Learning and Economic Policy Choices with an Application to IMF Agreements.

Covadonga Meseguer Yebra*
December 1999
New York University
cm74@is9.nyu.edu

Abstract

I discuss the role of learning in economic policy choices. I test whether choices of policies are driven by experience under them. A Bayesian approach is adopted to tackle this issue and hence, the paper is also an account of the possibilities and limits of this approach. The decision to sign IMF agreements is used as illustration. The model of rational learning suggests that governments are more likely to enter into agreements with the IMF after learning from their own and from the world experience that average results under agreements are better. Also, governments are prone to take risks when they observe very good performers under IMF agreements at the world and at the regional level. However, they are unlikely to continue under a policy that is costly as soon as growth resumes in the region. The rationalistic reading of this behavior is that governments avoid remaining under a costly policy in isolation. A more detailed research is needed in order to adjudicate among this account of switches and an alternative one that emphasizes deviations to the rational rule as the ultimate cause of choices.

***** Juan March Institute, Spain. Visiting Scholar at New York University, Department of Politics. Early steps of this research were carried out under the Training and Mobility of Researchers Program at the European Center for the Analysis of the Social Sciences (ECASS), Essex University, U. K, November-December, 1998. This paper is work in progress. Please, do not cite without the author's permission.

I. Introduction

I explore whether changes in policies are the result of a learning process. Is experience with policies informative when it comes to make future choices?

The argument that countries learn from their own past experience as well as from the experience of others is well established. This reasoning is used to explain, for instance, the "universal convergence" in economic policies witnessed during the 1980s and 1990s (Williamson, 1994; Krueger, 1993; Nelson, 1990; Harberger, 1993). The story of the sweeping wave of economic reforms that took place following the 1982 debt crisis could well be told along the following lines: past economic policy is seen responsible for major economic disruptions. As a result, governments' beliefs about the efficacy of that policy are revised downwards. Governments may not have experience with alternative policies. But they find in others' experience a pool of nonrival-nonexcludable information on which their beliefs about that alternative policy can be based. They compare and choose the policy that maximizes their expected utility. A new policy results from that comparison and it remains in place until this balance is reversed in favor of an alternative. Convergence in policy choices is the major prediction of this model in which governments rationally update their beliefs using a shared pool of evidence. This is because the full mobility of experiences is accompanied by a full mobility of ideas[1]. A Bayesian model of learning is used to test the hypothesis that policy choices are driven by experience. In this model, governments start having some prior beliefs about the outcomes of alternative policies. Priors are updated rationally as information about outcomes becomes available. The combination of prior beliefs and data generates posterior beliefs that become priors on which subsequent policy decisions are based.

If governments act as rational learners and expected utility maximizers, we should expect policy changes to occur in the same direction in different countries and within short lapses of time. This is because Bayesian learning assumes that the information all actors are confronted with is the same, as it is the way information is processed. Also,

[1] On the mobility of ideas internationally, see Stallings (1990), Valdes (1995) and Haas (1997).

since experience rapidly overwhelms prior beliefs, convergence in policy decisions should take place after a few periods.

However, these predictions match reality only partially: convergence is neither immediate nor complete. Learning may be delayed for long periods of time (Lordon, 1998). And even if the frequency of adoption of a certain policy points towards consensus, cases of dissidence exist as well as various degrees of discretion in the application of policies. Furthermore, a Bayesian model cannot explain the introduction for the first time ever of a policy never tried before. When experience is not available, there must be other factors capable of informing and shaping beliefs. Here, ideas or blueprints come to the forefront.

The influence of ideas may remain important beyond the founding moment of a policy innovation. They may intervene in the updating process shaping the way available information is processed. Also, they may influence the conclusions to be drawn from the updating process. In short, experience may not always dominate prior beliefs, what points at the prevalence of ideas. Actual behavior suggests that Bayesian rationality may rely on fairly simplistic assumptions; but at least, it allows the operationalization of a concept characterized by its elusiveness.

I proceed as follows. In section 2, I revise briefly the literature on the role of ideas and learning in explaining policy changes. I present the Bayesian model of learning as well as the main criticisms behind its logic in section 3. Section 4 is an empirical illustration using the decision to sign agreements with the IMF as the policy choice. I conclude in section 5.

II. Ideas and Learning.

The literature on policy change has witnessed a revival of interest in cognitive factors to explain policy choices. Both ideas and learning are considered instruments that come in handy for policy makers who have to make policy decisions under uncertainty. In other words, they are uncertainty-reduction devises.

Ideas are road maps or plans of action which contain both normative and positive information about goals and instruments to attain those goals (Weir and Skocpol, 1985; Sikkink, 1991; Hall, 1986, 1989; Goldstein and Keohane, 1993; Colander and Coats, 1993). The bulk of studies on the topic have explored the question of whether ideas have an autonomous role in policy decisions or whether they gain prevalence only due to the interests and institutions that support them. The success of this line of research is partial, as some excellent works have made clear (Woods, 1995; Jacobsen, 1995; Yee, 1996; Blyth, 1997). The latter claim that research on the autonomous role of ideas has failed to explain why they may be attractive for a decision maker, what makes an idea persuasive, and which features of ideas may be relevant to discriminate between competing ones.

These criticisms aside, what one can conclude from the debate is that ideas per se seem not to be enough to explain a change in policy. In other words, ideas seem to be a necessary but not a sufficient condition to cause a change. They may be circulating for a long while before they gain prevalence and become a policy equilibrium. Thus, their mere existence seems not to be able to induce a change in policy. It is also true, though, that all major policy changes have been articulated having some set of ideas as their source of inspiration. Some authors have interpreted the "necessary but not sufficient condition" as evidence against the argument of the power of ideas. To me, this is to miss the point of the discussion.

Ideas are crucial in forming beliefs (see below) about a policy never tried before. In the absence of any experience, beliefs about the results of a new policy are informed only by ideas. Regardless of whether these ideas are hunches that turn out to succeed or a fully elaborated body of theory, their influence is undeniable. Examples of the former are the so-called proto-Keynesian experiments in the early 1930s (Hall, 1989). The best example of the latter is the Monetarist Revolution that took place in Great Britain and in the United States (late 1970s and early 1980s respectively, Hall 1993). Also, the recent wave of economic reforms in the Third World found inspiration in another blueprint, the Washington Consensus. In all these examples, ideas may not have been the sole driving force behind the innovations. But they played a fundamental role in inspiring and rationalizing them.

Learning is the other uncertainty-reduction mechanism I have referred to. It has been defined as a change in behavior resulting from experience and from the accumulation of new concepts and new information (Odell, 1982; Pierson, 1993; Hood, 1994; North, 1994; March and Olsen, 1989). A nice review of the literature on the topic can be found in Bennet and Howlett (1992). They go over seminal pieces on learning to conclude that different authors agree neither on the concept, nor on the subject or the consequences of learning. Government learning as defined in Etheredge (1981) is carried out by state officials and it results in changes in the policy process. Sabatier and Rose's (1991) lesson-drawing implies changes in instruments. Policy networks are at the center of this type of learning. Finally, Hall's (1993) social learning leads to a revision of ideas on the part of the state and societal actors. This is the only type of learning that extends beyond processes and instruments to affect the ultimate goals of policy. The result of this type of learning is a shift in paradigm or policy regime.

The proliferation of rival concepts of learning poses important methodological problems and makes the empirical test of the learning hypothesis a difficult task. These works have also been criticized for not being able to answer the question of who learns what, when, and why, as well as under which circumstances learning may be a relevant variable for explaining policy choices (Adler and Haas, 1997). These shortcomings are the main reason why an alternative definition of learning is presented and used in the sections that follow.

Another element should be introduced in this presentation of cognitive factors in decision making: beliefs (Basu, 1993; Goldstein and Keohane, op. cit.). The distinction between ideas and beliefs is not always clear in the literature. Yet, I think the distinction is helpful and meaningful. Ideas will be considered here intellectual blueprints that contain cause and effect relationships between certain instruments and certain goals to be pursued. One has beliefs - for instance, about the expected results following the application of a blueprint - based both on ideas and on the lessons of experience. For example, an idea can state that government deficits cause inflation. Based on that causal relationship, a government may belief that an increase in the budget deficit by one point

will induce an increase in inflation of about half a point. Hence, I consider beliefs to be parameterizations of ideas.

When experience is not available, ideas are the only source of beliefs. When experience is available, beliefs are updated or are not. Learning takes place only in the first case. Learning can reinforce initial beliefs and thus validate the ideas on which beliefs are based. But experience may also lead to a revision in beliefs and to a subsequent questioning of the validity of those ideas. In other words, rational learning is not a mere adaptation but it implies some assessment of the ideas actors have in mind.

Ideas may be important not only in informing beliefs in the absence of experience. They can influence the way experience is analyzed and remain prominent all through the updating process.

I can think about the following (non-exhaustive) relevant scenarios.

(i) Take ideas as the starting point. Ideas inform beliefs and beliefs determine the policy choice. After observing the results of policy, beliefs are updated according to Bayes rule and ideas are evaluated accordingly. Governments will choose at all times the policy that maximizes their expected utility. Thus, such a scenario depicts the ideal case of rational behavior at all stages. It is assumed that governments are confronted with the same information, they process it in the same way and hence, they make the same choices. Convergence is the obvious implication. I provide a simple example for the sake of illustration. Imagine that a government thinks that having an independent Central Bank will keep inflation low (idea). It actually beliefs that having an independent Central Bank will bring inflation down by one point (belief). After choosing this strategy (policy), this government observes that inflation has actually increased (experience). Moreover, it observes that governments in control of monetary policy have kept inflation under control or have even reduced it (experience). Hence, this government will revise downwards its beliefs about the efficacy of Central Bank independence out of its own experience. Others' experience would lead to an upward revision of its

beliefs about the efficacy of the alternative policy (rational updating). As a result of the updating, the government changes policy.

(ii) Beliefs can also be updated following procedures that deviate from Bayes rule. In other words, the feedback from experience to beliefs takes place but it is not based on a Bayesian calculation. Empirical studies suggest that actors tend to incur in biases when they process information. For instance, more weight is given to information that is closer in time and space as well as more reliable and unanimous. If this is the case, we may witness decisions that are not rational in the narrow sense defined above.

(iii) Ideas themselves may contain information that lead to a particular interpretation of updated beliefs. Consider the case of market reforms. The ideas on which they are based suggest that a worsening of the economy is expectable, and even desirable, following their application. This worsening can in fact be observed but since in that case the observation would equal the prior belief, posterior and prior beliefs would not differ. The outcome is the same as if updating does not take place, but the scenario is qualitatively different: beliefs are in fact updated but they are not altered. Neither are ideas or policy choices. Note also that ideas may locate good policy results at some point in the future - say, five periods ahead. If that is the case, whatever happens in between may be disregarded in terms of evaluating ideas. Finally, recall that policy choice is a comparative exercise. Governments may stick to a policy that performs badly if the alternative policy performs even worse. In sum, rationality may be more sophisticated that simply persisting in a policy that succeeds and abandon a policy that fails. Governments may stick to policies that yield bad results if those were anticipated, if they are temporarily ignored and if an alternative policy performs even worse.

(iv) Of course, governments may also be dogmatic. They may update their beliefs rationally, conclude from the process that the causal relationships embodied in ideas are wrong and yet, pursue the same policy. Moreover, governments can choose to ignore experience altogether.

(v) Finally, ideas play no role if governments are adaptive learners. Governments may randomly drift adjusting their expectations to past results without supporting their decisions in any logical relation between means and goals. Policy is chosen only on the basis of experience.

The Bayesian model presented in section 4 tests the first scenario, which embodies a narrow definition of rationality. The rest are rival hypothesis to that of a rational learning. Notice that these rival hypothesis are not exhaustive. However, I think they are the most relevant. Of course, an observed behavior that deviates from the one prescribed by Bayesian learning can be caused by variables outside the cognitive realm, such as the political costs that may be attached to a policy change.

Some of the examples outlined above suggest that there may be reasons why a government deviates from the rational path of choices prescribed by a Bayesian model, without this implying dogmatism. Particularly, the behavior depicted in scenarios (ii) and (iii) can be considered perfectly rational if the narrow model of learning is enlarged to take into account that governments learn from those under similar conditions or that they have more than one temporal pattern of belief updating.

I clarify these points in the presentation and discussion of the Bayesian model of learning that follows.

III. A Bayesian Model of Rational Learning and Beyond.

III.1 Bayesian Model of Rational Learning[2].

I present here the basic fundamentals of Bayesian learning. I emphasize concepts in this section while Appendix I offers a more formal presentation of technical details.

[2] This section is inspired by Berger (1985), Leamer (1991), Gelman et. al. (1995) and West and Harrison (1997). On Bayesian decision theory, see De Groot (1970), Winkler (1972), Raiffa (1972), Coyle (1972), Gardenfors and Sahlin (1997) and Pericchi (n. d). Interesting applications to Political Science are Western and Jackman (1994), Gerber and Green (1998) and Western (1998). See also Meseguer (1998)

Some notational complexity is, however, unavoidable. Nevertheless, the essence should be clear: governments start with some prior beliefs about expected outcomes of alternative policies. At some point, observations are gathered regarding the results of policies. The average results and the dispersion of those results are combined with initial beliefs to generate posteriors about outcomes of each policy. Governments' utility is a function of those posteriors. At all times, governments choose the policy that maximizes their expected utility.

Imagine that governments try to anticipate the expected (economic) consequences of alternative policies. I assume that governments are interested in the rate of growth that follows the application of a policy[3]. Governments are uncertain about what outcomes will be generated by each alternative. But, they have some prior beliefs and maybe access to historical data regarding the results of them in the past or somewhere else.

The distinctive feature of Bayesian statistics is the operationalization of that initial uncertainty in a probability distribution reflecting subjective beliefs about expected outcomes. This is a major point of departure from classical statistics, which is based on a frequentist approach to probability. Prior beliefs are especially relevant when decisions are made "unique" events, that is, events whose repetition under the same circumstances is unfeasible. This is the case of most political phenomena.

Imagine that governments can express their uncertainty about the expected results of policy alternatives, $j=\{A,B\}$, by means of a probability distribution. Outcomes of policies, Xj, are assumed to be normally distributed with an unknown mean, Mj, and an unknown variance, Vj. Governments are interested in learning about the expected value of the mean and the expected value of the variance.

In their prior specification, the conditional distribution of the mean is normally distributed. The marginal distribution of the variance follows an Inverse-chi^2 distribution. In this conjugate prior Normal/Inv-chi^2, the distributions of the mean and the variance are interdependent. Thus for $j=\{A,B\}$

$$(1)\ X_j \sim N(M_j, V_j)$$
$$M_j|V_j \sim N(\mu_j,\ \sigma_j^2/\tau_j)$$
$$V_j \sim Inv\text{-}\chi^2\ (\nu_j,\ \sigma_j^2);\ j = \{A, B\}$$

During time t, governments observe the performance of policies A and B. Imagine that n_A countries followed policy A and that n_B countries followed policy B. The following experimental or sample data becomes available at t.

$$X^j_t = x^A_{t,1}, x^A_{j,2} \ldots, x^A_{t,nA,}\ x^B_{t,1}, x^B_{j,2} \ldots, x^B_{t,nB}\ ;\ j = \{A, B\}$$

Sample data are drawn from normal distributions as in (1). Also and importantly, it is assumed that these observations are independent and identically distributed (i.i.d.). The sample means, and the sample sums of squares are sufficient statistics to summarize sample data.

Data combined with prior beliefs produce posterior beliefs, i. e., updated beliefs embodying evidence. The useful feature of Bayesian statistics is that it offers a mechanism of rational learning inspired by Bayes theorem. The expression below states that beliefs conditional on data - posterior beliefs - are proportional to prior beliefs times the likelihood.

$$p(M_j|X_j) \propto p(M_j) * p(X_j|M_j)\ j = \{A, B\}$$

Bayesian learning provides updating equations for the parameters of interest after observing n^j outcomes of policy. In words, governments start with some prior beliefs about average results and variability of results for policies A and B at t. Information is gathered that period and, at the end of the year, governments update their beliefs about A and B using equations (2) and (3) below. These posteriors then become priors in year t+1. Based on them, a policy is chosen that year. Rational updating of beliefs proceeds

[3] There may be, of course, other outcomes of policy politicians would want to learn about, for instance, the rate of inflation or unemployment following the implementation of a

sequentially under the assumption that samples gathered consecutively are independent. Posterior beliefs one year are prior beliefs the following one. The latter combined with new data produces new posteriors and the subsequent choice.

With a normal/Inv-chi[2] prior and a normal likelihood, as assumed above, the posterior expected value of the mean and the posterior expected value of the variance have the following shapes. For each country i, time t and $j=\{A,B\}$

$$(2)\ \mu_{it} = (\tau_{it-1}/\tau_{it})\,\mu_{i,t-1} + (n/\tau_{it})\,\bar{x}_{it} = \rho\,\mu_{it-1} + (1-\rho)\,\bar{x}_{it};\ 0<\rho<1$$

$$(3)\ s^2_{it} = S_{it}/\nu_{it}$$

$\bar{x}_t$ is the sample mean, S_t is the posterior sum of squares and n is the sample size. The expressions for τ_t ν_t and *St* are given in Appendix I. The procedure to find out priors for all parameters is also explained there[4].

Equation (2) tells us that posterior beliefs are a compromise between prior beliefs and sample information. It is important to note that the bigger the sample size the more weight sample information receives in forming posteriors. In turn, if priors convey very little uncertainty, that is, if governments have very precise beliefs about the outcomes of policies, the contribution of experience to posterior beliefs will be minor. Also note that the dispersion of observed results affects the posterior expected value of the variance through (3).

The dispersion of observed results has an interesting conceptual meaning: politicians will find it more difficult to attribute responsibility to policy for observed outcomes if the latter vary a lot.

One may argue, as classical statisticians do, that the process described above will be strongly influenced by the way the researcher models prior beliefs[5]. However, the

---

certain policy. The model extends easily to those cases

[4] To be fully accurate, the expression for the expected value of the variance is E(V)= S/(v-2)

[5] Prior beliefs can be non-informative – flat, diffuse, reference priors - or informative. If priors are non-informative, the likelihood overwhelms the priors in the formation of posteriors. One may want to choose non-informative priors if the aim is to "let the data speak for themselves". This is especially relevant in those settings in which gathering new

dynamics of Bayesian learning implies that the weight of prior beliefs in posterior formation vanishes after a few periods providing sufficient information is gathered. Even if two governments start with different initial priors, as it is perfectly legitimate in this setting, posteriors will converge towards an "intersubjective probability distribution" as more and more information becomes available.

It is important to note that the rate of adaptation to new data, $(1-\rho)$ converges very fast and decreases if initial priors are vague. In practical terms, this means that learning takes place at a fast rate during the first few periods, after which new data hardly influence posteriors. Beliefs become enduring after a few pieces of evidence have been gathered. Thus, Bayesian learning is a fast process of enduring consequences. Of course, it is possible to open this model and allow for the introduction of shocks. These can be modeled with new vague priors that reflect the increase in uncertainty attached to the shock. As a consequence, the rate of adaptation to new data increases dramatically. However, to proceed like this in an explanatory, retrospective model would imply an *ad-hoc* exercise. Furthermore, any observed change in policy should then be attributed to the combination of the shock and the subsequent learning and not to learning alone.

I have described the main features of the Bayesian learning process under two strong assumptions: cross-sectional observations and samples over time are exchangeable. This allows a sequential updating procedure of the type described above. Learning is a compromise between initial priors and experimental observations. As a result of the process, beliefs about the expected outcomes of alternative policies are revised and a decision that maximizes expected utility is made.

Utility depends on the posterior parameters obtained following the procedure described above. Recall also that policy choice has been defined as a comparative exercise between expected outcomes of alternative policies. Thus, the decision rule is a function of the posterior difference in means and a function of the posterior difference in

---

information is costly. Also, it may be the case that there is "insufficient reasons" to model prior beliefs in a particular way. However, informative priors make sense when there is theoretical and empirical material on which prior beliefs can be substantiated, something that frequently happens in comparative research.

variability of outcomes following alternative policies. At all times t and for all governments i

$$U^j(\mu^j, s^j) = \mu^j - \beta s^j;\ j = \{A, B\}$$

where β reflects attitudes towards risk. The decision rule is that country i chooses policy A at t if and only if $U^A_{i,t} > U^B_{i,t}$. It chooses policy B otherwise.

Rational choice coupled with rational learning fully defines the model of rational behavior as suggested in section 2, scenario (i).

What follows is a discussion of the limits of Bayesian learning as a tool to explain actual behavior. From the above description, some of its drawbacks should already result obvious to the reader. In particular, one wonders to what extent the fact that Bayesian learning rules out surprises is compatible with an explanation of policy switches (Denzau and North, 1994). It may well be the case that Bayesian learning is better suited to explain continuities rather than changes.

III.2. Beyond Rational Learning.

In this section, I review some of the main criticisms that Bayesian learning confronts when evaluated as a description of real behavior.

As stated in section 2, this model of learning in which there is a complete feedback at all stages and a rational choice is made as a consequence is just one possibility among many.

Governments may update their beliefs in ways that depart from Bayesian rationality. Biases - departures from Bayesian rule - are frequent in the processing of information. And although biases are not rational, they may be perfectly reasonable.

Empirical studies have pointed that more weight is given to information that is closer in time and space. However, Bayesian learning consolidate beliefs at early stages and they may remain only marginally altered for long periods of time. Furthermore, Bayesian learning assumes that information is processed equally regardless of its source,

a contention that is also arguable. As Robinson (1995: 26) points out, "the early post-war success of Japan seems to have been very influential in determining policy orientation in South Korea and Taiwan, just as these countries' experiences seem to have had subsequent ripple effects on Indonesia, Malaysia and Thailand". However, he states that this experience "should have had little impact in Africa and Latin America" given the big differences in historical and cultural endowments in these regions. Thus, governments may not find useful information coming from sources that they judge as heterogenous. On this basis, it is possible to hypothesize that the least noisy information for one country is its own past experience with policies - although, on the other hand, this experience is based on only one piece of evidence at a time. Also, conditions tend to be more similar within regions than outside them. This could well imply that evidence coming from neighbors is more informative. Note that this may result in deviant behavior from the narrow concept of rationality, even though such behavior can be judged as reasonable. Posteriors are conditional on observed results but results may be conditional on certain structural conditions that mediate between policy and results. The learning model can be extended to account for the fact that a government learns more the less it judges the difference between its own structural conditions and the structural conditions of the country whose experience is under analyses. The argument that governments exercise a geographic discrimination of information is appealing and tested empirically in the illustration below.

Applied studies have also detected other cognitive biases that could explain deviant behavior from the rational calculation Bayes rule implies: information is treated in an asymmetric way depending on whether it reinforces or contradicts own ideas. Besides, it seems that posterior beliefs formed on negative evidence tend to be given more weight in decision making in comparison to that formed on positive one (Kyburg, 1997; Hacking, 1997; March and Olsen, op. cit.).

Even if beliefs are revised rationally, the assessment of ideas following that revision is problematic. Imagine that the implementation of policy A is followed by very negative results. One government may think that this is a clear sign that the theory inspiring policy was wrong, and abandon it. But this is only one possible line of

reasoning. It may think that ideas were right but badly implemented or insufficiently sustained. It could also think that some exogenous and temporal factor is to be blamed for bad outcomes, but not the theoretical foundations of policy. And it may also be the case that ideas themselves embody predictions of bad results, as it is the case in "radical surgery" and "bitter pill" policy prescriptions. If this is the case, actual bad outcomes would match prior beliefs. As a result, policy decisions would remain unaltered even if rational updating takes place.

In other cases, the content of policy is not revised until its performance reaches some intolerable level. Even if the utility associated to a certain policy is less than that of an alternative, a government may stick to it until the downward revision crosses a certain threshold. This may be the case if governments evaluate gains and losses with regard to some reference point instead of using absolute levels of utility. Such behavior is suggested in Kahneman and Tversky's (1997) Prospect Theory. These authors emphasize that the framing of the choice may determine its outcomes. Thus, if a decision is framed in a way that sets governments in the domain of losses, we may witness the abandonment of the policy responsible for bad results and the adoption of risky policies. In essence, this theory suggests that bad performance can coexist with continuity in choices as long as "bad" does not become "too bad". Weyland (1996, 1998) has adopted this approach to explain the introduction of radical reforms in the 1990s in Latin America, Africa and Eastern Europe.

In short, even if the experience all governments are confronted with is the same, and they process it according to Bayes rule, ideas may dominate the interpretation of posterior beliefs instead of being the other way round. If this is the case, governments may choose a policy that does not maximize their expected utility at all times.

In extreme cases, we may witness that governments persist in their choices even if they are convinced that they contain logical flaws. Moreover, they can simply ignore experience. Within a Bayesian framework, this will happen if governments are very certain in their prior beliefs[6]. If that is the case, there is little that experience can add to those initial priors.

[6] This would amount to specify prior beliefs with very low initial variance.

In sum, strict rational learning is only one possible way of analyzing experience. Governments may process available evidence in different ways. They may also draw different conclusions about ideas on the basis of learning. Finally, they can ignore experience altogether. We sometimes may witness "irrational" behavior that, however, has strong foundations. Rational learning can be accommodated to include it. Of course, the reasons for deviant behavior multiply if factors other than cognitive ones are introduced in the analysis. Think, for instance, about a more complex utility function dependent on economic but also political results of policies. This scenario is not considered here, although it is obviously crucial. For the time being, I assume that governments and voters agree on their evaluation of policies and that there are no political costs attached to a change in policy.

Even if there are reasons to think that governments deviate from rational learning more often than not, this is an open empirical question. Next section explores it.

IV. Illustration. Learning and IMF Agreements.

In this section, a Bayesian model of learning is used to explain the decision of governments to enter into agreements with the International Monetary Fund (IMF). The respective status of being under and not under an IMF agreement is associated with observed outcomes of growth attached to each of them. Governments rationally update their prior beliefs in the view of that information. Those rational posteriors are then taken as independent variables to explain the actual path of observed choices. This provides an assessment of the capacity of the model of rational learning and rational choice to explain actual choices.

The decision to enter into agreements with the IMF may not be the most suitable to test the potential of the learning variable. Beliefs about the outcomes of agreements are not likely to exhibit the endurance that it is possible to attach to other policies, say, privatization or export orientation. This is because, in general, IMF agreements are not "normal economics". Privatization or export orientation are likely to be more in accordance with the behavior Bayesian learning prescribes. Once adopted, these policies

are likely to persist. However, the decision to sign agreements shows a more volatile trend, even if a closer look at the data suggests that governments entering into agreements tend either to sign agreements that extend over more than one period or to sign consecutive agreements[7].

Imagine governments are trying to decide whether to sign an agreement with the IMF or not. Let us assume that the outcomes of the two policy choices follow a normal distribution with mean *Mj* and variance *Vj*, both unknown. The initial priors and general setting is as in (1) but with *j={U,N}*, where U stands for being under an agreement and N for not being under it. During t, governments observe the average performance of both policies, $\bar{x}_j$, and the variability of results under each of them, S_j. The updating proceeds as explained in section 3.1. This updating gives every government posteriors for μ_j and for s_j that become priors in period t+1. Also recall that governments compare and choose the policy that maximizes utility.

The analysis is based on the ACGLPV World Database (Alvarez et. al., 1997). The study comprises 135 countries starting in 1960 with a total of 3623 time-year observations. Seventy countries enter into the database later. Rates of growth and variability of results under and not under an agreement for the world and in the previous year to each country's entry were used as prior beliefs (see Appendix III). Using this available information reduces somewhat the arbitrariness that characterizes the specification of priors. Those priors convey considerable uncertainty at all times[8].

To account for the possibility of a geographic discrimination of information, experience with policies was structured at three levels: the own experience, the experience in the region a country belongs to[9] and the experience of the world excluding

[7] I do not give here background on IMF activities, history or policies. A good account can be found in Vreeland (1999), who also explores the explanatory capacity of several variables for governments' and IMF's decision to sign agreements. Among his relevant variables, it is interesting that the number of other countries under an agreement is relevant in the decision of governments to enter into agreements. This is the only one significant to remain under them. An interesting account of the effects of IMF agreements on growth can be found in Przeworski and Vreeland (1999)

[8] A good study with informative prior beliefs is Jackman and Western (op. cit.)

[9] Those regions are Sub-Saharan Africa (n=1176), South Asia (n=166), East Asia (n=98), South East Asia (n=185), Pacific Islands/Oceania (n=63), Middle East/North Africa

the country and the region. Governments update their beliefs about each policy using three sources of information and they compare the resulting posteriors under and not under an agreement. Some caveat is required regarding the shape of the posterior beliefs based on the own experience. Since only one observation is gathered at a time, it takes more periods to consolidate beliefs on the basis of it.

As suggested above, the fact that a government that enters into an agreement tends to remain under it for more than one period suggests that choices are not independent overtime. Only 14 agreements out of 942 happen in isolation. Due to this dependency, a dynamic probit model was estimated (Amemiya, 1985). The details of the model are presented in Appendix II.

The dependent variable is *UNDER*, a dichotomous variable that takes value 1 if country i is under an agreement at time t and 0 otherwise. The first six independent variables are lagged values of the regressors and they inform about the probability of a transition from not being under to being under an agreement, p_{NU}. *R* is a variable that reflects past status. The coefficients that result from adding up coefficients 2-7 to coefficients 9-14 inform about the probability of continuing under an IMF agreement, p_{UU}. Regressors are the differences in updated beliefs about expected average results and expected variability of results being under and not being under an IMF agreement. *I* stands for own experience, *R* stands for regional experience and *W* for experience in the world. *POSMIL* and *POSDIL* refer to those posterior differences out of the own experience. *POSMRL* and *POSDRL* are differences in updated beliefs about average results and dispersion of results based on the experience in the region a country belongs to. Finally, *POSMWL* and *POSDWL* are differences in posteriors formed on the basis of the experience in the world. All variables are lagged one period. The results of the Dynamic Probit estimation are shown in table 1.

$$POSMIL = \mu^{U}_{I} - \mu^{N}_{I}$$

$$POSDIL = s^{U}_{I} - s^{N}_{I}$$

(n=323), Latin America (n=558), Caribbean and non Iberic America (n=177), Eastern Europe/Soviet Union (n=186) and Industrial Countries (n=739)

$POSMRL = \mu^{U}_{R} - \mu^{N}_{R}$

$POSDRL = s^{U}_{R} - s^{N}_{R}$

$POSMWL = \mu^{U}_{W} - \mu^{N}_{W}$

$POSDWL = s^{U}_{W} - s^{N}_{W}$

Table 1. Results of a Dynamic Probit Test.

Variable	Coefficient	Z=b/St. Er	Mean
1. Constant	-1.20	-11.19	
2. POSMIL	0.026	3.17**	-1.794
3. POSDIL	-0.022	-1.02	-0.383
4. POSMRL	-0.009	-0.42	-1.146
5. POSDRL	0.055	1.97**	0.243
6. POSMWL	0.232	2.76**	-1.15
7. POSDWL	0.155	2.99**	0.186
8. R	1.93	10.6**	0.256
9. POSMILR	-0.031	-2.22	-0.394
10. POSDILR	-0.205	-0.61	-0.088
11. POSMRLR	-0.074	-1.90	-0.357
12. POSDRLR	0.025	0.495	0.092
13. POSMWLR	-0.23	-1.657	-0.287
14. POSDWLR	-0.012	-0.139	0.092

** 95% Confidence level

	Predicted		
Actual	**0**	**1**	**Total**
0	2406	159	2565
1	189	734	923
Total	2595	893	3488

The model is significant as a whole and it exhibits a good predictive capacity. According to these results, the probability to go under an agreement is positively related to the posterior difference in average results under and not under an agreement derived from the own experience *(POSMIL).* In other words, the bigger this difference is, the higher the probability to sign an agreement with the IMF. This same result applies to the world experience: the bigger the difference between results under and not under an agreement in the world *(POSMWL*), the higher the probability to enter into agreements with the IMF. Interestingly, governments are prone to take risks when updated differences

in the variability of results as observed in the region and in the world are big *(POSDRL, POSDWL).* The higher this difference in dispersions, the more likely governments are to enter into agreements with the IMF. Finally, the lagged status, R, is very significant to predict actual status.

These results suggest that governments are more likely to enter into agreements with the IMF when they have learnt out of the own and the world experience that agreements yield better average results. Also, the higher the posterior difference in expected variability, the more likely governments are to enter into IMF agreements. Hence, even if some governments perform very bad under IMF agreements, the existence of very good performers in the region and in the world under that same status may induce governments to take risks and change policy. Note that this result is in accordance with the fact that switches tend to happen in waves, not only when the own policy performs bad but when there is evidence substantiated in a good number of cases about what alternative policies may yield.

The coefficients below are the result of adding the coefficients of the lagged variables *(L)* to the coefficients of the lagged variables times lagged status *(LR).* These coefficients inform about the probability of remaining under an agreement having been under in the immediate past, p_{UU}. When it comes to explain continuities, regional experience results significant. The bigger the difference in updated average results under and not under an agreement in the region, the less likely governments are to remain under an agreement. Although this may appear counterintuive, an explanation for such behavior could be that governments abandon agreements as soon as growth resumes since agreements with the IMF are not politically cost-free. This abandonment of agreements can push other governments to do the same, simply because remaining under IMF agreements in isolation has, as Vreeland (op. cit.) has shown, high sovereignty costs. Note that seeing very good performers in the world under IMF agreements also increases the probability of remaining under them.

Variable	**Coefficient**	**Z**
POSMIL+ POSMILR	-0.004	-0.42

POSDIL+ POSDILR	-0.042	-1.66
POSMRL+ POSMRLR	-0.084	-2.66**
POSDRL+ POSDRLR	0.081	1.89
POSMWL+ POSMWLR	0.002	0.02
POSDWL+ POSDWLR	0.143	1.97**

Finally, I provide the average results for all transition probabilities and for the long term probability to be under an agreement, p_{U^*} , as predicted by the model.

Probabilities	**Mean**
p_{UN}	0.19
p_{UU}	0.80
P_{NN}	0.92
p_{NU}	0.07
p_{U^*}	0.29

The model predicts 29% of agreements in the long term, which is not far from the 26% of actual agreements in the database.

To sum up, the model of rational learning tells us that governments are more likely to enter into agreements with the IMF if their own results and the world results under them are comparatively better. They are also likely to enter into IMF agreements if out of the regional and the world experience they learn that results under IMF vary more than not being under them. This suggests that the view of some very good performers in the region and the world can induce a change. In other words, the view of miracles can make governments risk prone. Finally, governments are less likely to continue under an agreement if updated beliefs out of the regional experience tell them that growth is higher under than not under an agreement. A possible explanation for this counterintuive behavior could be that governments decide to leave agreements when they know that others in their region will do, since remaining under IMF agreements in isolation has high sovereignty costs. In other words, governments abandon emergency policies as soon as growth resumes. Note also that miracles in the world are not only relevant in the decision to enter into agreements with the IMF but they also matter in the decision to remain under them.

IV.1. Sensitivity Analysis

As mentioned before, the main bone of contention between Bayesian and Classicals lies in the introduction of subjective beliefs into the analysis. To what extent are priors driving the conclusions? Bayesians argue that as long as evidence is abundant, this concern should not be important. In practice, Bayesians usually overcome this problem by using flat priors. With flat priors, data rapidly overwhelms initial beliefs.

In the empirical results shown above, governments' priors have been modeled as considerably vague but one would like to know what the results would be if governments' prior beliefs were modeled as more dogmatic, that is, as conveying much less uncertainty. Also, what would happen if governments' prior beliefs were even more uncertain than in the illustration above? As long as there is enough data, these questions would be irrelevant, and results would be robust regardless of the initial priors.

There are reasons to be concerned about the robustness of the results out of the own experience, since posteriors in this case are based, at best, in just one observation. The series *UNDER* for the own experience is especially troublesome due to the high amount of missing values. According to equation (2) and (3), the posterior equals the prior when data is not available (n = 0). In practical terms this implies that countries do not update their beliefs out of the own experience for long periods of time. It is dubious that the influence of priors vanishes when this is the case. This concern should not affect the results out of the regional and the world experience since updating in those cases is based on a good number of observations.

I give below the results of a Dynamic Probit Test in which all governments start with the 1960 priors for under and not under an agreement. The variance of the two parameters, mean and variance, is however increased ten times. Thus, governments' beliefs are modeled as conveying a very important uncertainty (see Appendix III). The table shows that results, especially the ones based on big sample sizes, are robust. All variables that were significant before go on being significant under the new specification without observing changes in the signs. The only novelty is that miracles in the region are

now significant to explain continuities under IMF agreements. This variable was significant before at a 10% level. Hence, increasing governments' uncertainty in their prior beliefs hardly changes substantive results. Some more attention is paid to what the experience in the region can teach, though. Finally, the results regarding the series based on the own experience under an IMF agreement should be taken with a grain of salt taking into account the fact that experience, and thus updating, is comparatively infrequent in this case.

Table 1. Dynamic Probit Test with very Diffuse Priors.

Variable	Coefficient	z=b/St. Er	Mean
1. Constant	-1.19	-12.36	
2. POSMIL	0.019	2.84**	-1.54
3. POSDIL	0.006	0.41	-0.73
4. POSMRL	-0.007	-0.358	-1.04
5. POSDRL	0.063	2.83**	0.24
6. POSMWL	0.249	3.19**	-1.13
7. POSDWL	0.155	3.29**	0.19
8. R	1.95	11.97**	0.25
9. . POSMIL+ POSMILR	-0.003	-0.39	
10. POSDIL +POSDILR	-0.03	-1.92	
11. POSMRL +POSMRLR	-0.086	-2.99**	
12. POSDRL +POSDRLR	0.087	2.48**	
13.POSMWL+POSMWLR	0.046	0.45	
14. POSDWL+POSDWLR	0.149	2.22**	

** 95% Confidence level

I also explored sensitivity to the data by introducing dummy variables for each region. The summary is given in Appendix III. There are three dummies that are significant and two of them change the qualitative results obtained above. The three dummies that are significant are South Asia, Latin America and Industrial countries. The first two show the highest percentages of agreements in the database, 47% and 49% respectively. In both cases the dummies are significant and positive. The dummy for the Industrial Countries is significant and negative. Actually, there are only 48 agreements in this region out of 739 country-year observations.

While the dummy for South Asia does not change the qualitative results, the dummies for Latin America and the Industrial Countries alter them slightly. All variables that were significant in the first model (table 1) go on being significant albeit some of them at a 10\% significance level. However, the view of very good performers in the world under IMF agreements stops being significant in the presence of these dummies.

V. Conclusions

The argument that learning is an important variable in explaining the evolution of policy choices over time is well established. Testing this argument is, however, a difficult task given that the concept of learning remains elusive. A Bayesian approach provides a means to tackle this issue.

I have first discussed some conceptual problems regarding the role of ideas and learning in policy choices. I have presented the model of Bayesian learning and I have discussed its limits. The main conclusion of this critical review is that the narrow concept of Bayesian rationality in which governments make a positive evaluations of policies that yield good results and abandon the policies that perform bad is an inaccurate account of reality. Updating can be enlarged so that different geographic and temporal patterns of updating are allowed. By doing so, it is possible to characterize as rational behavior one that would not be so under an strict definition of Bayesian rationality.

A Bayesian updating model has been tested with regard to governments' decision to sign agreements with the IMF. In this model, governments start having some priors beliefs about expected average results and expected variability of results under each status. After observing their own experience, the experience in the region and in the world, governments update those beliefs according to Bayes rule and compare those posteriors. The differences in posterior beliefs have been used as independent variables to explain the actual path of choices in relation to IMF agreements.

The model suggests that governments are more likely to enter into agreements with the IMF when they have learnt from their own experience and the experience in the world that agreements produce better average results. Also, and interestingly, it seems that the view of miracles under IMF agreements in the region a country belongs to and in

the world can induce governments to take risks and change policy. Very good performers in the world are also relevant in governments' decision to remain under an IMF agreement. However, when average growth under IMF agreements is better than not under them in the region, governments are less likely to remain under agreements. I have suggested the abandonment of agreements as soon as growth resumes as a possible explanation.

Although these results are interesting in themselves, adjudicating between strict Bayes rationality and the rival hypothesis to explain actual choices (section 2) will require a more careful study of the conditions under which governments actually change policy. Also, the fact that Bayesian learners are difficult to surprise after a while questions whether Bayesian learning is not better suited to explain continuities rather than changes. Variations in the updating rule to include discounting of the past seems to be a reasonable next step.

Appendix I

Conjugate Families for Samples from a Normal Distribution. Sampling from a Normal Distribution with Unknown mean and Unknown Precision.

Based on De Groot (1970), Gelman et. al. (1995), Lee (1997) and Zellner (1997). Proofs available on the texts.

Suppose growth is a random variable that follows a normal distribution with an unknown value of the mean, μ, and an unknown value of the variance σ^2. Suppose that their prior joint conjugate distribution is as follows: the conditional distribution of μ given σ^2 is a normal distribution. The marginal distribution of σ^2 is scaled inverse-χ^2. With this specification, the marginal distribution of μ follows a t-Student distribution.

Thus,

$$\mu|\sigma^2 \sim N(\mu_0, \sigma_0^2/\tau_0)$$
$$\sigma^2 \sim Inv\text{-}\chi^2(\nu_0, \sigma_0^2)$$
or
$$(\mu|\sigma^2, \sigma^2) \sim N\text{- }Inv\text{-}\chi^2\ (\mu_0, \sigma_0^2/\tau_0;\ \nu_0, \sigma_0^2)$$

The parameters are the location and the scale of μ and the degrees of freedom and scale of σ^2 respectively. Note that this specification implies that μ and σ^2 are dependent in their prior specification. If σ^2 is large, a high variance prior distribution is induced for

μ. Prior beliefs about μ are calibrated by the scale of measurement of X and is equivalent to τ_0 prior measurements on this scale (Gelman, et. al, p. 71).

Suppose now that a sample, $x_{n,}$ of n i.i.d observations on growth also normally distributed is gathered.

1. The joint posterior distribution, $p(\mu, \sigma^2 | x_n)$.

The posterior parameters for the location and scale of the mean and the degrees of freedom and scale of the variance are as follows:

(5) $\mu_n = (\tau_0/\tau_n)\,\mu_0 + (n/\tau_n)\,\bar{x} = \rho\,\mu_0 + (1-\rho)\,\bar{x}_n;\ 0<\rho<1$

(6) $\tau_n = \tau_0 + n$

(7) $\nu_n = \nu_0 + n$

(8) $S_n = S_0 + \boldsymbol{S_t} + [\tau_0 n\,(\bar{x}_n - \mu_0)^2]/\tau_n$

(9) $\sigma_n^2 = S_n/\nu_n$

where ν_0 are the prior degrees of freedom, S_0 is the prior sum of squares and $\boldsymbol{S_t}$ is the sample sum of squares.

2. The Marginal Posterior Distribution of σ^2, $p(\sigma^2 | x_n)$

$\sigma^2 | x_n \sim Inv\text{-}\chi^2(\nu_n, \sigma_n^2)$

with ν_n and σ_n^2 as in (7) and (9).

3. The Conditional Posterior Distribution ofμ , $p(\mu|\sigma^2, x_n)$

$\mu|\sigma^2, x_n \sim N(\mu_n, \sigma^2/\tau_n)$

with μ_n,τ_n as in (5) and (6). One normal way to proceed to sample from the joint posterior distribution is to draw σ^2 from its marginal posterior distribution as in (6) and then draw μ from its normal posterior distribution, using the simulated value σ^2.

4. The Marginal Posterior Distribution of μ , $p(\mu|x_n)$

$\mu|x \sim t_{\nu n}\ (\mu_n,\sigma_n^2/\tau_n)$

with ν_n, μ_n,σ_n^2 and τ_n as in (7), (5), (9) and (6) above.

5. Specifying the prior parameters.

Since σ^2 follows an Inv-χ^2, the following formulas apply.

(10) $E(\sigma^2) = S_0/(\nu_0-2)$

(11) $Var(\sigma^2) = 2S_0^2/[(\nu_0-2)^2(\nu_0-4)]$

Thus, after specifying values for the mean of the variance and the variance of the variance, prior values for S and ν can be obtained solving those equations. Also, since μ marginally follows a t-Student distribution

(12) $E(\mu) = \mu_0$

(13) $Var(\mu) = S_0/(\nu_0\tau_0)$

From which τ_0 can be obtained after specifying the variance of the mean and having obtained S_0 and ν_0.

Appendix II

A Dynamic Probit Model

When it is considered that the decision taken by country i at time t is related to the decision that same country took at time t-1, the model to be used is a dynamic probit model (discrete state, discrete time model or Markov model. See Amemiya, 1985)

Let $S_{i,t-1}$ denote participation status of country i at time t-1. That status can be "Under" if country i was under an IMF agreement at time t-1 ($U_{i,t-1}$). Alternatively, it can be "Not Under" if country i was not under an IMF agreement at time t-1 ($N_{i,t-1}$). $U_{i,t-1}$ is equal to 1 if country i was under at time t-1 and 0 otherwise. Similarly, $N_{i,t-1}$ has value 1 if country i was not under at time t-1 and 0 otherwise.

The general specification is

$$\begin{bmatrix} p(U_{it} \mid S_{i,t-1}) \\ p(N_{it} \mid S_{i,t-1}) \end{bmatrix} = \begin{bmatrix} p_{UU,it} & p_{NU,it} \\ p_{UN,it} & p_{NN,it} \end{bmatrix} \begin{bmatrix} U_{i,t-1} \\ N_{i,t-1} \end{bmatrix}$$

Where participation status at time t conditional on past status - left hand side - is made equal to a transition probability matrix times lagged participation status. The transition probability matrix contains the following information: $p_{UU,it}$ denotes the probability that country i stays under an agreement at time t while $p_{UN,it}=1-p_{UU,it}$ denotes the probability that participation in an agreement ends at time t (that is, it goes from "being under" an agreement at time t-1 to "not being under" at time t). Similarly, $p_{NU,it}$ denotes de probability that country i enters an agreement at time t. The probability that country i goes on not being under an agreement at time t is $p_{NN,it}=1-p_{NU,it}$

Under this setting, the probability of signing an agreement at time t is the following

$$P(U_{it}|S_{i,t-1}) = p_{UU,it}U_{i,t-1} + p_{NU,it}N_{i,t-1} = p_{NU,it} + (p_{UU,it} - p_{NU,it})\, U_{i,t-1} \quad (1)$$

The same goes for $P(N_{it}|S_{i,t-1})$

In a Univariate Dynamic Probit setting, there is a theory on transitions and on continuities. Transitions and continuities are a function of the same set of lagged regressors. In other words, the same theory is used to explain both phenomena. In this paper,

$$p_{NU,it} = F[\beta'(\mu_{it-1}+s_{it-1})] = F\{\beta' [(\mu^{U}_{it-1}-\mu^{N}_{it-1}) + (s^{U}_{it-1}- s^{N}_{it-1})]\} \quad (2)$$

$$p_{UU,it} = F[\gamma'(\mu_{it-1}+s_{it-1})] = F\{\gamma' [(\mu^{U}_{it-1}-\mu^{N}_{it-1}) + (s^{U}_{it-1}- s^{N}_{it-1})]\} \quad (3)$$

where F(·) is the CDF of the standard normal distribution.

For convenience, we set $\gamma = \alpha+\beta$. Then, even if the explanatory theory is the same, its impact on probabilities differs as reflected in different coefficients.

$$p_{UU,it} = F(\gamma' X_{i,t-1}) = F[(\alpha+\beta)' (\mu_{it-1}+s_{it-1})] = F(\alpha' (\mu_{it-1}+s_{it-1})] + \beta' (\mu_{it-1}+s_{it-1})]) \quad (4)$$

Using (2) and (4) in (1) and rearranging terms, we get the following

$$P(U_{it}|S_{i,t-1}) = p_{NU,it} + (p_{UU,it}- p_{NU,it}) U_{i,t-1} = F(\beta' (\mu_{i,t-1}+s_{i,t-1})]) + [F(\alpha' (\mu_{i,t-1}+s_{i,t-1})])] U_{i,t-1}$$

Appendix III

Data arrangements and Sensitivity Analysis

Data Arrangements

Prior to statistical analysis, the database was rearranged so that for each country i and year t there were six pieces of information: the country's own experience under and not under an agreement, the average experience in the region a country belongs to under and not under an agreement and the experience in the rest of the world excluding the own experience and the regional experience. The same goes for the observed variability of results.

What follows is an illustration of how to get prior parameters out of the prior specification. In year 1959, average growth in the world being under IMF agreements was -.95 and the variability of those results were 14 (included the country, the region and the

rest of the world). The following specification was used: E(M) = -0.95, Var(M) = 14, E(V) = 14 and Var(V) = 56. Using equations (10) and (11) the prior degrees of freedom, ν_0 are 11 and the prior sum of squares, S_0, is 126. Using (13), the value for $\tau_0 =$ 0.8 is obtained. Of course, μ_0 =-0.95. Thus, since the mean follows marginally a t-Student distribution with 11 degrees of freedom in this prior specification, a prior 95% confident interval for the mean is [-0.95+-√14*2.20]=[-9.178, 8.22], which conveys a considerable amount of uncertainty.

Initial priors are based on observed average outcomes and variance of results in the world prior to the year in which some entry in the database is observed. Note the world include all available observations in a certain year. For instance, countries entering in year 1960 use as prior beliefs average results and variance of results under and not under an agreement in year 1959. Those entering in year 1961 use 1960 data and the same goes for the rest of countries and years. As for the variance, its expected value was taken to be the one observed each year. The variance of the variance was taken to be four times its expected value. For instance, in year 1960 the prior for the variance under an agreement had the following mean and the following variance [14,56]. As for not being under an agreement, it was [20,80].

Sensitivity Analysis

Regarding the sensitivity analysis to priors, all countries start with the 1960 priors variances multiplied by ten. The priors for the mean and the variance under an agreement are [*x*, 140] and [14, 560] respectively. For not under an agreement, the priors were made equal to [*x*, 200] and [20, 800]. Thus, uncertainty was increased ten times. By *x* I mean that each country was allowed to start updating its own prior on average results as in table 3.

As for the data, the summary results after including dummies for each region is provided below.

Table 3. Sensitivity of Results to Data

Dummy	Significance	Change in Results
Sub-Saharan Africa	No	No
South Asia	**Yes**	No
East Asia	No	No
South East Asia	No	No
Oceania/Pacific Islands	No	No
Middle East/ North Africa	No	No
Latin America	**Yes**	**Yes**
Caribbean/ Non Iberic Am	No	No
Eastern Europe/Soviet U	No	No
Industrial Countries	**Yes**	**Yes**

Bibliography

Adler, E., and P. M. Haas. 1997. Conclusion: Epistemic Communities, World Order and the Creation of a Reflective Research Program. In Knowledge, Power and International Policy Coordination, ed. P. M.Haas, 367-90. Columbia: University of South Carolina Press.

Amemiya, T. 1985. Advanced Econometrics. Cambridge, Massachusetts: Harvard University Press.

Basu, K. 1993. Comment on "Two Strategies for Economic Development: Using Ideas and Producing Ideas" by Romer. Proceedings of the World Annual Bank Conference on Development Economics.The International Bank for Reconstruction and Development.

Bennet, C and M Howlett. 1992. The Lessons of Learning: Reconciling Theories of Policy Learning and Policy Change. Policy Sciences, 25: 275-294.

Berger, J. O. 1985. Statistical Decision Theory and Bayesian Analysis. second ed. New York: Springer-Verlag, 1980. Reprint. New York: Springer- Verlag.

Biersteker, T. J. 1995. The Triumph of Liberal Economic Ideas in the Developing World. In Global Change, Regional Response. The New International Context of Development, ed. B. Stallings, 174-96. Cambridge: Cambridge University Press.

Blyth, M. M. 1997. Any More Bright Ideas? The Ideational Turn of Comparative Political Economy. Comparative Politics 29(2): 229-50.

Colander, D., and A. W. Coats. 1993. The Spread of Economic Ideas. Cambridge: Cambridge University Press.

Coyle, R. G. 1972. Decision Analysis. London: Nelson.

Denzau, A and D. North. 1994. Shared Mental Models: Ideologies and Institutions. Kyklos. Vol 47: 3-31.

Gardenfors, P., and N. Sahlin. 1997a. Introduction: Bayesian Decision Theory Foundations and Problems. In Decision, Probability and Utility, eds. P. Gardenfors, and N. Sahlin, 1-181. 1997 ed. Cambridge: Cambridge University Press.

Gardenfors, P., and N. Sahlin. 1997b. Unreliable Probabilities, Risk Taking and Decision Making. In Decision, Probability and Uncertainty, eds. P. Gardenfors, and N. Sahlin, 313-34. 1997 ed. Cambridge: Cambridge University Press.}

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 1995. Bayesian Data Analysis. London: Chapman and Hall.

Gerber, A., and D. P. Green. July 1998. Rational Learning and Partisan Attitudes. American Journal of Political Science 42(3): 794-818.

Goldstein J and R Keohane (eds). 1993. Ideas and Foreign Policy. Beliefs, Institutions and Political Change. Ithaca: Cornell University Press.

De Groot, M. 1970. Optimal Statistical Decisions. New York: McGraw-Hill.

Haas, P. 1992. Introduction: Epistemic Communities and International Policy Coordination. International Organization 46(1): 1-35.

Haas, P. M., and E. B. Haas. 1995. Learning to Learn: Improving International Governance. Global Governance 1: 255-85.

Haas, P. M. 1997. Introduction: Epistemic Communities and International Policy Coordination. In Knowledge, Power and International Policy Coordination, ed. P. M. Haas, 1-35. Columbia: University of South Carolina Press.

Hacking, I. 1997. Slightly More Realistic Personal Probability. In Decision, Probability and Uncertainty, eds. P. Gardenfors, and N. Sahlin, 118 135. 1997 ed. Cambridge: Cambridge University Press.

Hall, P. 1986. Governing the Economy. The Politics of State Intervention in Britain and France. New York: Oxford University Press.

Hall, P. 1989. The Political Power of Economic Ideas. Keynesianism Across Countries. Princeton: Princeton University Press.

Hall, P. A. 1993. Policy Paradigms, Social Learning and the State. The Case of Economic Policy Making in Britain. Comparative Politics 25(3): 275-96.}

Harberger, A. 1993. Secrets of Success: A Handful of Heroes. The American Economic Review 83(2): 343-50.

Etheredge, L. 1981. Government Learning: An Overview in S. L. Long. ed., The Handbook of Political Behavior, Vol2, New York: Pergamon.

Heclo, H. 1974. Modern Social Politics in Britain and in Sweden. New Haven: Yale University press.

Hood, C. 1994. Explaining Economic Policy Reversals. Buckingham. Philadelphia: Open University Press.

Jacobsen, J. K. 1995. Much Ado About Ideas. The Cognitive Factor in Economic Policy. World Politics 47(January): 283-310.

Jeffrey, R. C. 1997. Probable Knowledge. In Decision, Probability and Uncertainty, eds. P. Gardenfors, and N. Sahlin, 86-96. 1997 ed Cambridge: Cambridge University Press.

Kahler, M. 1990. Orthodoxy and Its Alternatives: Explaining Approaches to Stabilization and Adjustment. In Economic Crisis and Policy Choice, ed. J. Nelson, 33-62. Princeton: Princeton University Press.

Kahneman, D., and A. Tversky. 1997. Prospect Theory: An Analysis of Decision Under Risk. In Decision, Probability and Uncertainty, eds. P. Gardenfors, and N. Sahlin, 183-214. 1997 ed. Cambridge: Cambridge Univeristy Press.

Krueger, A. O. March 1997. Trade Policy and Economic Development: How We Learn? The American Economic Review 87(1): 1-22.

Kyburg, H. E. 1997. Bets and Beliefs. In Decision, Probability and Uncertainty, eds. P. Gardenfors, and N. Sahlin, 101-17. 1997 ed. Cambridge: Cambridge University Press.

Leamer, E. E. 1991. A Bayesian Perspective on Inference From Macroeconomic Data. Scandinavian Journal of Economics 93(2): 225-48.

Lee, P. M. 1997. Bayesian Statistics. An Introduction. Second ed. London: Arnold.

Lordon, F. 1998. The Logic and the Limits of Desinflation Competitive. Oxford Reviw of Economic Policy, vol 14(1): 96-113.

March, J. G. 1978. Bounded Rationality, Ambiguity, and the Engineering of Choice. Bell Journal of Economics 9: 587-608.

March, J. G., and J. P. Olsen. 1989. The Uncertainty of the Past: Organizational Learning Under Ambiguity. In Decisions and Organizations, ed. J. G. March, 335-58. Oxford: Basil Blackwell.

Meseguer, C. 1998. "Decisiones e Incertidumbre. Discusion Metodologica y Una Aplicacion.".Ms

Nelson, J. e. 1990. Economic Crisis and Policy Choice. Princeton: Princeton University Press.

North C, D. 1994. Economic Performance Through Time. The American Economic Review (June): 359-68.

Odell, J. S. 1982. U.S International Monetary Policy. Markets, Power and Ideas As Sources of Change. Princeton: Princeton University Press.

Pericchi, L. R. n.d. "Analisis De Decision, Inferencia y Prediccion Estadistica Bayesiana.". Ms.

Pierson, P. 1993. When Effect Becomes Cause. Policy Feedback and Political Change. World Politics 45(July): 595-628.

Przeworski, A. 1993. "Intertemporal Politics. The Support for Economic Reforms in Poland.". Ms.

Przeworski, A., and J.R. Vreeland. 1999. "The Effect of IMF Programs on Economic Growth". Ms.

Quattrone, G. A., and A. Tversky. 1988. Contrasting Rational and Psychological Analysis of Political Choice. American Political Science Review 82(3): 719-36.

Raiffa, H. 1972. Decision Analysis. Introductory Lectures on Choices Under Uncertainty. Reading: Addison-Wesley.

Robinson, J. 1997. Theories of Bad Policy. Policy Reform, Vol 1:1-44.

Rodrick, D. 1996. Understanding Economic Policy Reform. Journal of Economic Literature 34(March): 9-41.

Rose, R. 1991. What Is Lesson-Drawing? Journal of Public Policy 2(1): 3-30.

Sabatier, P. A., and H. C. Jenkins-Smith. 1993. Policy Change and Learning. An Advocacy Coalition Approach. Boulder: Westview Press.

Stallings, B. 1990. International Influence on Economic Policy: Debt, Stabilization and Structural Reform. In The Politics of Economic Adjustment, eds. S. Haggard, and R. Kaufman, 41-88. Princeton: Princeton University Press.

Stiglitz, J. 1999. Knowledge for Development: Economic Science, Economic Policy and Economic Advice. ABCDE. Website //www.worldbank.org.

Stokes, S. 1996a. Public Opinion and Market Reforms: The Limits of Economic Voting. Comparative Political Studies 29(5): 499-519.

Tinbergen, J. 1984. Development Cooperation As a Learning Process. In Pioneers in Developments, eds. G. M. Meier, and D. Seers, 315-31. Oxford: Oxford University Press.

Valdes, G. 1995. Pinochet's Economists. Cambridge: Cambridge University Press.

Vreeland, J. R. 1999. "The IMF: Lender of Last Resort Or Scapegoat.". Ms.

Weir, M., and T. Skocpol. 1985. State Structures and the Possibilities for the Keynesian Responses to the Great Depression in Sweden, Britain and the United States. In Bringing the State Back In, eds. P. B. Evans, D. Rueschmeyer, and T. Skocpol, 107-63. Cambridge: Cambridge University Press.

West M and J. P. Harrison .1997. Dynamic Bayesian Forecasting and Dynamic Models. New York: Springer Verlag.

Western, B.1998. Causal Heterogeneity in Comparative Research: A Bayesian Hierarchical Modelling Apprach. American Journal of Political Science 42(4): 1233-59.

Western, B., and S. Jackman. 1994. Bayesian Inference for Comparative Research. American Political Science Review 88(2): 412-23.

Weyland, K. 1996. Risk Taking in Latin America Economic Restructuring: Lessons From Prospect Theory. International Studies Quarterly 40: 185-208.

Weyland, K. 1998. The Political Fate of Market Reform in Latin America, Africa and Eastern Europe. International Studies Quarterly, 42: 645-674.

1988. Why Economic Policies Change Course. Paris: OCDE.}

Williamson, J. 1993. Democracy and the "Washington Consensus". World Development 21(8): 1329-1280.

Williamson, J., and S. Haggard. 1994. The Political Conditions for Economic Reform. In The Political Economy of Policy Reform, ed. J. Williamson, 525-96. Washington: Institute for International Economics.}

Winkler, R. L. 1972. An Introduction to Bayesian Inference and Decision. New York: Holt.

Woods, N. 1995. Economic Ideas and International Relations: Beyond Rational Neglect. International Studies Quaterly 39: 161-80.

Yee, A. S. 1996. The Causal Effects of Ideas on Policies. International Organization 50(1): 69-108.