Available online at www.sciencedirect.com

SCIENCE DIRECT®

Economics Letters 86 (2005) 167–172

economics letters

www.elsevier.com/locate/econbase

A simple approach to identifying the incentives for policy experimentation

Martin Ellison[a,b,*], Jouko Vilmunen[c]

[a]Department of Economics, University of Warwick, Coventry CV4 7AL, United Kingdom
[b]CEPR, United Kingdom
[c]Bank of Finland, Finland

Received 25 March 2003; accepted 9 July 2004
Available online 10 November 2004

Abstract

We present a simple concavity condition that describes the incentives for a policy maker to pursue policy experiments. Our approach avoids the computational cost of numerical methods that fully characterise optimal experimentation under learning.

Keywords: Experimentation; Learning

JEL classification: D81; D83

1. Introduction

Policy makers are often uncertain about the effects of their actions on the economy. If the uncertainty is due to incomplete knowledge of how policy actions translate into economic outcomes, it can be argued that policy should contain an element of experimentation. Policy experiments help the policy maker learn and so reduce future uncertainty about the policy transmission mechanism. In this paper, we propose a simple approach to identify whether there are incentives for a policy maker to make such

* Corresponding author. Department of Economics, University of Warwick, Coventry CV4 7AL, United Kingdom. Tel.: +44 2476 523471; fax: +44 2476 523032.
E-mail address: martin.ellison@warwick.ac.uk (M. Ellison).

0165-1765/$ - see front matter
doi:10.1016/j.econlet.2004.07.009

experiments. We show that a simple concavity condition captures the returns to learning and so defines the incentives for policy experimentation. The value of policy experiments has been examined recently in an important paper by Wieland (2000). The numerical techniques used are potentially very powerful and have a wide range of practical economic applications. However, the method is very computationally intensive and needs considerable computing resources to fully identify the optimal degree of experimentation in policy. Our approach does not rely on numerical methods and therefore avoids such computational costs.

2. Incentives for policy experiments

2.1. Model

We introduce our approach using a simple model in which the policy maker sets policy whilst faced with uncertainty about how policy actions affect economic outcomes. The loss function of the policy maker defined in Eq. (1) has a quadratic penalty for any deviations in the outcomes $\{y_t\}$ from a target level y^*. δ is the discount rate.

$$\mathcal{L}(\{y_t\}) = E_t \sum_{i=0}^{\infty} \delta^i (y_{t+i} - y^*)^2 \tag{1}$$

The policy transmission mechanism is shown in Eq. (2). The policy maker chooses x_t and observes the outcome y_t, but is unable to observe either β_t or ε_t. Without loss of generality, we assume that the parameter β can take one of two values, β_1 or β_2. The policy maker is uncertain about β_t and so forms a belief $p_t \in [0,1]$ that β_1 is the true value at time t. ε_t is an unobserved control error.

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

Beliefs are not static in this model but evolve over time as the policy maker learns about which is the most likely value of β_t. An application of Bayes rule describes how beliefs are updated on the basis of new information. In Eq. (3), initial beliefs p_t are updated to p_{t+1} at the end of the period, after the realisation of y_t. Under such Bayesian learning, p_{t+1} depends on the relative probability of observing outcome y_t with the two possible β values.

$$p_{t+1} = \frac{p_t P(y_t | x_t, \beta_1)}{p_t P(y_t | x_t, \beta_1) + (1 - p_t) P(y_t | x_t, \beta_2)} \tag{3}$$

2.2. Optimal experimentation

To derive the optimal degree of experimentation, the policy maker needs to minimise its loss function (1) subject to the policy transmission mechanism (2), learning (3), and initial beliefs p_t. By incorporating the learning mechanism in the minimisation problem, the policy maker fully internalises the return to experiments. However, the presence of learning means that the policy maker faces a *non*linear-quadratic control problem, for which no closed-form analytical solution exists. Wieland (2000) shows that standard dynamic programming algorithms can be applied to obtain an approximate numerical solution but these are computationally very intensive. Further difficulties arise due to the "curse of dimensionality". In more realistic models, the state space is more complex and numerical solutions rapidly become infeasible.[1]

2.3. A simple approach

From the point of the main focus of our paper, the most important aspect of the policy problem is whether the policy maker has the incentives to experiment. We therefore propose a simple approach focussing on the incentives for experimentation, which avoids the computational difficulties of fully characterising the optimal degree of experimentation. The key question we ask is whether a policy maker internalising learning has an incentive to perform experiments. In the terminology of the learning literature, we want to know whether an active learning policy leads to faster learning than a passive learning policy.

We start from the behaviour of the policy maker when it does not internalise the learning mechanism, i.e. the passive learning policy. In the model, the policy maker minimises the loss function (1), subject to the transmission mechanism (2) and initial beliefs p_t, but ignores learning (3). Since there are no dynamic links between periods, the problem reduces to minimising the one-period loss each period. The solution is shown in Eq. (4).

$$x_t = \frac{p_t\beta_1 + (1-p_t)\beta_2}{p_t\beta_1^2 + (1-p_t)\beta_2^2} y^* \forall t \tag{4}$$

Substituting the passive learning policy (Eq. (4)), together with the transmission mechanism (Eq. (2)), into the loss function (Eq. (1)) gives the ex ante expected one-period loss (Eq. (5)) under passive learning as a function of beliefs. The term in the variance of the control error ε_t is policy invariant and so omitted for notational simplicity.

$$\mathcal{R}(p_t) = \frac{p_t(1-p_t)(\beta_1-\beta_2)^2}{p_t\beta_1^2 + (1-p_t)\beta_2^2} y^{*2} \forall t \tag{5}$$

It is important to note that the policy maker still learns under the passive learning policy. Beliefs continue to be updated according to Bayes rule (Eq. (3)) and eventually tend towards one of the extreme (certainty) values, 0 or 1. However, the policy maker does not consciously internalise this mechanism. The question we pose is whether the policy maker has an incentive to speed up its learning through policy experiments. With policy experimentation, beliefs tend faster towards one of the extreme values. In essence, the choice is between two expected paths for beliefs, $\{p_t\}$ under passive learning and $\{p'_t\}$ with experimentation.

To identify the incentives for policy experiments, we consider a stylised experiment in which the policy maker increases x_t by a small amount above the passive learning policy level at time t. In all future periods, the policy maker returns to the passive learning policy. The experiment has no cost in period t since $\mathrm{d}\mathcal{L}_t/\mathrm{d}x_t = 0$ when evaluated under the passive learning policy.[2] The benefit of the

[1] To replicate the results of the illustrative example in Wieland (2000) it took the authors 4 h of processing time in GAUSS 5.1. In more complex models, we increased the number of state variables to two and three and found ourselves waiting first 2 days and then 3 weeks for the numerical algorithm to converge.

[2] This is an application of the envelope theorem.

experiment in terms of reduced losses in future periods depends on whether the path of beliefs it induces, $\{p_t'\}$, leads to a lower losses than the original passive learning path, $\{p_t\}$. In summary, experiments are beneficial if condition (6) holds.

$$E_t \sum_{i=0}^{\infty} \delta^i \mathcal{R}(p_{t+i}') \leq E_t \sum_{i=0}^{\infty} \delta^i \mathcal{R}(p_{t+i}) \tag{6}$$

The stylised experiment we consider has two important features. Firstly, the experiment leads to faster convergence of beliefs to 0 or 1 and so ex ante introduces greater variance in future beliefs. Secondly, since the effect of the experiment on the mean of future beliefs cannot be predicted, $E_t p_{t+I}' = E_t p_{t+_i}\ \forall i$. Taken together, these imply that the experiment creates a mean-preserving spread of beliefs from $\{p_t\}$ to $\{p_t'\}$. The effect of mean-preserving spreads is well known from the uncertainty literature. We apply the result of Rothschild and Stiglitz (1970), which in our case states that condition (6) is satisfied if and only if the function $\mathcal{R}(\cdot)$ is concave. Hence, we have Proposition 1.

Proposition 1. *The policy maker has an incentive for policy experimentation if the expected one-period loss function evaluated under the passive learning policy is concave with respect to beliefs.*

According to Proposition 1, the incentive for policy experiments in our simple model depends on whether the expected one-period loss function (Eq. (5)) is concave or convex.

3. Application

To illustrate our method, we apply it to the debate over optimal experimentation in monetary policy. Wieland (2000) argues that monetary policy should include experiments to help the central bank learn about the monetary transmission mechanism. In contrast, Ellison and Valla (2001) contend that this is not true since experiments destabilise the inflation expectations of the private sector. The concavity condition can help explain these viewpoints.

3.1. The Wieland (2000) model

The first illustrative model of Wieland (2000) has the loss function and transmission mechanism shown in Eqs. (7) and (8), respectively. There are two possible sets of parameter values in the transmission mechanism, (α_1, β_1) and (α_2, β_2). The parameter ω measures the relative weight of policy in the objective function.

$$u(y_t, x_t) = \sum_{i=0}^{\infty} \delta^i \left[(y_{t+i} - y^*)^2 + \omega x_{t+i}^2 \right] \tag{7}$$

$$y_t = \alpha + \beta x_t + \varepsilon_t \tag{8}$$

The model is calibrated with the parameter values (α_1=4, β_1=−1), (α_2=−1, β_2=1) and y^*=ω=0. Following the steps outlined above, we can derive the expected one-period loss as a function of beliefs. It is shown for the calibrated parameter values in Fig. 1.

Fig. 1 shows that the expected loss function is highly concave. By our concavity condition, there is a strong incentive for policy experiments.

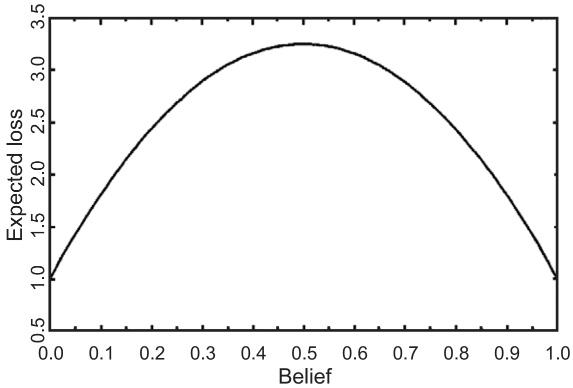

Fig. 1. Expected one-period loss in the Wieland (2000) model.

3.2. The Ellison and Valla (2001) model

The model of Ellison and Valla (2001) introduces private sector rational expectations into the transmission mechanism. Maintaining the notation of the previous section, the loss function of the policy maker and the transmission mechanism are defined in Eqs. (9) and (10). There are again two possible parameter values in the transmission mechanism, β_1 and β_2, but now it is the unexpected policy action $x_t - x_t^e$ which affects the outcome y_t. It is assumed that the central bank receives a private signal about ε_t before setting policy so there is a role for stabilisation policy.

$$u(y_t, x_t) = \sum_{i=0}^{\infty} \delta^i \left[(y_{t+i} - y^*)^2 + \omega x_{t+i}^2 \right] \tag{9}$$

$$y_t = \beta \left(x_t - x_t^e \right) + \varepsilon_t \tag{10}$$

The model is calibrated with β_1=3, β_2=0.5, y^*=0.01 and ω=9. The expected one-period loss in rational expectations equilibrium is shown in Fig. 2.

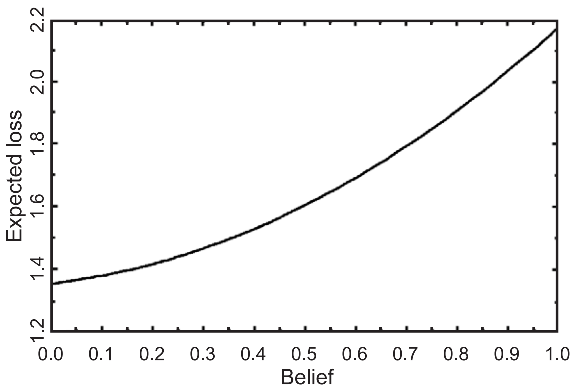

Fig. 2. Expected one-period loss in the Ellison and Valla (2001) model.

In this case, the expected loss function is convex so, by our concavity condition, the incentive for the policy maker is to avoid experimentation. In fact, policy should be designed to impede rather than promote learning.

4. Conclusion

We have demonstrated how a simple concavity condition can be used to examine the incentives to perform policy experiments under learning. We do not propose this as a replacement for numerical techniques that fully characterise the nature of optimal experimentation under learning. Rather, we see the approaches as complimentary. For example, the concavity condition can be applied first to identify the incentives for policy experimentation in different models and for a range of calibrated parameter values. Numerical methods can then be used to complete the full description of optimal experimentation.

Acknowledgements

We would like to thank N. Gregory Mankiw and Neil Rankin for helpful comments and suggestions.

References

Ellison, M., Valla, N., 2001. Learning, uncertainty and central bank activism in an economy with strategic interactions. Journal of Monetary Economics 48, 153–171.

Rothschild, M., Stiglitz, J.E., 1970. Increasing risk: I. A definition. Journal of Economic Theory 2, 225–243.

Wieland, V., 2000. Learning by doing and the value of optimal experimentation. Journal of Economic Dynamics and Control 24, 501–534.