an apt example. All the characteristics of private consumption break down. Firstly, no one person can claim direct ownership over the air, so it is a non-universal good. Secondly, any improvement in air quality will benefit a whole range of agents regardless of whether they switch to smokeless fuel or not. This means air does not enjoy exclusivity. While the pareto optimal situation is that everyone benefits by switching to smokeless fuel, the threat of the free rider means that non-cooperative Nash Equilibrium prevails. While this game theory framework is simplistic in its assumptions, useful insight is given into what has become to be known as the "tragedy of commons". In this case it seems to be the responsibility of Government to take action on behalf of society, geared with the knowledge that it is in at least the majority's best interest. Yet bituminous coal consumption in the Dublin area was only banned (1990) when the public's complaints became sufficiently vociferous for the government to take credible action.

Conclusion

This essay has shown that it is possible for substantial economic growth to be achieved in harmony with an almost untainted environment. In the long run, it is in both the industrialists' and the conservationists' interests that these two variables be maintained, after all, conservationists are consumers too, and entrepreneurs, as society members and as directors of labour, benefit from a healthy environment. However one must take account of the value the environment holds for society and adopt an anticipatory approach to implement 'greener' policies. This anticipatory approach, or any for that matter, will fail unless firms consider the consequences of non-compliance seriously. To start with, we could at least rectify our national income account anomalies which I mentioned at the outset. Under current practice, an environmental catastrophe will be recorded as a gain in national income accounts equal to the income earned by the factors used to clean up the aftermath. This kind of valuation is highly subjective, but its inconsistencies do not justify its complete abandonment. It at least provides a benchmark for achieving the goal of sustainable development using well grounded microeconomic foundations.

The Role of the Traded and Non-Traded Sectors in Economic Growth

Clare Mc Andrew (Senior Freshman)

The 1989 NESC report on the 1992 single market contended that 'the key to economic growth lies in the internationally exposed trading sectors. Those sectors of the economy which predominantly serve the domestic market, such as private non-traded service activities, cannot be regarded as an independent source of sustained economic growth.' While Clare McAndrew suggests the government should devote more resources to the non-traded sector, she reiterates the widely expressed view that it is the traded sector that provides the 'fulcrum' around which the domestic economy turns.

Introduction

Growth in any economy is determined by the accumulation of its stock of physical capital, the skills and education of its labour force, its natural endowment of resources and the technology it employs in turning these inputs into output. Investment in these stocks is of crucial importance in stimulating the economic growth process. In a small open economy like Ireland, the level of this investment depends on the attractiveness of the traded sector, which is therefore seen as the key to economic expansion. By comparing the roles of the traded and non-traded sectors, first section of this essay will establish the former's primary importance as the engine of growth in a small open economy. As trade is vital for the Irish economy, competition and growth policies have focussed on its open sector, while the role of the non-traded sector is considered of less importance. However, in the second section, I will present the argument that, due to the limited impact of domestic policy on variables within the traded markets, only the performance of the non-traded sector is within the scope of such policy and therefore holds significance as the sector that governments can influence. Finally, I will discuss how the distinctions between the two sectors are changing over time and portray the effects of a single market on accelerating this process.

The Importance of the Traded Sector

The domestic production structure can be decomposed into the traded-goods sector, containing all those production activities in which industries are exposed to international competition (shoes, cars and zinc, for example) and the non-traded sector, comprising all other goods (including construction, public administration and health services). Within the traded sector a further division can be made between a natural resource-based industry, such as agriculture, and a "footloose" sector (manufacturing). The importance of this subdivision lies in the fact that the location of the farmer is determined by the geographical location of natural resources, which imposes constraints on a nation's output, a limitation, which does not arise in the context of the non resource-based sector. While 63.2% of the Irish labour force were employed in the non-traded sector in 1985, the primacy of the traded sector can be acknowledged by the fact that a small open economy such like Ireland must trade its output in order to import goods from abroad. This crucial assumption implies that as national expenditure must equal the total value of the output produced in the economy, any deficits resulting from importation must be paid-for by exporting domestically produced goods to foreign markets. As the non-traded sector serves only internal markets, its output value is limited to the share of natural expenditure within this sector, as surpluses could not be exported.

In the Irish case, due to the limited potential for increasing exports from the main resource-based traded sector (agriculture), we must concentrate on raising production in internationally traded manufactured goods in order to increase domestic consumption and future output growth without pressure on our balance of payments. The vital importance of trade to Ireland is evident since we must import goods for production and consumption due to our limited supply of natural resources and therefore must exchange exports to pay our way. Yet even if we were to imagine an economy that could be totally self-sufficient, differences in consumer tastes and economies of scale would still make it uneconomical to remain isolated. Two important points can therefore be made concerning the relative roles of the traded and non-traded sectors of a small open economy.

Firstly the traded sector is vital if the economy needs to purchase imports, and secondly it is this sector that really determines the total level of national output. Total output is determined by production in both sectors, but since the value added to this total from the non-traded sector is determined by its allocated proportion of national spending, it is expansion of the traded sector and its multiplier effects on the protected sector that determine the overall expansion of national output.

Economic policy in Ireland has concentrated on facilitating expansion of output in this sector. As prices for traded good are exogenously given from international markets, policy has focused on attempting to increase the share of world and Irish investment that the domestic open sector can attract. One of the most obvious policy measures that discriminates between the sectors is the differential tax rates on profits, with the non-traded and traded sectors subject to a 38% and 10% tax on profits respectively. Government grants and subsidies have also been directed largely towards developing export industries which are seen as the primary means to aid the balance of payments and to indirectly enhance growth and employment in other sectors of the economy. The generally accepted view of policy makers, therefore has been that enhancing the competitive position of the economy to attract a greater share of world investment to its traded sector is the key to economic expansion.

A Role for the Non-Traded Sector?

The non-traded sector is crucial to attracting foreign investment into our economy. Some economists suggest that domestic policy has a limited influence on the traded sector as it cannot affect the prices of goods traded abroad or the prices of any imports into the economy for use in production. It could be argued that what will really attract footloose investors is the relatively inexpensive input costs in the non-traded sector. Therefore, if the government wants to increase the competitiveness of the traded sector and the economy as a whole, it should concentrate on improving the supply of non-traded inputs to traded-sector industries, and ensure that they are of a comparable price and quality to import substitutes.

Over 900 foreign firms operate in Ireland but less than one-third of their material input needs are met internally. By ensuring foreign firms can be competitively supplied by domestic markets, the output of the non-traded sector as well as the real value added by foreign firms to the economy can be raised. The infrastructure and institutional environment of a country may also serve as important determinants for potential investors, so it is desirable to maintain a supply of utilities. The relative price and quality of labour is also another key competitive factor, and as the most powerful Irish trade unions are based in the non-traded industries, national wage levels are determined to a large extent within this sector. Government efforts to moderate wage levels, therefore, must be directed towards the non-traded sector,and improving education and training services should be at the top of the government's agenda. Finally, it must be recognised that spending on non-traded goods and services accounts for around 60% of national output, which means that policies aimed at price reductions and quality improvements within this sector will not only enhance our competitive position, but also directly improve the living standards and incomes of the population. In essence, the non-traded sector is crucial for the nation's economic welfare and if policy can get the non-traded sector "right", the open sector will take care of itself.

The Traded and Non-Traded Sectors - A New Reality

Although the previous discussion may highlight some important considerations for economic policy, the procedure for comparing the roles of the two sectors and assessing their relative importance in facing a new set of challenges as we move towards the competitive environment of a single European market. The increased integration of markets and elimination of trade barriers has meant that traditionally non-traded industries, particularly services, are now being faced with similar competitive conditions to those in the traded sector. the removal of tariff barriers, quotas and legal regulations and the reduced transport costs within the community has left very few goods and services which are protected from foreign competition.

The effects of a single market are most prevalent in the previously unpenenetrated Irish services sector, which will force many firms to drop prices to remain competitive. In this decade, the services sector will face the same shake-out that took place in manufacturing when Ireland joined the EEC in 1973. Dramatic changes have already come to pass in the distribution, transport and financial services sectors, whilst moves to liberalise public purchasing are beginning to expose the traditionally protected public utilities, such as energy

and telecommunications to new levels of competition. The movement towards a common market for telecommunications services and equipment (as spelled out in the European Commission 1987 Green Paper) has coaxed Telecom Éireann to review its cost and pricing structures. Finally, with increased mobility of labour and capital in the EU, the prices of goods that remain non-traded will also become indirectly linked with prices in other countries are increasingly harmonised.

Therefore, it is important to realise that if a non-traded sector's performance is damaging Irish competitiveness, the government may be able to alter the sector into a traded one. Foreign competition may then reduce prices and improve quality if the reasons that the sector is non-traded have to do with government policy. The 1992 programme implies precisely such changes for many service industries that have been non-traded.

Human Capital: A Theoretical Outline

Andros Florides (Senior Sophister)

The debate over who for investment in human capital has come to the fore since the Minister for Education's decision in January to abolish third-level fees. Is it fair that the state should pay now to boost private individuals' earnings later? And exactly who benefits from extra training anyway? Andros Florides looks at the wider picture.

The aim of this essay is to convey to the reader the importance of investment in human capital. It shall be assumed that the major factor influencing the level of human capital is the degree of investment in education and training. By referring to a simple model of human capital, we can assess the costs and benefits of such investment. Care must be taken to distinguish between the social rate of return and the private rate of return. The discussion will also take into account the diminishing marginal rate of return of investment in human capital and capital market imperfections. The distinction between general and specific training and the subsequent consequences to the workers and firm shall also be made. To conclude, brief reference shall be made to other factors that may influence the level of human capital.

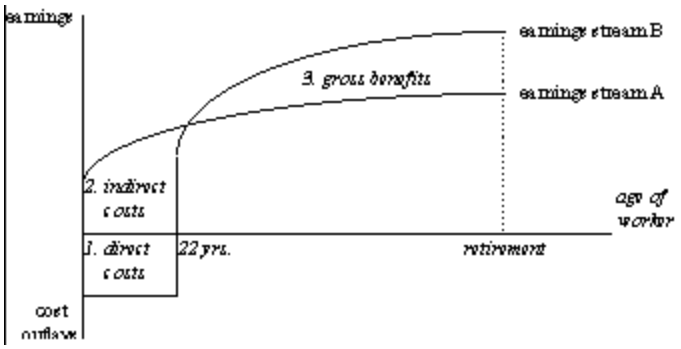
Investment in Education

The skills and knowledge embodied in an individual can be defined as human capital. All individuals attain a certain stock of human capital and this level is primarily influenced by education and training. Investment in human capital increases productivity. Adam Smith (1776)was the first to suggest that an educated worker could be likened to an expensive machine. The skills embodied in a person can be "rented out" to employers. The higher the level of skills a person has the higher this "rent" is likely to be. Thus, the expected returns on investment in human capital are a higher level of earnings and greater job satisfaction over one's working lifetime. Blaug, (1972) summarises the link between education and earnings by stating that "*the simplest explanation of the universal association between education and earnings across sectors, industries and occupational categories around the world is that the better educated are generally more flexible and more motivated, adapt themselves more easily to changing circumstances, benefit more from work experience and training, act with greater initiative in problem-solving situations, assume supervisory responsibility more quickly and, in short, are more productive than the less educated, even when their education has taught them no specific skills.*"

Investment in human capital may be analysed a by more "scientific" approach. By comparing the costs and benefits of an educational investment we are able to arrive at some conclusions as to the profitability of investing in education. Consider a simple model of human capital. Assume that a high school graduate is trying to determine whether to go to college. There are two general types of cost. Direct costs include, tuition, fees, books and supplies; the indirect costs of attending college are the foregone earnings of not entering the labour market after high school and the physical costs of studying and being examined. The costs must be compared to the economic benefits of investment in education, in other words, to the enlarged future flow of earnings. It is important to note that future benefits are worth less to us than the same benefits received today for two reasons. First, people prefer consumption today to consumption tomorrow because uncertainties make future enjoyments problematic. Second, interest can be earned by investing monetary benefits rather than using them for consumption. It is therefore necessary that the net present value (NPV), i.e. the discounted value of a financial sum arising at some future period, of the present and future costs and benefits of a college education be determined as

they accrue at different points in time. These costs and benefits can be represented diagramatically (figure 1). We must compare the costs (Areas 1 and 2) and the benefits (Area 3) in deciding if investment in education is profitable [1].

Figure 1.



The discounting formula for costs and benefits over a number of years can be formulated as follows:

$$Vp = E0 + E1(1+i)-1 + E2(1+i)-2 + ... + En(1+i)-n$$

The "E"s represent a stream of net incremental earnings; "n" is the duration of the earnings stream, i.e. expected working life; i is the interest rate. This tells us that the more distant the future earnings the greater the discounting. As with any other investment, educational investment should occur if Vp (the NPV) is greater than zero, as the discounted benefits exceed the discounted costs. Another method used in making the investment decision is to calculate the internal rate of return, r, and compare it with the interest rate, i. The internal rate of return (IRR) is the discount rate, r, at which the NPV is zero. Hence, the equation becomes

$$Vp = E0 + E1(1+r)-1 + E2(1+r)-2 + ... + En(1+r)-n = 0$$

The IRR, r, indicates the maximum rate of interest, i, that would allow investment to break even. If r exceeds the market rate of interest i, then investment is profitable. It is profitable to invest up to point at which i = r.

A number of generalisations and implications can now be made. First, the longer the expected working life, the more likely it is that the NPV of an investment in human capital will be positive, explaining why more young people than old people attend college. It is also a factor that explains the wage differentials between men and women, because female participation rates may be discontinuous on account of leaving the labour force to marry and raise children. Second, the lower the cost of investment in human capital, the larger the number of people who will find such an investment to be profitable. Thus, the lower the direct and indirect costs of attending college the higher the NPV of a college education, giving another reason why less older people attend college; the indirect costs (opportunity costs) of attending college are greater the older the individual. Third, the larger the college-high school earnings differential, the larger the number of people who will invest in college education, *ceteris paribus*. From the countless empirical studies which have estimated the returns on investment in human capital, there appears to be a general consensus that the rate of return (ROR) to an individual of a college education is between 5% and 15% above that of a non-college graduate.

The Social Rate of Return

Up to this point we have considered the rate of return to the individual; the private rate of return (SROR). As with any investment appraisal, all costs and benefits should

be included. We must take account of the effects of education on society as a whole; the social rate of return, (SROR). So, "*even if additional education did not raise the lifetime earnings, education might still be an investment from the social point of view.*"

There are certain biases in the estimated rates of return on education. This is because factors such as subsidies which are not paid by the individual are included in the SROR are not included in the PROR. Education benefits society in a number of ways. More educated workers tend to have lower unemployment rates and receive higher wages. Therefore society benefits by receiving more taxes (as more educated workers tend to work more continuously on average and the tax take from these workers are proportionally higher as they are taxed at higher marginal rates due to higher salaries). More poorly educated workers may also find crime an attractive means of supplementing their lower incomes. Society may benefit from investing in education by paying less for social welfare programmes and crime prevention/law enforcement. The children of more educated parents tend to receive better guidance and grow up in a more desirable environment.

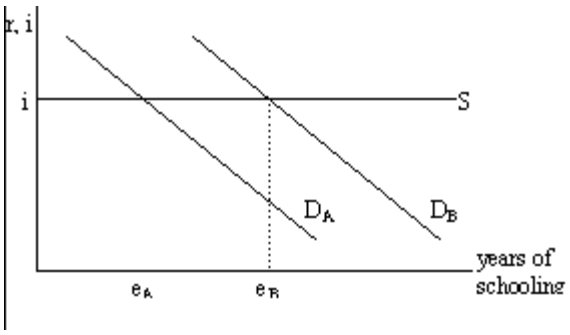
Blaug (1972) concludes that if the SROR exceeds the PROR then more investment in human capital should take place and vice-versa. The SROR also provides us with the rationale that education should be subsidised with public funds. McConnell and Brue (1989)state that "*the size of these public subsidies to education should be determined on the basis of the magnitude of the associated social benefit.*"

Diminishing Returns on Investment

As for other investments the Law of Diminishing Marginal Returns (DMR) applies and the marginal ROR on investments in human capital declines. This is because individuals have a fixed amount of time available to them, and the more years spent in education the less time spent in the workforce. This is analogous to the situation faced by women who may leave the labour force early so as to marry and raise children; there is a shorter period of future earnings. The Law of DMR may also explain why the ROR of an educational investment in Less Developed Countries (LDCs) exceeds that in Developed Countries (DCs) - the diminishing marginal ROR of further investment in human capital in DCs yields less than the essentially "new" investment in human capital in LDCs because further investment in DCs is typically investment in post-graduate courses, whereas investment in human capital in LDCs is often of the most basic form i.e. primary education. The diminishing RORs are also marginally lower due to the increased direct and indirect costs. Direct costs increase with further investment in education, as additional years of schooling typically cost more. Indirect costs, in the form of foregone earnings, also increase with additional education.

McConnell and Brue (1989) give three reasons why different people invest different amounts in human capital. First, consider two individuals, A and B, with different demand curves for human capital, DA and DB respectively, and a common supply curve, S, indicating that both individuals have the same access and terms to money capital for investment in education. DB is to the right of DA and this may be explained by B having greater natural abilities than A, where individual B can transform a given input of schooling into greater productivity, and hence greater earnings than A. Given the perfectly elastic supply curve of financial capital, individual B will invest in eB years of schooling. B's ROR on investment in education exceeds A's. The earnings differentials between B and A is further widened.

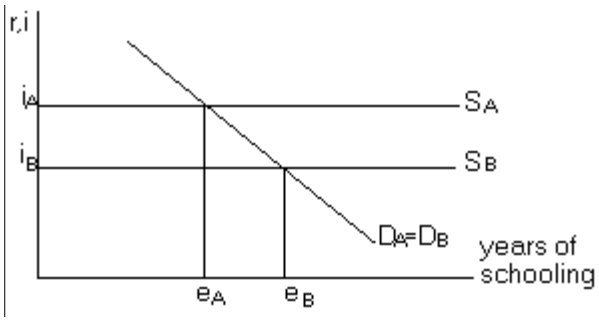
Figure 2.



Second, assume A and B are identical in terms of ability. However, their demands for human capital may not be the same due to discrimination. "A" may be black or female, for example, and therefore likely to encounter discrimination which reduces A's chance of transforming their human capital into incremental earnings. A's ROR on the same amount of education as B may be less due to this discrimination. A's demand for human capital is less than B's and this is why A invests less in education (eA) than B (eB). Discrimination in labour markets leads to less investment in education and further wage differentials.

Third, assume that A and B have identical abilities implying DA = DB. But now consider the situation where B obtains more favourable terms on acquiring money capital than A. B's superior credit rating may be explained if B is from a wealthier family than A and therefore has more collateral than A. The effect of this is shown below where SA and SB are A and B's respective supply curves for financial funds i.e. B obtains money capital at a lower rate of interest than A. Going back to the first equation, we find that the lower the rate of interest the greater the ROR. It is therefore rational for B to invest in more years of education (eB) than A (eA).

Figure 3.



Capital Market Imperfections

In our above analysis it was assumed that capital markets were perfect. We now consider the more realistic everyday situation facing students, that of imperfect capital markets. Imperfections may favour investment in physical rather than human capital because human capital is embodied in the individual and this means in the case of default, the financial institution has no collectable collateral on a loan. A physical asset, such as machinery, can be repossessed and sold to recover the loan, whereas human capital cannot. This means that a higher rate on interest is charged on loans intended for investment in human capital, so as to balance out the increased risk to the lender. More young people tend to invest in human capital, but, due to their lower credit ratings and lower set of collateral assets, fewer funds are made available by financial institutions for their investment in human capital.

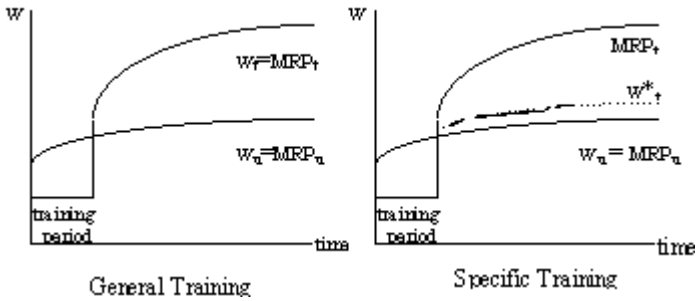
Capital market imperfections have important consequences. First, due to the increased rates in lending to students (especially those who are young) financial institutions may choose not to make funds available for education. Students from

better off families may still be able to afford a college education, while students from poorer families may not. The outcome of this is that the college/high school wage differential will tend to increase; the poor get poorer and the rich get richer. A second implication is that the government may attempt to offset capital market imperfections by subsidising education or by providing human capital loans [2].

General Vs. Specific Training

Training, just like education, is associated with an increased future earnings stream as it increases the worker's productivity. It is important to note the distinction between general training and specific training [3] and their subsequent effects on future earnings streams. General training of a worker by a firm refers to the creation of human capital which can be equally utilised by all firms in the industry. Specific training of a worker by a firm refers to the creation of human capital that is of use to that firm alone. The worker normally pays for general training since the skills acquired may be "rented" out to other firms in the form of higher wages when the person leaves the original firm. The firm will pay for specific training as these specific skills benefit that firm alone and are therefore not transferable.

Figure 4.



General training is paid for by the worker during the training period, where the worker typically receives a wage w<wu; wu is the wage of the untrained worker in the same firm. Training helps explain the convex age/earnings profile. Earnings rise quickly as the new skills are acquired, i.e. wt applies. Specific training is paid for by the firm since all benefits accrue to that firm alone. As in the above diagram, during training the worker receives a wage, wu, in excess of their Marginal Revenue Productivity (MRP). The employer is paying the worker more than his worth and is therefore losing out during the training period. However, once the skills have been acquired worker productivity is increased and the employer now gains as wu<MRPt. The employer may decide to pay an above competitive wage, w*t> wu, in order to reduce worker turnover.

Empirical Evidence

We now turn our attention to the empirical evidence. The real IRR is estimated at 14% for white American males in completing high school. Allowing for the ability of students reduces the average ROR on education so that the IRR is between 5% and 10%. Non-pecuniary benefits and costs of a job (job satisfaction, safety standards in the workplace, non-wage benefits etc.) also bias the IRR and the magnitude of such fluctuations are difficult to ascertain. Another factor complicating the measurement of the IRR is the quality of education. The human capital model implicitly assumes that each year of education is homogeneous. This is clearly not so. Layard's study of British males found that students who attend "selective entry secondary school" earned 11% more than those students who had not. A large differential exists between college and high school male graduates, illustrating that investment in education is profitable. The differentials are lesser between postgraduate and graduate earnings, explained by the diminishing marginal ROR of increased

investment. Earnings tend to level off and even decline due to the ageing process and the slow-down in training. Mincer (1974) showed that earnings peak at 33.75 years of work and then decline. It was briefly explained above that women and other discriminated groups such as blacks, have a more pronounced convex age/earnings profile than men caused by the greater discontinuity in their working careers.

The human capital model has been criticised by many economists on a number of points. First, the model assumes that all expenditures on education are investments. Blaug (1972) refutes this by saying that "*a years schooling for someone, invariably shares both consumption and investment aspects.*"8 By ignoring these consumption aspects, empirical research underestimates the ROR on educational investments. Second, non-wage benefits are also omitted from the model. The fact that college graduates obtain generally more pleasant and interesting jobs than high school graduates is also omitted and again tends to underestimate the ROR on educational investments. Third, "screening" has become a contentious issue in educational economics. Proving what schooling actually does is very difficult indeed. It is not easy to distinguish between higher wages caused by increased education, or by the fact that by grading and labelling a student, it is easier and more efficient in finding jobs that are suitable to their skills. This means that the ROR is overestimated because higher earnings may be due to credentials rather than increased productivity. Fourth, the model does not deal with the tendency that people with more innate ability (higher IQs etc.) go to college more and that they tend to do better in labour markets. Again, the ROR is overestimated. The most fundamental problem lies in the assumption that human capital can be observed in measurable units. The standard used for measuring human capital is the number of years of schooling. This statistic is by no means conclusive due to the large variation in the quality of education. Wealthier families can afford better education and this quality can not be readily compared to the quality of education afforded by a poorer family. This along with the fact that students from wealthier families tend to invest more in terms of years of education , widens the gap between the average levels of human capital of the rich and the poor. This further increases the earnings differentials between the rich and the poor, distorting a nation's distribution of income.

Other Influences

Studies have found that factors, other than education and training, are found to influence human capital and the levels of earnings. Consider religion, which may be an important dimension of family background and environment. Religion may influence the family values, morals, skills and goals of an individual which are inherited or acquired in childhood. These endowments may be important in the marketplace where honesty, diligence and reliability may affect the returns on human capital and thus increase earnings. Religious beliefs may also be of hindrance in the marketplace: the Amish rejection of modern technology excludes them from many activities for example. Canadian empirical studies have shown that Jews receive higher incomes than non-Jews. Jewish incomes for males are on average 16% higher than for non-Jewish males. The marginal benefit of education seems to be higher for Jews than for non-Jews. However, once account is taken of the fact that a large proportion of the Jewish population (95%) is urbanised, the Jewish/non-Jewish differential falls to 7.25%.

Cultural differences are also used to explain human capital and earnings differentials. Different ethnic views and cultures can lead to systematic differences in utility functions that lead to behavioural differences among women. For instance, black wives have a higher labour force participation rate than white women due to greater marital instability among blacks, extended black family households, black husbands' lower wages and less stable employment. These are just illustrations to show the diversity of factors influencing human capital and earnings differentials.

Conclusion

There are distributional undertones of investment in human capital that must be considered. In formulating policy on education it is crucial that these factors be taken into account. For instance, the tendency for wealthier families to invest more in education (generally speaking, education of higher quality) than poorer families, increases the wage differentials and hence the distribution of income between low and high income families. Subsidisation of education has been tried as a means of stimulating the increased participation of students from poorer families in third level education. The outcome of this policy was to in fact further stimulate the participation of students from wealthier families.Blaug's (1972) suggestion of educational vouchers seems to have a more positive effect in solving this problem.

The reality of the situation is that all individuals are not the same and do not possess the same skills and qualities. Individuals tend to be better at some activities than at others and it is these innate qualities that should be exploited. These innate qualities, such as cultural background, social class, religion and personal drive, are very hard to measure empirically, but it is these qualities that make separate us as individuals. Education, important as it is, is only one of a multitude of factors influencing the level of human capital.

Notes

[1] Note how the earnings differential between streams A and B increases over time.

[2] Siebert and Blaug (1985) in their studies of subsidies in education find that it is children from better-off families who tend to take up higher education subsidies. The targeting efficiency of subsidies is inadequate. Blaug is in favour of educational vouchers, whose value declines as parental income increases, as a means of evening-out the distribution of education among the different social classes.

[3] It is important to realise that pure general and pure specific training are only of theoretical use, as they are unrealistic in practice. Training of a mixed form is more likely.

Bibliography

Books

Blaug, M. (1972) : *Introduction to the Economics of Education* Harmondsworth, London.

McConnell, C. & Brue, S., (1989) : *Contemporary Labour Economics* , McGraw-Hill, New York.

Mincer, Jacob (1974): *Schooling, experience and earnings* New York National Bureau of Economic Research, London.

Siebert, D. et al, (1985) : *Surveys in Economics: Labour Economics.*

Smith, A., 1776 : *Wealth of Nations*, Book I.

IS THERE A "CORRECT" SHARE OF HEALTH CARE SPENDING IN GNP?

John Reynolds (Junior Sophister)

Spiraling health care costs are the sword of Damocles hanging over the head of many Western policymakers. Cuts too deep may hurt the poor and the elderly. Cuts too shallow may hurt future generations. John Reynolds asks in this essay whether or not it is possible to strike a happy medium. "The only truths which are universal are those gross enough to be thought so."

P.Valery

Introduction

In an unregulated competitive health care market it would be relatively easy to answer the question of what the correct share of health care spending of GNP should be - "*leave it to the market*". Unfortunately, most health care markets are (justifiably) heavily regulated. In such a system there is no automatic mechanism, such as the market, to guide it to the appropriate level of spending. This essay will examine the various techniques of economic analysis which have contributed significantly to setting the budget for health care and will conclude by outlining why there is no universal correct share of health care spending of GNP.

Determinants of health care spending

A widespread belief that an understanding of the fundamental determinants of health care spending may yield valuable insights into how such expenditure can be controlled, has led to a considerable volume of literature on these determinants. One of the most consistent conclusions in this literature is that the principle determinant of what a country spends as a share of GNP is income. Newhouse (1977) examined the relationship between medical care expenditure and income across 13 developed countries, regressing per capita medical care expenditures on per capita GNP. Consistent with an earlier study by Kleiman (1974) for a different set of countries, Newhouse reached two major conclusions;

(1) Firstly that GNP accounts for most of the variance in medical care expenditures across countries, and secondly that;

(2) the income elasticity of medical care expenditures across countries exceeds

one - by definition this implies that, at the margin, medical care is a luxury good.

Newhouse felt that in countries with high expenditure, the marginal unit of medical care is more likely to produce improvements in so-called subjective components of health, such as relief of anxiety and more accurate diagnoses, rather than improvements in morbidity and mortality rates. Could it be the case that countries spending more on medical care may well provide additional caring, but little additional curing ?

Table 1: Expenditure on Health in Low-Income and High-Income Economies

Country	%of Government Expenditure Spent on Health Care Services (1983)
Low Income Economies:	
Nepal	4.5%
Sri Lanka	5.1%
India	2.4%
Pakistan	1.0%
Industrial Economies:	
Germany, F R	18.6%
Australia	7.1%
Canada	6.3%
USA	10.7%

Source : Asian Development Bank et al. (1988)

Table 1 also supports the basic finding that national income largely determines the level of health care spending in an economy. In this table, low-income countries spend a smaller proportion of total Government Expenditure on health care : the percentage of public expenditure on health ranges from 1.0 to 5.1 per cent in low-income countries and from 6.3 to 18.6 per cent in the high-income (industrial) economies. In a break from the traditional line of thought, Ulf Gerdtham (1992) concluded that the age structure of the population of a country may be of prime importance in determining the level of health care expenditure. He discovered, in his research of health care expenditure in Africa, that the demand for medical services fluctuates with age - those under 15 years of age utilise medical services more than average. Yet in accepting Gerdtham's proposition, it is also of fundamental importance to note a number of problems associated with cross-national comparisons of health expenditure, as outlined by Robert Leu(1986); Definitions of health etc. are insufficiently standardised; exchange rate conversions always have a degree of randomness, and; input prices may be positively correlated with the level of national income.As a result, it is not surprising that, when based on international comparisons, a judgment of the effective impact of health care is elusive. To conclude, none of the cited studies of the determinants of spending on health care explicitly combine resource use (or cost) with enhanced or maintained well-being. They give no indication whatsoever of the 'right' level of spending. It is clear that in order to make progress on what to spend on health care, there must be more precise data on the productivity of health care interventions.

A Correct level of Spending?

On the whole issue of changes in government expenditure on health services, Sean Barrett, in his study of the "Social and Economic Aspects of the Health Services"(1979), points to several disquieting signs that increases in expenditure on the health services have been accompanied by a reduction rather than an improvement in the health of the community. He refers to Keating's article (1976) which reveals that the life expectancy of males of 30 years of age declined between 1960 and 1970 despite increases in government spending on the health service during that period. Moreover, in Tokyo in 1973 the International Economic Association concluded that healthcare is only one input into the process by which the health of the individual is improved. Income, education, lifestyle, work environment, work status, housing and health care all affect an individual's state of health. The following are the results of a U.S. study which analysed the potential changes in mortality rates which would be associated with a 10 percent increase in some variables;

Table 2: Percentage change in Age-specific Mortality Rates resulting from a 10 percent increase in several variables.

Income	Education	Cigarette Consumption	Per Capita Health Expenditure
+2.0%	-2.2%	+1.0%	-.065%

Source : Culyer (1976)As can be seen from the above table, the increase in per capita health care expenditure reduced the mortality rate by 0.65 percent, but a similar increase in education expenditure reduced the mortality rate by an even larger percentage of 2.2 percent, proving that an increase in healthcare expenditure is not necessarily the most effective means of reducing mortality and therefore, increasing the 'healthiness' of a community.Conclusion? In conclusion, little evidence can be drawn from economic studies on what is the most appropriate share of healthcare spending of GNP because there is no universal correct share. Health care is shaped by too many determinants - income, age structure, the structure of the health budgeting system (the more centralised the system is, the lower is the share of health care spending (Culyer 1988)), and each of these determinants vary in importance in different parts of the world.

To assume a universal share would be to deny all these international variations. For example, if GNP is the most important determinant of health care spending in a particular country, little can be done by way of direct health care policy in setting the budget. Governments have to aim to increase GNP if they want to increase the budget for health care. As I have already stated, GNP is the most important determinant of health care spending in most countries, but what about the African countries where the age structure of the population plays a key role in determining such spending. To assume there is a correct level of health care spending in GNP with respect to the GDP of a country would be to ignore these African countries. Therefore I believe there is no such 'correct' share of health care spending in GNP - there are no such "universal truths". I believe that it is up to the countries to decide for themselves what is the most suitable level of health care spending by use of the appropriate techniques of economic evaluation, namely cost benefit analysis, which should include an investigation of the determinants of health in their country. The result will not be a universal correct share but it will help the various countries to decide where increases and decreases in resources are best applied.

BIBLIOGRAPHY

Asian Development Bank et al., (1988) Economic Analysis of the Environmental Impacts of Development Projects, pp.39-40.

Barrett, Sean D., (1979) "Social and Economic Aspects of the Health Services", Irish Banking Review, March 1979.

Culyer, A., (1976) Need and the National Health Service, chapter 4.

Donaldson, C. & Gerard, K., (1993) The Economics of Health Care Financing, chapter 10.

Ekins, Paul, (1986) The Living Economy : A New Economics in the Making.

Gerdtham, Ulf, (1992) "Determinants of Health Care Expenditure in Africa : A Cross-Sectional Study" in World Developments 20(2), February 1992, pp.303-308.

Keating, W., (1976) "An Analysis of recent demographical trends with population projections for the years 1981 and 1986" in Journal of the Statistical and Social Inquiry Society of Ireland 1976-77, p.123.

Kleinman, E., (1974) "The determinants of National Outlay on Health" in M.Perlman (ed.): The Economics of Health and Medical Care.

Leu, Robert E., (1986) "The Public/Private mix and International Health Care Costs" in Culyer and Jonsson: Public and Private Health Services.

Newhouse, Joseph P., (1977) "Medical Care Expenditure : A Cross-National survey" in Culyer: The Economics of Health Volume 2.

What Proportion of Wage Differentiation is Just?

Ciarán Ó Ceallaigh (Senior Sophister)

Should some people be paid more than others? This seems a question central to economics and yet it is not one easy to answer. Ciarán Ó Ceallaigh tackles the problem in this essay, examining the theoretical arguments in favour of wage differentiation and evaluating their practical relevance.

Introduction

That there is variance in the amounts people are paid for supplying their labour is an obvious fact of economic life. This is not only true of market economies, but of all economies. There seems to be a consensus among economists that wage differentials are a good thing. If everyone was paid equal amounts, what incentive could be used to persuade some people to perform less pleasant tasks? Equal pay would effectively reward the lazy, as hard workers would receive no extra income in return for their efforts. Such objections would obviously make a critique of wage differentiation fruitless. However, the story does not end there. We must ask why people are paid different amounts. Having explained the main reasons for wage differentiation - and having worked towards a definition of 'just' - we can examine whether each of four reasons is a just reason for wage differentiation. We will then be much better placed to assess the ultimate justice of wage differentiation. It must be added that such an endeavour necessarily involves a degree of value judgement. As such, economics being a human science, we should not expect concrete conclusions. Such is the risk with going beyond observational discussion.

Justice

As an ambiguous moral value, the notion of justice has been under scrutiny for millennia. If we wish to arrive at a working definition, therefore, which we shall adopt as an assumption, we would be wise to narrow our search down to the question 'what is a just reason for wage differentiation?' It may be argued that the answer to this question involves two concepts, contribution and equality of opportunity.

The principle of contribution seems quite self-evident. Put simply one applies the maxim 'to each according to his contribution'. As such, it would be just for those who give more in terms of both production and utility to the employer and society respectively, to be paid more. This seems to be quite simple. However, according to John Rawls,

'[t]he marginal product of labour depends upon supply and demand. What an individual contributes by his work varies with the demand of firms for his skills, and this in turn varies with the demand for the products of firms. An individual's contribution is also affected by how many offer similar talents. There is no presumption, then, that following the precept of contribution leads to a just outcome unless the underlying market forces, and the availability of opportunities which they reflect, are appropriately regulated.'

What is Rawls saying here? It could be argued that he is saying that a wage system set according to contribution and efficiency may not be just unless such a competitive system allows for both choice of occupation and equality of opportunity.

So, how does our principle stand at this point? A difference in two wages is just if it is consistent with the principle of 'to each according to his contribution', allowing for certain demand and supply factors and for the principle of equality of opportunity. As 'contribution', in its relevant sense, is dependent on the scarcity of the relevant skill, the market dynamics must not be discounted. What, however, is equality of opportunity? Put simply, it means that '[I]n all sectors of society there should be roughly equal prospects of culture and achievement for everyone similarly motivated and endowed. The expectations of those with the same abilities and aspirations should not be affected by their social class.'

We have, it seems, arrived at something of a working definition. More coherent in its negative sense, an unjust reason for a wage difference is one that is in breach of the principle of 'to each according to his contribution' and/or the principle of equality of opportunity. This excludes any direct association of justice with efficiency. An efficient wage is to be found at the equilibrium point - the intersection of the demand and supply curves for labour. A just wage may or may not be at this point, though, as has been pointed out, the notion of 'contribution' is, to an extent, a market phenomenon.

It is probably regarded as quite unusual for an economist to write on justice- normally the reserve of the political theorist. Such a situation however, is merely unusual, not unknown. In reality, justice is an indefinable, even non-linguistic, concept. It is an emotional principle upon which the preservation of a community or society is based. The definition given above is quite narrow. In essence, we are comparing wage differentials to two basic assumptions economists make about society. We assume that it is fair, and we also assume that greater 'expenditure' of labour should, more or less, reap a greater return. It is not incorrect to describe such assumptions as 'justice'. Nor is it the full story of the moral philosopher and political theorist.

Basic Foundations

It is quite obvious that if every person and job were the same - if workers and jobs were homogeneous - wage levels would be uniform. It is also quite obvious that this is not the case. Jobs and workers are heterogeneous. The heterogeneous nature of jobs leads us to one reason for wage differential. If we think of a wage as compensation for a worker, i.e., for the loss of leisure time, it seems quite reasonable that some jobs require greater compensation. For example, some jobs may '...involve disagreeable working conditions - dust, noise, extremes of heat or cold, risk of injury, heavy physical effort - while others have light, clean, and pleasant surroundings.'

So some wage differentials may be a result of the work environment. The premium paid over the homogenous wage due to work conditions is known as the 'compensatory wage differential'. Is the unpleasant nature of a work environment a just reason for paying different wages? It seems that it is. By no means does it breach the principle of equality of opportunity. It could also be argued that, by performing an unpleasant task, a worker is contributing relatively more. As such, it seems that an employer would be justified in paying a compensating wage differential. The compensating wage differential, however, in no way accounts for wage differentials on the whole. Such differentials may generally be concentrated amongst unskilled workers. Even there, the evidence is not concrete:

'...studies generally find, for example, that wages are one-half to 2 percent higher for workers in manufacturing industries with the average risk of job fatalities (about 1 in 10,000 per year) than for comparable workers in industries with half that level of risk. If the death risk data are collected and correlated with wages by occupation instead of industry, however, the theory finds less support.'

As such, it would be best to look elsewhere.

Human Capital

According to P. T. Bauer, '...people are not equally motivated or endowed in their economic attitudes and attributes, and this explains substantial economic differences between them in open societies.' While not completely untrue, such reasoning could be seen as quite crude. There are many more concrete reasons for wage differentiation. Marvin Kostersof the American Enterprise Institute in Washington D.C. attributes roughly half of the variation in earnings to educational level and work experience.

In Ireland, the Economic and Social Research Institute stated, in a report of the 11th January, 1995, that a person with no qualifications is five times more likely to be 'poor' than a person with a Leaving Certificate. It also revealed that the risk of poverty was twenty times greater for the unskilled worker than for the professional or managerial worker. In the basic model, human capital is built up through a procedure of investment: '[w]hen a person makes a current expenditure on education or training, it is anticipated that one's knowledge and skills and, therefore, future earnings will enhanced.'

To put it simply, one incurs costs in the hope of reaping future benefits. Such hopes have a solid basis. For example, in the United States wage increases for college graduates out-paced those of high school graduates

by over 15% between 1979 and 1988. The same is true of on-the-job training. One incurs a short-term opportunity cost, in the form of a wage reduction, thus opening up the opportunity to earn high wages in the future. Such a system is true more often of general, as opposed to specific, training. General training gives the worker skills that may be applied outside the relevant firm. As such benefits accrue not only to the employer, but to the employee as well. Wage reduction, therefore, is a means of sharing the costs between the two beneficiaries. Specific training, on the other hand, leaves one with skills that are useful solely to the employer. As this prevents the worker from gaining a return, in the form of higher wages, from the training, the employer should logically bear the brunt of costs, i.e., should not reduce wages.

While based on a very simple principle, human capital theory is highly relevant to the explanation of wage differentials. However, when applying the principle of justice to human capital theory, it may be wiser to examine what McConnell and Brue call a 'modified human capital model'. Having established that greater schooling, etc., leads to higher wages - and, by implication, a greater divergence in schooling levels leads to greater wage differentiation - we should ask what causes some people to invest more in human capital than others.

'Many economists believe that we can better understand why the earnings distribution is skewed rightward by expanding the human capital model - that is, by taking a multifactor approach to earnings distribution - to include additional factors such as (1) ability, (2) family background.'

Put simply, ability is the inherent capacity to utilise one's human capital. It is the "power to do'. Economists who emphasise the significance of ability believe that differences in the attainment of human capital and the returns accrued can be reduced to differences in ability. So, for example,

'[p]eople who possess higher intelligence are more likely to choose to attend college than those with less intelligence. Even if these highly intelligent people did not go to college, they could be expected to have larger earnings than less intelligent people who did not attend college. In other words, if we could somehow control for the skills and knowledge gained during college, this high ability group still would have substantially higher earnings than their less able counterparts. Consequently, according to this view, much of the inequality of earnings normally attributed to differences in education and training is actually the result of differences in ability.'

A great deal of research has been performed into ability. Scholars have attempted to compare ability to earnings, holding such things as education equal. However most have come to the conclusion that the problem of 'ability bias' is small. It could be argued that the primary factor in family background is motivation. A child from a more wealthy family is much more likely to demand higher education than children from less wealthy families. A major aspect of the intergenerational poverty trap seems to be the lack of motivation to acquire human capital on the part of children who come from less wealthy backgrounds. Such people also don't share the financial advantages or contacts of the wealthy. It may even be the case that malnutrition (as opposed to undernutrition) may affect the ability of poor people for upward mobility. As such, earnings differentials may be a result of human capital gained through amongst other an inter-related mixture of ability and family background.

So, in this light, is human capital a just reason for wage differentiation? Being more of a sociological than an economic phenomenon, this question is very difficult to answer. However, as human capital differences account for in or around half of wage differences, it should be attempted. Ostensibly, it seems, rewarding people for having greater human capital is quite just. It is most definitely not in breach of the principle of contribution. People with higher skills are more productive. It does not seem to contravene equality of opportunity either as, at its basic level, everyone has the chance to increase their stock of human capital. However, if the above is true, and family background has an effect on acquisition of human capital, it is in breach of equality of opportunity. It means that, in part at least, one may have a greater wage by virtue of the fact that one was born into a wealthy family. The question then becomes one of degree. To what extent is family background relevant? If it is the greater reason for differences, then human capital is an unjust reason for paying differing wages. On the other hand, however, financial constraints are reduced by state intervention in the form of grants, scholarships, and student loans, all of which reduce the supply price for this group.

In Ireland the cost impact of education is lessened by a structure of grants and subsidies. In the Programme for Competitiveness and Work, the Social Partners promised '...a focusing of resources towards the

disadvantaged', in terms of education. Nevertheless, costs still seem to be high, and the wealthy also appear to have an advantage. So, again, is human capital a just reason for wage differentiation? All that can be said is that it is probably more just than unjust. As an investment, one should reap a return. However, it may be the case that there may be inequalities in who gets to make the investment. It seems that to know for sure we would have to answer some virtually unanswerable questions. What proportion of those with little in terms of human capital are like that due to lack of ability, and what proportion due to lack of motivation? Of these, what proportion lack the ability or motivation due to underprivileged backgrounds?

Discrimination

Discrimination may be given two definitions. Firstly, if one person is paid less than another person of equal ability for doing the same job, that person is discriminated against. In the second place, if a group in society, howsoever they may be defined, receives a per capita income that is smaller than that of a majority group, usually white males, we question whether that group is discriminated against. For example, in 1990, according to the American Bureau of the Census, white females earned 65 percent of the income of white males, while black females earned only 58 percent. This is despite the fact that '...the wage differential between continuously employed men and women is closing over time for almost every race-schooling group.' In order to discern the just of discrimination, we shall have to examine two aspects of the phenomenon: wage discrimination, i.e., what has been cited above, and human capital discrimination.

It is quite obvious that women are paid less than men. However, is this a result of unjust prejudice? Partly, but that is not the whole story. As a result of the traditional status of women in society, family obligations, while not dissuading women from working, 'may influence their choice of occupation and hours'. This narrowing of women's occupational horizons seems to lead to the distribution of the occupations of women being skewed towards those which are low paid. The narrowing of horizons, however, is not the only reason. Women seem to be endowed with less human capital than men: '[t]he advantage which white males have enjoyed compared to white females and blacks in obtaining a college education has been magnified through the greater access these white males have had to postmarked jobs training which has increased their productivity and earnings.'

However, differences in human capital stocks, no matter how much a result of prejudice, are not sufficient to explain discrimination. We must look at segregation between careers as regards women: highly paying professional careers can be over 90% male, while low paid clerical careers can be over 90% female. Here, both men and women have invested in their stocks of human capital, but the women have, or have been, focused towards careers where the return on human capital is relatively low. How much of this is a result of preferences and how much is a result of discrimination, it is impossible to say. However, the problem is serious enough for some analysts to suggest programmes of 'affirmative action', either give the disadvantaged a foothold in a previously discriminatory environment or explicitly promote such groups through the use of quotas. Though difficult to quantify, discrimination does exist as a reason for wage differentiation. As such it could be said to be totally unjust. It is a violation of both the principle of equality of opportunity and the principle of contribution. As such, many would argue that such intervention in the labour market as quoted above may be justified. That, however, is for policy-makers to decide.

Labour Unions

Among economists, it is generally agreed that labour unions increase wage differentials between members and non-members. If a union negotiates a wage increase for its members, the employer's demand for labour will fall. When this happens, workers who would have been demanded by the employer under the old wage, swell the ranks of jobs-seekers, thus putting downward pressure on non-union wages. So, when union wages rise, according to the theory, non-union wages fall. Indeed '...between 1979 and 1984 union wages rose 5.1 percent more than non-union wages in manufacturing and 5.6 percent more in non-manufacturing industries.'

An interesting aspect of this is the public service trade union movement. Here, pay increases have more far reaching effects. In a report by George Lee, it was shown that wage levels since 1987 in the Irish public service have '...far outstripped both average industrial earnings and inflation over the period.' That is, wage levels increased by just under 7 percent while inflation lay at just over 3.5 percent per annum. Since it is very difficult to perform rationalisation in the public sector, demand for labour cannot fall. The wage increases,

therefore, must be paid. However, a greater differentiation in wages does not come about in this direct manner only. As the wages must be (eventually) paid for out of taxes, taxes must increase to pay more. This means that net incomes will fall, thus increasing the wage differentiation. If taxes are not increased, public services must be cut back, thus, in theory at least, decreasing the welfare of the people. Although, where the effects of unions are concerned, there is much evidence that the tendency for earnings distributions to widen is offset by other factors.

Let us examine the justice of union-related wage differentials. It is quite obvious that the equality of opportunity principle is not breached, there are few restrictions on who can join a union. However, what about the contribution principle? Do union members contribute sufficient amounts to justify higher wage increases? Probably not. They may contribute more, often being skilled, but their contribution might not rise sufficiently to justify larger increases in wages. So, we may be able to say that, it is just that they receive higher wages, but unjust that they receive higher wage rises.

Conclusion

So, is wage differentiation just? We have discovered that stock of human capital and compensatory payments seem at least to be more just reasons than not for wage differentiation. On the other hand, discrimination is in most cases unjust. Union intervention is unfortunately probably too complex to assess. As such, it seems that wage differentiation is in general just, given that stock of human capital and compensatory payments probably make up half or more of the reasons. However, if it is just, it is not perfect. Many might recommend care to be taken through minimising unjust wage differences. The crucial point is to find a balance between intervention and interference. We must, it is argued, maximise the advantage s of the market, but minimise the inequities created by the people in the market. In short, we must moderate ourselves to searching for a golden mean.

Bibliography

Rawls, John *A Theory of Justice*, Oxford University Press, Oxford, 1973

Reynolds, Lloyd G., et al *Labor Economics and Labor Relations*, Prentice-Hall, London, 1991

Ehrenberg, Ronald G. and Smith, Robert S. *Modern Labor Economics*, Harper Collins, New York, 1994

Bauer, P. T.*Equality, the Third World and Economic Delusion*, Methuen, London, 1981

Kosters, Marvin H. 'Schooling, Work Experience, and Wage Trends', in *American Economic Association Papers and Proceedings* (Vol. 80, No. 2, May 1990), American Economic Association, Princeton (NJ), 1990

McConnell-Campbell, R. and Brue, Stanley L.*Contemporary Labor Economics*, McGraw-Hill, London, 1989

Programme for Competitiveness and Work, Government Publications, Dublin, 1994

Light, Audrey 'Gender Differences in Wages and Job Turnover Among Continuously Employed Workers', in *American Economic Association Papers and Proceedings* (Vol. 80, No. 2, May 1990), American Economic Association, Princeton (NJ), 1990

McClelland, Peter D*The American Search for Economic Justice*, Basil Blackwell, Oxford, 1990

Rent Seeking and the Social Costs of Monopoly

Alan Dunne (Senior Sophister)

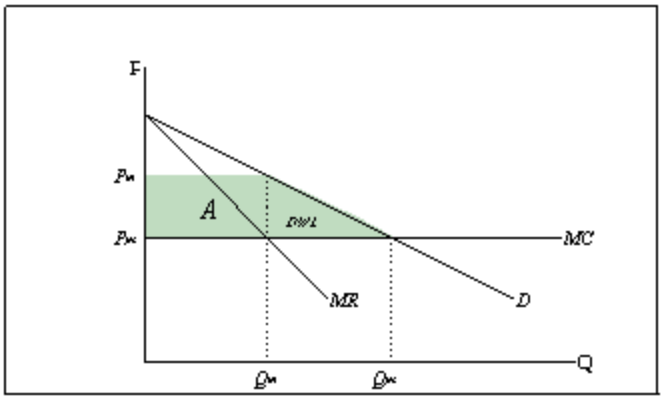
Attention to the social welfare loss under monopoly conditions was led by the pioneering work of Harberger in 1954. Subsequent writers have considerably expanded our understanding of the extent of the losses involved. Alan Dunne, in this paper, reviews this material, the focus on the application of rent-seeking theory to monopoly, its type, cost implications and solutions.

Economic theory suggests that monopoly results in a social loss because output is restricted below its optimal level, meaning that marginal benefit and marginal cost are not equated. Traditionally this social loss has measured in terms of the deadweight loss (DWL) of monopoly. However, this measure of social loss assumes that the monopoly is costlessly created and maintained. In fact, the opportunity to earn monopoly rents results in resources being invested in unproductive activities in their pursuit. In other words, rent seeking occurs. This essay examines the theory of rent seeking as applied to monopoly. The types, cost implications and solutions to rent seeking are discussed in turn. In conclusion, it will be evident that the costs of rent seeking are largely determined by the precise nature of the rent seeking game.

The Social Cost of Monopoly

The theory of monopoly states that a monopolist earns supernormal profits by restricting output and hence increasing prices above its perfectly competitive level.

Figure 1.



When price rises above this level, a transfer of income from consumers (who continue to consume the good) to the monopolist (measured by area A in the diagram) occurs. A further loss, known as the deadweight loss (shaded triangle), is incurred by people who stop buying the product. This refers to the consumer surplus that would have been generated by consumption of the good between Qm and Qpc, a quantity now neither produced nor consumed. However, this analysis hinges on the assumption of the monopoly being created and maintained costlessly. In fact, the deadweight loss underestimates the social cost of monopoly as the existence of an opportunity to earn monopoly profit (or rent) attracts resources into efforts to obtain and maintain

monopolies. This activity is known as rent seeking. Furthermore, resources may be expended wastefully by opponents to the creation of a monopoly; in other words, a reaction such as "rent protection" may be provoked.

Tullock (1967) employs the analogy of theft to explain the problem of rent seeking. The transfer of wealth from victim to thief involves no social loss; it is a direct transfer and, summing over all individuals, society's wealth remains unchanged. However, the opportunity for such transfers encourages the thief to invest his resources (human capital and tools) in theft, that is he engages in rent seeking. The potential victim, meanwhile, aware of the possibility of theft, will invest in locks, and alarms to prevent the transfer of wealth, i.e., rent avoidance. The fact that both parties have employed resources unproductively implies a social loss to society, regardless of the outcome. From a societal point of view, it would be much more beneficial if the threat of theft was absent and both parties invested in the production of goods merely for society's consumption. The net result is the inefficient use of resources by society, and therefore a location off the production possibilities frontier.

What types of costs?

Any cost incurred in the competition to obtain or maintain a monopoly is a cost of rent seeking. Types and magnitudes of costs will vary depending on the type of monopoly. Typically, monopolists will incur both strategic and administrative expenses. Strategic expenses may include research and development expenditure (R&D) in an attempt to obtain a patent, or the accumulation of capital as a barrier to entry.Tirole (1988) points out that since a patent is a winner-take-all game, firms may have a tendency to over-invest in research and development in order to obtain a patent and also invest in "risky" technology which has a potentially high profit. Both activities lead to the dissipation of monopoly rents.

Furthermore, a monopolist may invest in capital as a signal or precommitment to potential entrants of his long-term interest in the industry. Such investment may be necessary to make the threat to fight potential entrants credible. In theory, a firm may dissipate all profits in this manner, particularly in the short run, to maintain its position or to develop a reputation for fighting entrants.

Administrative expenses include the cost of lobbying, and also the substantial cost of human capital invested in a bid for monopoly. As part of their lobbying effort, firms may bribe or give favours to a government official (for example, one who allocates a licence). One could argue that there is no social cost incurred here as the result is simply a direct transfer. However, Krueger (1972) points out that individuals will have to invest time, energy and resources in competing for the position of licence official in the knowledge that success will be rewarded with bribes and favours, leading to a dissipation of rent in unproductive activities by individuals vying for such a position.

How large are the costs of rent seeking?

Posner (1975) argues that when the total expenditure by firms to obtain the rent is exactly equal to the rent, the expenditure has no socially valuable by-products, as the total cost of monopoly will equal the deadweight loss plus the monopolist's rent. Total rent dissipation occurs when competition for rents is perfectly competitive. Each firm invests in rent seeking to the point where the last dollar spent equals the improved probability of obtaining the rent, for example, in equilibrium ten firms with a 10% chance of getting a rent of £10m will each invest £1m. In the circumstances, the social cost of monopoly can be measured by measuring area A + DWL in diagram 1.

Early measures of the social cost of monopoly have been subject to much criticism. Posner criticised Harberger's deadweight loss for neglecting such effects of rent seeking. In particular, he criticised the rates of return used to calculate the monopoly

price increase. He argued that while many monopolists enjoy supernormal rates of return *ex post*, unless account is taken of expenditures by the firm in obtaining and maintaining its monopoly position, the percentage of the monopolists revenue that is attributable to monopoly pricing is underestimated. Hence monopoly profits, rates of return and the associated deadweight loss are all underestimated.

Posner overcame these problems by obtaining from industry studies estimates of demand at the relevant points along the demand curve. Using these estimates, Posner calculated that the social cost of monopoly was as much as 30% of industry sales in some industries. However, it must be recognised that the validity of these results depends crucially on the validity of his initial assumptions.

Are rents totally dissipated on socially wasteful expenditure?

Posner's results may be questioned on a number of counts. There are many reasons to suggest that the value of rent seeking expenditures will not equal the total value of rents. Firstly, monopolies can be obtained through mere luck or chance rather than foresight, as such a product may be invented and patented without excessive expenditure in R&D. Secondly, the contenders may not compete from an equal footing.

Some interests are better organised than others, ie. they may be in a stronger position, have strong contacts, more information and so on, and may be expected to have a higher probability of becoming a monopoly, In this case not all of the rent will be dissipated because firms will realise that one firm has a particular advantage and so they may not enter into the competition in pursuit of the rent. In general, assuming a limited number of players, and knowledge as to the probaility of success the total value of rent seeking expenditures will be equal or less than the total value of rents. This is because it would be irrational for a firm with only a 50% probability of obtaining a rent worth £10 million to invest more than £5 million in rent seeking activities. However, with larger numbers of players and uncertainty regarding probabilities of success, firms will not be able to make rational decisions based on perfect information. In this case expenditures could be less than , equal to or exceed the total value of rents. When one considers that many other firms may also engage in rent avoidance then this argument is plausible. Indeed, Posner argues that due to rent avoidance the social cost of monopoly may be high even for a monopolist earning only a normal return. If this occurs, the observed monopoly profits is an industry will underestimate the social costs. The cost of maintaining a monopoly may also be difficult to measure, for example an inefficient state monopoly may exist which would perform more efficiently if exposed to the rigours of the marketplace. However, workers and trade unions will lobby to prevent deregulation and may threaten industrial action. At the same time the government may be lobbied by private interests who wish to enter the market. In such cases, the social cost of the rent seeking is virtually impossible to measure.

Dissipation of rents is costly only to the extent that no socially valuable by-products are produced. However, are rent seeking expenditures socially wasteful ? In the era of a regulated monopoly position being allocated on the basis of lobbying influence, the expenses are socially wasteful. However, if the allocation is achieved by an auction the expenses are received by the government and are not wasted, (assuming the government employs the resources productively). There will also be many intermediate cases, perhaps R&D expenditure for a patent, advertising or capital accumulation as a barrier to entry. Thus, the extent to which any socially valuable by-products result from the process depends on the nature of the game involved.

Monopoly rents may be partially transferred to input suppliers. Assuming the input supply does not increase in response to this transfer then it is not considered to be socially wasteful. However, rent seeking activities by individuals trying to gain employment in industries with monopoly rents would be described as socially

wasteful.

Thus there are strong reasons to suggest that the assumption in Keynes' analysis does not hold. However, the value of rents are as likely to overestimate as to underestimate the value of rent seeking expenditures. Also, in many of the cases we analyse, particularly the case where a licence is allocated on the basis of lobbying, there are few socially valuable by-products. This implies that, while Posner's estimates of the cost of monopoly may be crude and inaccurate, they do depict the general picture in many industries in that the cost monopoly are evidently quite substantial.

Conclusion

Rent seeking greatly increases the social cost of monopoly. However, the true cost of rent seeking is dependent on the particular "game" or determination of monopoly. However, from some empirical evidence, it is apparent that in many industries the social cost of monopoly is significant. In general, prevention of that social cost can only be achieved by restricting entry into the activity for which a rent has been created. However, Krueger (1972) points out that doing so would have serious negative political implications, in that such a process would engender the scepticism of the electorate if the government was deemed to be showing favouritism. Efforts at reducing rent seeking costs should focus on altering the game so as to promote efficiency, such as allocating a licence by auction rather than by traditional lobbying. Therefore while the costs of rent seeking can in theory be significant, costs can be reduced by focusing on reducing inefficiencies in individual cases.

Bibliography

Books

Tirole, J., (1988) : *The Theory of Industrial Organisation*, MIT Press, Cambridge (MA).

Tullock G., (1967) : *The Welfare Costs Of Tariffs, Monopoly and Theft*.

Articles

Krueger, A. (1972) : "The Political Economy of Rent Seeking" in *American Economic Review*.

Posner, R., (1975) : "The Social Costs of Monopoly and Regulation" *Journal of Political Economy*.

Monopoly and X-Efficiency

Justin Morton, Senior Sophister

The concept of X-Efficiency rejects the technical efficiency notion of profit maximising and cost minimising. Liebenstein (1966) argued that individual workers are free to choose their level and interpret their own jobs . The equilibrium position for a firm price is when every individual maximises utility. In this essay, Justin Morton relates the theory of X-Efficiency to the monopolistic market structure.

INTRODUCTION

Adam Smith refers to monopoly as a "*...great enemy to good management*", and to competition as the medium for "*...new divisions of labour, and new improvements of art, which never otherwise would have been thought of*" Hayek claims that "*It is only through the process of competition that the facts will be discovered*." Cournot, after whom Cournot competition is named, finds that "*the result of competition is to lower the price.*" If increasing social welfare or well-being is an objective, then we should have some notion of which market structure is the most desirable. With this we can design competition policy and competition law so as to maximise the economic cake. Hence in this essay, I will examine the market structure of monopoly and its associated costs, concentrating on the theory of X-Efficiency.

Competition Theory

Any study of monopoly would be vacuous without firstly outlining the underlying (neo-classical) theory of competition. It starts by assuming perfect competition in the goods market. This involves infinite buyers and infinite sellers, each with perfect information regarding costs, profits and demand, freedom of entry and exit into/from the marketplace, all selling an homogeneous good. The factor market is also perfectly competitive, which means the marginal productivities of both capital and labour are known, and all factor contracts are complete. In both spheres agents are profit and utility maximisers subject to constraints (budget, leisure, ability cost and technology). Given that these conditions are fulfilled in all markets consumer welfare is maximised (or more technically the economy is in a pareto equilibrium). Economic resources are allocated in the precise way consumers wish, wishes being reflected by the price system. As well as allocative efficiency, perfect competition leads to productive efficiency (minimum average cost production, in other words), since above minimum average cost selling would mean zero sales.

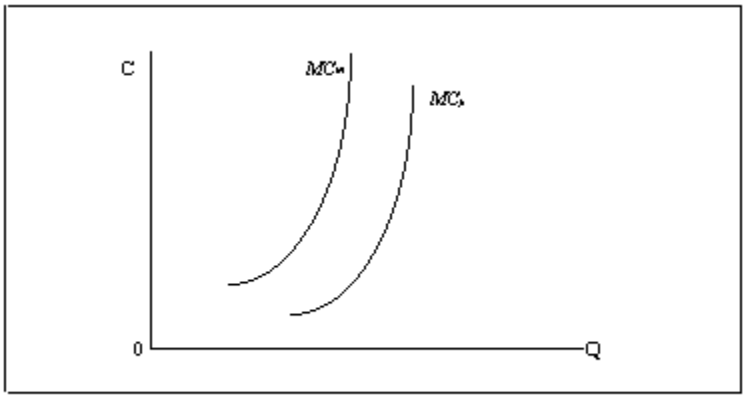
If the market structure has only one seller, rather than an infinite amount, and a barrier to entry (or exit) which guarantees only one market player, there is allocative inefficiency. Due to extra market power, the monopolist restricts quantity, sells at a higher price and earns supernormal profits. This allocative inefficiency is referred to as the dead-weight loss triangle of non-competition. [1] However, early estimates of the dead weight loss were small compared to intuitive estimates of the costs of non-competition. One possible explanation is that monopolies waste resources by rent seeking. A second explanation is what is termed X-inefficiency.

The Concept of X-Efficiency

Leibenstein introduced this theory of inefficiency generated from non-competition. Since it was not allocative and he was unable to characterise it as motivational or

technical, he named it X-efficiency. As a concept it may be summarised as follows: "*for a variety of reasons people and organisations normally work neither as hard or as effectively as they could. In situations where competitive pressure is light, many people will trade the disutility of greater effort, or search for the utility of feeling less pressure and of better interpersonal relations.*"[2]

Figure 1: Differing Costs of Monopoly and Perfect Competition

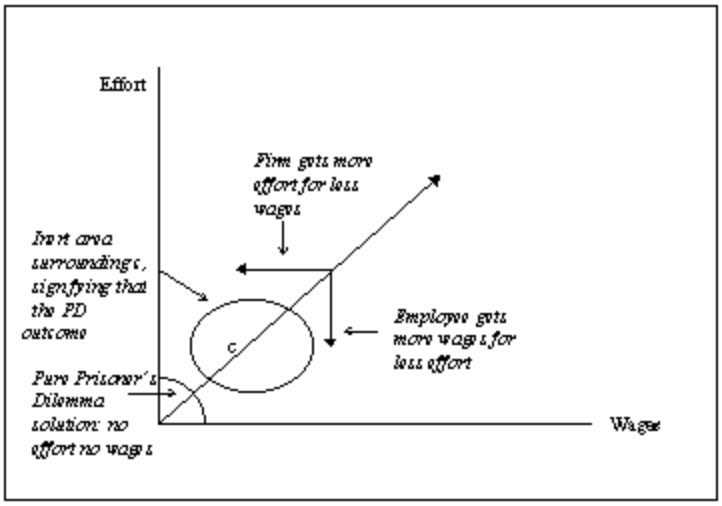


Essentially, since extra costs do not mean immediate bankruptcy for a monopolist, they will be slack in cost control and in the amount of effort put in by management and workers. This concept of X-efficiency leads to the existence of different cost structures associated with different market structures, higher costs being associated with non-competition. It seems intuitively quite attractive. However, Leibenstein's related theory of X-(in)efficiency, by which he explains the higher cost phenomenon is more controversial.

The Theory

Leibenstein enters what is termed micro-micro theory; which is "*the interactive, but somewhat constrained, economically bargained decision among 'atomistic' individuals within the firm*." [3] In examining the molecular make-up of the firm, which is treated as a maximising "black-box" in neo-classical theory, he finds that the internal agents are non-maximisers. Invoking the Yerkes-Dodson Law, at low pressure levels, individuals will not put much effort into carefully calculating decisions, but as pressure builds they move toward more maximising behaviour. He identifies an inert area, probably due to the incomplete nature of labour contracts. Although payment is specified, effort generally is not. Variation in effort is due to the discretion which employees have in choosing effort levels and discretion which top management have with regard to working conditions. Clearly a Prisoner's Dilemma type outcome could exist with effort and wages both at a minimum. However Leibenstein rules this out due to conventions, which ensure equilibrium within the inert area surrounding the point c.

Figure 2: Leibenstien's Theory of X-Efficiency



As pressure mounts, the circle reduces in size and wages reflect effort more and more accurately; with perfect information and honesty, the equilibrium is on the 45 line, where wages reflect effort.

Criticisms Of The Theory

The theories of monopoly and X-efficiency are not without criticism. Some authors argue that a monopoly may generate higher social welfare than perfect competition. With the opportunity of profits, monopolists will innovate and invent since the extra guaranteed rent will not be competed away. These supernormal profits can be invested in new product development and new technological advances, which are not necessary in the perfectly competitive world of horizontal demand curves. Perhaps, safety and general working conditions may not be adhered to in the cut-throat perfectly competitive world. This is not, however the predominant thinking.

Certain critics also question the existence of X-efficiency. Since all economic agents are rational, any slack is a rational leisure-income trade-off. Higher costs, therefore, are not a symptom of inefficiencies, but the effect of fully rational workers' preferences for leisure. According to Stigler (1976) "*increased output due to (say) increased effort is not an increase in 'efficiency', but a change in output.*"[4]

Another criticism of X-efficiency theory comes in the empirical evidence of motivational slack in competitive industries. Leibenstein himself refers to an example of two petroleum plants in Egypt only half a mile apart. One transpired to have been X-inefficient for years, after a management change increased output substantially without changing inputs. Why did this persist for so long in a competitive environment? It may be that the internal pressure is a greater influence than the external pressure. Internal pressure has been described as "*inner prodding, be it religious, moral, or cultural*" which motivates the individual to cost minimise for his employer.[5] Leibenstein (1966) refers to a domino type effect - if a top manager is X-inefficient for whatever reason, this lack of motivation will in turn affect all those below him. Hence it is possible to explain X-Inefficiency in the competitive market place via focusing on internal pressure.

General existence is not the key criticism, however. The main issue is that X-efficiency theory is not compatible with neo-classical microeconomic theory. Leibenstein's rejection of the black box firm would be an interesting advancement in non-competitive markets, if it were to agree theoretically with the general thinking. However, Leibenstein's idea of non-maximisers conflicts with the whole basis of economics as we know it. It is not logical or perhaps not possible to have an economic system based on non-maximising individuals.

Vickers (1993) provides us with a theory compatible with neo-classical economics which manages to explain differing degrees of inefficiency which constitute X-efficiency theory. He uses performance comparisons as incentives for efficiency, in his principal-agent, for instance firm owner-firm manager model. Pay is related to performance which is a function of effort, ability and luck. Whereas the distributions of luck and ability are known, the amount of effort is not. Hence there is a trade-off between the cost of slack and the cost of risk. Where there are other managers available for comparison, yardstick competition can be invoked, where Payi = f(Perfi, Perfj - Perfi). If performance related pay is impossble, but only pay related to expected future performance is possible in the case of academics for example, Vickers refers to the signal to noise ratio as the essential factor determining effort. "*Effort incentives are better the less noise in the luck element*." [6] If pay in the future is based on a performance ratio, it may be optimal for the manager to act inefficiently now, so as to retain future earning potential at a suitable effort level. Hence Vickers has isolated three effects; the insurance, the reputation and the ratchet effect. Each of these effects is compatible with maximising behaviour of non-competitive markets, and put together successfully explain differing degrees of efficiency with differing degrees of competition.

The Empirical Evidence

Theoretical issues aside, has there been any empirical studies suggesting the existence of X-efficiency? Many simple cases have been cited, for example the case of the Ford Motor Company with almost identical plants in England and Germany. The German plant managed to produce 50% more cars with 22% less labour. According to Leibenstein in the New Palgrave, "*despite identical plant design, the differing effort conventions help to explain the X-efficiency result in the UK plant*."[7] Of course many more complex econometric studies have been done. One of the usual econometric problems, that of missing variables, is especially important in this case. Frantz (1990) claims that there are many econometric studies which measure only external (market) pressure, while offering explanations that include internal constraints. He refers to "*approximately 60 empirical studies consistent with the implications of the theory*."[8] Button and Weyman-Jones (1992) note that two of the three approaches to measuring X-efficiency are based on maximising behaviour. The third is a non-parametric programming approach known as data-envelopment analysis (DEA). On studying a set of DEA based studies, the two authors were able to draw conclusions suggesting that bureaucratic or publicly administered industries were on average less efficient than their competitive counterparts.

Conclusions

In summary I have looked at the theories of monopoly and of X-efficiency. I raised some of the critical issues regarding the existence of X-efficiency and some of the theoretical objections to Leibenstein's explanations of the concept. Finally, I examined some of the empirical work done in the area. The million dollar question is whether there exists a gain other than Harberger's triangle in moving from monopoly to competition. The theory of X-efficiency provides us with an intuitive concept within the neo-classical world of maximisers predating Vickers and strengthened by positive empirical evidence. This I believe to be sufficient to guide us in the area of competition policy and law.

Notes

[1] Harberger, 1954

[2] Leibenstein, 1966

[3] Leibenstein, 1966

[4] Stigler, 1976

[5] Leibenstein, 1966

[6] Vickers, 1993

[7] Leibenstein, 1987

[8] Frantz, 1990

Bibliography

Books

Frantz, R., (1990): "X-Efficiency: Past, Present and Future" in Weiermair and Perlman, *Studies in Economic Rationality*, University of Michigan, Ann Arbor.

Leibenstein, H., (1987): "X-Efficiency Theory" in *The New Palgrave*, Mamillan. London.

Perlman, M., (1987) : "The Evolution of Leibenstein's X-Efficiency Theory" in Weiermair and Perlman, op. cit.

Vickers, J., (1993): *Concepts of Competition*, Clarendon Press, Oxford.

Articles

Button, K. and Weyman-Jones, T., (1992): "Ownership Structure, Institutional Organisation and Measured X-Efficiency" in *American Economic Review*.

Frantz, R., (1992): "X-Efficiency and Allocative Efficiency: What Have We Learned?" in *American Economic Review*.

Leibenstein, H., (1966): "Allocative Efficiency vs. 'X-Efficiency'" in *American Economic Review*.

Stigler, G., (1976): "The Xistence of X-Efficiency" in *American Economic Review*.

Competition Policy: Policy Approaches and the Relevant Market Defined

Nicholas McDonagh

Junior Sophister

One of the cornerstones of the development of the Single European Market has been the desire to hasten the demise of national monopolies through increased competition from other eu countries. However, as Nicholas McDonagh discusses, one not only has to deal with alternative competition policies but also with differing definitions as to the basic concepts needed to derive any such policy.

Competition policy is a contentious issue among policy makers today. Policies differ considerably: some may take a structuralist approach, others a more 'pragmatic' or 'control' approach. In this paper I wish to firstly discuss both. Secondly, I wish to show how recent work has derived a full definition of the 'relevant market' as a tool of policy. In this essay I shall draw extensively from the reports of Fishwick (1983)for the Commission of the European Communities.

The Structuralist Approach

The structuralist approach could be considered the pro-active approach to competition. Policy is vigorously directed towards the maintenance of a competitive market structure, in preference to a single market structure which is concentrated or under a dominant monopoly. Active measures are taken to disarm potential and actual monopolies, the logic being that monopolies are not in the public interest, as well as being sub-optimal in the economic sense. The policy of the Commission of the European Communities (now eu) has been largely of the structuralist type since 1963. There has been a gradual trend towards structures as a means of preserving competition in this country. Other countries such as France and, to a certain degree, the uk have also adopted the structuralist approach like the eu. Does this suggest that the policy of a fully integrated Union will be adverse to monopolies and concentrations? The case study evidence would suggest the affirmative. The Commission has been acting as a watchdog on anti-disbanding of cartel-like agreements among dominant firms as called for under Article 85 of the Treaty of Rome which states that there is 'a prohibition on all agreements between undertakings, decisions by associations of which have as their objective the prevention, restriction or distortion of competition within the common market' and Article 86: 'any abuse by one or more undertakings of a dominant position within the common market or in substantial part of thereof is prohibited as incompatible with the common market in so far as it may effect trade between member States' [1]. Commission policy has, however, in the past conflicted with decisions of appeal by the European Court of Justice concerning cases of alleged abuses of dominance, and of agreements between ventures that allegedly distorted competition or trade in the relevant market.

The Control Approach

In contrast to the structuralist approach, American and, to a certain extent, German policy has adopted a less adverse view to monopolies and concentrations. This control approach some would claim is more natural than the structuralist one, the control approach being seen as leaving the market to decide its own course. If this leads to a concentration in the relevant market then the only grounds for intervention are to control abuses of the concentration or monopoly that act against consumer welfare. Concentrations are seen as the zenith of an efficient competitive structure. Monopolies and concentrations can bring with them economies of scale, increased research and development, and so on . According to Fox (1983), us policy has been changed over the past 20 years from a structuralist approach to an approach which realises that 'gains in efficiency from mergers or increased concentrations may outweigh the adverse effects on competitive structure' [2]. We could thus conclude here that the 'control' approach refers only to supervision of

concentrations, monopolies and the like, in contrast to the structuralist approach which in effect frowns upon and avoids such development in the first place.

The Relevant Market Concept

Having looked at the two main approaches to policy formation in competition policy, I will now examine some of the 'technicalities' of enacting policy. When we are discussing the issue of competition in a certain market, we must ask what exactly the market in question is. In other words, what is the good or service in question (the relevant product)? What is the geographical boundary of the market (the relevant geographic market)? What is the relationship to the relevant product of other goods or services; are they substitutes, complements, interdependent goods or otherwise? All these variables are grouped under the term 'the relevant market'. The question remains: how do we define such a broad definition of variables under one parameter? Do we need to define the relevant market at all? Relevant market analysis can prove detailed and at times seemingly obvious. For example, in a particular case, United Brands Corporation[3] held a dominant position in the market for bananas and was accused by the Commission of abusing this position. The subsequent investigation compared the functional and reactive substitutability of bananas to other fruit under the subjects of texture, shape, flavour and type of consumer! For the purpose of analysis, we can in most cases define the relevant market in its broadest sense.

The Relevant Product Market

This is the actual good or service itself, defined in the narrowest terms (x). However, we must broaden this to encompass products which are perfectly substitutable to x. Substitution can be based on the following three criteria :

(i) Functional Interchangeability: does good y, the substitute, have the same physical or technical properties which enable it to do the task of x?

(ii) Reactive Interchangeability: Does y hold the same functionability and utility as a substitute to x, in the mind of the consumer?

(iii) Barriers to Substitution: do barriers to the substitution of x exist (special distribution channels, sunk costs and complementary goods for example.

Indirect tests such as the similar advertising or marketing of both goods may also indicate substitutability. Although neither the Horowitz test[4] nor the Stigler-Sherwin tests[5] for indirect substitution, shall be addressed directly, both are applicable. Direct tests such as cross price elasticity and substitution elasticity are examples of suggested tests of direct substitution. To sum up, though, all substitutes for x must be included in the relevant market. Potential substitutes will be dealt with later.

The Geographical Market

In order to define the relevant market we must outline the actual physical market in terms of size and scope, and whether it is a national or international market. Two goods could be considered in the same geographical market if suppliers cannot discriminate between them, thus ensuring that prices are equal. We can test for an integrated geographical market by considering whether or not there exists barriers to the area in question, as regards the supply and demand transferability of the relevant product between two geographical markets, and the level of transport costs.

Potential Supply/Competition

When we suggest potential supply/competition we are not advocating the view that any Tom, Dick or Harry is a potential supplier of the relevant product. We are specifically concerned with the maximum amount that could potentially be supplied where possible. This would include the excess capacity of actual suppliers now, the output of actual suppliers of the relevant product not already sold, and the transferability capacity of 'like-minded' production suppliers, in other words the ability and ease of transfer of production from one good to that of the relevant product. If the relevant product is produced in tandem with another good, as occurs often

in the chemical industry for instance, then an increase in the profitability of this second good may bring about an increase in the production of the relevant product. Backward/forward vertical integration may also prove to be a potential supplier. Short term barriers to entry to the market may induce the misperception that there is no potential supply outside the current suppliers. Therefore, these barriers to entry must be assessed on a broad external scale. Such barriers may include geographical barriers (transport costs), sunk costs, legal barriers, and so on. These may all change in the long run, thus increasing actual competition.

Relevant Product Interdependence

In order to achieve a fuller, broader picture of the relevant market we may also extend our view to encompass products which may affect the relevant product either directly or indirectly. Vertical interdependence, where the relevant product is an intermediary good is a good example. Here a good, a, may be dependent on another good, b, along the vertical chain of production: for example, a drug for cancer (a) may be dependent on a raw material necessary to produce the drug (b). In order to acknowledge dependence we must consider whether 'the buyer' of the relevant product (the drug) is able to substitute away from 'the supplier' to a different supplier of the raw materials. This has been a contentious issue in its own right among competition case study history[6]. The question of whether or not the relevant product is also a complementary good also surfaces. If so then there will be a need to examine the cross elasticity of demand in order to gauge the strength of the interdependence.We should note here the concept of 'Partenaire Obligatoire'[7]. Briefly this states that the above step-by-step analysis may not always be suitable if there exists dependence in the case of a vertical relationship. In the case of an alleged abuse of dominance, such as an infringement of Article 86, the definition of the relevant market may not be of use in uncovering the details of the abuse. 'Partenaire Obligatoire' studies the dependence of an obligatory trading partner where no transferable substitution is possible.Interdependence at a horizontal stage would suggest a case where the supplier of the relevant product was in some way dependent on the buyer of the relevant product. This is simple monopsony or oligopsony. These are rare examples and are to an extent mirror images of an abuse by a seller who holds a dominant position. In these cases the definition of the relevant market is necessary, as is perhaps the use of 'Partenaire Obligatoire'. Test for the strength of such a relationship are somewhat more difficult to procure. However, test for substitution regarding substitue outlets for the supplier to supply to, such as those based on functional and reactive interchangeability are applicable here. The dependence of the independent supplier of door panels specifically for Ford motor cars, on the Ford motor company and subsequently on the sales of Ford cars is a good example.[8]

The Definition of Relevant Market Sales

Finally in order to measure the relative market power of an undertaking in the relevant market a measurement of the relevant market sales may be a good indicator. This, combined with the above analysis of the dominant or interdependent relationships that the undertaking may possess can give a good indication of the strength of the firm in the relevant market.

Conclusion

In this paper we have taken a look at the two main types of policy in contemporary competition policy. Under the guise of various national approaches. We then went on to discuss the rather thorny issue of how one goes about defining the relevant market in cases concerning the competition authorities and undertakings etc. A step-by-step approach was thus derived. The definition of such a broad topic is bound to lead to debate, as indeed it has. It is likely that such a debate will continue into the future, and that the definition of the relevant market will sway in line with those arguments regardless of how broad or narrow they may be. What is important to note is that care must be taken not to ignore factors that will have considerable bearing on the actual market. In this light the steps taken by Fishwick and others, notably the Commission, are valid.

Bibliography

Books

Fishwick, F., (1986): Definition of the Relevant Market in Community Competition Policy, Commission of the European Communities.

Glais, M. & Laurent, P. (1983): Traite d'Economie et de Droit de la Concurrence, PUF, Paris.

Articles

Fox, E.M., (1983) : 'Abuse of a Dominant Position Under the Treaty of Rome: A Comparison with US Law', Annual Proceedings of the Fordham Corporate Law Institute.

Horowitz, L, (1981) : 'Market Definition in Antitrust Analysis: A Regression Based Approach', Southern Economic Journal, Vol. 46.

Stigler, G.J. & Sherwin, R.A., (1985) : 'The Extent of the Market', Journal of Law and Economics, Vol. XXVIII.

Cases

Commercial Solvents Corporation, (EEC Case), 1972.

Hugin v Lipton, (EEC Case), Dec. 1977.

Replacement Body Panels for Ford Cars (UK Case), Feb. 1985.

United Brands (Bananas - EEC Case), 1985.

Treaties

Treaty Of Rome, European Economic Communities

The Contestability of Transport Markets

Maura O'Connell (Senior Sophister)

Successful management of the transport sector has become one of the most difficult challenges for governments world-wide. Maura O'Connell portrays the inefficiencies of the Irish government's regulation programme and argues the case for a more deregulated, procontestable market.

Introduction

A contestable market is one in which the positions of incumbent firms are easily contested by entrants. Entry is completely free and not impeded by fear of retaliatory price alterations, and exit is costless. New and existing firms compete on perfectly symmetric terms. In this essay, I propose to outline the necessary conditions for a contestable market. The benefits of a regulated and openly competitive market are then discussed. In light of this, I will conclude by explaining the post-deregulation barriers to contestability which may continue to affirm the supremacy of existing firms over new entrants.

Features of a Contestable Market

The condition of free entry will be satisfied if entry does not require any sunk costs. This implies that there may be no cost discrimination against entrants in favour of incumbents.Baumol (1982) indicated that a perfectly competitive market is necessarily perfectly contestable, but not vice-versa. The reason behind this stipulation is that perfect competition is applicable as a guide only where scale economies are absent, so that a large number of rival firms can survive and prosper. Contestable markets, however, contain only a few large firms, and, since there is no regulation, cost differences arise only from distinctions in price or quality. Indeed, perfect contestablility has been hailed as a 'model of socially desirable firm behaviour.'

Such a market can be shown in equilibrium to enjoy all the welfare benefits usually associated with perfect competition. Perfect contestability precludes the earning of supernormal profits, which would expose the industry to 'hit and run' entry. Even a very transient profit opportunity need not be neglected by a potential entrant, who can go in, collect his gains, and leave the market without cost, should the climate turn hostile. It is this threat of competition which forces monopolists and oligopolists to offer consumers the benefits which competition would otherwise bring.

Perfect contestability also rule out productive inefficiency, either through sheer waste, a misallocation of resources, or poor organisation of the industry. Furthermore, the price of each product cannot exceed the marginal cost under conditions of pareto optimality.; hence cross-subsidisation in multi-product firms is unfeasible. The logic that leads to these conclusions can be illustrated in the case of an existing firm earning excess profits. In this circumstance a new entrant (Ryanair, for example) can punish the over-compensated and pareto inefficient incumbent (Aer Lingus) by offering lower airfares. Cross-subsidisation is prevented since a firm earning only normal profits will go bankrupt if it sells any of its products (fares in this instance) at prices that do not cover its costs. Predatory pricing is therefore abolished as a weapon of unfair competition.

The Rationale behind Uncontestable and Contestable Markets

The nature of government intervention is a contentious issue. Thomson and Hunter (1973) put forward a series of arguments for nationalisation of the transport sector. They suggest that technical economies could be gained only by integration and central control of the public utility industries and that competition has lead to a wasteful under-utilisation of capital equipment. Rivalry in the haulage sector, for example may lead to the bidding down fo average loads and hence the raising of unit costs to an unnecessarily high level. In the bus or airline industries, this wasteful competition argument refers to the duplication of an existing service with adequate capacity. They also stress the benefits of a nationalised transport utility to those who has not

previously benefited from basic public services and question whether finance would be forthcoming if the private capital market was the only source.

Nash (1982) emphasises the need for regulation in order to ensure safety in transport markets (granting licences only if drivers and vehicles are deemed safe, for example). He also draws attention to the dangerous nature of direct competition between vehicles on the same route at the same time. The historical market failure postulated in this case is one of imperfect information. The user, it is argued, is incapable of correctly perceiving the risks involved and is therefore incapable of signalling his willingness to pay more for the operator who adopts safer equipment and work practices. Glaister (1982) cautiously proposes that the monitoring of service quality may be necessary because individuals need to be protected by weights and measures from unscrupulous procedures

However there is a tendency to see things from the eye of an established operator who has something to lose from a new competitor, to the neglect of the user who might have been offered a better quality service at low cost. In fact, there arguments are invalidated by the available evidence.

Will the market, left to itself, seriously under-invest in transport? The historical evidence suggests otherwise. Kelsey revealed that in 1885 there were 467 transport companies listed on the London Stock Exchange, with a nominal value of IRL750 million (IRL250 billion at 1985 prices), 14 percent of the total of all companies quoted. By 1985, however, this figure stood at IRL2.8 billion and accounted for just 1.3 percent of the aggregate. Government policies of nationalisation, not market forces, wiped out investors interest in the transport sector. With the exception of the railways, this pattern has been repeated down the years in all sectors of the transport market. Posner, in response to the public interest theory of regulation, cites the theory of regulatory capture: 'Regulation is not about the public interest at all, but is a process, by which interest groups seek to promote their private interest.... Over time, regulatory agencies come to be dominated by the industries regulated.'

The case of CIE is a prime example of the inefficiency of government regulation. The Natural Prices Commission found that most of the thirty-seven licensed provincial bus companies operating in 1972 charged significantly lower fares than CIE. In 1982 the CIE Expressway provincial bus fare was 82 percent higher than Midland Bus (Athlone-Mullingar) and 69 percent higher than the average of Midland, Wharton, and Kavanagh. Public policy, despite deregulation, prevents the operation of a level playing field int the independent bus service in Ireland. CIE has an annual subsidy in excess of IRL100million and its borrowings are part of the Public Sector Borrowing Requirement. By contrast, many independent bus companies use hire purchase. CIE receives first option on the school transport contract, worth IRL26.2 million in 1989. Where competitors seek an extension to a route, or a new one altogether, proposals are sent to CIE rather than being independently assessed; there is no right to appeal against the decisions and CIE has the benefit of its competitors' market research. Despite these huge advantages, CIE remains an inequitable service.

The airline industry is perhaps the strongest case for contestability. The 1944 Chicago convention agreed that the right to operate air services should be governed by bilateral deals between governments. In the spirit of the times, politicians wished to maintain control of their airspace. Fifty years on, this $220 billion industry is one of the biggest blots on free trade. The deregulation of air transport between Ireland and the United Kingdom in recent years has become a classic illustration of improved efficiency following the termination of a cartel. The Anglo-Irish air cartel under-supplied the market as passenger numbers dropped from 1.95 million in 1978 to 1.85 million in 1985. The cartel had restricted the supply of discounted fares by use of the computer yield management programme. The market was then opened to competition in May 1986. Passenger numbers increased from 2.7 million in 1987 to a phenomenal 4.2 million in 1989, while prices on the London to Dublin route have decrease by as much as two-thirds. The passenger is evidently much better-off under deregulation.

There can be no salvation for those who cannot compete, no matter whose flag they carry. Baumol 1982 argues that '...we must reject as perverse the propensity of regulators to resist the closing down of unprofitable lines of activity.' Jan Carlzon. president of the SAS group, prophesies that '...those in the airline business who persist in harbouring protectionist leanings will clearly expire'.

Barriers to Contestability

The gains from deregulation on both sides of the Atlantic have been immense, yet incumbents still hold a number of structural advantages over potential entrants. The continuation of 'grandfather rights', whereby the allocation of slots is decided by incumbents through scheduling committees, severely prohibits the fulfilment of the condition of free entry. This barrier to contestability must be replaced with a market for slots, either by lottery or by auction. Ground-handling monopolies, the preserve of established operators discriminate against new entrants by not allowing them to carry out their own check-ins and baggage handling. Furthermore, the computer reservation systems of incumbents must be used by fledgling firms due to economies of scale. These systems may be used to sway travel agents' displays against smaller airlines.

In addition, larger airlines can practice predatory pricing on competitive routes by cross-subsidising it with gains from uncontested routes, an action intended to bankrupt the smaller airlines. They also cut out competition by taking over minor operators (Air France with respect to Air Inter, UTA for example. Finally, subsidies to principal national airlines have given them an unfair advantage over potential entrants, enabling carriers to sell below cost. American Chapter 11 bankruptcy laws have allowed weak airlines to continue flying without having to make interest payments or pension-fund payments, thus inhibiting a proper share-out of the industry.

Conclusion

Deregulation is procontestable and must be pursued at EU level. It leads to higher productivity, improved quality and lower prices. Yet one must remember that, in the words of Baumol 1982 '...perfect contestability serves not primarily as a description of reality, but as a benchmark for desirable industrial organisation which is far more flexible and is appreciable far more widely than the one which was available to us before.'

Bibliography

Glaister, S, 1982: *Fundamentals of Transport Economics*

Nash, C.A., 1982 *Economics of Public Transport*

Thomson and Hunter, 1973 *The Nationalised Transport Industries*

Baumol. W, 1982 ;'Presidential Address to the AEA' in the American Economic Review.

Agricultural Overproduction: Time to Diet?

Úna Clarke (Senior Freshman)

The efficient and equitable allocation of factors of production has always posed a problem for economists, nowhere more so than in the agricultural sector. In this essay, Úna Clarke highlights the problem of excess supply and analyses the solutions now being prepared.

The agricultural sector enjoys considerable importance in Ireland, not least for historical and cultural reasons, but due also to its impact on the economy. In total it supplies 16 percent of total employment in the primary sector, which increases to 30 percent when account is taken of food industry linkages. It amounts to 15 percent of GNP and is a very significant component of our export earnings. Thus, European Union (EU) agricultural policies are of vital importance to Irish agriculture, with 85 percent of total spending on agriculture accounted for by the then EC in 1989. Therefore, any changes at EU level reverberate throughout the Irish economy. EU spending on agriculture has arisen through its price and market support of agricultural producers, with the dual aim of ensuring food supply stability and of maintaining agricultural incomes. This Common Agricultural Policy (CAP) has however resulted in the creation of excess capacity in the sector because of the distortion of demand and supply signals. In this essay, I intend to discuss how this excess supply is to be dealt with. Prices and markets policy will be examines first before an analysis of rural development is entered into.

CAP has loomed large within Irish agriculture, since our joining the EU in 1973 and has contributed to growth in GNP and rising farm incomes. However, to a large extent the seeds of its downfall have been sown by its oven success, with extensive overproduction and EU budgetary pressures leading to the need for a substantial overhaul of the system. CAP, operating via market intervention, export subsidies, import controls and direct aid, had in recent years become increasingly difficult to justify. It had contributed to, *inter alia*, higher consumer food prices, polarisation of the agricultural sector between large and small farms, the capitalisation of the support system with purchase of quotas and higher land prices, environmental pollution and land degradation. Budgetary support realistically could not continue at such a high level, as funds were needed for cohesion payments in advance of the Single Market and the prospective membership of Central and Eastern European countries (CEECs). Large proportions of these economies' GNP are derived from agriculture and therefore the maintenance by the EU of support at previous levels would have led to its bankruptcy. International demand for reform were also growing, emanating mainly form the US in the GATT negotiations. America's own price support system was under pressure from falling world prices. Reform of the world systems, in the New World Order, allowed an opportunity to bind the developing world and ex-COMECON economies both politically and economically into market-oriented policies.

CAP reforms have resulted in reductions in both price levels and output levels, penalties for overproduction and compensation for non-production. Further reforms are in the pipeline to accommodate the accession of CEECs to the EU. As these countries are engaged in the production of commodities already oversupplied in the EU, further cuts will be necessary, leading to the disappearance of the present CAP system by the end of the century in an enlarged EU. If the cost of food production subsidisation is shifted away from the EU, it will fall back on national budgets. The Irish taxpayer hasn't to date had to bear this burden, being a net recipient of EU funds. The issue of overproduction is thus due to become a very real and contentious issue in Ireland, and, one imagines, one that will highlight the urban-rural divide. The GATT agreement, taking effect fully, means in essence, world free trade, the opening of European markets to world competition, and loss of export subsidies. Food industry linkages, meanwhile, are obvious areas to develop as they create markets for agricultural output. However, the agricultural industry must begin to take seriously the concerns of consumers and environmental groups, use of inputs and pollution.

Rural development has become increasingly more important at both EU and national level. Issues such as distinct urban and rural infrastructures aside, rural development is viewed as a method of reducing excess supply by encouraging movement of resources to non-agricultural activity. Such measures would encourage rural diversification into agri-tourism, forestry, development of industrial bases and infrastructure in rural communities. But one wonders if the opportunity for rural development has already been missed. Has the

dependency culture of the last twenty years, which has been engendered by massive subsidisation, sapped the entrepreneurship within the agricultural community and indeed outside it. In addition, any proposed initiatives will be enacted against a backdrop of reduced public services in rural communities, rationalisation of hospitals, schools, creamery outlets, garda stations and post offices. Rural populations are falling due to migration from the land to either non-agricultural work in urban areas or emigration abroad.

To conclude, it is evident that the problem of excess capacity in the agricultural sector has to be tackled. Attempts to solve the problem, whilst maintaining rural populations and standards of living are proving difficult. The urban population, who account for 80 percent of the total EU population, have borne for years the cost of excess supply in higher food prices. One seriously doubts their willingness to pay higher taxes in order to support non-production in the long-term The urban-rural divide has been fuelled by the predominant view of a farming community paying little tax and yet benefiting from huge aid packages. However, the division within the sector itself means that the larger, highly capitalised farmer benefits from EU polices whilst the marginal farmer becomes more and more dependent on non-farm income. One must wait to see if the new rural development programme will maintain that divide and simply result in an excess of heritage parks interpretative centres and golf courses.

Bibliography

Matthews, Alan: 'Agriculture and Rural Development' in *The Economy of Ireland* by J.W. O'Hagan (ed.), IMI, Dublin, 1991.

The Irish Times Jan 5th, 1995.

Derivatives and Farming: Fate Brought Them Together

John Bohill (Junior Sophister)

Price support has been the prevalent method of subsidisation of farmers' incomes over recent decades. A movement is underway, however, to dismantle these protectionist structures. In this paper, John Bohill examines the present treatment of risk in the agricultural sector and looks at the prospects for a competitive future.

'We should have thought of it a million years ago - in the nineties'

Samuel Beckett, Waiting for Godot

Introduction

The Common Agricultural Policy (CAP) has succeeded for many years in obscuring market realities from the average European farmer. A cosy protective cushion has been placed around him, and he is unwilling to let it go. However, with the reform of this function of the European Union (EU), risk is about to re-enter the European agricultural arena. This paper will outline these risks, and analyse current methods for dealing with them in existing free-market systems. Actuarial and financial advances suggest new avenues in the search for risk hedges in agriculture; these will be evaluated. Finally, implications for a newly unstable EU farming market and the resulting impact on Irish farmers will be proposed.

Farming - A Risky Business

'Risks to farm revenues come from two sources: prices and yields. When both prices and yields are insured, so is the product of the two, farm revenues.'

President Clinton, Economic Report to Congress, February 1995 [quoted in Financial Times, February 15th 1995]

In the absence of all support, the farmer absorbs, over a substantial time-period, all the uncertainties in getting the seed to harvest. This may be mostly contingent on weather conditions, since the farmer may eliminate other risks through diligent work in the fields. However, it has often been seen that a farmer cannot rely on hard work alone, and so must provide for risk. This paper divides price and product risk in the following terms.

Price Risk

Consider a farm producing corn, which has just experienced frost and has 25 percent damage. If all farmers - or even a significant area producing corn - face the same turn in the weather, the price will be positively affected (the quantity supplied has gone down, assuming normal demand-supply characteristics). Of course, not all farmers face the same weather conditions, and so our farmer's future income clearly depends on its relative producing performance over the year. Weather movement is a stochastic process: its value changes over time in an uncertain way. It is generally accepted that the farmers trade is in growing corn, and not in risk/return economics; he would therefore attempt to remove the stochastic variable from his income calculations. Strategies for doing this are explained below.

Production Risk

Complete inability to produce in a specific area due to *'acts of God*' must be treated differently, even though it may be considered an extreme form of 'price risk'. Since nothing is supplied, the level of price is irrelevant;

the farmers income is nil. The 'production risk' comes under the umbrella of conventional insurance, since the occurrence of such catastrophes is rare, but the consequences are sometimes vast (as seen in the Netherlands recently). Low risk, high (negative) return occurrences must have policy implications also, since the welfare of the nation could be at stake. The potato famine for example, whilst a repeat occurrence is implausible for other reasons, has important lessons for modern agri-business.

To sum up, therefore, if it is assumed that it is in the farmers interest to remove stochastic variable from his income, we may proceed to develop policy on this basis.

Existing Treatment of Price Risk

The Common Agricultural Policy (CAP) has been the major form of price support (and by association, income support) for the farming community in recent European history. It has been a considerable drain on resources, and has almost driven the Community to bankruptcy. Its budgetary weight is a result of the European love affair with the pastoral culture, and it is only in recent years that reason has been heeded. The reforms proposed in 1992 promised a reduction of 30 percent in guaranteed farm commodity prices over three years, to be compensated by area aid, which would be unrelated to production. Since then, results have been promising: British farmers have seen an increase in real income of 15 percent since the mid-eighties, stocks of beef and arable land have decreased, and there was a budget surplus at the end of 1994. Despite a large drop in prices in 1993, the stabilisation achieved last year is set to continue. The foreseen hike in expenditure by the Agricultural Department of the Commission is mostly due to the enlargement of the Union. The picture of reform may seem promising, for both farmers and taxpayers. However, in the case of UK farmers, some of the increase in income may be derived from the devaluation of Sterling, since they receive their support payments in ecu. Greater structural change is needed. Efficient farmers produce too much, while those who would be pushed out in a competitive market are encouraged to remain. In 1991, 80 percent of the support funds were going to 20 percent of the farming community, showing a scheme flawed in its welfare implications. Land set-aside projects, tax co-responsibility levies, and the lowering of inflation-adjusted support payments only remove farming from market realities. Similarly, the movement from price support to direct payments per area farmed, or livestock quality owned serve only to place the burden of subsidy on the taxpayer, rather than the consumer.

In the United States, price supports are also under attack. This year Congress cut US Dollar10 billion of the budget, and in polls, tax-wary farmers have even welcomed reforms (only 37 percent of them are in favour of the current system). Most observers believe it would be cheaper to guarantee 70 percent of normal crop revenues: a saving of US Dollar4.2 billion over five years is estimated. Since the subsidy overhaul, there has been no decline in farm income. However, a more graduated re-structuring of the industry is taking place. Production is more centralised and more idle land is used. Farms are diversifying and an increase in grain exports is expected.

There are already price guarantee techniques in place in the US which require no intervention. Futures and option are traded on the Chicago Board of Trade (CBoT), the Chicago Mercantile Exchange (CME) and the New York Mercantile Exchange (NYMEX) based on agricultural denominators. All are sufficiently liquid - the CBoT is the world's largest exchange - to be a viable replacement for price supports for the farmer. Consider again our corn farmer. If he wishes to hedge risk for his September harvest, he may go short on 5000 bushel futures contracts at CBoT. The selling price is pre-set, with an obligation to transact on the settlement date. Alternatively, he may put a long option on corn. If the price is lower than the selected strike price at September (a 'European' option), he will receive a cash premium. If the price is significantly higher, he may decide not to exercise his option, and receive the higher market price regardless.

Under multilateral agreements derived from the Uruguay round of GATT, a world market price for commodities is on its way. Export subsidies on the EU side will disappear, and the market-influenced regime of the United States should begin to step in. It is therefore likely that 'conventional' derivatives will enjoy greater prominence in European markets, and liquidity should improve. Farmers and farming groups/co-ops should understand the technology prior to this, ensuring a painless transition.

Existing Treatment of Production Risk

There have been significant bubbles in the re-insurance markets in the past. Lloyds has exposed its 'names' to risks of which they have been ignorant and the government is frequently called in as a re-insurer of last resort, when the market breaks down. It has been well-documented that people generally under-estimate the risk of catastrophe, and so never 'get round' to insuring themselves. The government, which is sometimes seen as an outside benevolent trust, face pressure to move in and 'bail out' the unfortunates. On the other side of the Atlantic, Congress now requires farmers to have United States Department of Agriculture (USDA) insurance, before any price support is paid. This will increase crop insurance liabilities in existing re-insurance markets from US Dollar13 billion to US Dollar40 billion. The need for re-insurers is greater but they cannot be found, except behind the closed doors of Lloyds. In some instances in 1993, farmers were grossly under-insured for catastrophes. As a result, the USDA decided to give aid to farmers with insurance losing 35 percent of their crop, and to those without insurance losing 40 percent. The corruption caused by such governmental aid schemes has also been cause for alarm: in Georgia, for instance, it was found that farmers sat on their own compensation committees! It has therefore been necessary for the Government to search for an adequate escape route.

The Answer?

The market-based solution has been a long time coming. On October 18, 1994, the committee of CBoT agreed on the wording of a new kind of contract, to be termed 'area yield options contracts'. There would initially be four contracts traded: Illinois soya bean, Iowa corn, Kansas winter wheat, and North Dakota spring wheat, based loosely on what are known as 'Catastrophe Insurance Contracts'. The catastrophe contracts are priced according to a loss ration based on accumulated claims calculated by the Insurance Services Office (ISO). If the loss ratio moves up by 1 percent, the settlement price (thevalue of the contract) for each contract goes up by US Dollar250. Despite a slow start, the contracts are now enjoying steady growth, with six thousand contracts traded between May and September 1994. This figure is ten times that traded in 1993. Richard Sandor, the derivatives guru, has said that the form 'is past its infancy and is starting to grow' [quoted in Financial Times, September 8th, 1994]. A shortfall in liquidity is to be overcome with modification of the contract. Since damage claims are often unquantified until a year after the catastrophe, there is a move to change the contracts to yearly, rather than the conventional three month period. This change should attract more of the over-the-counter (OTC) market, which is currently flourishing.

The techniques learned from the above contracts are to be applied to crop insurance. A holder of such an 'area yield option' would select a strike yield, rather than strike price, say 100 bushels of corn per acre. A put would therefore net the holder a cash payoff if the actual yield was below this figure. Presumably, existing option pricing techniques would only require minor modification to be applied here. CBoT and the USDA are counting on the concentration of the agriculture to form the basis of the success of the contract. If railroads, co-operatives and processors are dependent on the location of crop growth, then it will also be in their interest to trade. The hope, then is that the straining insurance market could find a new avenue in distributing production risk: there may be more 'players' in underwriting.

Conclusion

In an intervention-free world, farmers' income is to a large extent outside their control. Existing remedies for this uncertainty have all but bankrupted governing bodies. Hence, there is a large movement away from bureaucracy, and a growing reliance on market-based solutions. Agricultural risk is conveniently divided between price and production risk. The former already has a large market devoted to dividing risk among willing parties in the United States, namely the commodity futures and options exchanges. There will be an inevitable move towards this system in Europe as we search for CAP substitutes and enter a world trading under the jurisdiction of the World Trade Organistation. Production risk, however, does not have a proven market-traded solution. The emergence of derivatives finance has inevitably led to a possible candidate. The initial trepidation which led CBoT to postpone the launch, has become optimism due to the forthcoming Republican Farm Bill (1995). The exchange has thus decided to move ahead with the launching of one of the contracts, Iowa corn, later this year. Only time will measure its usefulness, but the logic is unquestionable.

The implications for the Irish farmer will be twofold, and will occur at different times. As the price support mechanism of CAP evaporates, world futures and option quotations will become more pertinent in the lives of farmers and co-operatives. Secondly, the doubts places over Lloyd's ability to police world insurance

coverage mean a broader market is needed, one which is independent of tax-sensitive voters. The Area Yield contract is positioning itself in an immature market prior to these structural changes, but its eventual global impact could be enormous.

Bibliography

Chicago Board of Trade : Commodity Trading Manual, Chicago, 1989

Hull, John C.: Options, Futures and Other Derivative Securities, 2nd edn., Prentice Hall, New Jersey, 1993

Matthews, A.: 'Agriculture and Rural Development' in The Economy of Ireland, by John W. O'Hagan (ed), IMI, Dublin, 1991.

The Economist February 11th, 1995: ' Old MacDonald had an Option', London

Financial Times February 14th 1995: 'Blowing Britain's Animal Welfare Trumpet', London

Financial Times, February 15th 1995: 'Brussels Upbeat on Farm Policy Reforms', London

Financial Times, November 16th, 1994: 'Survey of Derivatives-Reaping Rewards from Catastrophes'

Financial Times, July 14th 1994: 'CBoT Hopes for Big Yield in Crop Insurance Futures', London

Financial Times, September 8th 1994; London

International Business Week, January 9th 1995: 'Full Granaries Don't Mean Full Pockets', McGraw-Hill, New York

International Business Week, January 9th, 1995: 'For Swingers Only - A Wild and Crazy Market', McGraw-Hill, New York

Wall Street Journal, November 21st 1994: 'The Outlook - Some Farmers Favour Ending Crop Subsidies', New York.

Wall Street Journal, August 8th, 1994: 'House Passes Bill to Cut Farm Disaster-Aid Costs', New York.

Wall Street Journal, October 21st, 1993: 'Disaster Aid Helps Some Even Without Insurance', New York

Wall Street Journal, October 31st, 1994: 'Aid Raid - Crop Disaster Program has Georgia Farmers Raising Lots of Squash' New York.

Income Tax: An Appraisal of the Irish Situation

Suzanne O'Neill (Senior Freshman)

In recent years, government budgets have been subjected to three competing forces: fiscal rectitude, fiscal reform and increasing unemployment. The clamour for fiscal rectitude has given way to calls for reform that will also contribute to solving the unemployment problem. In this paper, Suzanne O'Neill examines the impact of the latest round of direct taxation reforms, questioning whether they will have the desired effect of increasing the number of individuals entering salaried employment.

'Ultimately taxes will be chosen on the foundation of how well they are consistent with other objectives of government policy' *Charles Allan*

Introduction

The 1995 budget will be remembered most notably for its treatment of income taxation and more specifically for its treatment of the stated central objective of rewarding work. Thus this paper seeks not only to examine the post-budget income taxation system, but also to question any effects on work effort and behaviour that these changes would induce. Therefore I propose to outline initially three considerations which one must be aware of when discussing taxation in general, followed by a practical breakdown of how one moves from gross to disposable income, incorporating this year's budget changes. I will conclude with a discussion as to whether an increase in the numbers participating in salaried work will materialise.

Taxation: Some General Issues

It is worth knowing firstly that taxation is the principal means of transferring resources to the government. These resources are used to provide essential services in the economy, such as health and education, and also to provide some basic needs to members of the community deemed to be unable to provide for themselves. Taxation is the direct and inevitable consequence of government expenditure, the level of which is determined by public demand. Therefore while it is the community who decide the level of taxation, albeit indirectly, it is the government who must implement the structure of taxation that will satisfy revenue needs.

Secondly, since the focus of this essay is on the taxation of income, it is essential to note that PAYE-related ('pay as you earn') social insurance, health and employment levies all fall under the umbrella of income taxation. Thus the practice of fiscal illusion, where one concentrates solely on income tax rates, is seen to understate the actual amount of tax deducted at source from an individual's salary.

Thirldly, it is important to note that not all individuals are liable for tax under the same schedule. The self-employed are assessed differently from those employed in the private or public sectors. Furhtermore, public sector workers are liable for pay-related social insurance (PRSI) at a lower rate than any other group of employees.

Income Tax in Practice

Not all of a person's income is liable for tax due to the presence of allowances, relieves, exclusions and exemptions in the Irish system. Tax-free allowances are administered uniformly, regardless of one's gross income, and currently stand at £2,500 for a single person. The system provides additional allowances for widows and single parents. In the determination of taxable income one has traditionally been allowed to deduct relieves from gross income, thereby reducing taxable income. One has been allowed to offset certain expenditure against tax liability. Two types of relieves exist: business and non-business. Although the conflict that this topic raises between the self-employed and the PAYE sector has provided many hours of entertaining debate, a more serious issue is the appropriateness of their existence in a progressive and equitable income tax system. Certain relieves, such as mortgage interest relief and voluntary health insurance relief, are balanced in favour of the more affluent in our society. With regard to exclusions, certain types and sources of

income such as imputed income and investment returns are not liable for taxation. Finally income from state-run gambling institutions, social welfare payments and artists' incomes have traditionally been exempted from tax altogether. Having derived one's taxable income, this is then subject to a progressive system of tax rates. At present the tax rate for the initial £8,900 earned is 27% with a higher 48% rate levied on income above this threshold. On top of this PRSI and various government levies add to the income tax actually paid by the individual.

The 1995 budget saw the government increasing personal allowances and increasing the threshold income above which the higher income tax rate applies. The rationale for this is to increase an employee's take-home pay. PAYE allowances were left unaltered. In contrast to its generous nature in other areas, Mr. Quinn chose to gain revenue by reducing mortgage interest relief, voluntary health insurance relief and abolishing covenants. In addition the government introduced an exemption on PRSI payments on the first £50 for the private sector and £10 for the public sector, earned per week. PRSI allowances were also reduced.

The combination of these changes according to the Minister for Finance, reward work. In order to assess further the significance of these income taxation changes, an analysis of the effect of direct taxation on work behaviour is needed.

Work Effort and Behaviour

It is widely accepted that the higher the marginal rates of income taxation, the greater the disincentive to work and the greater the incentive to avoid and evade tax. However, one must distinguish between the influences of income taxation on (i) those currently in paid employment and (ii) on the choice of ones occupation and work pattern.

(i) Theory dictates that as the rate of income taxation increases two effects occur; an income and a substitution effect. In principal, a higher tax rate reduces take-home pay and in order for an employee to maintain his original level of income, he/she must work additional hours. The substitution effect operates on the basis that the amount one earns for an extra hour of labour is declining as the marginal tax rate increases, thus one substitutes towards leisure and away from labour. This is as a result of the decreasing opportunity cost of leisure as measured by ones declining take home pay per hour of labour. However, it is in reality very difficult to assess the actual magnitude of either the income or substitution effects thereby making it difficult to identify a predictable outcome resulting from a change in direct taxation rates. Hence, the favourable income tax changes in the 1995 budget, due to changes in allowances and relieves rather than the tax rates per se, may or may not have an effect on work behaviour. The problem lies largely with ones ability to work overtime. By working the normal week in a given industry, one may benefit from increased income, however should working additional hours push an employees income above the 27% threshold value then that employee may not be better off when the financial and opportunity costs in terms of leisure are considered.

(ii) The income taxation system has a clearer impact on the choice of occupation and work pattern. Certain groups of individuals may not enter paid employment if they consider that the taxation system would leave them financially worse off. Individuals in this category include married women and in many cases the unemployed, who find themselves caught in unemployment traps. High and indefinite social welfare payments coupled with the perception of punitive income taxation, may encourage some of the community to evade taxation altogether by working in the black economy. The taxation system may also act as a disincentive to new enterprises and the self employed who often find their burden of income taxation coupled with employers pay related social insurance contributions, too great to maintain viable businesses. income taxation also has the ability to affect where people work, leading to emigration of highly skilled and motivated entrepreneurs. The 1995 budget with the increase in childrens allowances should provide additional incentive for mothers and unemployed fathers to enter paid employment. The reduction in employer's PRSI contributions also provided a welcome indication of support for native Irish entrepreneurship.

Conclusion

As one noted at the outset of this paper, taxation will always have to be paid to satisfy public demand for government expenditure. However in recent budgets, one has witnessed the realisation that the collection of revenue must not be to the detriment of other government objectives. Thus one has concluded that the

lessening of the direct taxation burden in the 1995 budget, especially for those earning below the average industrial wage, will on balance increase the incentives and thus the number of individuals entering paid employment or providing additional such employment.

Bibliography

Books

Allan, C. : *Theory of Taxation*, Penguin, Harmondsworth, 1991.

O'Hagan, J.W. : *The Economy of Ireland*, (6th edn.),IMI, Dublin, 1991.

IS THERE A "CORRECT" SHARE OF HEALTH CARE SPENDING IN GNP?

John Reynolds (Junior Sophister)

Spiraling health care costs are the sword of Damocles hanging over the head of many Western policymakers. Cuts too deep may hurt the poor and the elderly. Cuts too shallow may hurt future generations. John Reynolds asks in this essay whether or not it is possible to strike a happy medium. "The only truths which are universal are those gross enough to be thought so."

P.Valery

Introduction

In an unregulated competitive health care market it would be relatively easy to answer the question of what the correct share of health care spending of GNP should be - "*leave it to the market*". Unfortunately, most health care markets are (justifiably) heavily regulated. In such a system there is no automatic mechanism, such as the market, to guide it to the appropriate level of spending. This essay will examine the various techniques of economic analysis which have contributed significantly to setting the budget for health care and will conclude by outlining why there is no universal correct share of health care spending of GNP.

Determinants of health care spending

A widespread belief that an understanding of the fundamental determinants of health care spending may yield valuable insights into how such expenditure can be controlled, has led to a considerable volume of literature on these determinants. One of the most consistent conclusions in this literature is that the principle determinant of what a country spends as a share of GNP is income. Newhouse (1977) examined the relationship between medical care expenditure and income across 13 developed countries, regressing per capita medical care expenditures on per capita GNP. Consistent with an earlier study by Kleiman (1974) for a different set of countries, Newhouse reached two major conclusions;

(1) Firstly that GNP accounts for most of the variance in medical care expenditures across countries, and secondly that;

(2) the income elasticity of medical care expenditures across countries exceeds

one - by definition this implies that, at the margin, medical care is a luxury good.

Newhouse felt that in countries with high expenditure, the marginal unit of medical care is more likely to produce improvements in so-called subjective components of health, such as relief of anxiety and more accurate diagnoses, rather than improvements in morbidity and mortality rates. Could it be the case that countries spending more on medical care may well provide additional caring, but little additional curing ?

Table 1: Expenditure on Health in Low-Income and High-Income Economies

Country	%of Government Expenditure Spent on Health Care Services (1983)
Low Income Economies:	
Nepal	4.5%
Sri Lanka	5.1%
India	2.4%
Pakistan	1.0%
Industrial Economies:	
Germany, F R	18.6%
Australia	7.1%
Canada	6.3%
USA	10.7%

Source : Asian Development Bank et al. (1988)

Table 1 also supports the basic finding that national income largely determines the level of health care spending in an economy. In this table, low-income countries spend a smaller proportion of total Government Expenditure on health care : the percentage of public expenditure on health ranges from 1.0 to 5.1 per cent in low-income countries and from 6.3 to 18.6 per cent in the high-income (industrial) economies. In a break from the traditional line of thought, Ulf Gerdtham (1992) concluded that the age structure of the population of a country may be of prime importance in determining the level of health care expenditure. He discovered, in his research of health care expenditure in Africa, that the demand for medical services fluctuates with age - those under 15 years of age utilise medical services more than average. Yet in accepting Gerdtham's proposition, it is also of fundamental importance to note a number of problems associated with cross-national comparisons of health expenditure, as outlined by Robert Leu(1986); Definitions of health etc. are insufficiently standardised; exchange rate conversions always have a degree of randomness, and; input prices may be positively correlated with the level of national income.As a result, it is not surprising that, when based on international comparisons, a judgment of the effective impact of health care is elusive. To conclude, none of the cited studies of the determinants of spending on health care explicitly combine resource use (or cost) with enhanced or maintained well-being. They give no indication whatsoever of the 'right' level of spending. It is clear that in order to make progress on what to spend on health care, there must be more precise data on the productivity of health care interventions.

A Correct level of Spending?

On the whole issue of changes in government expenditure on health services, Sean Barrett, in his study of the "Social and Economic Aspects of the Health Services"(1979), points to several disquieting signs that increases in expenditure on the health services have been accompanied by a reduction rather than an improvement in the health of the community. He refers to Keating's article (1976) which reveals that the life expectancy of males of 30 years of age declined between 1960 and 1970 despite increases in government spending on the health service during that period. Moreover, in Tokyo in 1973 the International Economic Association concluded that healthcare is only one input into the process by which the health of the individual is improved. Income, education, lifestyle, work environment, work status, housing and health care all affect an individual's state of health. The following are the results of a U.S. study which analysed the potential changes in mortality rates which would be associated with a 10 percent increase in some variables;

Table 2: Percentage change in Age-specific Mortality Rates resulting from a 10 percent increase in several variables.

Income	Education	Cigarette Consumption	Per Capita Health Expenditure
+2.0%	-2.2%	+1.0%	-.065%

Source : Culyer (1976)As can be seen from the above table, the increase in per capita health care expenditure reduced the mortality rate by 0.65 percent, but a similar increase in education expenditure reduced the mortality rate by an even larger percentage of 2.2 percent, proving that an increase in healthcare expenditure is not necessarily the most effective means of reducing mortality and therefore, increasing the 'healthiness' of a community.Conclusion? In conclusion, little evidence can be drawn from economic studies on what is the most appropriate share of healthcare spending of GNP because there is no universal correct share. Health care is shaped by too many determinants - income, age structure, the structure of the health budgeting system (the more centralised the system is, the lower is the share of health care spending (Culyer 1988)), and each of these determinants vary in importance in different parts of the world.

To assume a universal share would be to deny all these international variations. For example, if GNP is the most important determinant of health care spending in a particular country, little can be done by way of direct health care policy in setting the budget. Governments have to aim to increase GNP if they want to increase the budget for health care. As I have already stated, GNP is the most important determinant of health care spending in most countries, but what about the African countries where the age structure of the population plays a key role in determining such spending. To assume there is a correct level of health care spending in GNP with respect to the GDP of a country would be to ignore these African countries. Therefore I believe there is no such 'correct' share of health care spending in GNP - there are no such "universal truths". I believe that it is up to the countries to decide for themselves what is the most suitable level of health care spending by use of the appropriate techniques of economic evaluation, namely cost benefit analysis, which should include an investigation of the determinants of health in their country. The result will not be a universal correct share but it will help the various countries to decide where increases and decreases in resources are best applied.

BIBLIOGRAPHY

Asian Development Bank et al., (1988) Economic Analysis of the Environmental Impacts of Development Projects, pp.39-40.

Barrett, Sean D., (1979) "Social and Economic Aspects of the Health Services", Irish Banking Review, March 1979.

Culyer, A., (1976) Need and the National Health Service, chapter 4.

Donaldson, C. & Gerard, K., (1993) The Economics of Health Care Financing, chapter 10.

Ekins, Paul, (1986) The Living Economy : A New Economics in the Making.

Gerdtham, Ulf, (1992) "Determinants of Health Care Expenditure in Africa : A Cross-Sectional Study" in World Developments 20(2), February 1992, pp.303-308.

Keating, W., (1976) "An Analysis of recent demographical trends with population projections for the years 1981 and 1986" in Journal of the Statistical and Social Inquiry Society of Ireland 1976-77, p.123.

Kleinman, E., (1974) "The determinants of National Outlay on Health" in M.Perlman (ed.): The Economics of Health and Medical Care.

Leu, Robert E., (1986) "The Public/Private mix and International Health Care Costs" in Culyer and Jonsson: Public and Private Health Services.

Newhouse, Joseph P., (1977) "Medical Care Expenditure : A Cross-National survey" in Culyer: The Economics of Health Volume 2.