University of Strathclyde Business School

STRATHCLYDE

DISCUSSION PAPERS IN ECONOMICS

Regional Policy Evaluation: Ignorance, Evidence and Influence

BY

Eric McVittie and J Kim Swales

No. 03-06

DEPARTMENT OF ECONOMICS
UNIVERSITY OF STRATHCLYDE
GLASGOW

Regional Policy Evaluation: Ignorance, Evidence and Influence

Eric McVittie
University of Plymouth Business School

and

J. Kim Swales
Fraser of Allander Institute, University of Strathclyde

16 February 2003

The research presented in this paper is linked to the ESRC-funded project *An Analysis of National and Devolved Economic Policies*, Grant no. L219252102. The authors would like to thank Alex Christie and participants at the ESRC Urban and Regional Economics Seminar Group, Cardiff, 2003 for comments on earlier versions of this work.

1. Introduction

The present UK government has adopted a "New Localism" in the operation of regional policy (Balls, 2002). This involves the delegation of discretionary powers to locally elected bodies, such as the Scottish Parliament and the Welsh Assembly, and a host of devolved agencies, such as the Regional Development Agencies (RDAs). We fear that because of standard asymmetric information problems, such policies will be extremely difficult to evaluate, which has implications for both their efficiency and democratic accountability.

However, even without these particular extended principal agent problems, the effectiveness of regional policy is hard to measure. We illustrate this by discussing the issues raised in evaluating and appraising more conventional policies. In our approach we aim to be complementary with more practical surveys of regional policy evaluation (Taylor, 2002). We argue that evaluation raises issues of allocation, co-ordination, information and motivation. The optimal policy is difficult to identify, even with a very simple policy framework, because some of the outcomes are necessarily qualitative and others are conflicting. We hope that this discussion will be used to lay the ground for a more focussed future examination of the evaluation issues raised by the present and proposed regional policy framework.

Section 2 considers the classic welfare approach to regional policy evaluation in which the policy problem involves constrained maximisation. In Section 3, the general strengths and weaknesses of this approach are discussed. Section 4 presents policy evaluation as though it were a control engineering problem. This is what is implied in taking the standard welfare approach. Sections 5 and 6 explore the implications of introducing informational problems. Section 5 investigates the identification of the Social Welfare Function and Section 6 the identification of the true model of the economy. In Section 7 we consider how uncertainty in these two areas can interact to further complicate and confuse policy evaluation. Section 8 explores attempts to evaluate without theory. We see no way out along this route. Section 9 is a short conclusion.

2. The Classic Welfare Approach to Regional Policy Evaluation

The standard welfare economics approach to policy evaluation and appraisal comprises five elements. First the evaluation must identify the key policy instruments, *p,* and exogenous non-policy variables, *n*. By policy instruments we mean any variable under the direct influence of the government or government agencies. This could include the provision of infrastructure, services and advice; setting tax and subsidy rates; and the imposition of physical and financial controls. Exogenous non-policy variables would typically include economic variables such as foreign demand for the economy's exports, and physical factors such as the incidence of natural disasters.

Second, the vector of policy-relevant endogenous target variables, *q*, needs to be selected. In principle, all endogenous variables in the economy will be affected by a change in one or more exogenous policy variables. However, for practical purposes a limited set of target variables needs to be chosen. There are two theoretical reasons why a variable might be excluded from those identified in vector q: the impact on the variable might be very small or the variable might have no welfare significance. Examples of the kind of variable that might be included amongst the target variables for regional policy would be employment and output for individual regions.

Of course, in so far as the operation of regional policy significantly affects more general policy areas, such as, for example, national output, the level of inflation and environmental indicators, these should also be included in the vector of target variables. This is an element of what the present government refers to as "joined-up policy".[1] Also, in practice the choice of target variables has to take into account the requirement for accurate and timely measurement.

Third, the evaluation then requires a theory that links the target variables to the policy instruments and non-policy variables. This can be represented as:

$$q = q(p, n) \tag{1}$$

We are thinking of theory here as a mechanism for understanding the causal relationships between policy instruments and target variables. Equation (1) should be

understood to incorporate behavioural relationships and, where appropriate, macroeconomic constraints. Equation (1) should therefore be seen as the representation of the overall economic constraint within which regional policy operates.

Fourth, evaluation generally entails numerical precision. Not only the qualitative changes in target variables, but also quantitative measurements of the size of such changes are normally required. Therefore knowledge of the appropriate functional form and key parameter values is necessary. Assuming that this function is differentiable, the change in the ith policy relevant variable, dq_i, resulting from a small change in the jth policy instrument, dp_j, is calculated via the following expression[2]:

$$dq_i = \frac{\partial q_i}{\partial p_j} dp_j \tag{2}$$

Finally a Social Welfare Function - less formally, a set of evaluation weights – is required to judge whether changes to the economy generated by the application of policy have brought about a net benefit. The Social Welfare Function is simply a mechanism for ranking different economic states.

$$SWF = S(n,q) = S(n,q(p,n)) \tag{3}$$

The standard cost-benefit procedure is an example (Schofield, 1987; Swales, 1997).

If we have perfect knowledge about the exact form of expressions (1) and (3), the nature of the criteria that should be used for *ex ante* appraisal and *ex post* evaluation is relatively straightforward. Again, assuming the functions are differentiable, the *ex ante* acceptance criterion for a policy package that changes each element j of the vector of policy instrument by dp_j is given as:

$$dS = \sum_j \sum_i \left[\frac{\partial S(q,n^e)}{\partial q_i} \right] \left[\frac{\partial q_i(p,n^e)}{\partial p_j} \right] dp_j > 0 \tag{4}$$

where n^e is the expected value of the exogenous, non-policy variables.

Although expression (4) looks a little daunting, it is easily unravelled. The first bracketed term inside the double sigma signs expresses the rate of change of social

welfare with respect to target variable q_i. It measures the relative marginal weight of variable q_i in the Social Welfare Function. The second bracketed term is the rate of change of the target variable q_i with respect to the policy variable p_j. Equation (4) therefore simply identifies the net effect on the vector of target variables, dq, of the policy package, dp, and then values these changes using marginal weights from the Social Welfare Function.

Equation (4) can similarly be used for *ex post* policy evaluation by replacing the expected values of the non-policy variables, n^e, by their actual values, n. Again a positive value for the expression in equation (4) indicates that the policy generated an increase in welfare. Note that if the target variables are defined such that they enter the Social Welfare Function with a positive marginal weight, then policy must at least have the possibility of generating negative effects on some target variables. If not, *ex ante* appraisals and *ex post* evaluations of these sorts are trivial exercises.[3]

It is a short step from appraising policy proposals to identifying the optimal policy. This is determined by setting

$$\sum_i \left[\frac{\partial S(q,n^e)}{\partial q_i}\right]\left[\frac{\partial q_i(p,n^e)}{\partial p_j}\right]=0 \qquad \forall_j \tag{5}$$

This is the welfare maximising policy as long as the appropriate second order conditions hold (Chiang, 1974, ch. 11). Equation (5) simply states that the jth policy variable should be increased until its marginal impact on social welfare is zero. This marginal impact is calculated by multiplying the impact on each target variable by the marginal weights supplied by the Social Welfare Function.

3. General Strengths and Weaknesses of the Conventional Evaluation Approach

The major strength of this conventional approach to economic evaluation is the clear separation of the economic constraints (equation 1) and the policy objectives - the Social Welfare Function (equation 3). This separation is useful both conceptually and practically. Conceptually, the policy choice can be set up as a constrained optimising problem, a type of challenge economists are well equipped to tackle. Practically, it is beneficial to separately identify the two very different broad

influences on policy choice. That is to say, this approach distinguishes between the objectives of policy, which reflect citizens' preferences and are articulated through the political process, and the economic relationships between the key variables, which are determined by physical laws, factor supplies and those ruling institutional arrangements outwith government's control. Moreover, the optimising decision draws attention to the importance of marginal trade-offs both in production and the Social Welfare Function. It provides a conceptually rich analysis in contrast to the one-dimensional "pledges" and "targets" approach currently favoured by the present government.

However, the weakness in the analysis portrayed in Section 2 is that it implies perfect knowledge about policy objectives and the way in which policy operates. That is to say, the specific forms of equations (1) and (3) are assumed to be common knowledge, at least amongst the evaluation community. Where this is the case, the aim of evaluation is to improve co-operation and co-ordination amongst the government and other economic agents in improving allocation. But our contention is that regional policy evaluation has very different aims. In point of fact, the UK government has never identified the objectives of regional policy very clearly and certainly not the marginal policy weights. Also there is much disagreement concerning the operation of the UK economy, and *a fortiori* the operation of economic policy, across space. These informational deficiencies transform the nature of the evaluation process and this is explored in the remainder of the paper. Not surprisingly, information generation, revelation and motivational issues feature strongly.

4. Evaluation as a Control Engineering Problem

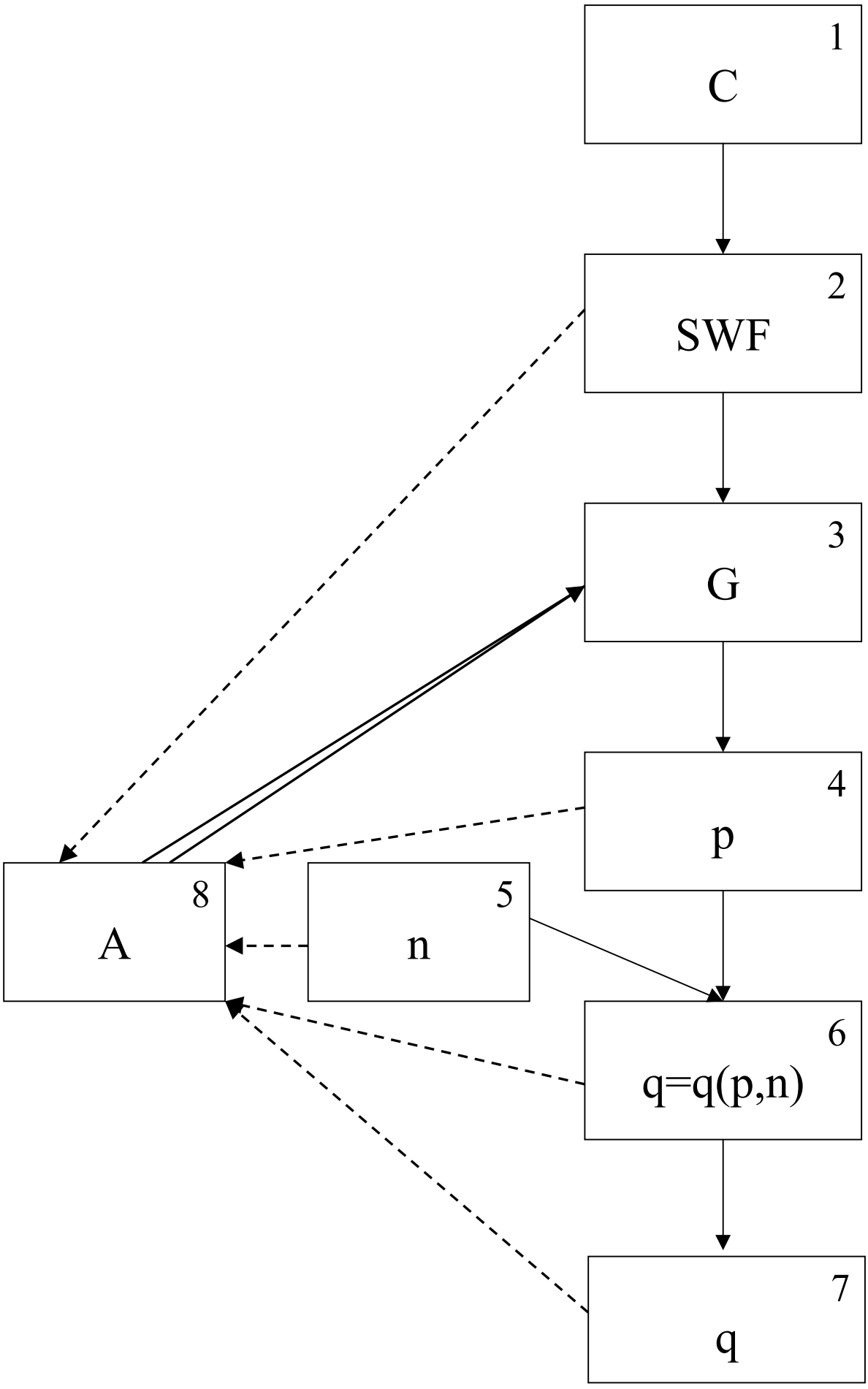
With perfect information, appraisal, evaluation and optimal policy choice become purely technical problems, almost akin to those in control engineering (Dixit, 1996; Stiglitz, 2002). However, this is not the typical perception of economic evaluation in general, and certainly regional policy evaluation in particular. With perfect information, evaluation would be denoted by the loop shown in Figure 1. In this diagram, information is represented by dotted lines, actions by hard lines and signals by double lines.

In Box 1 we have the citizens (C). Box 2 identifies the Social Welfare Function (SWF), which is taken to be a true representation of citizens' preferences over different economic states. Box 3 introduces the government (G), which acts as the citizens' agent and attempts to maximise this social welfare through policy intervention. Box 4 gives the chosen values for policy variables and Box 5 the levels of the exogenous non-policy variables. Box 6 represents the true relationship between the policy and non-policy variables and the target variables. Box 7 is the resulting values for the target variables. Box 8 contains the appraiser/evaluator (A).

The function of appraiser is to collect information and then send a signal to the government. The information concerns: the SWF, from Box 2, the exogenous non-policy and policy variables, Boxes 4 and 5, the economic constraints, Box 6, and the outcome for the target variables, Box 7. The signals could be either or both suggestions for future policy (*ex ante* appraisal) or judgements over past policy (*ex post* evaluation).

From Figure 1 a number of key points can be made concerning the role of the evaluator. The first is that with perfect information, her main function would be in *ex ante* appraisal, rather than *ex post* evaluation. Two main skills are involved. The first is judgement concerning the expected future (or, with reporting lags, present) values of the non-policy variables. This judgement would be made using the information coming from Box 5. The second is undertaking the maximising calculations. The information relating to the effectiveness of past decisions, the conventional *ex post* evaluations, would only be a check. It would be one, but not generally the only, information source concerning the relationship in Box 7.

Figure 1: Evaluation with Complete and Accurate Information



In this scenario the evaluator has, as Dixit and Stiglitz separately correctly maintain, essentially an engineering control role. The evaluator acts in a way similar to a forward-looking thermostat that sets radiator controls in response to estimated present and expected future outside temperatures. The position of the evaluator in Figure 1 has many similarities to the operation of the UK Monetary Policy Committee (MPC). This committee sends a signal to the Bank of England every month concerning changes in the interest rate policy to maintain inflation levels within a target band. It is very much concerned with *ex ante* appraisal, is operating to a well defined Social Welfare Function concerning inflation and with a well researched theory linking inflation to interest rates. However, the operation of the MPC is unusual amongst government policy evaluation and certainly differs markedly from regional policy evaluation.[4]

From the perspective of Figure 1, there is no necessity for the evaluator to be separate from the government. Economies of specialisation might favour the government outsourcing this service, but in principle it could be done by civil servants. Moreover, Figure 1 implies that there are no motivational problems: governments operate solely in the interests of the citizens and the appraiser acts diligently and truthfully. Similarly, with perfect information evaluation merges into policy formulation, as reference to the Monetary Policy Committee suggests. Although *ex ante* appraisal is usually conceived as a process whereby the evaluator appraises individual policy schemes submitted by the government, with a well-defined SWF and set of constraints, the evaluator could be given the job of choosing the optimal policy. Finally, with the perfect information scenario, there is no reason for the evaluations to be public knowledge, but again there is no reason for them not to be.[5]

5. Evaluation as an Information Problem: The Social Welfare Function

What is rather curious about the process outlined in Table 1 is that although the central aim of policy evaluation is the revelation of information, there is assumed to be no conflicts of interest involved. All agents act truthfully and diligently and fully reveal what they know. Perhaps with perfect information, everyone is a potential evaluator so that

there is no possible gain in attempting to distort or withhold information. However, we know that where information is costly to gather and/or process, and in particular where there is asymmetric access to information, co-ordination and motivational problems are likely to arise. What is implied here is that it is in the interests of some (or all) individuals and groups to withhold or distort information in their own interests.

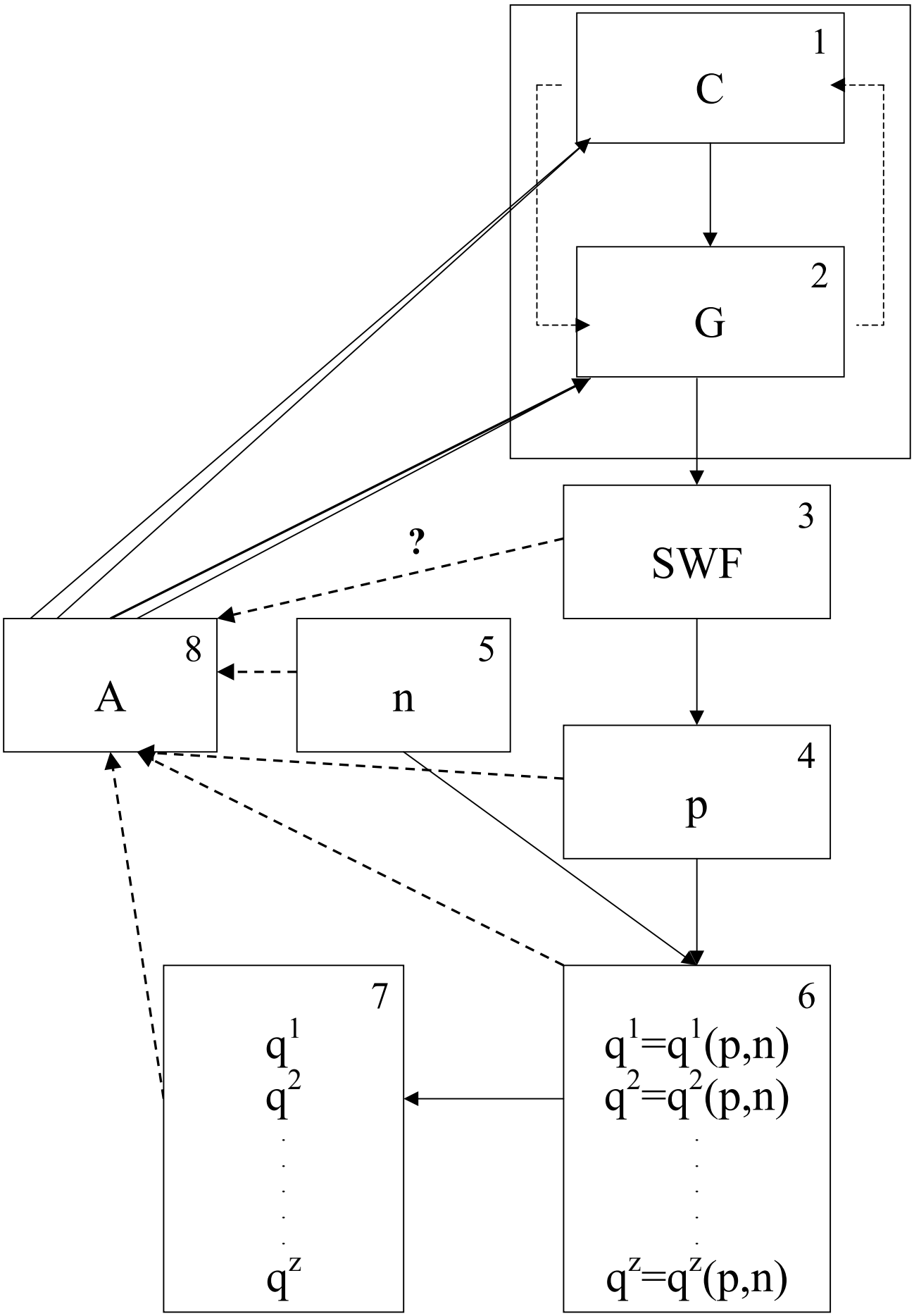
As stated in Section 3, in practice two major informational issues beset regional policy evaluation. These are illustrated in Figure 2, which amends the representation of the evaluation process given in Figure 1. It illustrates the evaluation process with incomplete information. One issue concerns the agency problems surrounding the identification of the Social Welfare Function. This is represented in Box A, which incorporates Boxes 1 and 2 for the citizen and the government. The other is the competing set of weakly parameterised theories in Box 6. Further, these informational difficulties interact in the evaluation process.

We begin with the specification of the Social Welfare Function. Since the inception of UK regional policy, governments of all political persuasions have been very vague in spelling out their policy goals and particularly the weights attached to the individual elements making up their regional policy objectives. In practice, for regional policy evaluation, only the most crude target variables are usually identified - in the past, net regional employment generated and net exchequer costs incurred have been the most popular – with no marginal trade off rates explicitly specified. More recently, the Public Service Agreement targets for regional policy over the period 2003-2006 were to

> [m]ake sustainable improvements in the economic performance of all English regions and over the long term reduce the persistent gap in growth rates between regions, defining measures to improve performance and reporting progress against these measures by 2006 (HM Treasury, 2002, p. 25)[6].

The Social Welfare Function is supposed to drive appraisal and evaluation. Why is the government so reticent in revealing the Social Welfare Function and what are the implications for evaluation?

Figure 2: Evaluation with Incomplete and Inaccurate Information



The first point to be recognised is that in Figure 2 we have removed the fiction that the citizens have a self-evident Social Welfare Function that the government simply attempts to maximise. Rather, we take it that one of the functions of the government is to articulate such a welfare function. That is to say, in the implicitly democratic account outlined in Figure 2 the citizens elect the government, partially based upon promises concerning policies that they will carry out, and the government then sets about implementing a SWF of its choice. The day-to-day political process implies that information will flow from citizens, or representatives of particular groups, to the government and from the government to the citizens. That is to say, in Box A we present a stylised depiction of the political process.

The position of the government involves at least three potentially troublesome informational aspects. The first is that formally the government acts as the agent to the principal, the electorate, and one of its functions is to organise the provision and financing of public goods, which includes the operation of appropriate economic policies.[7] There are standard moral hazard problems associated with this role, involving both potential hidden actions and hidden information available to the government. The operation of the government clearly need some monitoring and, at the bottom line, evaluation has a role in countering government corruption and fraud (Rose-Ackerman, 1999).

However, the evaluation process clearly has a more general potential impact on the likelihood of the government's being re-elected. Whilst governments praise the bracing effects of evaluation for others (teachers, doctors etc.), it is not necessarily unreasonable that they should attempt to evade evaluation themselves. One argument is that if the government's objective function, their Social Welfare Function, were to be known too precisely, then their performance might be too straightforwardly identified. Perhaps governments see an advantage in selectively releasing information *ex post* - showcasing good news and burying bad news – rather than setting up a measuring rod *ex ante* against which their performance can be measured.

There is a major problem with this approach: propaganda is not an effective communication strategy. The notion that citizens will believe available reports on the effectiveness of government policies to be unbiased whilst these reports have in fact been selectively doctored so as to screen out negative elements and highlight positive ones is not credible. And once citizens begin to question the objectivity of the available evidence on government performance, then even genuine good news is undermined. This is the problem facing the present Labour government concerning so-called "spin".

It seems to us that the argument that the government will fail to make clear its Social Welfare Function in order to shield itself from adverse criticism is not a strong one. It is important for the credibility of a particular government that it should be subject to searching scrutiny by independent institutions such as the press and academia. Whilst there might be persistent short-run temptation to massage evidence on policy impacts, such behaviour damages long-term credibility. This having been said, political decisions do often seem to be driven by short-term considerations and perhaps governments are only just realising the corrosive long-run effect of excessive information manipulation.

However, a second, rather different, issue follows from the government's position as agent to the electorate. This is that the government is not serving a single principal but rather multiple principals. But it is well known that where these principals differ in their interests, it might be impossible to identify a single Social Welfare Function within a democratic setting (Arrow, 1963). Faced with this problem but a necessity to act, it might be the optimal strategy for a government trying to build a majority for a particular policy to keep the aims of the policy vague. This might be particularly important for spatial policy where it is likely that there would be spatial losers as well as winners following any particular policy and where these spatial groups will be relatively easy to organise and activate.[8]

Whilst the government can be seen as the agent of the electorate, in many situations - when it directs policy - it will also act as the principal. In these circumstances private sector organisations - firms or unions for example - might take the role of agents

in the delivery of policy goals. However, for successful co-ordination, it might be important that the government's "type" (its specific preferences) be known only imprecisely. For example, in "cheap talk" games players can choose more effective strategies if they know one another's type. However, unless they have perfectly aligned preferences, there is no equilibrium disclosure of information if each is forced to reply with total precision (Crawford and Sobel, 1982; Farrell and Rabin, 1996). It is in the interests of each to attempt to mislead the other player. But because the other player knows this, this misrepresentation is taken into account. But the first player can anticipate this so that escalation in information distortion occurs with no final equilibrium. However, if each player has only to identify their type imprecisely, within a wider group, an equilibrium is potentially possible with some information passing between players. Similarly, in the bilateral trading problem with agents of different types, Myerson and Satterthwaite (1983) show that exchange will not take place if, and only if, it is efficient. Essentially, it is not incentive-compatible for either agent to be honest about their type. This equally applies in cases where one of the agents is the government.

An example from regional policy might be the refusal by the Treasury to identify a precise shadow wage for regional employment. If this were known, it could weaken the government's position in negotiations over combined actions with unions or firms, so that it would be in the government's interest to misrepresent its true value for the shadow wage or present the information in a very vague form.

The discussion over the identification of the Social Welfare Function raises two key issues concerning the role of evaluation. The first relates to the potential disjuncture between the government and the electorate. In Figure 1 we imply that the interests of citizens straightforwardly translate into a well-defined Social Welfare Function. The government then attempts to undertake policies that will improve the position of its citizens through increasing their welfare. However, in Figure 2 we introduce two ideas that question this. First, citizens almost certainly have interests that differ one with another, so that the determination of the Social Welfare Function becomes problematic. Second, members of the government are likely to have interests that differ from the

electorate as a whole. Therefore, it is likely that to counter moral hazard problems, the government needs to be monitored.

From these perspectives the role of policy evaluation is to increase the information available to the electorate. This information concerns both the nature of the operation of particular policies and the diligence of the government in pursuing these policies. These facts should improve the democratic process in that they give the electorate more information that can be used to more effectively choose between different policies and different governments (Stiglitz, 1999). Also, the fact that its policies are being monitored should also increase the effort that any incumbent government puts in to policy delivery. Essentially this implies that in democratic countries, evaluation should play a role in the democratic process – the process of the policy choice and scrutiny of the government. However, how central to the role of evaluation should these considerations be? Who is the evaluation for and how widely should the results be distributed?

The second key issue is that there might be conflict between the effectiveness of policy and the rigour of the evaluation. We have argued that the effectiveness of policy might rely on the government's objectives being rather opaque whereas the effectiveness of the policy evaluation depends on their being transparent. Here there is a tension between the role of policy evaluation in increasing government accountability and its role in improving direct government efficiency.

We do not have a watertight argument with which to navigate these difficult waters; we present here no all-encompassing model whereby these competing roles for evaluation are traded off or somehow rendered compatible. However, our own instincts follow those of Stiglitz (2002, p. 488) who writes: "Without unbiased information, the check that can be provided by the citizenry [on the behaviour of the government] is limited; without good information, the contestability of the political process can be undermined." It is therefore of concern when the government attempts to control the research agenda for academics, implicitly asserts that its own perception of policy

concerns are necessarily those of the population as a whole and finances evaluation reports that are not freely available to the public.

.

6. Evaluation as an Information Problem: The Link between Policy Instruments and Target Variables

In the conventional, perfect information, account of evaluation, as represented by Figure 1, we know precisely the link between the vectors of policy instruments, non-policy variables and the target variables. However, in practice in the UK there is no consensus over the appropriate theory that underpins the operation of regional policy and certainly no agreed precisely parameterised model. Part of the evaluation process therefore becomes necessarily bound up with model choice and hypothesis testing.

One reason why there is disagreement over the underlying economic processes causing variation in activity over space might simply be the complexity of the phenomena. This would particularly apply to analyses that focus on a region's development as the qualitative evolution of economic activity over time. Testing such theories might simply be inherently difficult. However, a key element in the uncertainty surrounding the true nature of spatial economic processes within the UK is the lack of regional data.

It will be useful to focus on data problems with measures that potentially make up the policy-relevant endogenous target variables, *q,* and the exogenous non-policy variables, *n*. For most of the relevant regional variables, measures are imprecise. That is to say, the true accuracy is unknown. For some variables that we would expect to be in the *n* set, estimates are simply not available. We begin by taking the example of the regional policy regime of 1960s and 1970s.

In the accepted policy impact model used to evaluate regional policies in this period, the main target variable was employment. In practice, operational definitions, and accurate measures, of regional employment are needed. These would include both

employees and self-employment and would, at a minimum, allow distinctions between part-time and full-time jobs. Reasonably robust estimates of employee jobs are available for UK regions. However, there are more substantial problems with the self-employment data and information on employment status, etc.

These problems are magnified in the context of newer policy regimes that take productivity measures as target variables. HM Treasury's preferred productivity measure is GDP per hour. Data issues are here much more serious, given that there are problems in estimating both the numerator (regional GDP) and denominator (regional employment) in 'productivity' measures.

The process of estimating GDP at the national level is highly complex and is likely to generate various types of errors, so that the final accuracy of these estimates is problematic (Akritidis, 2002). Since the accuracy is unknown, it is difficult to judge whether changes in estimated GDP reflect real economic events or variations in the estimation process. These problems are much more severe for regional estimates. This is due to a number of factors. First, smaller sample sizes in the main survey sources increases sampling errors, particularly for the smaller regions. Second, modelling is used to 'regionalize' results from national surveys that are not stratified at a regional level. Third, for the regional data we lack the triangulation checks between the product, expenditure and income estimation methods. GDP estimates for UK regions use only the income approach (Lacey, 2000).

ONS have not provided an assessment of accuracy for UK regional accounts since 1990 (ONS, 1990). Recent statements merely comment that "the overall margin of error of the estimates … is very difficult to judge" (Lacey, 2000). Data accuracy is likely to be substantially less with regional than with national estimates due to additional inaccuracies generated by the process of allocating national estimates by region. These errors might entail systematic biases as well as a mere lack of precision, with potentially serious policy implications (Cameron and Muellbauer, 2000).

As we have already commented, there are also errors in the denominator of productivity measures, employment. This generates additional uncertainty concerning the accuracy of the resulting productivity measures, since errors in the numerator and denominator can interact in complex ways. The result is that we have fairly poor information, even on the target variables for regional policy. There is concern that these data might not be sufficiently precise to allow identification of impacts from feasible regional policy interventions. that is to say, the margin of error is likely to be large in comparison with realistic policy impacts.

Data problems also extend to non-policy variables within the impact model. Take again a simple demand-driven impact model with an employment target. Policy impact depends on intra- and inter-regional Input-Output (I-O) linkages, etc. The nature of these linkages is poorly understood for most UK regions, due to a lack of regional I-O accounts. For example, a key characteristic of regional, as against national, economies is that they are very open - regions typically trade heavily with other regions. However, in the UK, apart from Scotland, there are no official regional trade data. Further, the openness of the region makes price competitiveness an important issue. However, there is a complete lack of regional price data. Again, a stylised fact concerning regional economies is that factor mobility between regions is much higher than between nations. But the UK has very limited information about regional migration and capital stock. The overall effect of this lack of regional data is to seriously limit scope for testing alternative theories of the determinants of regional economic activity and accurately parameterising such models.

The lack of an agreed theoretical framework has two important implications for evaluation. First, any empirical work on regional data potentially benefits appraisal and evaluation performance if it improves information about the operation of regional economies. Second, *ex post* evaluation should be closely linked to the hypothesis testing literature. In fact policy changes should provide relevant exogenous shocks that can be used to reveal the comparative static properties of the regional economy.

This is in contrast to the view that evaluation involves a choice from a cookbook of methods where the particular context suggests the most appropriate technique. For an economist, the approach of researchers such as Ashcroft and Taylor (1979), Harris (1991), Moore *et al* (1986) and Wren and Waterson (1991) has much to recommend it.[9] Using this method, the evaluation jointly tests the policy effectiveness hypotheses together with other hypotheses about the operation of the regional economy. Our main criticisms of this kind of work are simply that not enough has been done and sometimes too much emphasis is placed on identifying policy effects and insufficient in attempting to interpret the results concerning other elements of the regional economy.

It is of interest to compare this kind of analysis with the "industrial survey" approach that accompanied the shift towards discretionary policy in the 1980s (PACEC, 1993). This is evaluation based upon interview and questionnaire surveys of recipient firms. The impacts of policy are identified under three headings: additionality, displacement and multiplier effects. In practice, this type of evaluation focuses on the measurement of additionality. This is the accuracy with which civil servants can identify marginal projects; projects that would not have gone ahead without aid. The measurement of displacement only attempts to take into account product market effects, and the impact is asserted not tested. No labour market displacement is typically identified at the regional level. Further, multiplier effects are often determined by applying simple rules of thumb. Importantly, these evaluations provide few insights into how regional economies operate or how regional policy actually effects the relationship between regional aggregates.

A number of useful points can be made about the implications for policy evaluation of the underdeveloped state of empirical regional models in the UK. The first is that regional policy evaluation will improve with any improvement in our empirical knowledge of the operation of the UK economy over space. Such improvements might come from work that is not directly policy orientated. Second, where evaluation methods imply or impose a particular untested model, then this should be made explicit. Third,

that evaluation itself should attempt to improve our understanding of the economic processes that operate within and across regions.

7. Interaction between the Uncertainties Concerning the SWF and the Appropriate Empirical Model

The uncertainty in one part of the evaluation process can interact with uncertainty in other parts. For example, uncertainty over the true model of the economy permits uncertainty over the SWF. If the relationship between policy instruments and target variables were common knowledge, it would be difficult for the government to conceal the marginal weights in its SWF. A government maximising its SWF will equate the marginal rate of transformation between target variables in the economy with the marginal rate of substitution in the SWF. Under these circumstances, a correct model of the economy, together with optimising behaviour by the government, will reveal the marginal weights that enter the SWF. However, uncertainty over the appropriate empirical model for the economy allows the government to be vague over the implicit marginal weights in its SWF.

A rather more concrete example of interaction can occur if a particular empirical model from Box 6 in Figure 2 is specified for use in an evaluation. As will be argued at greater length in the next section, the characteristics of the specific model used in evaluation might well affect whether a particular policy initiative will be identified as efficient or not. Under these circumstances the government or a pressure group might attempt to impose a particular (hidden) SWF through the choice of the evaluation model to be used.

A regional authority or pressure group, whose SWF will give greater weight to activity in its own region, might attempt to influence the national government by employing such an approach. Imagine an economy in which there are two regions, North, N, and South, S. Also assume that only employment, e, enters the government's or regional authority's SWF and that these are given as:

$$SWF_G = \ln e_N + \ln e_S \quad (6)$$

$$SWF_N = e_N \quad (7)$$

where subscript G stands for government and subscripts N and S for the corresponding region.

Equation (6) represents the government's SWF. This is increasing in employment in each region and, for a given national employment level, favours a more equal regional distribution of employment. Equation (7) is the SWF for the Northern regional authority. This is simply increasing in Northern employment and disregards (gives a zero weight to) the employment level in the South.

Imagine that a particular policy change (Δp) generates a change in the target variables (Δq), whose true value, represented by the T superscript, in this case is given by $[\Delta e^T_N, -\Delta e^T_S]$. What is implied here is that employment increases in the North are accompanied by employment decreases in the South. Using the government's SWF, the efficiency of the policy is uncertain until the absolute size of the employment changes and the initial employment levels in the two regions are known. However, under the Northern regional authority's SWF, the policy is unambiguously efficient.

Suppose that the Northern regional authority commissions an evaluation but specifies that a model should be used which only identifies the impact on the Northern region. Many regional impact models are of this type, as the spatially concentrated direct impacts of regional policy can be easily identified but the spatially diffuse indirect impacts are more difficult to quantify.[10] Imagine that the measured impacts on Northern and Southern employment, Δe^M_N and Δe^M_S are therefore given by:

$$\Delta e^M_N = \alpha \Delta e^T_N, \quad \Delta e^M_S = 0 \quad (8)$$

where $\alpha > 0$. This means that the change in measured Northern employment is positively related to the true change in Northern employment and the measured change in Southern employment is zero.

The key point here is that if the Northern regional authority can get the government to accept this evaluation method, then the government's "measured" SWF becomes observationally equivalent to Northern regional authority's true SWF. The change in the governments measured SWF, ΔSWF^{M}_{G}, can be expressed as a function of the true change in the Northern regional authority's SWF value, ΔSWF^{T}_{N};

$$\Delta SWF_G^M = \ln\left(1 + \frac{\alpha \Delta SWF_N^T}{e_N}\right) \tag{9}$$

where

$$\Delta SWF_G^M = 0 \;\; iff \;\; \Delta SWF_N^T = 0$$

and

$$\frac{\partial \Delta SWF_G^M}{\partial \Delta SWF_N^T} > 0$$

Expression (9) means that whenever the change in the value of the true SWF for the Northern regional authority is positive, the change in the government's measured SWF will also be positive. Moreover, where two policies are to be compared, the ordering using the government's measured SWF will be the same as that using the Northern regional authorities true SWF.

This discussion implies that the distinction between the SWF and the economic constraints can become blurred where there is uncertainty over the nature of these functions. That is to say, disputes concerning which particular theory or method to adopt in an evaluation might in fact be disputes over the appropriate SWF. That this notion is not fanciful is supported by the Treasury directive that evaluations of regeneration policy should take place under the assumption that no additional employment will be generated in the economy as a whole (PACEC, 1993; HM Treasury, 1997). That is to say, the Treasury evaluation rules impose:

$$\Delta e_N^M = -\Delta e_S^M$$

This rule counters the particular strategy outlined above, but if that is its role, it again clouds the distinction between the economic constraints and the SWF: a judgement about

the nature of the constraint is adopted in order to reduce argument fuelled by differences over the nature of the SWF.

8. Theory-lite Evaluation: Hillclimbing

One crucial truth implicit in the discussion up to now concerns the central role played by theory in all aspects of the evaluation process. Theory directs us to look for certain relationships and not others. It tells us that certain variables are fundamental, others can be ignored. It supplies a causal narrative linking the policy instruments with the target variables. These causal narratives allow us to understand how policy works and to explain its operation to others. Where the theory that underlies evaluation is not stated, it remains implicit.

Certain key elements in the evaluation process outlined in Section 2 - that is the identification and parameterisation of the model for the relationship between policy instruments and target variables – are clearly theory dependent. However, the particular theoretical framework adopted will play a part in all of the elements that make up the evaluation process.

The appropriate exogenous variables and what are taken to be effective policy instruments will be affected by the chosen theoretical approach. For example, as argued in the previous section, for regional evaluation at least, the Treasury's official position is that the natural rate hypothesis operates at the national level. This implies that total UK employment is exogenous to the operation of this form of policy (HM Treasury, 1997).[11] Further, theory will direct us to the appropriate target variables and their weight in the Social Welfare Function. The retro theoretical framework that identified regional problems with failure in the labour market points to employment as a key target variable. However, the recently adopted spatial-co-ordination-failure approach promotes regional productivity as the key target variable (HM Treasury, 2001).

To a certain extent there is a degree of circularity here. The availability of data is likely to reflect the theoretical ideas current in the recent past. Where data collection is costly, the government targets its information gathering activities to those that are perceived to be particularly useful. This can limit the evaluation of policies driven by a new theoretical agenda. An example is cluster policy, where the identification of cluster performance is often problematic because data are not collected at the level of the cluster but rather by outdated sectors (Learmonth *et al*, 2003).

We have claimed that uncertainty about which theory to use for policy evaluation is problematic. However, it is sometimes claimed that we can cut this particular Gordian knot through the use of theory-lite evaluation techniques. These techniques are thought to replicate, either directly or indirectly, the role of controlled experiments in the natural sciences. One method is to try to get matched samples of companies. For example, one sample of companies can be chosen which did, and another which did not, receive the policy (Hill, 2001). The same approach could be applied to individuals, areas or some other appropriate aggregation of policy recipients. There are at least two problems with this method. The first is that it does not get around the problem of requiring a prior theory because one needs to match the sample companies on some criteria and these criteria will be theory specific. Second, it is difficult to identify such samples of firms in the normal operation of policy: such an exercise would therefore have to be set up experimentally (Taylor, 2002).

A second, at one time influential, approach that was claimed to be atheoretical involves the use of the shift-share technique (Fothergill and Gudgin, 1979). The underlying notion here is that shift-share removes the national effects operating through the region's industrial structure, leaving only a region-specific residual. The method was then further extended to remove other elements of regional structure (specifically its spatial structure). In all of this, shift share was described explicitly as not being a theory but merely a standardisation procedure. The problem with this claim is that shift-share will only accurately identify the impacts of national effects working at the level of the region if it accurately replicates those effects. In order to know that we need the true

theory of how those effects work. Therefore shift-share cannot be seen as an atheoretical technical device, and further its theoretical foundations are weak (Holden *et al*, 1989).

A more radical response to the discussion up to now might be to question the need for such intellectual rumination at all. There are many activities that seem conceptually fraught with difficulty, such as learning to read or riding a bicycle, which are relatively straightforward once attempted. Many practitioners might see evaluation in the form of a control engineering solution, as identified in Figure 1, but implementing a backward-, rather than forward-, looking policy-adjustment process.

The form of evaluation suggested here is essentially a trial and error, iterative, hill-climbing procedure. In this type of approach, the distinction between *ex ante* appraisal and *ex post* evaluation is blurred. Policy becomes a repeating cycle whereby future policy directions depend upon the perceived effectiveness of past policy initiatives (Batterbury and Hill, 2003). Policies are used as long as they are producing positive changes and adjusted when they are not. In the context of the devolved delivery of policy, the efficiency of this trial and error procedure might be enhanced through benchmark competition (Tirole, 1994). With devolution, the effectiveness of a number of policies could be compared simultaneously.

A number of points need to be made about such a general approach. The first is that with this method policy goals need to be modest and policy instruments simple and robust, rather than complex and sophisticated, if the government is operating with such a lack of information. For example, the authorities need rapid feedback if individual policies are ineffective. Policies that necessarily produce results over an extended time period are less efficiently operated in this way.

Second, for the actual climbing of hills, the strategy has some major drawbacks - this is the reason that ramblers purchase maps, guides and compasses. The method might only lead to a local, rather than global optimum and might be achieved by a very inefficient route. However, the hill walker has the advantage of knowing that any gain in

height has occurred through her own actions: the physical characteristics of the hill are, except under catastrophic conditions, fixed. But this cannot be said for the process of regional policy evaluation, producing the well-known counter factual problem. To use this method efficiently implies that we can accurately model the operation of the economy without policy but cannot model the operation of the economy with policy. But, as we have argued already, we have difficulty in modelling the regional economy, either with or without policy intervention.

A third problem is that a strategy of changing policy iteratively in response to evidence on effectiveness generates problems of time consistency in the setting of policy (Kydland and Prescott, 1977). Essentially, the government cannot credibly commit to future policy actions, and this might seriously limit policy effectiveness. This is important where investment decisions are to be taken by the private sector which depend on aid levels in the future.[12]

9. Conclusions and Challenges for the Future

Four main conclusions can be drawn from the discussion in this paper. The first is that all existing evaluation results for UK regional policy must be tentative, because of the weakness of regional data and our lack of knowledge about the underlying spatial economic model. Perhaps accompanying this conclusion should be the advice that policy aims should therefore be relatively modest and the temptation to micro-manage regional policy should be resisted. Second, wherever possible, policy evaluation should be used to reveal more information about the operation of the regional economy. That is to say, evaluation in the future should focus more on understanding spatial interaction and less on monitoring policy provision. Third, evaluation should be regarded not as an abstract procedure but part of a process through which policy choice is made. As such, the evaluation outcomes can have important implications for the interests of individuals and groups. In this setting, issues of asymmetric information are important. These must be recognised and the implications for the efficiency of evaluation understood. Finally,

evaluation is part of the democratic process. As such it is dangerous for the government to dominate the evaluation agenda.

In future, we would expect the evaluation of regional policy to be much more challenging. First, current policy innovations are motivated by novel theories that are likely to be more difficult to test. Traditional regional policy instruments are based around very basic economic theory. Automatic factor subsidies applied to a range of industries reduce unit costs, increasing regional competitiveness and output but lead to the substitution of the unsubsidised inputs by the subsidised ones. However, discretionary aid and schemes aimed to improve company dynamic efficiency through the generation of positive spill-over effects are less easily conceptualised and tested. Second, the data requirements for policy evaluation in the future will be more demanding yet the availability of time series of reliable regional data is very limited. Third, the devolution and delegation of powers of regional policymaking and delivery to multiple assemblies and agencies will almost certainly increase the problems of asymmetric information. Finally, although there is a rhetorical commitment to evidence-based policy, the present emphasis on selling the government's policies to the electorate seems to be accompanied by a low tolerance for alternative perspectives.

References

Akritidis, L. (2002) "Accuracy assessment of national accounts statistics", *Economic Trends*, No. 589, London: TSO.

Arrow, K. (1963) *Social Choice and Individual Values, 2nd edition*, Wiley, New York.

Ashcroft, B. and Taylor, J. (1979) "The Effect of Regional Policy on the Movement of Manufacturing Industry in Great Britain", in D. Maclennan and J.B. Parr eds., *Regional Policy: Past Experience and New Directions*, Martin Robertson, Oxxford.

Balls E. (2002) "The New Localism: Speech by the Chief Economic Adviser, Ed Balls to the CIPFA Annual Conference in Brighton", HM Treasury, 2002.

Batterbury, S. and Hill, S. (2003), "Evaluation and Policy Development", *mimeo*, paper delivered at the ESRC Urban and Regional Economics Seminar Group, Cardiff.

Cameron, G. and Muellbauer, J. (2000) "Earnings Biases in the United Kingdom Regional Accounts: Some Economic Policy and Research Implications", *Economic Journal*, vol.110, pp. F412-F429.

Chiang, A.C. (1974), *Fundamental Methods of Mathematical Economics, 2nd. ed.,* McGraw-Hill Kogakusha, Ltd., Tokyo.

Crawford, V. and Sobel, J. (1982) "Strategic Information Transmission", *Econometrica*, vol. 50, pp. 1431-1451.

Dixit, A.K. (1996) *The Making of Economic Policy: A Transaction-Cost Politics Perspective*, MIT Press, Cambridge, Mass.

Farrell, J. and Rabin, M. (1996), "Cheap Talk", *Journal of Economic Perspectives*, vol. 10, pp. 103-118.

Fothergill, S. and Gudgin, G. (1979) "In Defense of Shift-Share Analysis", *Urban Studies*, vol. 12, pp. 246-255.

Harris, R.I.D. (1991) "The Employment Creation Effects of Factor Subsidies: Some Estimates for Northern Ireland Manufacturing Industry 1955-83", *Journal of Regional Science*, vol. 31, pp. 49-64.

Hill, S. (2001) "Do Regional Development Agencies Make Any Difference?", Paper delivered to the International regional Science Association, British and Irish Section, Annual Conference, Durham, September 2001.

HM Treasury (1997), *Appraisal and Evaluation in Central Government: Treasury Guidance*, The Stationery Office, London.

HM Treasury (2001), *Productivity in the UK: 3 – The Regional Dimension*, www.hm-treasury.gov.uk

HM Treasury (2002), *Public Service Agreements: 2002 Spending Review*, www. hm-treasury.gov.uk

Holden, D., Nairn A. G. M. and Swales J. K. (1989) "Shift-Share Analysis of Regional Growth and Policy: A Critique", *Oxford Bulletin of Economics and Statistics*, Vol. 51, pp. 15-34

Kydland, F.S. and Prescott, E.C. (1977) "Rules Rather Than Discretion: The Inconsistency of Optimal Plans", *Journal of Political Economy*, vol. 85, pp. 4473-490.

Lacey, D. (2000) "UK Regional Gross Domestic Product (GDP): Methodological Guide", *Office for National Statistics*.

Learmonth, D.L., Munro, A. and Swales, J.K. (2003) "Multi-Sectoral Cluster Modelling: The Evaluation of Scottish Enterprise Cluster Policy", *European Planning Studies*, vol. 11, forthcoming

Moore, B., Rhodes, J. and Tyler, P. (1986) *The Effect of Government Regional Economic Policy*, HMSO, London.

Myerson, R. and Satterthwaite, M. (1983) "Efficient Mechanisms for Bilateral Trading", *Journal of Economic Theory*, vol. 29, pp. 1-21.

Office for National Statistics (1990) "Regional Accounts, 1989 – Part 1", *Economic Trends*, No. 444, London: HMSO.

PACEC (1993) *Regional Selective Assistance 1985-88*, HMSO, London.

Rose-Ackerman, S. (1999), *Corruption and Government: Causes, Consequences and Reform*, Cambridge University Press, Cambridge, UK.

Schofield, J.A. (1987), *Cost Benefit Analysis in Urban and Regional Planning*, Unwin Hyman, London.

Stiglitz, J.E. (1999), "On Liberty, the Right to Know, and Public Discourse: The Role of Transparency in Public Life", *mimeo*, Oxford Amnesty Lecture, Oxford.

Stiglitz, J.E. (2002), "Information and the Change in the Paradigm in Economics", *American Economic Review*, vol. 92, pp. 460-501.

Swales, J.K. (1997), "A Cost-Benefit Approach to the Evaluation of Regional Selective Assistance", *Fiscal Studies*, Vol. 18, pp. 73-85.

Taylor, J. (2002) The Evaluation of UK Regional Policy: How much Progress has been Made?" in B. Johansson, C. Karlsson and R.G. Stough eds. *Regional Policies and Comparative Advantage*, Edward Elgar, Cheltenham, UK.

Tirole, J. (1994), "The Internal Organisation of Government", *Oxford Economic Papers*, vol.46, pp. 1-29.

Wren, C. and Waterson, M. (1991) "The Direct Employment Effects of Financial Assistance to Industry", *Oxford Economic Papers*, vol. 43, pp. 116-138.

Footnotes

[1] Problems of bounded rationality and asymmetric information seriously limit the scope and desirability of "joined-up" policy making (Tirole, 1994).

[2] Where equation (1) is not differentiable a comparable, but less compact, relationship can be produced which recognises possible discontinuities in the function represented by equation (1).

[3] Policies will, in general, produce a stream of impacts over time. Therefore part of the Social Welfare Function involves identifying the appropriate time discount rates, attitude to risk, etc..

[4] Balls (2002) seems to equate plans to devolve and delegate regional policy with the decision to delegate monetary policy in the UK to the MPC. In a future paper we intend to contrast more thoroughly the operation of devolved regional policy as against the operation of the MPC.

[5] In fact because the government has delegated policy in order to deal with the time inconsistency problem, for the operation of the MPC it is important that appraisal and implementation are open.

[6] This is as part of the Public Service Agreements to HM Treasury, the Department of Trade and Industry and the Department of the Deputy Prime Minister. Also the Department for Work and Pensions has, as part of its Public Service Agreement, agreed to "increase the employment rates of … the 30 local authority districts with the poorest initial labour market position, and significantly reduce the difference between their employment rates and the overall rate" (HM Treasury, 2002, p. 31).

[7] In this paper we use the terms "citizens" and "the electorate" as synonyms.

[8] We conjecture that this problem will become more prominent for regional policy with the move towards more geographically devolved economic policy making. Essentially, as suggested in Section 7, regions will have individual, conflicting, SWFs.

[9] This list is taken to be indicative and is in no sense exhaustive.

[10] For example, Taylor (2002) argues that one fundamental question for UK regional policy evaluation that remains to be determined is the impact of regional policy on non-assisted areas.

[11] This is independent of the fact that regional policy generates supply-side changes and therefore if effective could change the natural rate of unemployment.

[12] For example, it might be that Locate in Scotland was effective in attracting Foreign Direct Investment because the agency signalled a long-term commitment by Scottish Enterprise and the Scottish Office to aid incoming firms.