LEARNING AND MODEL VALIDATION

IN-KOO CHO AND KENNETH KASA

Abstract. This paper proposes a selection criterion for adaptive learning models. Instead of assuming that agents revise beliefs about a fixed model, we allow agents to test the specification of their models, and to select new ones that appear to strike a better balance between fit and complexity, as measured by estimated Kullback-Leibler Information Criteria. This combined process of model revision and model selection is called *validation dynamics*. We prove that as the agent validates the model more frequently, the validation dynamics converges to a single model, which the agent uses almost always. This model is called the *dominant* recursive learning model, which we can characterize as a recursive model with the largest rate function, using the parlance of the large deviation theory. We illustrate the concept of validation dynamics using examples from Sargent (1999) and Evans and Honkapohja (2001). Possible extensions of the analysis to robust testing and selection are also discussed.

JEL Classification Numbers: C120, E590

1. Introduction

Recursive learning models are effective tools for selecting and computing Rational Expectations Equilibria. They can also be used to explain certain kinds of dynamics that are difficult to explain using rational expectations models (e.g., Bray and Savin (1986), Marcet and Sargent (1989), Sargent (1993), Evans and Honkapohja (2001) and Cho, Williams, and Sargent (2002)). A typical learning model starts by endowing a decision maker with a specific parametric model, often called the Perceived Law of Motion, and a specific algorithm for recursively estimating its parameters. This recursive learning algorithm specifies how the agent perceives the underlying dynamics of his environment, and how the agent's beliefs evolve over time in response to new information. Because selection and computation of an equilibrium relies on asymptotic stability of the learning algorithm, it is important to select a sensible learning algorithm when modeling the behavior of boundedly rational agents.

Unfortunately, the existing literature does not address the question of how agents actually select their models. In fact, agents in these models are not permitted to express any doubts about the validity of their models, even if observed data would support an obvious alternative. In this paper, we develop a new approach to adaptive learning, which offers a selection criterion for sensible learning algorithms. Instead of imposing a particular model

Date: August, 2006.

We thank Lars Hansen, Seppo Honkapohja, Albert Marcet, and Tom Sargent for helpful discussions. Financial support from the National Science Foundation (ECS-0523620) is gratefully acknowledged. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

on the agent, we assume that the agent has a *menu of models*, from which he selects the best possible approximation of the unknown true data generating process. We impart a form of bounded rationality to agents by assuming they are endowed with a class of "simple" models, such as a linear models with Gaussian perturbations.

Following the traditional learning literature, we view agents as econometricians, who attempt to fit their models to observed data by recursively estimating their parameters. In contrast to the existing learning literature, however, we assume agents are aware that models are only approximations, and can therefore be misspecified. Thus, agents run specification tests to ensure that their models provide good approximations of the true data generating process. Good learning models should survive repeated specification testing but also, should be selected over the alternative models, if the present model fails the specification test.

The combination of the estimation and testing procedure will be called *model validation*. If a model is validated, agents continue to use the existing functional form, while recursively updating its parameters to better fit the data. On the other hand, if a model is invalidated, agents select a new better fitting model before making current policy decisions.

Econometricians have made substantial progress during the past two decades on the problem of testing and comparing misspecified models.[1] A unifying theme in this work is to view models as inducing probability distributions over the data, and to then compare models using information-theoretic measures of the difference between two probability distributions. Most recent work on comparing misspecified models is based on the Kullback-Leibler Information Criterion (KLIC), which measures the relative entropy between two probability distributions (Zeitouni and Gutman (1991a) and Zeitouni and Gutman (1991b)). Models with smaller estimated KLICs are preferred. A key result in this literature is that consistent estimates of the difference between two KLICs can be obtained *without* prior knowledge of the true data-generating process. As a result, models can be compared without having to assume that any of them are correctly specified. Basing model selection on relative entropy also has the virtue that it naturally delivers a measure of model complexity, thereby striking a balance between bias and variance, and allowing models with different dimensions to be sensibly compared.

The decision maker in our model exploits these recent developments. However, traditional econometric methods (for both estimation and inference) presume the data-generating process is exogenous. In our model this is not the case. Our agent actually uses his model to make decisions, and as a result, the data-generating process *responds to* the agent's own estimation and testing efforts. That is, the data are *endogenous*. When it comes to inference, this means there will be a difference between the nominal and actual sizes of our agent's test statistics. We assume this discrepancy goes undetected, which is an important aspect of the bounded rationality of our agent.

While one could conceivably formulate a variety of alternative validation processes, we adopt the following information-theoretic approach:

[1]A highly selected sample includes: White (1982), Vuong (1989), Hansen and Sargent (1993), Sin and White (1996), Kitamura (2001), Rivers and Vuong (2002), and Marcellino (1999). White (1994) and Burnham and Anderson (2002) contain textbook treatments.

(1) An agent is endowed with a collection of potentially misspecified linear Gaussian models, each with a finite number of unknown parameters.
(2) Each period the agent tests the specification of his model based on its estimated KLIC. The agent accepts the hypothesis that the present model is the best model if and only if its estimated KLIC does not exceed the minimum estimated KLIC by more than a pre-specified threshold, reflecting the importance of Type I errors.
(3) If the present model survives the test, then the model is updated and used to formulate a current policy function.
(4) If the difference exceeds the threshold, the current reference model is rejected, and a new model is selected, based on the model with the minimum estimated KLIC statistic.

We believe that this validation process reflects an important aspect of the actual behavior of policy makers, who often have access to a large pool of econometricians that continuously test and improve their models. However, the analysis of validation dynamics poses unique challenges. In particular, the evolution of the reference model is stochastic, and the duration of a reference model is a random variable.

Following the lead of Yin and Krishnamurthy (2005), we approximate the validation process by a system of ordinary differential equations (ODE), each associated with a particular model, along with a set of endogenous transition probabilities between these ODEs. Once a model is adopted, its parameters are estimated using a standard recursive algorithm, which can be approximated by ODE, following the standard technique (e.g., Kushner and Yin (1997)). However, because each model is continuously tested, there is a positive probability that the present model is replaced by another model from a different class. Our focus is on determining the model which is used most often in the limit.

We define a *dominant* recursive learning model to be the model that has the smallest asymptotic rejection probability. This turns out to be the model with the smallest probability that the outcome escapes from its stable self-confirming equilibrium. Such a model is most likely to survive repeated specification testing around the self-confirming equilibrium. Our main result is that validation dynamics converge to the dominant recursive learning model. This result provides a criterion to select a particular learning model, which typically entails a self-referential feature arising from the interaction between the decision and the belief formation of the agent.

After establishing the main convergence result, we examine two well known models as examples. The first example is based on a model from Evans and Honkapohja (2001). It shows that some learning models need not be resilient.

The second example is based on Sargent (1999) and Cho, Williams, and Sargent (2002). While the general theory says that the validation process converges to the dominant recursive learning model, it is difficult to characterize the dominant model without imposing additional structure on the problem. In this example, we show that the dynamics of simpler models can be represented as projections of the dynamics of larger models. Then, the dominant model is the smallest model that converges to the self-confirming equilibrium.

Our paper focuses on parametric model uncertainty in order to illuminate the key idea. We have essentially added a few unknown (discrete) parameters to a standard learning exercise. One route toward genuine model uncertainty would be to consider nonparametric

alternatives, as in Chen and White (1998). This can be problematic, however, given the sample sizes usually available to macroeconomists. Nonparametric methods are also ill-suited to situations where the model is being used to solve a control problem. An attractive alternative, better suited to a control context, has been proposed by Hansen and Sargent (2006b), whose robust control methods adopt a (constrained) worst-case approach to model uncertainty. Hansen and Sargent (2006a) extend these methods to allow filtering of underlying hidden states. Using discrete, time invariant, hidden states to index unknown models allows agents to respond to model uncertainty in a way that is robust to general forms of distributional misspecification. Hansen and Sargent (2006a) approach this problem from the perspective of model averaging, which could easily be adapted to model selection and would deliver. a form of *robust* model validation. We argue that it would also deliver an endogenous reference model, which is exogenously given in Hansen and Sargent (2006b) and Hansen and Sargent (2006a).

There has been some prior work attempting to link learning with model validation, which we should briefly mention. First, the early work of Bray and Savin (1986) touched on this issue. They ask whether agents could use standard diagnostics, like Chow tests and Durbin-Watson statistics, to detect the time variation in parameters that their own learning behavior generates. Bray and Savin (1986) found that when convergence is slow, agents are generally able to detect the misspecification of their models. In a repeated game context, Foster and Young (2003) allow players to construct, test, and revise simple models of their opponent's behavior. Hypothesis testing produces convergence to a Nash equilibrium in a relatively strong sense, although testing errors produce rare but recurrent experimentation phases. Perhaps closest in spirit to our analysis is recent work by Branch and Evans (2005). They also study a situation where agents not only update the coefficients of a given model, but also select among alternative parametric models based on their recent forecasting performance.

The remainder of the paper is organized as follows. Section 2 provides a formal description of the validation dynamics. Section 3 describes the general properties of the validation dynamics. Section 4 explores the validation dynamics in a couple of examples. Section 5 briefly discusses how our model validation approach could be adapted to deliver an endogenous reference model in robust control problems. Section 6 provides a few concluding remarks.

2. Learning and Validation Dynamics

The concept of model validation has a long history in the statistics and engineering literatures. Our goal is to apply and adapt some of the language and methods of this well developed literature to macroeconomic learning models. This is not a straightforward exercise in translation, mainly because in our setting the data are endogenous to the validation process. Another problem concerns the real-time nature of our exercise. Most validation procedures are retrospective, or are based on hold-out samples, which are ill-suited to ongoing decision making problems. Yet another problem, again related to the decision making aspect of the problem, relates to the nature of the alternative hypothesis. Many existing real-time validation procedures, e.g., those based on CUSUM statistics, fail

to specify an explicit alternative.[2]. Unfortunately, these methods are of little use to our agent, who, at the end of the day, must make a decision based on *some* model. Simply concluding a model is rejected is not very useful. Instead, our agent must adopt the perspective of "it takes a model to beat a model," which means that validation must be defined relative to explicit alternatives.

This section begins by describing a fairly general validation framework. It first discusses how to recursively estimate a given parametric model, and then goes on to discuss how the concept of conditional relative entropy can be used to compare models.

2.1. **General Description.** A representative agent is endowed with an action space A and state space X, with probability space $(A \times X, \nu)$, which represents the actual law of motion. In particular, $\nu(a, x)$ is the distribution over next period's state as a function of the current state and the agent's current action. At same time, the agent has a *perceived* distribution over next period's state, $\mu(a, x)$, which can differ from ν. This perceived distribution is generated by a parametric model, which belongs to a finite model class $\mathcal{M}$, which is is assumed to be compact, but not necessarily convex.[3] $\mathcal{M}$ embodies restrictions and beliefs about the economic environment via functional form specifications and assumed error term distributions.

The agent seeks to maximize the expected present discounted value of the strictly concave payoff function

$$u : A \times X \to \mathbb{R}.$$

This problem is characterized recursively by the following Bellman equation,[4]

$$V(x) = \max_{a \in A} \left\{ u(a, x) + \delta \int V(x') d\mu(x' : a) \right\}$$

where $V(x)$ is the agent's value function. Notice that expectations are computed using the perceived law of motion while the state evolves according to the actual law of motion ν.

Because beliefs and decisions interact, a pair of actions and belief that are consistent with each other is a natural focal point.

Definition 2.1. *A self-confirming equilibrium is* (a^*, μ^*) *where*

$$a^* \in \arg\max_{a \in A} \left\{ u(a, x) + \delta V(x') d\mu^*(x' : a) \right\} \tag{2.1}$$

$$d^*(\mu^*(\cdot : a^*), \nu(\cdot : a^*)) = 0 \tag{2.2}$$

where d^* *is the metric induced by the weak topology.*

[2]See, e.g., Chu, Stinchcombe, and White (1996)

[3]Convexity does not necessarily apply when models are non-nested. Non-convexity is mainly an issue for computation. However, with a relatively small number of candidate models, agents can simply search over all models to find the best current model.

[4]Note, in a traditional Bayesian learning context current beliefs become part of the state, and we would add an equation describing how beliefs evolve as a function of actions. In keeping with the standard adaptive learning literature, we impose a form of bounded rationality here by supposing that the agent thinks the current distribution (i.e., model) will remain in force in all future periods.

Since (2.2) is imposed only on a^*, μ^* can differ from ν conditioned on other actions. Thus, a self-confirming equilibrium need not be a rational expectations equilibrium.[5]

A recursive learning algorithm is a mapping

$$\mathsf{R} : X \times \mathcal{M} \to \mathcal{M},$$

such that

$$\mu_{t+1} = \mathsf{R}(x_t, \mu_t)$$

A stationary point of R is a measure μ^{**} satisfying

$$\mu^{**} = \mathsf{R}(x, \mu^{**})$$

for all x in the support of the marginal distribution of ν over X.

We say that R is an error correction process if $\mu_{t+1} - \mu_t$ is a monotonic transformation of $\nu_t - \mu_t$ where ν_t is the probability distribution over X induced by $\nu(Z_t, a_t)$. It is easy to show that if a recursive learning algorithm is an error correction process, (a^{**}, μ^{**}) is a self-confirming equilibrium where a^{**} solves (2.1) with respect to μ^{**}.

Throughout this paper, we focus on the class of recursive learning algorithms which can be represented as error correction processes. Thus, we treat a self-confirming equilibrium as a stationary point of the recursive learning algorithm.

If $\nu_t = \mu_t$, then the decision maker has no reason to suspect the validity of the model. However, along the transition path toward the self-confirming equilibrium, $\nu_t \neq \mu_t$. In conventional learning models, the agent does not ask questions about whether μ_t is a reasonable description of ν_t, but simply updates μ_t according to R.

A boundedly rational agent with limited computational capacity may have access to only a small class of models, each parametrized by handful of explanatory variables along with a simple functional form. If $\mathcal{M}$ is relatively small, then $\mu \in \mathcal{M}$ is typically misspecified, as it may not assign a positive probability to a state which influences the true law of motion. If the decision maker is aware of this fact, then it would be natural for him to test whether μ_t is a reasonable approximation of ν_t.

We describe a testing procedure following Dembo and Zeitouni (1998). Consider a null hypothesis

$$H_o : \quad \nu_t = \mu_t.$$

Let X_t be a state in period t generated by probability distribution $\nu(a_t, x)$. Define $X^t = (X_1, \ldots, X_t)$.

Definition 2.2. *A test is a mapping*

$$\mathsf{L}_T^t : X^t \times \mathcal{M} \to \mathbb{R}$$

and a set $\mathcal{A} \subset \mathbb{R}$ such that one accepts H_o if and only if $\mathsf{L}_T^t \in \mathcal{A}$.

The performance of L_T^t is generally measured by the two types of errors:

$$\alpha_t = \mathsf{P}\left(\mathsf{L}_T^t(x_1, \ldots, x_t, \mu_0) \notin \mathcal{A} : \mu_0\right)$$

and

$$\beta_t = \mathsf{P}\left(\mathsf{L}_T^t(x_1, \ldots, x_t, \mu_0) \in \mathcal{A} : \mu_1\right)$$

[5]See Fudenberg and Levine (1993).

for some alternative model μ_1. We define a universally optimal Neyman-Pearson test (Dembo and Zeitouni (1998)) as follows,

Definition 2.3. *A test L_T^t is optimal for a given threshold $\rho > 0$ if among all tests satisfying*

$$\limsup_{t\to\infty} \frac{1}{t} \log \alpha_t \leq -\rho,$$

L_T^t maximizes

$$-\limsup_{t\to\infty} \frac{1}{t} \log \beta_t$$

against any alternative hypothesis μ_1.

If μ_t survives the test, then it is updated according to R. Otherwise, the agent must find a new model. The model selection process is to choose μ^* based on $X^t = (X_1, \ldots, X_t)$ according to $\mathsf{L}_S^t(X^t, \mu)$, which represents the gap between the observed data and model μ:

$$\mu_t^* \in \arg\min_{\mu\in\mathcal{M}} \mathsf{L}_S^t(X^t, \mu).$$

Note that μ_t^* is constrained by $\mathcal{M}$, which may not contain the true model ν_t. There is no presumption that the new model is closer to the true model. Rather, the agent is fully aware that the model is only an approximation of the true law of motion, and merely searching for a best fit model in order to guide his decision subject.

Definition 2.4. *A validation process is a combination of the testing and the model selection processes, $(\mathsf{L}_T, \mathsf{L}_S)$. A validation process is consistent if $\mathsf{L}_T = \mathsf{L}_S$.*

If a validation process is inconsistent, then testing and model selection are based on two different criteria, and it is possible that a best fit model is immediately rejected by the testing procedure. For the rest of the paper, we focus on the consistent validation process. Unless specified otherwise, by a validation process, we always refers to the consistent validation process.

2.2. **Special Case.** While the testing and model selection process can be non-parametric, we focus on parametric models for a number of reasons. First, parametric $\mathcal{M}$ capture a form of bounded rationality, reflecting agents with limited computational capacity. Second, comparisons of model complexity are important in our analysis. In parametric models, complexity can be easily measured by counting the number of explanatory variables. Third, since the agent uses his model to solve a control problem, simple models have additional advantages.

In order to illustrate the key ideas of the validation process and its implications for learning models, let us focus on a special, yet important, case where both the testing L_T^t and the selection L_S^t are based on relative entropy. A formal description of each component of the validation process is in order.

2.2.1. *Model Class.* We focus on the class of linear quadratic Gaussian (LQG) problems. We assume that $\mathcal{M}$ consists of probability distributions induced by a class of linear models that can be represented as

$$y_t = X_t\beta_k + \epsilon_t \tag{2.3}$$

where $X_t \in X \subset \mathbb{R}^k$ and ϵ_t is i.i.d. Gaussian. Let $\mathcal{M}^k$ be the class of linear models that have k parameters: the dimension of β_k is k. Define

$$\mathcal{M} = \mathcal{M}^1 \cup \cdots \cup \mathcal{M}^{\overline{K}} \qquad \overline{K} < \infty$$

as the model class, which is endowed to the agent. We call $\mathcal{M}^k$ model k in order to emphasize the number of parameters. A specific element of $\mathcal{M}^k$ represents the agent's beliefs. For this reason, we call a generic element of a model a belief. Loosely speaking, a model corresponds to a functional form of (2.3), while beliefs correspond to a particular specification of β_k. When necessary, we write μ_{β_k} to represent the probability distribution over the state induced by belief β_k. To emphasize the dimension of X_t, we often write $X_{k,t}$ in place of X_t. In order to prevent estimates of β_k from drifting to infinity, we also assume that β_k is selected from a large convex compact set $\mathcal{B}_k$ for $k = 1, \ldots, \overline{K}$. We assume that $\mathcal{B}_k$ is sufficiently large. We shall formalize the notion of being sufficiently large in a moment. We parametrize the belief of the agent by β_k. When necessary, we write μ_{β_k} to represent the probability distribution over the state induced by belief β_k.

Assuming the distribution of ϵ is known is an important restriction. However, it does greatly simplify matters, which then allows us to focus on various issues arising from functional form misspecification. In Section 5, we briefly explore the key issues and challenges associated with relaxing this assumption.

Given β_k, the agent solves a dynamic optimization problem. We assume this problem has an LQR structure, so that the best response is a linear function of the state. Let $g_k(\beta_k)X_t$ be the optimal response of the agent induced by his belief β_k. We assume that $g_k(\beta_k)$ is a Lipschitz continuous function. This assumption ensures that the actual law of motion is a continuous function of the perceived law, as parametrized by β_k.

2.2.2. *Least Square Learning.* We assume that β_k is estimated by recursive least squares:

$$\begin{aligned} \beta_{k,t} &= \beta_{k,t-1} + \eta_t R_{k,t-1}^{-1} X_{k,t}'(y_t - X_{k,t}\beta_{k,t-1}) \\ R_{k,t} &= R_{k,t-1} + \eta_t \left(X_{k,t}' X_{k,t} - R_{k,t-1}\right) \end{aligned} \tag{2.4}$$

where $\eta_t > 0$ is the gain sequence. Since $\beta_{k,t} \in \mathcal{B}_k$ which is convex and compact, we need to project back $\beta_{k,t}$ if the right hand side of (2.4) falls outside of $\mathcal{B}_k$. Thus, our learning algorithm should be written as

$$\begin{aligned} \beta_{k,t} &= \pi_k \left[\beta_{k,t-1} + \eta_t R_{k,t-1}^{-1} X_{k,t}'(y_t - X_{k,t}\beta_{k,t-1})\right] \\ R_{k,t} &= R_{k,t-1} + \eta_t \left(X_{k,t}' X_{k,t} - R_{k,t-1}\right) \end{aligned}$$

where

$$\pi_k(\beta_k) = \begin{cases} \beta_k & \text{if } \beta_k \in \mathcal{B}_k \\ \beta_k^* & \text{otherwise} \end{cases}$$

and β_k^* is a fixed point in the interior of $\mathcal{B}_k$.

We follow econometric convention by assuming that the first component of X_t is 1. Throughout this paper, by a recursive learning algorithm, we refer to (2.4). If the agent is suspicious about the stationarity of the data generating mechanism, it is natural to discount past observations. This feature can be easily incorporated by setting $\eta_t = \eta > 0$.

In this case, (2.4) is called a constant gain recursive learning algorithm. For the rest of this paper, we consider the constant gain algorithm.

The asymptotic properties of the least squares learning algorithm can be analyzed through the associated ordinary differential equation (ODE), which is essentially the trajectory of the mean of $\beta_{k,t}$ when η_t is small:

$$\dot{\beta}_k = \Psi_k(\beta_k) \tag{2.5}$$

where

$$\Psi_k(\beta_k) = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathsf{E}\left[R_{k,t-1}^{-1} X_{k,t}'(y_t - X_{k,t}\beta_{k,t-1}) \ : \ \beta_{k,1} = \beta_k\right].$$

We assume that Ψ_k is continuous for $k = 1, \ldots, \overline{K}$, and has a unique trajectory from any initial condition. Let β_k^s be a stationary point of ODE: $\Psi_k(\beta_k^s) = 0$. We say that β_k^s is locally stable, if there is a small neighborhood around β_k^s such that for any initial condition in the neighborhood, the trajectory remains in the same neighborhood.

We can always choose $\mathcal{B}_k$ sufficiently large to include the stationary points of (2.5) in its interior. Although the projection facility π_k is an integral part of the learning algorithm, we shall omit π_k in writing the learning algorithm to simplify the exposition. The only reason we need to confine β_k inside a compact set is to ensure that the neighborhood of β_k^s is recurrent with respect to the stochastic process induced by (2.5). Containing $\beta_{t,k}$ inside a compact set is a traditional way of ensuring recurrence, but can be replaced by milder and more natural condition as in Borkar and Meyn (2000).

Employing standard methods from the stochastic approximation literature (Kushner and Yin (1997)), we can show that if $\eta_t = t^{-1}$ and β_k^s is locally stable, then there is a small neighborhood around β_k^s such that for any initial condition in this neighborhood, $\beta_{k,t} \to \beta_k^s$ with probability 1. If $\eta_t = \eta > 0$, then the recursive formula converges weakly to a stationary distribution.

The key conditions are the smoothness of the best response of the agent, and the recurrence of the neighborhood of β^s. Throughout the paper, we assume the recursive learning algorithm has a globally stable stationary point, which is a self-confirming equilibrium.

Assumption 2.5. *$\forall k \in \{1, \ldots, \overline{K}\}$, there exists β_k^s which is a globally stable stationary point of* (2.5).

We use a stationary point and a self-confirming equilibrium interchangeably. The stationary point of the associated ODE provides a natural focal point of the learning dynamics. If a learning algorithm does not have a locally stable point, it is difficult to identify the focal point of the learning dynamics. The general case where the recursive model admits multiple locally stable stationary points can be analyzed by the same logic, although the notation becomes considerably more cumbersome.[6]

While the properties of the center of the distribution of $\beta_{k,t}$ can be analyzed through (2.5), we need to examine the large deviation properties (LDP) in order to understand the tail portion of the distribution of $\beta_{k,t}$ around the self-confirming equilibrium β_k^s.

[6]Even if the learning dynamics does not have a stable point, one may extend the analysis along the lines of Benaim (1996).

Define a fictitious time scale $\tau > 0$ as follows. Interpret $\eta_t = \eta$ as the time between the two subsequent rounds and t as the calendar time. Thus, ηt represents the real time (or fictitious time) to reach round t. For $\tau > 0$, define $m(\tau) = \lceil \tau/\eta \rceil$ as the smallest integer that exceeds τ/η. Given $\{\beta_{k,t}\}_{t=1}^{\infty}$ induced by (2.4), let $\{\beta_k^\eta(\tau)\}_{\tau>0}$ be the continuous time process obtained by linear interpolation of $\{\beta_{k,t}\}_{t=1}^{\infty}$ with respect to the fictitious time τ. Let $\{\beta_k(\tau)\}$ be the pointwise limit of $\{\beta_k^\eta(\tau)\}_{\tau>0}$ as $\eta \to 0$. We need the tightness of $\{\beta_k^\eta(\tau)\}_{\tau>0}$ to ensure the existence of the limit, which is implied by the compactness of $\mathcal{B}_k$.

Given $\rho > 0$, define

$$N_\rho(\beta_k^s, \tau^*) = \{\beta_k : \beta_k(0) = \beta_k^s, \quad \exists \tau \leq \tau^*, \quad \|\beta_k(\tau^*) - \beta_k^s\| \geq \rho\} \tag{2.6}$$

as the collection of sample paths starting from β_k^s which exits the ρ neighborhood of $\beta_k^s \in \mathbb{R}^k$ before rime $\tau^* > 0$. It is easy to see that $\forall \tau^* > 0$,

$$\lim_{\eta \to 0} \mathsf{P}(N_\rho(\beta_k^s, \tau^*)) = 0$$

which implies that $N_\rho(\beta_k^s, \tau^*))$ is a rare event. We are interested in the rate at which $\mathsf{P}(N_\rho(\beta_k^s, \tau^*))$ vanishes. In particular, we shall use Sanov's Theorem (Dembo and Zeitouni (1998)) to characterize the following limit

$$\limsup_{\eta \to 0} \eta \log \mathsf{P}(N_\rho(\beta_k^s, \tau^*) : \beta_k^s)$$

Let $\mu_{\beta_k^s}$ be the stationary distribution of (2.4) with a fixed $\eta > 0$. Similarly, let $\nu_{\beta_k(0)}$ be the probability distribution of $\beta_{k,t}$ evolving according to (2.4) with initial condition $\beta_k(0) = \beta_k^s$. The (conditional) relative entropy between the perceived distribution $\mu_{\beta_k}^s$ and the current actual distribution ν_{β_k} is defined as,

$$\mathsf{H}(\nu_{\beta_k(\tau)} \| \mu_{\beta_k^s}) = \begin{cases} \int \log \frac{d\nu_{\beta_k(\tau)}}{d\mu_{\beta_k^s}} d\nu_{\beta_k(\tau)} & \text{if } \nu_{\beta_k(\tau)} \ll \mu_{\beta_k^s} \\ \infty & \text{otherwise.} \end{cases} \tag{2.7}$$

Note that relative entropy is just an expected log-likelihood ratio statistic. In the statistics literature, relative entropy is referred to as the Kullback-Leibler Information Criterion (KLIC), and we use these terms interchangeably.

By invoking Sanov's theorem (Dembo and Zeitouni (1998)), we have

$$\limsup_{\eta \to 0} \eta \log \mathsf{P}(N_\rho(\beta_k^s, \tau^*) : \beta_k^s) \leq - \inf_{\beta_k \in N_\rho(\beta_k^s, \tau^*)} \mathsf{H}(\nu_{\beta_k(\tau)} \| \mu_{\beta_k^s}) \equiv -\mathsf{H}^*(\mu_{\beta_k^s}) \tag{2.8}$$

and moreover, $\forall \rho > 0$,

$$\mathsf{H}^*(\mu_{\beta_k^s}) > 0$$

which is called the rate function of $\beta_k(\tau)$.

We also know that the weak inequality holds as an equality for a particular path, called the *dominant* escape path. Standard results from large deviations theory imply that as $\eta \to 0$, the probability distribution over $N_\rho(\beta_k^s, \tau^*) : \beta_k^s)$ conditioned on the event that $\exists \tau^*$ such that $\beta_k(\tau^*) \in \partial N_\rho(\beta_k^s, \tau^*)$ is concentrated at the dominant escape path. Let β_k^e be the dominant escape path, and τ^e be the first exit time of the dominant escape path out of the ρ neighborhood of β_k^s. Then, we can conclude that as $\eta \to 0$ $\beta_k^e(\tau^e) \in \partial N_\rho(\beta_k^s, \tau^*)$ is the most likely exit point out of the ρ neighborhood of β_k^s.

The rate function $\mathsf{H}^*(\mu_{\beta_k^s})$ tells us how difficult it is to escape from the self-confirming equilibrium β_k^s. If the rate function increases, the probability of escape decreases at an exponential rate proportional to H^*. Thus, if $\beta_k(\tau)$ has the largest rate function among all the models in $\mathcal{M}$, the recursive learning algorithm is most likely to stay in the neighborhood of the self-confirming equilibrium associated with $\beta_k(\tau)$.

Definition 2.6. *β_k^s is a dominant self-confirming equilibrium, if*

$$\mathsf{H}^*(\mu_{\beta_k^s}) = \max_{1\le k'\le K} \mathsf{H}^*(\mu_{\beta_{k'}^s}). \tag{2.9}$$

If there are multiple k that satisfies (2.9)*, then we choose the smallest k. The recursive learning model associated with the dominant self-confirming equilibrium is called the dominant recursive learning model.*

2.2.3. *Validation.* As we focus on consistent validation, we choose H as the basis for both testing and model selection. We do this for a couple of reasons. First, if μ^* is selected by

$$\mu^* = \arg\min_{\mu\in\mathcal{M}} \mathsf{H}(\nu\|\mu),$$

then μ^* is a penalized maximum likelihood estimator, which is widely used in practice. Note that the selection of the estimator is bounded by model class $\mathcal{M}$. Second, the testing procedure to accept

$$H_o : \nu = \mu$$

if and only if

$$\mathsf{H}(\nu_{\beta_t}\|\mu_{\beta_{k,t}}) < \rho \tag{2.10}$$

is an optimal test (Dembo and Zeitouni (1998)), where $\rho > 0$ is a pre-specified threshold, ν_{β_t} is the actual distribution of β_t evolving according to (2.4). Let s_t be a random variable that represents the model being used in period t.

We are interested in the stochastic process of $\{X_t, \beta_{s_t,t}\}$ induced by the following recursive validation procedure. In period t, the agent tests his present model $\beta_{k,t}$ according to (2.10). If $\beta_{k,t}$ passes the test, then the agent solves his control problem using this model. In the following period, after a new observation is realized, he updates $\beta_{k,t}$ to $\beta_{k,t+1}$ according to (2.4).

If $\beta_{k,t}$ fails the test, the agent chooses a new model $\beta_{k',t}$ by solving

$$\min_{\beta\in\mathcal{M}} \mathsf{H}(\nu_{\beta_t}\|\mu_\beta). \tag{2.11}$$

Based on $\beta_{k',t}$, the agent chooses his action and updates $\beta_{k',t}$ to $\beta_{k',t+1}$ based on the new observation. In period $t+1$, the agent repeats the same procedure. By the validation dynamics, we mean the dynamic properties of $\{X_t, \beta_{s_t,t}\}$.

By definition, the validation process consists of two distinct procedures: estimation *and* testing. The estimation process can be represented recursively as in (2.4). The testing procedure induces a Markov chain over models. It is most convenient to write down the validation process as a recursive least square estimator with Markov regime switching process (Yin and Krishnamurthy 2005).

Suppose model k is characterized by a parameter vector, β_k, and a state vector, X_k. In the class of linear quadratic Gaussian (LQG) models, the best response is a linear function of the present state. Thus, we can write the best response as

$$b_k^r(\beta_k, X_k) = g_k(\beta_k)X_k.$$

Still, the actual law of motion can depend on the entire state profile. We write (2.4) in a slightly more general form as

$$\beta_{k,t+1} = \beta_{k,t} + \eta_t h_k(\beta_{k,t}, X_t, y_t) \tag{2.12}$$

for the least square learning algorithm for β_k. Note that X_t instead of $X_{k,t}$ is in the argument of h_k, since the actual realization y_t can be influenced by the entire profile of the state variables, even though the agent's control is conditioned only on $X_{k,t}$.

Let $s_t \in \{1, \ldots, K\}$ indicate the model number in period t. Let p_t be a probability distribution over $\{1, \ldots, K\}$ and $\mathsf{P}(\beta_{k,t}, X_t, y_t)$ be the transition function induced by the testing procedure:

$$p_{t+1} = p_t \mathsf{P}(\beta_{k,t}, X_t, y_t). \tag{2.13}$$

We can stack (2.12) to form

$$\begin{bmatrix} \beta_{1,t+1} \\ \vdots \\ \beta_{\overline{K},t+1} \end{bmatrix} = \begin{bmatrix} \beta_{1,t} \\ \vdots \\ \beta_{\overline{K},t} \end{bmatrix} + \eta_t \begin{bmatrix} h_1(\beta_{s_t,t}, X_t, y_t) \\ \vdots \\ h_{\overline{K}}(\beta_{s_t,t}, X_t, y_t) \end{bmatrix}. \tag{2.14}$$

The algorithm appears to indicate that the coefficient of every model is synchronously updated, based on the realized data by the k-th model being used in period t. However, the agent actually updates the coefficient asynchronously. While model k is being used, the agent updates only $\beta_{k,t}$ in period t. If model k' is selected after the model k is rejected by the testing procedure, the estimation of $\beta_{k'}$ is based on the history of outcomes in the past, some of which were generated by different models. This is another aspect of bounded rationality of the agent who underestimate the impact of his action on the evolution of the state variables. By the time when model k' is selected, the updating behavior can be described *as if* the agent has been updating model k' in the background while model k is being used, and then, continues to update $\beta_{k',t}$ after model k' is selected.

2.2.4. *Numerical Implementation.* For a general distribution, $\mathsf{H}(\nu_t \| \mu_t)$ is difficult to calculate. However, in the context of linear Gaussian models, the validation process can be stated in a computationally tractable form, which offers useful insights into the inner working of the validation process.

In the context of Gaussian linear regression, where models are "good" in the sense that White's Information Matrix Equality (White (1982)) is approximately satisfied for a large t,[7] we have

$$\mathsf{H}(\nu_t \| \mu_t) \approx t \log(\hat{\sigma}_t^2) + 2K_t \tag{2.15}$$

[7]When this is not the case, minimizing expected KLIC produces the so-called Takeuchi Information Criterion. See Burnham and Anderson (2002) for more details.

where K_t is the number of parameters in model μ_t, and $\hat{\sigma}_t^2 = t^{-1}\sum \hat{\varepsilon}_t^2$, where $\hat{\varepsilon}_t$ is just the time t regression residual. The right hand side of (2.15) is the well known Akaike Information Criterion (AIC).

Suppose that the agent is comparing two models μ_1 and μ_2 after the previous model was invalidated. We can obtain an asymptotically unbiased estimate of two models' relative (expected) KLICs:

$$\begin{aligned} \mathsf{H}(\nu_t\|\mu_1) - \mathsf{H}(\nu_t\|\mu_2) &\approx AIC_2 - AIC_1 \\ &= [\log(\hat{\sigma}_2^2) - \log(\hat{\sigma}_1^2)] + \frac{2}{t}(K_2 - K_1). \end{aligned} \tag{2.16}$$

This expression provides a very natural, theoretically motivated, basis for model comparison and selection. In particular, positive values favor Model 1, while negative values favor Model 2. Notice that the first term on the right-hand side captures relative model fit, while the second captures relative model complexity. Hence, equation (2.16) resolves a trade-off between bias and variance that is present in all model selection problems.

Schematically then, the agent uses the following recursive testing and model validation process:

- Let $\mathcal{A}_t$ be the time-t value of the recursively estimated AIC statistic for the current reference model.
- Let $\mathcal{A}_t^*$ be the minimum AIC statistic across all models at time-t. All models are always being fit to the data, even if they are not currently being used for making a decision.
- If $\mathcal{A}_t - \mathcal{A}_t^* < \rho$, then the current model is validated and updated to guide the decision of the agent.
- If $\mathcal{A}_t - \mathcal{A}_t^* \geq \rho$, then the current reference model is replaced by the model corresponding to $\mathcal{A}_t^*$, and the decision is instead based on this new model.

Calibrating ρ to given Type I error probability requires knowledge of the sampling distribution of the difference between two AIC statistics. Vuong (1989) did this for the case of i.i.d. data. It turns out that the sampling distribution depends on whether the models are nested or non-nested. These results were subsequently extended to the dynamic case by Rivers and Vuong (2002), for the case of non-nested models, and by Marcellino (1999), for the case of nested models. In principle, we could allow our agent to apply these tests. However, this would still lead to incorrectly sized tests, due to both the endogeneity of the data and the recursive nature of the tests.

Instead, we assume the agent adopts the following quasi-Bayesian calibration strategy, based on the notion of 'Akaike weights' (see, e.g., Burnham and Anderson (2002)). Suppose that each period the agent computes the following value for each model:

$$\pi_{it} = \frac{\exp(-.5\Delta_{it})}{\sum_j \exp(-.5\Delta_{it})} \tag{2.17}$$

where $\Delta_{it} = AIC_{it} - AIC_{min,t}$ and the summation runs over all models in $\mathcal{M}$. Notice that π_{it} can be interpreted as a frequentist estimate of the probability that model-i is currently the best KLIC model. We assume the agent retains the current reference model as long as $\pi_{max,t} < \rho \cdot \pi_{it}$, where $\rho > 1$. For example, if $\rho = 1.1$ then the agent retains the

current reference model as long as the probability of the best model does not exceed its probability by more than 10%.

3. Analysis

We focus on the constant gain algorithm where $\eta_t = \eta > 0$. Let $O(\eta)$ be a function satisfying

$$\lim_{\eta \to 0} O(\eta) = 0.$$

Lemma 3.1.

$$\mathsf{P}\left(\beta_{k,t+1} \text{ fails the test} \mid \beta_{k,t} \text{ passes the test}\right) \leq O(\eta)$$

Proof. Since the fictitious time is in the same scale as η, it suffices to show that $\forall \mu > 0$, $\exists \tau' > 0$ such that $\forall \tau \in (0, \tau')$,

$$\mathsf{P}\left(\mathsf{H}(\nu_\tau \| \mu_{\beta_k(\tau)}) \geq \rho : \mathsf{H}(\nu_0 \| \mu_{\beta_k(0)}) < \rho\right) \leq \mu.$$

Under the Gaussian assumption on the perturbation term, $\mathsf{H}(\nu \| \mu)$ is a continuous function with respect to both arguments. Moreover, the actual law of motion is a continuous function of $\beta_k(\tau)$, because the best response function is assumed to be a continuous function of $\beta_k(\tau)$. Thus, if $\mathsf{H}(\nu_0 \| \mu_{\beta_k(0)}) < \rho$, then there exists $\mu^* > 0$ such that if $|\beta_k(\tau) - \beta_k(0)| < \mu^*$, then $\mathsf{H}(\nu_\tau \| \mu_{\beta_k(\tau)}) < \rho$. Hence, it suffices to show that there exists $\tau' > 0$ such that $\forall \tau \in (0, \tau')$,

$$\mathsf{P}\left(|\beta_k^\eta(\tau) - \beta_k^\eta(0)| > \mu^*\right) = O(\eta).$$

Recall that

$$\beta_{k,t+1} = \beta_{k,t} + \eta R_{k,t}^{-1} X_{k,t} \left(y_t - X_{k,t}\beta_{k,t}\right) = \beta_{k,t} + \eta h_k(\beta_{k,t}, y_t, X_{k,t}).$$

Define

$$\bar{h}_k(\beta_{k,t}) = \mathsf{E} h_k(\beta_{k,t}, y_t, X_{k,t})$$

where expectations are taken with respect to the stationary distributions of y and X evaluated at $\beta_{k,t}$, and

$$\xi_{k,t} = h_k(\beta_{k,t}, y_t, X_{k,t}) - \mathsf{E} h_k(\beta_{k,t}, y_t, X_{k,t})$$

which is a martingale difference.

Given a fixed fictitious time $\tau > 0$, we can decompose the path of $\beta_k(\tau)$ as follows,

$$\beta_k(\tau) - \beta_k(0) = \int_0^\tau \Psi_k(\beta_k(s)) ds + \eta \sum_{t=1}^{\lceil \tau/\eta \rceil} \xi_t + O(\eta)$$

where the last $O(\eta)$ term arises because τ may not be an integer multiple of $\eta > 0$. By Doob's martingale inequality, $\forall \mu'' > 0$,

$$\mathsf{P}\left(\left| \eta \sum_{t=1}^{\lceil \tau/\eta \rceil} \xi_t \right| > \mu''\right) \leq \eta^2 \left\lceil \frac{\tau}{\eta} \right\rceil \max_{t=1,\ldots,\lceil \tau/\eta \rceil} \frac{\mathrm{var}(\xi_t^2)}{(\mu'')^2} = O(\eta)$$

since

$$\sup_{t \geq 1} \max_{t=1,\ldots,\lceil \tau/\eta \rceil} \mathrm{var}(\xi_t^2) < \infty.$$

Since Ψ_k is a continuous function, we can choose $\tau' > 0$ such that $\forall \tau \in (0, \tau')$,

$$\left| \int_0^\tau \Psi_k(\beta_k(s))ds \right| \leq \frac{\mu^*}{2}.$$

If we then set $\mu'' = \frac{\mu^*}{2}$, we have

$$\mathsf{P}\left(|\beta_k^\eta(\tau) - \beta_k^\eta(0)| > \mu^*\right) = O(\eta)$$

as desired. □

Lemma 3.1 indicates that the probability of switching from the current model vanishes as the gain sequence converges to 0. Hence, its impact on the mean dynamics of $\beta_{k,t}$ vanishes at the rate of η^2, which allows us to ignore its impact on the mean dynamics as $\eta \to 0$.

We are interested in those models which are selected by the decision maker infinitely often. For the sake of simplifying the exposition, let us assume that every model is selected infinitely often along the validation path. This assumption implies that the Markov chain induced by the validation process is irreducible.

Define the transition probability

$$q_{t,k,k'} = \mathsf{P}(s_{t+1} = k' : s_t = k, \beta_{k,t}, X_t, y_t).$$

Lemma 3.1 implies that

$$\sum_{k' \neq k} q_{t,k,k'} = O(\eta) \qquad \forall t \geq 0.$$

Fix a large t when $s_t = k$. The evolution of $\beta_{k,t+1}$ can be written as

$$\begin{aligned}\beta_{k,t+1} &= \beta_{k,t} + \eta R_{k,t}^{-1} X_{k,t}' \left((1 - \sum_{k' \neq k} q_{t,k,k'}) h_k(\beta_{k,t}, X_t, y_t) + \sum_{k' \neq k} q_{t,k,k'} h_{k'}(\beta_{k',t}, X_t, y_t) \right) \\ &= \beta_{k,t} + \eta(1 - \sum_{k' \neq k} q_{t,k,k'}) R_{k,t}^{-1} X_{k,t}' h_k(\beta_{k,t}, X_t, y_t) + \eta \sum_{k' \neq k} q_{t,k,k'} R_{k,t}^{-1} X_{k,t}' h_{k'}(\beta_{k',t}, X_t, y_t).\end{aligned}$$

Let us write $\eta_t' = \eta(1 - \sum_{k' \neq k} q_{t,k,k'})$ and $\eta_t'' = \eta \sum_{k' \neq k} q_{t,k,k'}$. We define the fictitious time scale according to η_t'. In particular, define $m(\tau) = \min\{T : \sum_{t=1}^T \eta_t'' \geq K\}$ for $\tau > 0$. Let $\beta_k^\eta(\tau)$ be the continuous time process obtained by the interpolation of $\{\beta_{k,t}\}$. Then, $\forall K$, $\forall s > 0$

$$\lim_{\eta \to 0} \frac{\sum_{t=m(\tau)}^{m(\tau+s)-1} \eta_t''}{\eta} = 0$$

since $\sum_{k' \neq k} q_{t,k,k'} = O(\eta)$. By employing standard results from the stochastic approximation literature, we then have, $\forall \tau, s > 0$,

$$\lim_{\eta \to 0} \beta_k^\eta(\tau + s) - \beta_k^\eta(\tau) = \int_\tau^{\tau+s} \Psi_k(\beta_k(\tau'))d\tau'$$

in probability.

Once model k is selected, it remains the model of choice for an extended period. Since model k is selected infinitely many times, we conclude that

$$\lim_{t\to\infty} \beta_{k,t} = \beta_k^s$$

weakly as $\eta \to 0$.

Given $\eta > 0$, the duration of model k is dictated by the escape probability of $\beta_{k,t}$ from the neighborhood of β_k^s, which is completely characterized by the rate function $\mathsf{H}^*(\beta_k^s)$ defined as (2.8). Define

$$k^* = \arg \max_{k\in\{1,\ldots,\overline{K}\}} \mathsf{H}^*(\beta_k^s)$$

which is precisely the index of the dominant self-confirming equilibrium. As $\eta \to 0$, s_t stays exponentially longer period around k^* than any other k. Thus, the validation dynamics converges to the dominant self-confirming equilibrium.

Proposition 3.2. *Suppose that* $\forall \eta > 0$,

$$\lim_{t\to\infty} \frac{\#\{t : s_t = k\}}{t} > 0 \qquad \forall k \in \{1, \ldots, \overline{K}\}.$$

Then,

$$\lim_{\eta\to 0} \lim_{t\to\infty} \mathsf{P}\left(\beta_{k,t} = \beta_{k^*}^s\right) = 1.$$

While Proposition 3.2 says that one should choose model k^* and focus on the dominant self-confirming equilibrium, it does not identify the structure of the dominant self-confirming equilibrium, especially k^*. In many applications, however, we can identify interesting properties of the dominant self-confirming equilibrium. In the following section, we consider a couple of well known examples to examine what can be learned about dominant self-confirming equilibria.

4. Examples

This section illustrates the role of model validation in a sequence of alternative environments. The novel feature of our validation approach is the endogenous nature of the data-generating process. In general, this endogeneity arises from both expectational feedback and from the agent's use of the model to solve a control problem. The examples in this section are designed to highlight how control and feedback influence validation dynamics. Accordingly, we begin by discussing model validation *without* feedback or control. Since this is well trodden territory, we merely summarize the main results. Next, we isolate how control influences model validation. We do this by summarizing known results from the adaptive control literature on recursive model order selection. We then study an example from Evans and Honkapohja (2001) to illustrate how expectational feedback influences model validation. Finally, we put all the pieces together by incorporating model validation into an example from Sargent (1999), which features both control and expectational feedback.

4.1. **Model Validation With Exogenous Data.** If our agent were 'small', in the sense that his actions did not influence the data-generating process he was attempting to learn about, then we could avail ourselves of a huge literature on model selection. A unifying theme in this literature is the need to strike a balance between model fit and model complexity or, in statistical terms, between bias and variance. Model selection methods differ according to how complexity is defined, how it is measured, and how it is penalized.

Suppose an agent must select one of two models. Denote their time-t quasi-log-likelihoods as $f_{it}(\cdot|\theta_i)$, where $\theta_i \in \Theta_i$ for $i = 1, 2$. Assume $\Theta_1 \subseteq \mathbb{R}^p$ and $\Theta_2 \subseteq \mathbb{R}^q$ where $p > q$, so that Model 2 is more 'parsimonious' than Model 1. Choice is based on the following penalized quasi-likelihood ratio statistic

$$IC_n(\cdot) = \sum_{t=1}^{n} f_{1t}(\cdot|\hat{\theta}_{1n}) - \sum_{t=1}^{n} f_{2t}(\cdot|\hat{\theta}_{2n}) - \hat{c}_n(\cdot) \tag{4.18}$$

where $\hat{\theta}_{in}$ are quasi-maximum likelihood estimators of the unknown parameters θ_i, and $\hat{c}_n(\cdot) \geq 0$ is a complexity cost function. Most of the model selection literature focuses on how to specify $\hat{c}_n$. As noted earlier, Akaike (1974) argued that $c_n = (p - q)$ provides an asymptotically unbiased estimate of two model's relative KLIC when they good approximations of the underlying DGP (in the sense that White's (1982) Information Matrix Equality is approximately satisfied). Schwarz (1978) adopted a Bayesian approach, and showed that under certain conditions $c_n = (p-q)\log(n)/2$ provides an asymptotic estimate of posterior model probabilities. Rissanen (1989) exploited the connection between codes and probability distributions to develop a *minimum description length* (MDL) approach to model selection. The key idea behind MDL is that data regularities can be used to *compress* data. Since codes are just algorithms for compressing data, data with a lot of regularity can be compressed a lot so that they have short minimum code (or description) lengths. In keeping with Occam's Razor, Rissanen argued that model selection should be based on MDL, where a model's description length consists of two parts: (1) its implied code length for describing the data, which is inversely related to model fit, and (2) the code length of the model itself that is the number of bits it takes to describe the model. This second part delivers a natural measure of model complexity. Under general conditions, minimum code lengths for a given model and data set are achieved by maximum likelihood estimators. On the other hand, there are many ways to encode a given model. Depending on how you do this, you can obtain complexity cost functions that resemble either AIC or BIC (see Hansen and Yu (2001)). Rissanen advocated a so-called 'uniform encoder', where each parameter is encoded with a precision of $1/\sqrt{n}$, which is natural since this is the MLE convergence rate. This implies that the complexity cost of each parameter is $-\log(1/\sqrt{n})$ (nats), so that if a model has p parameters, its complexity cost is $-p\log(1/\sqrt{n}) = p\log(n)/2$, which is just the BIC penalty. Rissanen (1989, p. 71) also proved that, outside a set of measure zero, the minimum achievable average KLIC distance between a model and a p-dimensional DGP is $p\log(n)/2$. In other words, there is a limit to how well we can expect to approximate the data. Ploberger and Phillips (2003) extend this result to nonstationary data, and show that the bound *increases* when there are trends in the data.

Much of the early model selection literature focused on consistency. That is, suppose there is a 'true model' out there and that it lies within $\mathcal{M}$. Under what conditions on c_n will we (asymptotically) select the true model with probability 1? The main results along these lines are: (1) AIC is inconsistent, i.e., it has a nonvanishing probability of selecting overparameterized models, and (2) BIC is consistent. Unfortunately, these results are of little use to us, given that we want to allow for the possibility that $\mathcal{M}$ does *not* contain the true DGP. Instead, our agent just wants to find the minimum KLIC model. More relevant is the work of Nishii (1988). He shows that if $\hat{c}_n = o_p(n)$ and $P[(\log\log(n))^{-1}\hat{c}_n \to \infty] = 1$ then basing model choice on (4.18) will consistently select the minimum KLIC model. Notice that both AIC and BIC satisfy the first condition, but only BIC satisfies the second. Sin and White (1996) extend these results to the case of dependent data. Not surprisingly, sufficient conditions for consistent minimum KLIC model selection still include $\hat{c}_n = o_p(n)$, but the rate of increase is strengthened to $P[n^{-1/2}\hat{c}_n \to \infty] = 1$. However, if models are strictly nested, then $P[\hat{c}_n \to \infty] = 1$ is sufficient.

In practice, samples are finite, and judicious model selection recognizes Type I and II errors, particularly when there are costs to switching models. Hence, it is desirable to *test* whether two models have the same KLIC, allowing for the possibility that either, or both, may be misspecified. This issue was studied by Vuong (1989) for the case of i.i.d. data. Here c_n can be related to the test's significance level. Imposing $c_n = o(n)$ then expresses the common sense intuition that significance levels should decrease with sample size. Vuong shows that the asymptotic distribution of (4.18) depends on whether the models are nested. If not, then QLR scaled by a consistent estimate of its variance converges at rate $\sqrt{n}$ to a $N(0,1)$ distribution. When models are nested, then QLR converges in general to a weighted sum of chi-squares. These results were extended to the case of dynamic models and dependent data by Rivers and Vuong (2002) (for the non-nested case) and by Marcellino (1999) (for the nested case).

4.2. **Model Validation With Control.** When agents *use* their models to solve a control problem then the data-generating process becomes endogenous, and it is not at all clear to what extent the previous results continue to apply. For example, if an agent's control variable settles quickly to a constant, then this is obviously going to limit his ability to estimate parameters describing how his control actions influence the evolution of the state. Of course, this is not an issue from a Bayesian perspective, where agents recognize their influence over the DGP and run experiments. In fact, as noted by Easley and Kiefer (1988), with discounting it may not even be *optimal* for agents to learn the true DGP. However, from the passive-learning/certainty-equivalence perspective of adaptive control, the question arises as to whether standard control policies can be modified to ensure eventual learning of the true model.

Suppose an agent wants to solve the following problem,

$$\inf_{u_t}\left\{\limsup_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}(Q_1 x_i^2 + Q_2 u_i^2)\right\} \tag{4.19}$$

subject to

$$x_{i+1} = \sum_{k=0}^{p} a_k x_{i-k} + \sum_{k=0}^{q} b_k u_{i-k} + w_{i+1}$$

This is a standard (steady state) LQR problem, and is assumed to satisfy the usual regularity conditions. The only twist is that not only are the coefficients (a_k, b_k) unknown, *but so are the lag orders* (p, q). Model selection in this context can be interpreted as estimating (p, q). The question is: Can a control policy be formulated that both achieves the control objective in (4.19) and delivers consistent estimates of (a_k, b_k, p, q)?[8] This question has been addressed by Hemerly and Davis (1989) and Chen and Zhang (1990).

Under the important assumption that *a priori* upper bounds for (p, q) are known, they formulate a recursive control and order selection strategy that is exactly analogous to the one we described in Section 2, the only difference being that they allow for a general complexity cost function, c_n, as in (4.18), as opposed to our Akaike inspired penalty. They show that if: (1) the control policy satisfies a so-called 'attenuating excitation' condition, ensuring enough data variability to deliver consistency, but not so much as to degrade steady state control objectives, and (2) c_n satisfies appropriate order conditions, then optimal steady state policy delivers almost sure convergence to the true model.[9] For the case of (4.19), these order conditions can be described as follows. Assume (p^*, q^*) are the known upper bounds on the lag orders. Let $\gamma = p^* + \max(p^*, q^*)$, and $\varepsilon \in (0, 1/[2(\gamma+1)])$ and $\delta \in [0, (1 - 2\varepsilon(\gamma+1))/(2\gamma+3))$. Then if

$$\frac{(\log n)(\log\log n)^d}{c_n} \longrightarrow 0, \qquad d > 1 \qquad \text{and} \qquad \frac{c_n}{n^{1-(\gamma+1)(\varepsilon+\delta)}} \longrightarrow 0$$

the agent discovers the true model almost surely. Notice that this imposes a larger complexity cost than in the exogenous data case. Loosely speaking, more complex models take more time and effort to learn, which has added costs in a control setting. Hence, you want to penalize this more severely.

Unfortunately, these results are not directly useful for us, for a number of reasons. First, with discounting it us no longer clear that it is reasonable to impose consistency. Second, the analysis of Hemerly and Davis (1989) and Chen and Zhang (1990) presumes the model is stationary, and so employs decreasing gain learning algorithms. Third, we want to allow for the possibility that $\mathcal{M}$ does *not* contain the true model. Next, we discuss an example where these issues are addressed. All models are misspecified, the agent discounts the future, and he uses constant gain learning algorithms. Rather than employ the brute force probabilistic methods of Hemerly and Davis (1989) and Chen and Zhang (1990), we employ the powerful analytical methods of large deviations theory.

---

[8]Note, consistency of (a_k, b_k) and (p, q) needs to be analyzed separately. Since (p, q) are discrete integers, their convergence rates will in general be different. In particular, consistency for $(\hat{p}, \hat{q})$ implies exact equality after a finite time interval.

[9]Interestingly, this 'attenuating excitation' condition is analogous to the sort of 'natural experiments' that El-Gamal and Sundaran (1993) identified as being important in ensuring consistency of Bayesian learning with endogenous data.

4.3. **Model Validation With Expectational Feedback.** Many macroeconomic models take the form of expectational difference equations. Sometimes these result from explicit control problems, in which case there is no essential difference between 'control feedback' and 'expectational feedback'. However, sometimes these equations feature no explicit control other than the need to forecast future realizations of an endogenous variable. One prominent example is the famous cobweb model. This model has long been a useful laboratory for analyzing various issues in dynamic economics, first with constant coefficients adaptive expectations, then with rational expectations, then with adaptive least-squares learning, and most recently, with misspecified adaptive least-squares. We continue this tradition by using it to study model validation dynamics. In particular, we follow-up on an example studied by Evans and Honkapohja (2001, pgs. 318-20). They used the cobweb model to analyze the E-stability properties of so-called Restricted Perceptions Equilibria (RPE), in which agents (exogenously) omit relevant variables from their PLMs. In their analysis, *any* model can be a RPE, as long as its estimated coefficients adjust to account for the omitted variable bias. Here we allow agents to *test* their models, and ask whether some RPE are more resilient than others.

Consider then the following reduced form model,

$$p_t = \alpha E_{t-1} p_t + \beta_1 w_{1,t-1} + \beta_2 w_{2,t-1} + \varepsilon_t \tag{4.20}$$

where $(w_{1,t}, w_{2,t})$ are zero mean exogenous variables, and ε_t is an i.i.d. shock. Let us assume that $w_{i,t}$ and ε_t are Gaussian. This model has a trivial Rational Expectations equilibrium,

$$p_t = \frac{1}{1-\alpha}\left(\beta_1 w_{1,t-1} + \beta_2 w_{2,t-1}\right) + \varepsilon_t$$

From Evans and Honkapohja (2001), we also know that it has a pair of RPE. Let RPE_1 denote the RPE obtained by including $w_{1,t}$ and excluding $w_{2,t}$. Let RPE_2 denote the reverse case. We then have:

$$RPE_1: \qquad p_t = (1-\alpha)^{-1}(\beta_1 + \Omega_{11}^{-1}\Omega_{12}\beta_2) w_{1,t-1} + \varepsilon_t \tag{4.21}$$

$$RPE_2: \qquad p_t = (1-\alpha)^{-1}(\beta_2 + \Omega_{22}^{-1}\Omega_{12}\beta_1) w_{2,t-1} + \varepsilon_t \tag{4.22}$$

where Ω_{ij} are the elements of the second moment matrix, $\Omega = E(ww')$. Both RPE are E-stable if and only if $\alpha < 1$.

Now suppose an agent simultaneously entertains *both* models, continuously testing one against the other. For simplicity, suppose that he does not consider the correctly specified larger model. Further suppose that the agent wants to guard against parameter drift within each model, so he adopts a constant gain recursive learning algorithm. What can we say about the model validation dynamics in this case? From the analysis in Section 3, we can say that the dominant model will be the one with the bigger large deviations rate function.

Thanks to the linear Gaussian structure of the problem, we can derive the closed form equation for the large deviation rate functions:

$$\mathsf{H}_1(\beta_1) = \left(\frac{2(1-\alpha)}{\eta\Sigma_1(\bar{\beta}_1)}\right)(\beta_1-\bar{\beta}_1)^2 \tag{4.23}$$

$$\mathsf{H}_2(\beta_2) = \left(\frac{2(1-\alpha)}{\eta\Sigma_2(\bar{\beta}_2)}\right)(\beta_2-\bar{\beta}_2)^2 \tag{4.24}$$

where $\Sigma_i(\cdot)$ is the variance of the least-squares orthogonality condition associated with Model-i (evaluated at the RPE).

The dominant recursive learning model will be the model with the smaller $\Sigma(\cdot)$, which precisely means that the model that fits better. Consequently, from the neighborhood of the self-confirming equilibrium, it will be harder to escape from, and so will be the dominant recursive learning model. If the shocks are Gaussian, this will be entirely determined by the relative variances of w_1 and w_2, along with the relative magnitudes of β_1 and β_2. All else equal, Model 1 will be dominate if $\Omega_{11} > \Omega_{22}$.

4.4. **Sargent (1999).** Assume the government selects a target inflation rate, x_t, to minimize the social cost of realized inflation, y_t, and unemployment, u_t:

$$\min_{x_t} \mathsf{E}(1-\delta)\sum_{t=1}^{\infty}(y_t^2+u_t^2)\delta^{t-1} \tag{4.25}$$

for a given discount factor $\delta \in (0,1)$. The true law of motion governing $\{u_t, y_t\}$ is

$$\begin{aligned} u_t &= u^* - \theta(y_t - x_t^e) + v_{1t} \\ y_t &= x_t + v_{2t} \\ x_t^e &= x_t \end{aligned} \tag{4.26}$$

where x_t^e is the private sector's expectation of x_t. The last condition is referred to as the 'Fed watcher' assumption.[10] It is motivated by the idea that the private sector knows the government's policy rule. We assume that the perturbation terms $\{v_{1t}, v_{2t}\}$ are i.i.d. Gaussian over time with variance σ_i^2 $(i=1,2)$.

Suppose that the government's model class $\mathcal{M}$ consists of linear Phillips curve models. A generic element of $\mathcal{M}$ can be written as follows:

$$u_t = \gamma_{0,t} + \gamma_{1,t}y_t + \sum_{\ell=1}^{k_y}\lambda_{y,\ell,t}y_{t-\ell} + \sum_{\ell=1}^{k_u}\lambda_{u,\ell,t}u_{t-\ell} + \xi_t \tag{4.27}$$

where $k = k_y + k_u$ for some $k \in \{0,1,\ldots,\overline{K}\}$, or more compactly,

$$u_t = \mathsf{Z}_{k,t}\beta_{k,t} + \xi_{k,t}$$

where

$$\begin{aligned} \mathsf{Z}_{k,t} &= (1, y_t, y_{t-1}, \ldots, y_{t-k_y}, u_{t-1}, \ldots, u_{t-k_u}), \\ \beta_{k,t} &= (\gamma_{0,t}, \gamma_{1,t}, \lambda_{y,1}, \ldots, \lambda_{y,k_y}, \lambda_{u,1}, \ldots, \lambda_{u,k_u})', \end{aligned}$$

[10] In Cho and Kasa (2006), we show that this assumption does not alter the model's large deviations properties.

and ξ_t is a regression residual. In addition to the functional form, the government assumes that

$$\xi_t \sim N(0, \sigma_\xi^2)$$

and independent over $t \geq 1$.

Note that $\beta_{0,t}$ corresponds to the short term Phillips curve with no lagged variables. One can include a model with only a constant (vertical short term Phillips curve) in $\mathcal{M}$ without changing the key properties of the validation dynamics. In what follows, we focus on the restricted model class that excludes lagged unemployment from the right-hand side of (4.27). We shall come back to discuss dynamic Phillips curve models after completing the analysis of static Phillips curves.

The government recursively estimates the regression coefficients using the least squares estimator:

$$\beta_{k,t+1} = \beta_{k,t} + \eta R_{k,t}^{-1} \mathsf{Z}_{k,t} \hat{\epsilon}_{k,t} \tag{4.28}$$

$$R_{k,t+1} = R_{k,t} + a\left[\mathsf{Z}_{k,t}\mathsf{Z}_{k,t}' - R_{k,t}\right] \tag{4.29}$$

where

$$\hat{\epsilon}_t = u_t - \mathsf{Z}_{k,t}\beta_{k,t}$$

and $\eta > 0$ is a constant gain parameter. Sargent and Williams (2005) discuss conditions under which constant gain recursive learning approximates Bayesian updating. Given $\beta_{k,t}$, the government sets x_t by solving (4.25).

Let

$$\dot{\beta}_k = \Psi_k(\beta_k) \tag{4.30}$$

be the associated ordinary differential equation (ODE), and β_k^s be the stationary solution of the ODE, which is the self-confirming equilibrium.

Due to the LQR structure of the government's problem, its policy function takes the following linear feedback form

$$x_t = b_k(\beta_{k,t}) \cdot \mathsf{Z}_{k,t}$$

where x_t is the target inflation rate when using model k. Realized inflation is then

$$y_t = b_k(\beta_{k,t}) \cdot \mathsf{Z}_{k,t} + v_{2t}.$$

For $k > k'$, define $\mathbf{B}_{k'k}$ as a linear operator that selects the first $k'+2$ components out of $k+2$ elements: $\forall \beta \in \mathbb{R}^{k+2}$ $\mathbf{B}_{k'k}\beta$ is the first $k'+2$ components of β. Let $(\mathbf{B}_{k,k}\beta_k, 0, \ldots, 0) \in \mathbb{R}^k$ be the vector obtained by replacing the last $k - k'$ elements in β_k by zero's.

A straightforward calculation shows that

$$b_{k'}(\beta_{k',t}) = b_k(\mathbf{B}_{k'k}\beta_{k,t}, 0, \ldots, 0)$$

if $\beta_{k',t} = \mathbf{B}_{k'k}\beta_{k,t}$. Note that (2.4) for $\beta_{k',t}$ is

$$\begin{aligned} \beta_{k',t} &= \beta_{k',t-1} + \eta R_{k',t-1}^{-1} \mathsf{Z}_{k',t}'(u_t - \mathsf{Z}_{k',t}\beta_{k',t-1}) \\ R_{k',t} &= R_{k',t-1} + \eta\left(\mathsf{Z}_{k',t}'\mathsf{Z}_{k',t} - R_{k',t-1}\right). \end{aligned}$$

If $\beta_{k',1} = \mathbf{B}_{k'k}\beta_{k,1}$, then we can write the recursive formula as

$$\begin{aligned} \mathbf{B}_{k'k}\beta_{k,t} &= \mathbf{B}_{k'k}\beta_{k,t-1} + \eta R_{k',t-1}^{-1}\mathbf{B}_{k'k}\mathsf{Z}_{k,t}'(u_t - (\mathbf{B}_{k'k}\mathsf{Z}_{k,t})(\mathbf{B}_{k'k}\beta_{k',t-1})) \\ R_{k',t} &= R_{k',t-1} + \eta\left([\mathbf{B}_{k'k}\mathsf{Z}_{k',t}]'[\mathbf{B}_{k'k}\mathsf{Z}_{k',t}] - R_{k',t-1}\right). \end{aligned}$$

Thus, the associated ODE for $\beta_{k'}$ $\dot{\beta}_{k'} = \Psi_{k'}(\beta_{k'})$ can be written as $\dot{\beta}_{k'} = \Psi_{k'}(\mathbf{B}_{k'k}\beta_k)$. The dynamics of model k' can be viewed as the dynamics of model k with a constraint that the remaining $k - k'$ variables are set to 0.

Note also that $\forall k \in \{1, \ldots, \overline{K}\}$, $\dot{\beta}_k = \Psi_k(\beta_k)$ has a unique globally stable point β_k^s. Moreover, $b_k(\beta_k^s) = b_{k'}(\beta_{k'}^s)$ for $k \neq k'$. One can easily calculate

$$\beta_k^s = (\gamma_0^s, -\theta, 0, \ldots, 0)$$

where $\gamma_1^s = -\theta$ and $\gamma_0^s = \theta u^*/(1+\theta^2)$. Thus, $\beta_{k'}^s = \mathbf{B}_{k'k}\beta_k^s$. At the self-confirming equilibrium, (u_t, y_t) is distributed at the same center, and its distribution can differ only in the tails.

These two properties - (1) that the dynamics of smaller models are "nested" within the dynamics of larger models with constraints, and (2) that every model has the same self-confirming equilibrium outcome, allow us to characterize the dominant recursive learning model as the smallest model that converges to the self-confirming equilibrium outcome.

Proposition 4.1. *The dominant recursive learning model contains no lagged variables variables* ($k = 0$).

Proof. Fix k and the self-confirming equilibrium β_k^s. Fix a small positive number ρ, and let $N_{k,\rho}(\beta_k^s)$ be the ρ neighborhood of β_k^s. Recall (2.6), which is a rare event. We use an alternative characterization of the rate function $\mathsf{H}^*(\mu_{\beta_k^s})$ through S-functional.

Define the H-functional as

(4.31)

$$\Lambda(\alpha, \beta_k^s, t) = \lim_{\tau \to 0} \sup \lim_{\eta \to 0} \sup \frac{\eta}{\tau} \log \mathsf{E}\left[\exp\langle \alpha, \sum_{k=1}^{\lceil \tau/\eta \rceil} (R_{k,t})^{-1} X_{k,t}(u_t - X_{k,t}\beta_{k,t}\rangle \mid \beta_{k,t} = \beta_k^s, \mathcal{H}_t \right],$$

where $\mathcal{H}_t$ is the sigma algebra generated by information at t. The Legendre transform of the H-functional is defined as

$$\Lambda^*(\beta, \zeta, t) = \sup_{\alpha} \left[\langle \alpha, \zeta \rangle - \Lambda(\alpha, \beta, t) \right] \tag{4.32}$$

and the action functional is then defined as

$$S(\beta, T, \phi) = \int_0^T \Lambda^*(\dot{\phi}, \phi, t) dt \tag{4.33}$$

where $\phi(0) = \beta^q$ and ϕ is absolutely continuous; otherwise, $S(\beta, T, \phi) = \infty$. Recall the definition of $N_\rho(\beta_k^s, T)$ in (2.6). The key large deviation parameter is obtained by solving

$$S^*(\beta_k^s, T) = \inf_{\phi \in N_\rho(\beta_k^s, T)} S(\beta, \phi, T). \tag{4.34}$$

Dupuis and Kushner (1989) proved that

$$S^*(\beta_k^s, T) > 0$$

which is precisely the rate function.

Note that if $k' < k$, then the set of escape paths is precisely the projection of $N_\rho(\beta_k^s, T) \subset \mathbb{R}^{k+2}$ onto $\mathbb{R}^{k'+2}$. Thus, the set of possible escape paths becomes smaller as the model

dimension becomes smaller. As a result, the solution of the associated minimization problem (4.34) becomes larger as the model dimension becomes smaller: if $k' < k$, then

$$S^*(\beta_k^s, T) \leq S^*(\beta_{k'}^s, T).$$

□

Since the decision maker has only a limited computational capability, he behaves as if he prefers simpler models as long as they fit reasonably well.

Note that dominance is evaluated only around the self-confirming equilibrium, because $\beta_{k,t}$ spends most of its time around the self-confirming equilibrium. Along the path toward the self-confirming equilibrium, the validation dynamics may selects a model other than the dominant one, which eventually switches to the dominant model near the self-confirming equilibrium.

Cho, Williams, and Sargent (2002) show that along the dominant escape path the constant gain algorithm of $\beta_{0,t}$ (the dominant recursive learning model) experiences long excursions away from the self-confirming equilibrium. Along the escape path from the high inflation self-confirming equilibrium outcome to the low inflation Ramsey outcome, the decision maker finds it beneficial to add more lagged variables in order to soak up the model's autocorrelated residuals. The same reasoning applies to the path converging to the self-confirming equilibrium as depicted in Figure 1.

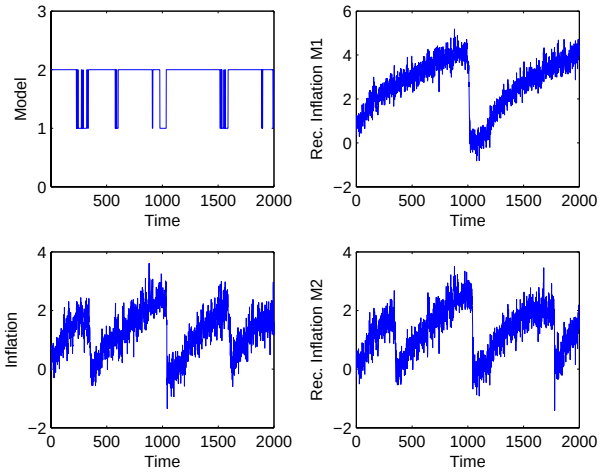

Figure 1: Validation and Recursive Learning Dynamics: $\eta = .02$

In the simulation we set $\overline{K} = 2$, so that the model contains at most two lags of inflation. We also set the threshold for the testing procedure so that the present model is maintained unless the the probability of the alternative is at least 10% higher. The left-half of Figure 1 plots the validation dynamics. The upper half plots the sequence of selected models and the bottom half plots the resulting inflation path. The right-hand side of Figure 1 then reports the standard constant gain recursive learning dynamics. The upper half plots

inflation when a simple static model is used, and the bottom half plots inflation when a more complex dynamic model is used.

Around the self-confirming equilibrium, the government realizes that even the smallest model fits the data well. Once the smallest model is selected around the self-confirming equilibrium, the government continues to use the model, because of its resiliency, until the dynamics experience another escape episode from the self-confirming equilibrium.

Interestingly, it turns out that the escape dynamics are completely dictated by the dominant model, which is selected only around the self-confirming equilibrium, although more complex models are used most of the time. Because the dominant model has the largest rate function, escapes are considerably less likely than in larger models. Put differently, if the government is using a smaller model, then escapes will appear to occur more slowly.

This observation delivers an alternative explanation of a puzzle identified by Cogley and Sargent (2004). There is an ample evidence that the Fed actually began to suspect the short-term Phillips curve was misspecified by the mid-1970s, several years *before* the Volcker disinflation. Cogley and Sargent (2004) attribute this policy inertia to Bayesian model uncertainty. The Fed stuck to a high inflation policy because it risk-dominated the policy implied by the better-fitting model.

Our result provides an alternative interpretation of policy inertia. Even if a model is rejected, it is typically the case that a similar model is the best fitting model. That is, even though a tail-sensitive test, like our relative entropy test, may detect a change, estimators like least squares or maximum likelihood, which focus on fitting the center of the distribution, will typically dictate a modest model revision. Drastic policy changes only take place at self-confirming equilibria, where model rejections are surprising, and therefore, informative. In particular, if the monetary authority constructs a model by judiciously selecting the explanatory variables, then the escape from the self-confirming equilibrium would be slower than otherwise. In our model, the inertia is a consequence of bounded rationality of the decision maker who is trying to economize his limited computational capability in building and forecasting the short term Phillips curve.

5. Robust Model Validation

Recently there has been a debate between Sims (2001) and Cogley and Sargent (2001) about the presence of regime changes in U.S. inflation data. Sims points out that if agents fit models with homoskedastic error terms when in fact the data are heteroskedastic, they may be fooled into inappropriately inferring that there have been breaks in the data, and as a result, inappropriately reject their models. Sims' point is relevant for us too. Until now we have assumed the decision maker knows the model's error distribution, even if he doesn't know its parameters. What if his beliefs about this distribution are wrong? If he ignores this possibility then he exposes himself to the kind of error that Sims highlighted. Our goal here is not just to model policymakers as econometricians, but to model them as *good* econometricians. Good econometricians worry about *robustness*.

If the only thing he had to worry about was parameter estimation and distributional uncertainty, there would be a straightforward response - use GMM rather than maximum likelihood. However, the agent we model is not just an econometrician, he is a decision

maker who *uses* his model to devise a control policy. Moreover, this control policy influences the data-generating process.[11] This means that he must make enough assumptions about his environment that he can solve his control problem, and this necessarily exposes him to greater specification risk than if he just needed to estimate parameters.

Our approach to this problem is to blend the recent literature on robust control and filtering (Hansen and Sargent (2006b)) with the recent literature on robust inference in moment condition models (Kitamura and Stutzer (1997)). We do this by building on Pandit (2004) and Pandit and Meyn (2004). As in the econometric literature on information-theoretic GMM and empirical likelihood, we define models by parameterized *moment conditions*. However, for us, these moment conditions do not come from economic theory; they define a permitted class of *model perturbations*, within which Hansen and Sargent's 'evil agent' can select a model to subvert the agent's model validation and control efforts. The more moment conditions there are, the less freedom the evil agent has, and hence, the less robust will be the outcome.

An important by-product of a robust model validation approach is that it endogenously delivers a 'robustified' reference model. Existing work on robust control is silent about where the robustness-seeking agent's reference model comes from. It only considers perturbations to a *given* reference model. In addition, due to the links between KLIC and Type I and Type II error rates, our approach endogenously generates *detection errors*. Current work on robustness specifies these errors exogenously, as a device to calibrate 'reasonable' amounts of robustness. Instead, our approach exogenously specifies a set of moment conditions.

Let us now drop the assumption that the decision maker knows the exact distribution of ξ_t so that the decision maker faces some form of model uncertainty. In this case, it is natural for the decision maker to pursue some form of robustness in the validation and decision making process. We formulate the robust validation process following the framework of Pandit (2004) and Pandit and Meyn (2004). Let ϕ be the marginal distribution of X. If the decision maker knows the distribution of ξ_t, he can calculate the probability distribution

$$\sum_{\ell=0}^{\overline{\ell}} Z_{t-\ell}\beta_{t-\ell}.$$

However, we now assume the decision maker has only partial information about the distribution. Instead of the exact distribution, the decision maker knows only a few moments. For example, consider the following moment-constrained set of parameterized models

$$\mathbb{P}(\beta) = \left\{\phi \mid \mathsf{E}^{\phi} Z\beta = 0, \quad \text{and} \quad \mathsf{E}^{\phi}(Z\beta)^2 = \sigma_{\xi}^2\right\}$$

where $\sigma_{\xi}^2 = \mathsf{E}\xi_t^2$. In this case, the decision maker does not know the exact distribution of the regression disturbance. He only knows the first two moments. The number of restrictions imposed on the moment class can be interpreted either as an expression of the decision maker's bounded rationality or as an expression of his preference for robustness. If he knows the correct distribution of ξ, he must know every moment of ξ, and therefore,

[11] Remember, however, that in keeping with our assumption of bounded rationality, we assume the agent ignores this feedback.

the moment class is subject to an infinite number of constraints. The finiteness of the constraints can be interpreted as a bound on his capacity to process information, which exposes him to model uncertainty.

For $\rho_1 > 0$ and a probability distribution ϕ on $\mathbf{X}$, define

$$\mathcal{Q}_{\rho_1}(\phi) = \left\{\phi' \mid \mathsf{H}(\phi'\|\phi) < \rho_1\right\}$$

and

$$\mathcal{Q}_{\rho_1}(\mathbb{P}) = \bigcup_{\phi\in\mathbb{P}} \mathcal{Q}_{\rho_1}(\phi).$$

For a small $\rho_1 > 0$, we can interpret $\mathcal{Q}_{\rho_1}(\phi)$ as the class of models which are difficult to differentiate from ϕ. Similarly, $\mathcal{Q}_{\rho_1}(\mathbb{P})$ is the class of models that are difficult to distinguish from models in $\mathbb{P}$.

The robust validation process can now be defined as follows. Let ϕ_k be the probability distribution over $\mathbf{X}$ induced by the present reference model parameterized by γ_k. Given a "smooth" empirical distribution $\hat{\mathcal{M}}_t$, define

$$\mathcal{L}(\hat{\mathcal{M}}_t, \gamma_k) = \inf_{\phi\in\mathbb{P}(\gamma_k)} \mathsf{H}(\hat{\mathcal{M}}_t\|\phi)$$

as the "worst case" relative entropy over $\mathbb{P}(\gamma_k)$. If

$$H(\hat{\mathcal{M}}_t|\gamma_k) < \rho_2,$$

then the decision maker uses γ_k to solve

$$\sup_{\mathbf{u}_t} \inf_{\phi\in\mathbb{P}(\gamma_k)} \mathsf{E}^{\phi}(1-\delta)\sum_{k'=1}^{\infty} \delta^{k'-1} U(u_{t+k'}, Z_{t+k'}) \tag{5.35}$$

where $U(\cdot)$ is the one period payoff, and $\mathbf{u}_t = (u_t, u_{t+1}, \ldots)$ is the sequence of controls. If

$$\mathsf{H}(\hat{\mathcal{M}}_t|\gamma_k) \geq \rho_2,$$

then γ_k is discarded, and a new reference model γ_{k+1} is constructed by solving

$$\sup_{\gamma} \inf_{\phi\in\mathbb{P}(\gamma)} \mathsf{H}(\hat{\mathcal{M}}_t\|\phi).$$

With γ_{k+1} in place of γ_k, the decision maker solves (5.35).

6. Concluding Remarks

This paper has attempted to model macroeconomic policymakers as econometricians. We've done this by combining recent work in both macroeconomics and econometrics. From macroeconomics, we've borrowed from the work of Sargent (1993) and Sargent (1999) on boundedly rational learning dynamics. From econometrics, we've borrowed from recent work on robust hypothesis testing (Zeitouni and Gutman (1991b), Pandit (2004) and Pandit and Meyn (2004)) and the analysis of misspecified models (Vuong (1989), Rivers and Vuong (2002), and Hansen and Sargent (1993)). As it turns out, this produces a rather difficult, and as yet unconsummated, marriage.

From a macroeconomic standpoint, it is difficult because we abandon the Rational Expectations Hypothesis, thereby putting ourselves into the 'wilderness of bounded rationality'. We do this not because we like to analyze difficult and ill-posed problems,

but simply because of the casual observation that, as econometricians, macroeconomic policymakers do not spend their time refining estimates of a known model, but instead spend most of their time searching for new and better models. Of course, it is not *necessary* to abandon Rational Expectations and traditional Bayesian decision theory when confronting model uncertainty.[12] However, we think there are good reasons to explore alternative approaches.[13]

The marriage between macroeconomics and econometrics is difficult from an econometric standpoint because, presumably, policymakers have some influence over the data-generating processes they are attempting to learn about. The econometric analysis of misspecified models with endogenously generated data is truly uncharted territory.

We make progress on this problem by relating it to a problem that *is* relatively well understood, namely, the dynamics of constant gain recursive learning algorithms. We prove that as the government employs an increasingly stringent specification test, the dynamics generated by a process of testing and model revision, which we call *validation dynamics*, converge in a very strong way to the dynamics generated by recursive learning models. This is a useful connection to make, because it enables us to apply the results of Williams (2001) and Cho, Williams, and Sargent (2002) on escape dynamics to help us understand a wide range of Markov-switching macroeconomic dynamics. Looking at it from the other side, a second payoff from making this connection is that it provides a more secure behavioral foundation for recursive learning models.

Although we feel this paper takes a significant step forward in understanding the interplay between macroeconomics and econometrics, there are certainly many loose ends and unexplored avenues remaining. Perhaps the most promising one is to follow-up on the connections between robust validation, robust control, and robust inference in moment constrained models.

[12]See Brock, Durlauf, and West (2004) for Bayesian model averaging.

[13]See Hansen and Sargent (2006b), Kreps (1998), and Bray and Kreps (1987).

References

Akaike, H. (1974): "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control*, 19, 716–723.

Benaim, M. (1996): "A Dynamical System Approach to Stochastic Approximations," *SIAM Journal of Control and Optimization*, 34, 437–472.

Borkar, V. S., and S. P. Meyn (2000): "The O.D.E. Method for Convergence of Stochastic Approximation and Reinforcement Learning," *SIAM Journal of Control and Optimization*, 38(2), 447–469.

Branch, W. A., and G. W. Evans (2005): "Model Uncertainty and Endogenous Volatility," University of Oregon.

Bray, M., and N. Savin (1986): "Rational Expectatons Equilibria, Learning, and Model Specification," *Econometrica*, 54, 1129–60.

Bray, M. M., and D. M. Kreps (1987): "Rational Learning and Rational Expectations," in *Arrow and the Ascent of Modern Economic Theory*, ed. by G. R. Feiwel, pp. 597–625. New York University Press.

Brock, W. A., S. N. Durlauf, and K. D. West (2004): "Model Uncertainty and Policy Evaluation: Some Theory and Empirics," University of Wisconsin.

Burnham, K., and D. Anderson (2002): *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer, 2nd edn.

Chen, H.-F., and J.-F. Zhang (1990): "Identification and Adaptive Control for Systems with Unknown Orders, Delay, and Coefficients," *IEEE Transactions on Automatic Control*, 35, 866–877.

Chen, X., and H. White (1998): "Nonparametric Adaptive Learning with Feedback," *Journal of Economic Theory*, 81.

Cho, I.-K., and K. Kasa (2006): "Learning Dynamics and Endogenous Currency Crises," Simon Fraser University.

Cho, I.-K., N. Williams, and T. J. Sargent (2002): "Escaping Nash Inflation," *Review of Economic Studies*, 69, 1–40.

Chu, J., M. Stinchcombe, and H. White (1996): "Monitoring Structural Change," *Econometrica*, 64, 1045–1065.

Cogley, T., and T. J. Sargent (2001): " Evolving Post World War II U.S. Inflation Dynamics," in *NBER Macroeconomics Annual, 2001*, ed. by B. Bernanke. MIT Press.

——— (2004): "The Conquest of U.S. Inflation: Learning and Robustness to Model Uncertainty," UC Davis and New York University.

Dembo, A., and O. Zeitouni (1998): *Large Deviations Techniques and Applications*. Springer-Verlag, New York, 2nd edn.

Dupuis, P., and H. J. Kushner (1989): "Stochastic Approximation and Large Deviations: Upper Bounds and w.p.1 Convergence," *SIAM Journal of Control and Optimization*, 27, 1108–1135.

Easley, D., and N. Kiefer (1988): "Controlling a Stochastic Process with Unknown Parameters," *Econometrica*, 56, 1045–64.

El-Gamal, M. A., and R. K. Sundaran (1993): "Bayesian Economists...Bayesian Agents: An Alternative Approach to Optimal Learning," *Journal of Economic Dynamics and Control*, 17, 355–383.

Evans, G. W., and S. Honkapohja (2001): *Learning and Expectations in Macroeconomics*. Princeton University Press.

Foster, D. P., and H. P. Young (2003): "Learning, Hypothesis Testing and Nash Equilibrium," *Games and Economic Behavior*, 45, 73–96.

Fudenberg, D., and D. K. Levine (1993): "Self-Confirming Equilibrium," *Econometrica*, 61, 523–545.

Hansen, L. P., and T. J. Sargent (1993): "Seasonality and Approximation Errors in Rational Expectations Models," *Journal of Econometrics*, 55, 21–55.

——— (2006a): "Recursive Robust Estimation and Control Without Commitment," University of Chicago.

——— (2006b): *Robustness*. Princeton University Press.

Hansen, M. H., and B. Yu (2001): "Model Selection and the Principle of Minimum Description Length," *Journal of the American Statistical Association*, 96, 746–774.

Hemerly, E., and M. Davis (1989): "Recursive Order Estimation of Stochastic Control Systems," *Mathematical Systems Theory*, 22, 323–346.

Kitamura, Y. (2001): "Comparing Misspecified Dynamic Econometric Models Using Nonparametric Likelihood," Yale University.

Kitamura, Y., and M. Stutzer (1997): "An Information-Theoretic Alternative to Generalized Method of Moments Estimation," *Econometica*, 65(4), 861–874.

Kreps, D. M. (1998): "Anticipated Utility and Dynamic Choice," in *Frontiers of Research in Economic Theory: The Nancy L. Schwartz Memorial Lectures, 1983-1997*. Cambridge University Press.

Kushner, H. J., and G. G. Yin (1997): *Stochastic Approximation Algorithms and Applications*. Springer-Verlag.

Marcellino, M. (1999): "Model Selection for Non-Linear Dynamic Models," Bocconi University.

Marcet, A., and T. J. Sargent (1989): "Convergence of Least Squares Learning Mechanisms in Self Referential Linear Stochastic Models," *Journal of Economic Theory*, 48, 337–368.

Nishii, R. (1988): "Maximum Likelihood Principle and Model Selection when the True Model Is Unspecified," *Journal of Multivariate Analysis*, 27, 392–403.

Pandit, C. (2004): "Robust Statistical Modeling Based On Moment Classes With Applications to Admission Control, Large Deviations and Hypothesis Testing," Ph.D. thesis, University of Illinois at Urbana Champaign, University of Illinois, Urbana, IL, USA.

Pandit, C., and S. P. Meyn (2004): "Extremal Distributions and Worst-Case Large-Deviation Bounds," Submitted for publication.

Ploberger, W., and P. C. Phillips (2003): "Empirical Limits for Time-Series Econometric Models," *Econometrica*, 71, 627–673.

Rissanen, J. (1989): *Stochastic Complexity in Statistical Inquiry*. World Scientific.

Rivers, D., and Q. Vuong (2002): "Model Selection Tests for Nonlinear Dynamic Models," *Econometrics Journal*, 5, 1–39.

Sargent, T. J. (1993): *Bounded Rationality in Macroeconomics*. Clarendon Press.

——— (1999): *The Conquest of American Inflation*. Princeton University Press.

Sargent, T. J., and N. Williams (2005): "Impacts of Priors on Convergence and Escapes from Nash Inflation," *Review of Economic Dynamics*, 8, 360–391.

Schwarz, G. (1978): "Estimating the Dimension of a Model," *Annals of Statistics*, 6, 461–464.

Sims, C. A. (2001): " Comment on Sargent and Cogley's 'Evolving Post World War II U.S. Inflation Dynamics'," in *NBER Macroeconomics Annual, 2001*, ed. by B. Bernanke. MIT Press.

Sin, C.-Y., and H. White (1996): "Information Criteria for Selecting Possibly Misspecified Parametric Models," *Journal of Econometrics*, 71, 207–225.

Vuong, Q. H. (1989): "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometica*, 57(2), 307–333.

White, H. (1982): "Maximum Likelihood Estimation of Misspecified Models," *Econometrica*, 50, 1–25.

——— (1994): *Estimation, Inference and Specification Analysis*. Cambridge University Press.

Williams, N. (2001): "Escape Dynamics in Learning Models," Ph.D. thesis, University of Chicago.

Yin, G. G., and V. Krishnamurthy (2005): "Least Mean Square Algorithms with Markov Regime Switching Limit," *IEEE Transactions on Automatic Control*, 50(5), 577–593.

Zeitouni, O., and M. Gutman (1991a): "Correction to: "On universal hypotheses testing via large deviations"," *IEEE Trans. Inform. Theory*, 37(3, part 1), 698.

——— (1991b): "On universal hypotheses testing via large deviations," *IEEE Trans. Inform. Theory*, 37(2), 285–290.

Department of Economics, University of Illinois, 1206 S. 6th Street, Champaign, IL 61820 USA

E-mail address: inkoocho@uiuc.edu

URL: http://www.business.uiuc.edu/inkoocho

Department of Economics, Simon Fraser University, 8888 University Drive, Burnaby, B.C. Canada. V5A 1S6

E-mail address: kkasa@sfu.ca

URL: http://www.sfu.ca/~kkasa