Econometric Evaluation of Rational Belief Models *

Victor Chernozhukov[†] Sergei Morozov

STANFORD UNIVERSITY

Version: September 10, 1999,
Revised: January 31, 2000
Very Preliminary and Incomplete [‡]

Abstract

The paper proposes a method for construction, estimation, and testing the Rational Beliefs (RB) models. RB models, due to Kurz (1994b), allow agents' beliefs to *differ* from the Rational Expectations (RE), but require that beliefs cannot be contradicted by past data. By implication, RB and RE must agree in strictly stationary worlds, while a disagreement is allowed in non-stationary setting. The estimation method involves sample counterparts to the conditional and unconditional moment restrictions formed from the Euler equations and rationality conditions. In essence, the method deduces systems of conditional beliefs consistent with the conditional moment restriction posed by the Euler equations. Consistent test statistics then discriminates the rationality from non-rationality. The attractive features are (i) the estimation and testing procedures are implemented without solving explicitly for RB equilibria, (ii) learning is permitted, and (iii) both the econometrician and the economic agents are put on the "equal footing" in the sense of Muth (1961), and "down to earth". Under flexible regularity conditions, the test statistics are shown to converge in distribution to the continuous functionals of generalized Brownian bridges, whose coordinates are projections on the space of moment functions that are used to phrase the rationality conditions. As a result, the limit distributions are non-standard or standard, depending on whether the test statistic is itself a function of finite-dimensional projection or a functional of the whole process, respectively. The resampling and simulation methods allow for valid approximation of either distribution. A simple estimated model of aggregate consumption and stock market behavior, populated by investors with rational beliefs, points to the variation in agents' sentiments as a dominant source of asset price volatility.

*We thank Mordecai Kurz for energetic support and invaluable advise and comments. We are indebted to seminar participants at Stanford who contributed comments and suggestions.

[†]Corresponding author. E-mail address: vchern@stanford.edu, Mail address: Victor Chernozhukov, Department of Economics, Stanford CA 94305-6072

[‡]Please, do not circulate or cite without permission of the authors. Comments are greatly appreciated.

JEL classsfication: C52, D8, E21
Keywords: Rational Beliefs, Sentiments

1 Introduction

The Rational Beliefs (RB) models, introduced in Kurz (1994b), allow agents to hold beliefs that *differ* from the Rational Expectations (RE). These beliefs, however, should satisfy the key coherency condition, termed the Rationality condition: the beliefs can not be contradicted with the past data. The principle does not imply Rational Expectations.

Precisely, non-stationarity of the world allows a Rational Belief and Rational Expectation to disagree. In RB models, RE represent an RB, as a special case. The principal sources of non-stationarity of the economy can be either exogenous, such as technological shocks, or endogenous, such as the perceived non-stationarity in the beliefs[1].

The RE models were designed to represent the situation in which the knowledge of economic agent is superior: the true probability law of motion is assumed to be known or approximately known by the agents. Such requirement turns out to be highly problematic when the economies are *significantly* non-stationary. Non-stationarity, *in principle*, precludes the very possibility of learning the true probability law of motion, even if the infinite history is available.

Driven by the failure of RE to explain puzzling features of macroeconomic dynamics, such as equity premium, consumption and stock price volatility, forward discount bias, to name a few[2], many attempts at rigor have been made to "drift" away from RE in various ways. Models of learning, for example, recognized the approximate nature of Rational Expectations Equilibrium (REE)[3]. In the limit of the learning dynamics, REE is typically recovered. Analysis of investor behavior and patterns of stock price reaction to news such as earnings announcements has lead some researchers to develop models of investor sentiment[4] which, coupled with Bayesian learning, were consistent with evidence of *underreaction* to news but *overreaction* to a *series* of good or bad news. In spirit, the literature on investor sentiment is close to RB approach but for the motivation it is derived from: sentiment models are rooted in elements of psychology while the central requirement of RB

[1]An example due to Kurz (1994a) shows that even if the exogenous environment is completely stationary, if agents *believe* in non-stationarity, the resulting dynamics of prices *is* non-stationary. This implies a history from which we can not learn the true probability law of motion.

[2]Long series of empirical papers have documented the issue. See, for example, Hansen and Singleton (1982), Mehra and Prescott (1985), Hansen and Jagannathan (1991), Froot and Frankel (1989), Burnside (1994), Geweke (1999).

[3]This includes Marcet and Sargent (1989), Anderson, Hansen, and Sargent (1999), Brock and LeBaron (1996), Epstein and Wang (1994), Cochrane (1989), Hansen and Sargent (1993), Krusell and Smith (1996). The list is by no means exhaustive. In Marcet and Sargent (1989), agents use the estimated transition laws of correct functional forms which they mistakenly take as non-random and time-invariant. In Cochrane (1989) and Krusell and Smith (1996), consumers use decision rules that are perturbed by small amounts in arbitrary directions form optimal ones.

[4]See Barberis, Shleifer, and Vishny (1998).

theory is rationality with respect to the observed past. Another brand of literature deals with the 'robust' decision-making approach[5]. In particular, Anderson, Hansen, and Sargent (1999), suggested a "robust" version of the dynamic decision-making, in which the decision-maker recognizes that the RE reference model is akin to an estimate provided by econometrician, and he therefore guards against the estimation uncertainty by choosing beliefs that are least-favorable to him in a neighborhood of the estimate. We argue later on that such a paradigm, with additional reinterpretation could be fit into a very specific class of the empirical RB models that we provide here. Indeed, 'robust' rules for choosing the *most pessimistic* beliefs is only one special way of assigning beliefs to the agents. Such beliefs could be rational in the sense of not being rejected by *rationality* specification tests we pose here, but the imposition of pessimism seems unwarranted. We like to allow for less gloomy worlds. Further, Anderson, Hansen, and Sargent explicitly prohibit learning to keep the misspecification error constant, in a sense[6]. The RB models, on the other hand, that we consider here, expressly allow for agents to learn both in the long run and short run (adaptation). Overall, deviating from RE formulations has allowed the models to enjoy some success in explaining one or the other puzzle.

On the other hand, the RB modeling, representing a significant departure from the RE framework, has been considerably successful. *One* unified and simple model of economy, with agents having Rational Beliefs, is able to accurately explain such seemingly unrelated and diverse phenomena as excessive volatility of prices (Kurz and Motolese, 1998; Kurz, 1998), risk and equity premia (Kurz and Beltratti, 1997), forward discount bias (Kurz, 1997b; Black, 1997) , dynamic money non-neutrality (Motolese, 1998)

[5]A sample of other recent papers include Gilboa and Schmeidler (1989), Hansen, Sargent, and Tallarini (1999), Anderson, Hansen, and Sargent (1999), and Epstein and Wang (1994). Gilboa and Schmeidler, and Epstein and Wang offer an axiomatic approach, in which the belief is represented by a family of probability measures and the 'minmax' criterion is utilized in decision making. In the formulation of Hansen, Sargent, and Tallarini, Epstein and Wang (1994), and Anderson, Hansen, and Sargent (1999), there is the "true" or reference model, around which a family of permissible misspecifications is stated. The agents then use a 'minmax' criterion to select the decision rule that would reflects the most pessimistic 'belief' in the class of beliefs allowed by *misspecification* . The families of misspecifications are stemmed from the robust control and information theory literature. In particular, these include H_∞, H_2, and maximum entropy methods. Equivalent formulations are derived from the risk-sensitive recursive formulation.

[6]In Anderson, Hansen, and Sargent (1999), the size of misspecification is interpreted as the preference for robustness parameter. On the other hand, in the stationary world with infinite data, learning can recover the true law of motion, and misspecification cannot arise. Hence, alternative view of the misspecification size in the context of Anderson, Hansen, and Sargent is to relate it primarily to the data size. To accommodate the size of misspecification varying with the size of the data, future developments would need to augment the model by some learning mechanism, perhaps at the expense of mathematically elegant recursive formulation. The econometric "empiricist" model developed below already incorporates the above model as a special case, but does not impute to the agents the "minmax" rules for the choice of their beliefs. Our econometric modeling avoids explicit solutions and thus avoids technicalities.

and other “puzzles” (see Kurz (1997a)). In simulation studies, simple models of RB Economies perform better than the models of RE Economies by the order of magnitudes.

To date, however, all of the above studies employed simulation, or simulation combined with calibration, as a main tool. Simulation approaches require explicit solution for RB equilibria and the computational and informational burden of doing so can be excessive. This limits the computational complexity of implementable models much beyond the finite state Markovian specification, put together by M. Kurz and his colleagues [7]. The approach requires the modeler to have knowledge of the true probability law and of the agent's beliefs on the infinite sequences. In the present work, we supply *simple* econometric weaponry to RB modelers in an effort to help better reveal the empirical content of RB models. This methodology can be applied to highly complex models, as the methodology avoids explicit solutions for RB equilibria. Importantly, it does not require the knowledge of the true law of motion and the beliefs on the infinite sequences. This is plausible, because, in the general nonstationary settings, the econometrician, even armed with the infinite data, can not deduce the true law of motion, as it requires of him to know the number of parameters at least of the order of cardinality of the data. Henceforth, in manufacturing the econometric methods, we pose the following principles, that the proposed econometric method must have.

The method should represent a joint economic and econometric paradigm in which:

(i) Agents and the econometrician do not know the true probability law, and they recognize this.

(ii) An agent has access to finite but increasing data. As an econometrician himself, he is allowed to draw inferences and accumulate knowledge, i.e. learn.

(iii) Beliefs of agents should to be rational, i.e. they should imply the same quantification of history as the history itself. The rationality conditions should be simply phrased.

(iv) The method should avoid the closed form solutions of stochastic equilibria, since these can only be obtained in special simple cases and with a knowledge of the true probability law of motion.

The proposed qualities (i) – (iv) agree with the principles of modeling advocated by Kurz (1994a) in requiring the above notion of rationality, and by Muth (1961), which states that the econometrician and the agents should be put on “equal footing” in the sense that both should have the same data and structural knowledge. Furthermore, the principle

[7] For the finite state Markov specification, the Kurz group has successfully harnessed the supercomputer power by adopting some advanced parallel computing techniques that allow their simulation studies to run on more than 50 Unix workstations simultaneously.

(ii) strips both the agents and econometrician of other god-like qualities, such as access to the infinite data.

It should be said, however, that the theoretical model suggested by Kurz implicitly postulates that the data is infinite by requiring the agents to know the stationary measure[8] (the learnable part of true law) and to coordinate their beliefs with it. In practice this translates into regarding the estimated stationary measure as the true stationary measure, and ignoring the estimation uncertainty (as irrelevant or very small). We shall refer to this model as to the *canonical* RB model, and this model will receive the *priority* in the empirical implementation and discussion. Explicitly built into our construct is, however, another model that incorporates and accounts for the additional estimation uncertainty resulting from the finiteness of the data. Importantly, it manifests itself in the asymptotic approximations to the distribution of the test statistics that checks the rationality of the agents. We shall refer to the second model as to the *empiricist* RB model. Aside minor subtleties, at the heart of both econometric models is the principle of compatibility with the data, the rationality principle proposed by Kurz (Kurz, 1997a).

Of course, we, the econometrician, always have to make some prior simplifications and assumptions about the structure of the economy and beliefs, e.g. allow ergodicity (to be able to conduct estimation and inference), assume bounded heterogeneity when characterizing non-stationarity (to enable derivation of the stochastic limit results for estimators and test statistics), assume some functional forms for relevant functions and environments. These sacrifices ought to be made to extract something intelligible from empirical data and make the theory generate testable implications. Apart from such general *a priori* assumptions, the methodology proposed here agrees with the described earlier principles.

Overview

The paper is organized as follows. The discussion of Rational Beliefs Models and its translation to finite (but increasing data) is given in section 2. The canonical and the empiricist RB models are phrased there in econometric terms. The section also presents the key basic building blocks of the empirical RB models.

Section 3 describes the novel econometric estimation and modeling methodology. This is the first methodology, to our knowledge, that offers ways of empirical investigation of RB models (other than the simulation methods). The approach avoids explicit solutions for the equilibria.

Section 4 is an application to a representative agent economy with a rational belief. The choice of representative agent situation is motivated by the availability of data, ease of implementation, and the resulting multitude of applications, such as security pricing,

[8]See definition in the next section.

decomposition of uncertainty, and the market price of risk.

To wit, subsection 4.1 formulates the modeling approach as applied to the standard consumption and asset pricing model. Furthermore, we develop a *quantitative* decomposition of uncertainty ("market price of risk") into two parts: one arising from variation in the intertemporal rates of substitution, as in the usual case, and the second due to "endogenous" uncertainty, arising from the inability to learn the true law of motion and having to use a Rational Belief in lieu of the true probability measure. This offers the quantitative and qualitative explanation of the "equity premium". The econometric methodology then applies to propose estimates of market price of risk and facilitate the comparisons with the observed empirical values. A particularly simple "log-normal" reference example is worked out in subsection 4.2, with two following subsections, 4.3 and 4.4, implementing prototypical empirical evaluation of RB using U.S. economy data. Section 5 offers concluding remarks. Appendix covers various technicalities that are glossed over to ease the exposition. In particular, it contains all formal definitions and technicalities related to the presence of learning in the data. It defines all estimators and test statistics. The null hypothesis of rationality is formalized, and the test statistics are offered. These test statistics are continuous functionals (with respect to the uniform metric) of the rationality score processes that represent the empirical processes whose finite dimensional projections form moment conditions for checking rationality. The resulting test statistic is either a "sup" statistics, in the spirit of Kolmogorov-Smirnov statistic, over all moment functions or a quadratic form of any finite-dimensional projection on the space of moment functions. The asymptotic distributions of such test statistics are derived under the null and the alternative, and the resulting test decisions have been demonstrated to be asymptotically consistent in discriminating the rationality vs. non-rationality of a deduced system of beliefs. The corresponding asymptotic distributions can be viewed either as the "sup" of a generalized Brownian bridge to which the rationality score converges or the quadratic form of the finite-dimensional projection of the above process.

2 The Basic Building Blocks of RB models

Basic Notation

Let $\{x_t\}$ be a stochastic process, describing the economy, on probability space $\left(\left(\mathbb{R}^N\right)^\infty, \mathcal{B}\left(\left(\mathbb{R}^N\right)^\infty\right), \mathbb{P}\right)$. $\mathbb{P}$ is the *true probability law*. A proper *belief* of an agent is the probability space $\left(\left(\mathbb{R}^N\right)^\infty, \mathcal{B}\left(\left(\mathbb{R}^N\right)^\infty\right), \mathbb{Q}\right)$, or simply $\mathbb{Q}$.

Generally, $\mathbb{Q}$ and $\mathbb{P}$ can be taken as distinct from each other probability measures. The measure $\mathbb{P}$ is assumed to generate the so called stationary or ergodic measure, which is defined by the limits of empirical distribution functions over all finite horizon events (see definition below). Denote this measure by $\mathbb{M}$. Throughout the paper we assume that the

$\mathbb{P}$ is *nonstationary* and *ergodic*, in the usual sense employed in the literature. Sequence x_t is assumed to be stable, i.e. detrended etc.

Decision Making and the Optimality Conditions
Discrete time models of optimal behavior of economic agents often lead to the first order conditions of the form

$$E_{\mathbb{Q}_t^{t+k}} h(x_{t+k}, \theta) = 0, \quad k = 1, 2, \ldots \tag{2.1}$$

where x_{t+k} is a finite-dimensional vector of the variables observed by the agents and econometrician as of date $t + k$, θ is finite-dimensional parameter unknown to econometrician, characterizing underlying preference or technology, h is a finite dimensional mapping and $E_{\mathbb{Q}_t^{t+k}}$ is an expectation operator with respect to probability measure $\mathbb{Q}_t^{t+k}$, which is the belief of an agent about the events k periods forward from the date t, conditional on his/her current information set I_t. Equations of this type can emerge from the first order necessary conditions of an agent's utility maximization problem in an uncertain and non-stationary dynamic environment. A detailed example is exhibited in the section 4. Throughout we assume that the econometrician observes the actions of the agent within last T_e periods of history, whose total length is assumed to be H (in the canonical model $H = \infty$). T_e here is allowed to be less than or equal to H in the empiricist RB model, whereas in the canonical RB model, it is negligible in relation to H.

Rationality: Compatibility with the Data
Kurz (1974; 1997a) has proposed to consider those beliefs Rational that are *compatible* with data. The belief, $\mathbb{Q}$, should not ignore the past history. In the theoretical and in the simulation studies of RB models the past history is always assumed to be infinite, the assumption we surely abandon here.

The notion of compatibility proposed by Kurz requires the *observed empirical frequencies to agree with the theoretical frequencies that the beliefs generate*. Henceforth, this is the *key principle* of our modeling. The rigorous definition given by Kurz is stated below.

Definition of Rationality – (Kurz, 1997a)

(I) *Stability* . A *dynamic system* is a pair of probability space and a shift transformation T (for $x^t \equiv (x_t, x_{t+1} \ldots)$, $Tx^t = x^{t+1}$, and $T^{-1}S = x : Tx \in S$). A dynamic system, $\left(\left(\mathbb{R}^N\right)^\infty, \mathcal{B}\left((\mathbb{R}^N)^\infty \right), \mathbb{J}, T \right)$, is *stable* if, for $\mathbb{M}_n(B)(x) \equiv n^{-1} \sum_{j=1}^{n-1} 1_B(T^j x)$, and any cylinder B

$$\tilde{\mathbb{M}}_{\mathbb{J}}(B) = \lim_{n \to \infty} \mathbb{M}_n(B)(x) \text{ exists } \mathbb{J} \text{ a.e.} \tag{2.2}$$

$\tilde{\mathbb{M}}_{\mathbb{J}}(B)$ can be uniquely extended to the *stationary* probability measure $\mathbb{M}_{\mathbb{J}}$ on all sets $\mathcal{B}\left((\mathbb{R}^N)^\infty\right)$.

(II) *Compatibility with Infinite Data*. $\mathbb{Q}$ is said to be compatible with the data generated by $\left(\left(\mathbb{R}^N\right)^\infty, \mathcal{B}\left((\mathbb{R}^N)^\infty\right), \mathbb{J}, T\right)$ if $\mathbb{M}_{\mathbb{Q}} = \mathbb{M}_{\mathbb{P}} \equiv \mathbb{M}$

A belief $\mathbb{Q}$ is said to be *Rational*, in the sense of Kurz, if the above conditions are satisfied together with the certain other technicalities. The reasons for conditions (I -II) are appealing : real world economy appears to be generally stable (I); whenever the limits in 2.2 are known, a rational believer forms his belief so as to not contradict the $\mathbb{M}_{\mathbb{P}}$ (II).

We do not employ directly the notion of compatibility posed, since it is very hard to work with econometrically. Our modelling approach will avoid (i) specification of the beliefs on the infinite sequences and (ii) specification of unconditional beliefs. It will only model the conditinal beliefs $\mathbb{Q}_t^{t+k}$ and conditional stationary measure $\mathbb{M}_t^{t+k}$, and therefore the rationality implications should be phrased in terms of restrictions on such systems:

Let H denote the length of history of the economy. As we take it to the limit, for a rich class of functions $\mathcal{G}$ (called convergence-determining class), for any $g \in \mathcal{G}$,

$$\begin{aligned}\lim_{H\to\infty} \frac{1}{H}\sum_{t=1}^{H} E_{\mathbb{Q}_t^{t+k}} g(x_{t+1},\ldots,x_{t+k}) &= E_{\mathbb{M}} g(x_{t+1},\ldots,x_{t+k}) \\ &= \lim_{H\to\infty}\frac{1}{H}\sum_{t=1}^{H} g(x_{t+1},\ldots,x_{t+k}), \quad \mathbb{P}\text{ a. e.}\end{aligned} \tag{2.3}$$

The above roughly states that the data generated by the system of conditional beliefs should yield the same averages as those historically observed. Under technical regularity condition this condition is equivalent to that of Kurz. These technical regularity conditions require $\mathcal{G}$ to be fairly broad, for example, $\mathcal{G}$ has to be a class of continuous or Lipshitz bounded functions. In some practical applications it is possible to relax regularity conditions and employ a narrower and even finite class of functions. For instance, in the "log-normal" example of section **4.3**, it is sufficient to consider a very narrow class of functions that consists of all polynomials of the second order, i.e $\mathcal{G} = \{g : x \to g(x) = x \text{ or } g(x) = xx'\}$. In this example, therefore, only the long term means and covariances have to agree with those generated by the the beliefs to guarantee that the above rationality condition holds for all g in the class of continuous and bounded functions. In general settings, however, $\mathcal{G}$ has to be very broad. This broadness makes the verification of the rationality condition by the agents himself difficult or impossible, as we explain below.

There are two key difficulties that an agent (and also econometrician) encounter when attempting to verify (2.3). First, consider an agent who has beliefs over the entire sequence of historical events. If H, the length of history, is finite, it is possible to compute the sample counterparts of the rationality condition (2.3) for any fixed g. This means that the estimation uncertainty *is* present when characterizing the long term moments (the learnable part of the history). In the canonical RB model, this estimation uncertainty is required to be zero or negligible.

Therefore, the proposition, that the data is finite, large, and increasing, forces us to phrase the *asymptotic* notions of the rationality condition (2.3). Furthermore, we have to phrase such conditions from the point of view of econometrician who observes the agent making choices in the last T_e periods of time.

This rephrasing is rather simple if the convergence determining class $\mathcal{G}$ is finite. Complications may arise when $\mathcal{G}$ is infinite. If the class is too broad, e.g. all continuous and bounded functions, we may look at sample counterparts of the above rationality conditions with many and possibly all functions g. It is then natural to use some metric to make the key statistical decision of discrimination between rational and non-rational belief systems. For example, take the "sup" over the difference between the sample averages generated by the beliefs and the observed historical averages. Unfortunately, this is useless econometrically even if the beliefs are rational, since such statistic, generically, would not converge to zero (in probability, and hence not a.e.)[9, 10]. Furthermore, suppose we want to consider the above "sup" statistic and its asymptotic distributions so as to employ the asymptotic statistical decision theory in discriminating rationality. The class $\mathcal{G}$ of all bounded continuous functions is not Donsker (again see van der Vaart and Wellner (1996)) , therefore eliminating the very possibility of existence of the asymptotic distribution[11].

The "equal footing" postulate requires the agent to be an econometrician, just as us. This principle and the above considerations lead us therefore to curb the class $\mathcal{G}$ and make it finite or "small", precisely, requiring $\mathcal{G}$ to be Glivenko-Cantelli or Donsker, depending

[9]This is because , the above class $\mathcal{G}$ of all bounded continuous functions is not Glivenko-Cantelli – see van der Vaart and Wellner (1996).

[10]This observation suggests, that rationality notion of canonical model can not be obtained, generically, in the limit of learning, i.e. through forcing the agent to construct the beliefs on the sequence that respect the sample counterpart of rationality conditions by taking sup over the functions. See importantly Hoffmann-Jørgensen (1994), van der Vaart and Wellner (1996), Hoffmann-Jørgensen (1991) for treatments of convergence, and generally the treatment of the probability "with a view towards statistics", where probability theory is built so that it can be always viewed as a limit of convergence of a statistical experiments. Similar observations and certain theoretical results as pertains to RB are obtained in Nielsen (1997a) and Nielsen (1997b). (Kurz, 1997a), p.10 responds to the critique, by saying that the class of sets over which the compatibility is to be defined can be shrunk. This is equivalent to the above observation prompting to "shrink" the class of functions. Furthermore, such the largest class obtains by considering exactly the Glivenko-Cantelli class. Some sufficient conditions for a function to belong to this class under non-stationarity and serial dependence are given in the appendix.

[11]The statistic then is not measurable, and is not asymptotically measurable. It is then impossible to assign distribution to it – see van der Vaart and Wellner (1996).

on the need[12].

Thus the econometric considerations lead us to the following much less general rationality requirements, which are phrased from the econometrcian's point of view. Suppose that the econometrician observes the agent's decisions that respect optimality conditions for T_e periods and is able to deduce the sample counterpart to the average 'generated' by the beliefs (the lhs of (2.3)). If class $\mathcal{G}$ is finite, then the rationality requirement is

$$\frac{1}{T_e-k}\sum_{t=H-T_e+1}^{H-k} E_{\mathbb{Q}_t^{t+k}} g(x_{t+1},\dots,x_{t+k}) \approx E_{\mathbb{M}} g(x_{t+1},\dots,x_{t+k}), \quad k=1,2,\dots,\forall g\in\mathcal{G}$$

Precisely, we only require that $\frac{1}{T_e-k}\sum_{t=H-T_e+1}^{H-k} E_{\mathbb{Q}_t^{t+k}} g(x_{t+1},\dots,x_{t+k})$ be consistent[13] for $E_{\mathbb{M}} g(x_{t+1},\dots,x_{t+k})$. In terms of probability limits (as opposed to a.e. limits) we require that updating of the conditional beliefs is done so that:

$$\begin{aligned}\mathbb{P}\lim_{H\to\infty,T_e\to\infty}\frac{1}{T_e-k}\sum_{t=H-T_e+1}^{H-k} E_{\mathbb{Q}_t^{t+k}} g(x_{t+1},\dots,x_{t+k}) &= E_{\mathbb{M}} g(x_{t+1},\dots,x_{t+k})\\ &= \mathbb{P}\lim_{H\to\infty}\frac{1}{H-k}\sum_{t=1}^{H-k} g(x_{t+1},\dots,x_{t+k}), \quad k=1,2,\dots\end{aligned} \tag{2.4}$$

For infinite class $\mathcal{G}$, e.g. the class of all polynomials or class of smooth functions, we require the Glivenko-Cantelli property to hold, that is, (2.4) to hold *uniformly for* $\forall g\in\mathcal{G}$: i.e., for any $\epsilon>0$[14] [15]:

$$\begin{aligned}\lim_{H\to\infty,T_e\to\infty}\mathbb{P}^*\Bigg(&\sup_{g\in\mathcal{G}}\left|\frac{1}{T_e-k}\sum_{t=H-T_e+1}^{H-k} E_{\mathbb{Q}_t^{t+k}} g(x_{t+1},\dots,x_{t+k}) - E_{\mathbb{M}} g(x_{t+1},\dots,x_{t+k})\right| > \epsilon,\\ &\sup_{g\in\mathcal{G}}\left|E_{\mathbb{M}} g(x_{t+1},\dots,x_{t+k}) - \frac{1}{H-k}\sum_{t=1}^{H-k} g(x_{t+1},\dots,x_{t+k})\right| > \epsilon\Bigg) = 0\end{aligned} \tag{2.5}$$

The assumption above is a *bare minimum* needed in order to obtain the key estimation and testing (discrimination of rationality) consistency results. Note that if $\mathcal{G}$ is finite, the prop-

[12]Generally, if class $\mathcal{G}$ has finite entropy integral (a measure of "complexity"), then it is Donsker (and therefore Glivenko-Cantelli), but the requirements are more stringent under dependence conditions (see appendix and e.g. (Andrews, 1989; Andrews, 1993)).

[13]The econometric regularity conditions make further assumptions to derive limit distributions for test statistics and estimators.

[14]To conduct asymptotic inference, we shall require the Donsker property to hold.

[15]$\mathbb{P}^*$ is the outer probability measure (see (van der Vaart and Wellner, 1996)) (in case the event is measurable, $\mathbb{P}^*$ is replaced by $\mathbb{P}$). Using $\mathbb{P}^*$ above is a measure-theoretic subtlety and should be ignored.

erty (2.5) holds from (2.4) automatically. The above condition applies to both the canonical model and the empiricist model[16]. We also stress that considering infinite classes is highly practical, contrary to what it may seem: indeed, we may take the class of functions parameterized by a parameter in a Euclidean space, e.g. $g(x,\delta) = x^{\delta}, \delta \in [a,b]$, then the check of the rationality condition would involve finding the least favorable solution δ^* in an infinite class of possibilities. The resulting distribution is certainly non-standard, but it can be approximated by bootstrap and simulation methods.

This section has introduced the basic building blocks of econometric RB models. The essence of the proposed methodology is to explore both the optimality conditions (2.1) and the rationality conditions (2.4) and (2.5) . In particular, by assuming parametric functional forms of the conditional beliefs and the conditional stationary measure, the method first explores the restrictions on the system of beliefs implied by the optimality conditions. By identifying plausible sub-families, the method offers the test statistics that can be used to test the rationality conditions (2.4) and/or (2.5). Further extensions are in section 4.

3 Structure of the Econometric Procedure

3.1 Specification of Conditional Beliefs and Conditional Stationary Measure

First of all, we shall only work with finite horizon events, meaning that only the Euler equation up to K periods ahead are considered, and K is fixed[17].

Canonical RB modeling requires the agents to offer non-stationary probability belief measure $\mathbb{Q}$ over infinite sequences. Furthermore, similar requirement is imposed on $\mathbb{P}$.

[16] The key issue distinguishing the two comes in when asymptotic distributions of the relevant test statistics that check the rationality are considered. The key difference is whether the variation of terms $\frac{1}{H}\sum_{t=1}^{H} g(x_{t+1},\ldots,x_{t+k})$ is considered as negligible or not. A careful reader would note that we omit another serious technical issue that distinguishes the formulation of rationality condition from the econometric point of view in the canonical vs. empiricist model. The correct form of the rationality condition in the empiricist model should in principle include additional index of H, i.e. the conditional beliefs should have the structure of the form $\{Q_{tH}^{t+k}, t = T_e,\ldots,H\}$. This, in principle, should be different from what the econometrician can conclude by observing the actions of the agents in the empiricist model, i.e. $\{Q_{th}^{t+k}, (t,h) = (T_e,T_e),\ldots,(H,H)\}$. Under additional assumptions, requiring the learning of the conditional stationary measure to be weakly progressive in relation to T_e, all concepts become equivalent. These technical issues are treated in the appendix.

[17] In principle, one can set K to grow logarithmically in the econometric sample size T_e. One can go even further and allow for polynomial growth, but this requires to work with concrete dependence settings. We are not going to explore this direction.

In our modeling we shall avoid such specification primarily to agree with the principles posed in the introduction. Instead, the approach involves specifying the conditional beliefs and stationary measure over *finite* horizon events. Furthermore, this modeling should allow non-stationarity. Therefore, the rationality conditions should be posed so that to form a testable implication in terms of such conditional measures. This has already been accomplished in section 2.

In our approach, the conditional beliefs are modeled as deviations from the conditional stationary measure. Such an approach is plausible, because, with availability of large data set, an econometrician can deduce the conditional stationary measure, and hence characterize an important part of the belief. Furthermore, we hope that the deviations from the conditional stationary measure would then be deducible by econometrician from the economic actions of agents. Even if deviations were not fully recoverable, certain more restricted systems of deviations could be deduced and used in the subsequent check of rationality and further economic analysis.

Henceforth, we suppose that the conditional belief $\mathbb{Q}_t^{t+K}$ on all events that happen in periods from $t+1$ to $t+K$ is given by or rather *well approximated by* a parametric probability measure

$$\mathbb{Q}_t^{t+K}(\cdot) = \mathbb{B}(\cdot, \tilde{\alpha}_t^K + l^K(I_t, \lambda_t^K)), \tag{3.6}$$

where $\mathbb{B}(\cdot, \pi)$, for any π, is a probability measure on the state space Ω. For any given t, and conditioning variables I_t (length of which is assumed to be bounded by J, for practical purposes[18]), the belief of an agent is characterized by two components: $\tilde{\alpha}_t^K$ and $l^K(I_t, \lambda_t^K)$, which we term here as the *subjective* and the *objective* components, or the *sentiment* and the *learning* structures, respectively.

The conditional belief is assumed to be tied to the conditional stationary measure as follows:

$$\mathbb{M}_t^{t+K}(\cdot) \equiv \mathbb{B}(\cdot, l^K(I_t, \lambda^K)),$$

In the canonical model, λ_t^K is equal to λ^K, and we distinguish the notation for the sentiment by using α_t^K in place of $\tilde{\alpha}_t^K$.

Note that the objective structure represents the estimator of the conditional stationary measure in the empiricist model. In that model $l^K(I_t, \lambda_t^K)$ represents the parametric learning and forecasting structure. λ_t^K is the estimator that the agent uses to assess this structure. The estimator is unobserved by the econometrician. To give an example, suppose $\mathbb{B}$ is a normal family with unit variance. Take $l^K(I_t, \lambda_t^K) = I_t'\lambda_t^K$ as the estimate

[18]This is justified, for example, if we assume that structural break arrival time is bounded by some constant J a.s.

of the mean, where λ_t^K is $t-$th iterate of the recursive least squares. Alternatively, the estimator λ_t^K used by the agent could be something else, e.g. an estimator derived by recursive least absolute deviations schemes. By using these alternative schemes, the agent could guard against misspecification errors (such as possibility of an ϵ-contamination of the Normal by the "witch" Cauchy distribution, which breaks down, for example, least squares based learning mechanisms). In any case, the choice of estimator is a personal matter, and it could be rationalizable by different rationality considerations: minimax, invariance, equivariance, and robustness principles. Therefore, we must allow for subjective, but *consistent* choice of the estimator. That is, we only require an additional constraint to hold:

$$\lambda_t^K \xrightarrow{\mathbb{P}} \lambda^K,$$

and also impose other technical assumptions, all delegated to the appendix. Thus, the *aim of long-term learning* in the empiricist model is to characterize the *conditional stationarity measure*.

The above considerations give also the precise interpretation of the "sentiment" parameter $\tilde{\alpha}_t^K$: it represents the subjective correction to the history-based forecast, in particular it absorbs:

Model or Endogenous Uncertainty. This involves pessimism and also optimism, relative to conditional stationary measure. Its source is a lack of knowledge of the true conditional probability that generally is not equal to the conditional stationary measure.

Adaptation. Short-term (non-convergent) learning, that we call *adaptation* is embedded in $\tilde{\alpha}_t^K$. In the settings where the economy is given by stationary environments, perturbed by the structural breaks, adaptation leads to learning of the new parameters. The structural breaks prevent the complete learning. This learning could have classical or Bayesian flavors – we do not impose any particular structure on it.

Other effects may also reflect systematic pessimism and adaptation e.g. found by certain minimax ("robust") rules in the neighborhood of the conditional stationary measure or by imputing to the agent specific rules of learning. Our analysis does not impute such rules[19].

3.2 Construction of Econometric Estimators and Test Procedures

In this procedure, we, the econometrician, ultimately aim at describing the beliefs via characterization of conditional stationary measure and also the sentiments through ob-

[19]In principle, it could be an interesting extension of the present paper to impute and test the plausibility of such rules. At present, we leave this possibility unexplored.

serving the agents making the economic choices according to the Euler equations. We then want to examine the rationality of these beliefs by proposing a test statistic.

We first explore the conditional moment restriction, that the Euler equations represent. For measurable functions m^i, we must have

$$E_{\mathbb{Q}_t^{t+k}} h(x_{t+k}, \theta) m^i(x_{t+j}, j \in I) = 0, \quad I = \{0, \ldots, k-1\} \quad (1 \leq k \leq K)$$
$$(i = 1, \ldots, M)$$
$$(t = H - T_e, \ldots, H)$$

This, together with the proposed in the previous section specifications of the functional forms for conditional beliefs $\mathbb{Q}_t^{t+k}$, gives the system of equations:

$$E_{\mathbb{B}(\tilde{\alpha}_t^K + l^K(I_t, \lambda_t^K))} h(x_{t+k}, \theta) m^i(x_{t+j}, j \in I) = 0, \quad I = \{0, \ldots, k-1\} \quad (1 \leq k \leq K)$$
$$(i = 1, \ldots, M) \tag{3.7}$$
$$(t = H - T_e, \ldots, H)$$

This allows to form a large number of equations that put restrictions on the system of conditional beliefs, implied by the optimality of the agent's choices. We hope then to identify a subfamily of beliefs that, with the additional restrictions, are fully described by the above equations. The estimation procedure[20] substitutes $\hat{\theta}$ and $\hat{\lambda}^K$ in the above equation in place of θ and λ_t^K. Estimator $\hat{\lambda}^K$ depends usually on the specification of the conditional stationary measure (see appendix). This allows to proxy *consistently* the sentiment parameters $\tilde{\alpha}_t^K, t = H - T_e, \ldots, H$, as $T_e \to \infty, H \to \infty$ for either the canonical or empiricist model.

It should be noted that the class $\mathbb{B}$ can be made semi- and non-parametric. This would amount to specifying a parametric family with growing (with T_e) number of parameters, e.g. $\mathbb{B}$ could be modeled through a composition of analytically convenient basis functions that can recover $\mathbb{B}$ in any target class[21].

The second set of econometric restrictions is derived from the rationality conditions. It is imposed then as a valid sample counterpart of restriction (2.4):

$$\frac{1}{T_e} \sum_{t=H-T_e+1}^{H} E_{\mathbb{Q}_t^{t+K}} g(x_{t+1}, \ldots, x_{t+K}) - \frac{1}{H} \sum_{t=1}^{H} g(x_{t+1}, \ldots, x_{t+K}) = 0$$

[20]Under regularity conditions imposed in the appendix, preference parameter θ can be estimated by the usual method of moment procedure, as that defined in Hansen and Singleton (1982). However, such validity comes as a result of an unconditional moment restriction arising from rationality condition.

[21]For example, the class of smooth differentiable functions can be recovered via the use of orthogonal polynomials.

for $g \in \mathcal{G}$[22], i.e. equivalently,

$$\frac{1}{T_e}\sum_{t=H-T_e+1}^{H}\int g(z)\mathbb{B}(dz,\tilde{\alpha}_t^K+l(I_t,\lambda_t^K))-\frac{1}{H}\sum_{t=1}^{H}g(x_{t+1},\ldots,x_{t+K})=0,\quad g\in\mathcal{G} \tag{3.8}$$

In summary, therefore we have two sets of conditions:

$$\begin{aligned} f_{1m}(\tilde{\alpha}_t^K+l^K(I_t,\lambda_t^K),\theta)=0 \quad \forall t\in(H-T_e,...,H) \\ \frac{1}{T_e}\sum_{t=H-T_e+1}^{H} f_{2g}(\tilde{\alpha}_t^K+l^K(I_t,\lambda_t^K),w_t^K)=\frac{1}{H}\sum_{t=1}^{H}g(w_t^K),\ \ m\in\mathcal{M},\ g\in\mathcal{G} \end{aligned} \tag{3.9}$$

where $w_t^K \equiv (x_{t+1},\ldots,x_{t+K})$, f_{1m} is the left hand side in (3.7), $f_{2g} \equiv \int g(z)\mathbb{B}(dz,\tilde{\alpha}_t^K + l(I_t,\lambda_t^K))$.

The procedure then involves formation of the test statistics based on the proposed sample counterpart of the rationality conditions in a form of a metric over the moment functions g. Indeed, substituting $\hat{\alpha}_t^K$, $\hat{\lambda}^K$ in the above rationality conditions, we then can form the following two types of test statistics.

If $\mathcal{G}$ is finite, form a metric of the quadratic form:

$$\Gamma' W \Gamma$$

where $\Gamma = \left\{\frac{1}{T_e}\sum_{t=H-T_e+1}^{H} f_{2g}(\hat{\alpha}_t+l^K(I_t,\hat{\lambda}^K),w_t^K)-\frac{1}{H}\sum_{t=1}^{H}g(w_t^K), g\in\mathcal{G}\right\}$, and W is a positive definite matrix.

If $\mathcal{G}$ is infinite, form the sup metric as follows:

$$\sup_{g\in\mathcal{G}}\left|\frac{1}{T_e}\sum_{t=H-T_e+1}^{H}\int f_{2g}(\hat{\alpha}_t^K+l^K(I_t,\hat{\lambda}_t^K))-\frac{1}{H}\sum_{t=1}^{H}g(w_t^K)\right|$$

e.g. if g is a (well-behaved) parametric class $g(\cdot,\delta), \delta\in\Delta$, the above test statistics is computed as

$$\sup_{\delta\in\Delta}\left|\frac{1}{T_e}\sum_{t=H-T_e+1}^{H} f_{2g}(\hat{\alpha}_t^K+l^K(I_t,\hat{\lambda}_t^K))-\frac{1}{H}\sum_{t=1}^{H}g(w_t^K)\right| \tag{3.10}$$

[22] K can be made to grow logarithmically or perhaps even polynomially. See Koenker and Machado (1997), for the treatment of some estimation problems with the growing number of parameters/moment restrictions. None of those results apply to our settings for several reasons: first, the estimation method is more complicated than any usual method of moments estimator, further, these treatments concern i.i.d. settings.

The above statistic consistently (in the usual sense employed in the asymptotic statistical decision theory) discriminates the rationality from non-rationality for either the canonical or empiricist model, as shown in the appendix. The full range of appropriate asymptotic results is derived for estimators and the test statistics, under the null of rationality and the (fixed) alternative of non-rationality. The test statistics are shown to converge in distribution to the continuous functionals of generalized Brownian bridges, whose finite dimensional projections are viewed as projection on the space of moment functions that are used to phrase the rationality conditions. As a result, the limit distributions of the test statistic are either non-standard or standard, depending on whether the test statistic is itself a function of finite-dimensional projection (as in the quadratic form above) or a functional of the whole process (as the sup statistics above). The resampling and simulation methods allow for valid approximation of either distribution. The regularity conditions are particularly flexible, making them amenable to any new LLN and CLT results should these become available. The sufficient conditions for the estimators include near-epoch-dependence and α-mixing. Near-epoch dependence is one of the most general concepts of dependence.

The recovered time series of sentiments can be useful for further applied analysis. Next section includes the quantitative decomposition of market uncertainty into the components arising from variation in the intertemporal marginal rate of substitution and from the model or endogenous uncertainty. Other potentially interesting issues may include the analysis of sentiments as functions of socio-political environments.

4 Application to a Representative Agent RB Economy

This section considers a representative agent economy. Subsection 4.1 formulates the modeling approach as applied to the standard consumption and asset pricing model. Further, we develop a *quantitative* decomposition of uncertainty ("market price of risk") into two parts: one arising from variation in the intertemporal marginal rates of substitution, as in the usual case, and the second due to "endogenous" uncertainty, arising from the inability to learn the true law of motion and having to use a Rational Belief in lieu of the true probability measure. This offers the quantitative and qualitative explanation of the "equity premium" by the RB models. A simple "log-normal" example is worked out in subsection 4.2, and section 4.3 offers its empirical evaluation using U.S. economy data.

4.1 Designing RBE and Testable Restrictions

To interpret and motivate the developments of the previous sections, we construct the representative agent economy along the lines of Hansen and Singleton (1982), Lucas

(1978), Prescott and Mehra (1980) and Breeden (1979). Analysis of this model in RBE framework can be found in Kurz (1994a).

Suppose that a representative infinitely lived consumer chooses consumption $\{c_\tau\}$, and short-term investment plans $\{Z_\tau\}$ so as to maximize

$$E_{\mathbb{Q}_t}\left[\sum_{\tau=t}^{\infty}\beta^{\tau-t}u(c_\tau)\right] \tag{4.11}$$

where subscript t reflects conditioning on the current information set I_t. Usual assumptions on the utility kernel and the discount factor are made throughout. $\mathbb{Q}_t$ denotes the conditional belief of an agent.

The feasible consumption and investment plans must satisfy the sequence of budget constraints:

$$c_\tau + \sum_{j=1}^{N} q_{j\tau}Z_{j\tau} \leq \sum_{j=1}^{N} R_{j\tau}Z_{j\tau-1} + W_\tau \tag{4.12}$$

where $Z_{j\tau}$ are the purchases of the security j at the date τ, priced at $q_{j\tau}$, with date $\tau + 1$ payoff of $R_{j\tau+1}$, $j = 1, \ldots, N$. W_τ is the agent's endowment process. All prices are denominated in units of consumption good.

Euler Equations

Given the representative agent environment above, the necessary first-order optimality conditions (Euler equations) are as follows:

$$E_{\mathbb{Q}_t^{t+K}}\frac{u'(c_{t+K})}{u'(c_t)}\frac{\beta R_{jt+K}}{q_{jt}} = 1 \tag{4.13}$$

Conditional Beliefs and Stationary Measure

We give an example of the conditional belief and conditional stationary measure concerning the logs of x_t, the economic variables that enter the Euler equations above (we specify them exactly in the empirical section). Let $\mathbb{Q}_t^{t+K}$ be characterized by normal c.d.f.:

$$N(\mu_t^K + m^K(I_t, \lambda_t^K), \Sigma_t^K + S(I_t, \lambda_t^K)) \tag{4.14}$$

where I_t contains current and J lagged values of the conditioning variables, m^K is vector s.t. $dim(m^K) = dim(\mu^K) = dim(x_{t+1}, \ldots, x_{t+K})$, $dim(\Sigma_t^K) = dim(S^K) = dim(x_{t+1}, \ldots, x_{t+K}) \times dim(x_{t+1}, \ldots, x_{t+K})$, $\Sigma_t^K + S^K(I_t, \lambda_t^K)$ is assumed to be positive semi-definite.

Note that:

$$\begin{aligned}\mu_t^K &\text{ is the location sentiment vector variable,}\\ \Sigma_t^K &\text{ is the volatility sentiment matrix variable,}\\ m^K(I_t,\lambda_t^K) &\text{ is the location learning vector variable,}\\ S^K(I_t,\lambda_t^K) &\text{ is the scale learning matrix variable.}\end{aligned}$$

In earlier notation, therefore, $\alpha_t^K = \begin{pmatrix}\mu_t^K \\ \mathrm{vec}(\Sigma_t^K)\end{pmatrix}$ and $l^K(I_t,\lambda_t^K) = \begin{pmatrix} m^K(I_t,\lambda_t^K) \\ \mathrm{vec}(S^K((I_t,\lambda_t^K))\end{pmatrix}$. Note also that in the canonical model, $\lambda_t^K = \lambda^K$. Finally, the conditional stationary measure is given by

$$N(m^K(I_t,\lambda^K), S(I_t,\lambda^K)) \tag{4.15}$$

Rationality Restrictions

Given that $\mathbb{M}$ is stationary and ergodic by assumption, it follows that the unconditional measure

$$\mathbb{M}^{t+k} = E_{\mathbb{M}}(\mathbb{M}_t^{t+k}) \text{ is independent of t, i.e } \mathbb{M}^{t+k} = \mathbb{M}^k$$

(a) Finite class $\mathcal{G}$

Suppose that $\mathbb{M}^k$ is defined by a c.d.f from a (log) normal family. Then the posed earlier rationality condition can be reduced to:

$$\begin{aligned}\mathbb{P}\lim_{(T_e,H)\to\infty}\frac{1}{T_e}\sum_{t=H-T_e+1}^{H}\left(\mu_t^K + m^K(I_t,\lambda_t^K)\right) &= m^K,\\ \mathbb{P}\lim_{(T_e,H)\to\infty}\frac{1}{T_e}\sum_{t=H-T_e+1}^{H}\left(\Sigma_t^K + S_T^K(I_t,\lambda_t^K)\right) &= S^K\end{aligned} \tag{4.16}$$

where

$$\begin{aligned} m^K &= \mathbb{P}\lim_{T\to\infty}\frac{1}{H}\sum_{t=1}^{H} z_t^K \\ S^K &= \mathbb{P}\lim_{T\to\infty}\frac{1}{H}\sum_{t=1}^{H}\left(z_t^K - m^K\right)\left(z_t^K - m^K\right)'\end{aligned}$$

where $z_t^K \equiv (\ln x_{t+1}, \ldots, \ln x_{t+K})$

(b) Infinite class $\mathcal{G}$

If the unconditional measures are non-normal, the number of conditions put forward has to be larger. We may consider infinite-dimensional Glivenko-Cantelli classes.

4.2 Decomposition of Uncertainty

Let us go back to the Euler equations (4.13). For one period returns[23], rewrite them in the form:

$$E_{\mathbb{M}_t^{t+1}} \frac{d\mathbb{Q}_t^{t+1}}{d\mathbb{M}_t^{t+1}} \frac{u'(c_{t+1})}{u'(c_t)} \beta R_{jt+1} = q_{jt}, \ j = 1, \ldots, N \tag{4.17}$$

where $\frac{d\mathbb{Q}_t^{t+1}}{d\mathbb{M}_t^{t+1}}$ is a Radon-Nikodym derivative[24].

The random variable $\delta_t = \frac{d\mathbb{Q}_t^{t+1}}{d\mathbb{M}_t^{t+1}} \frac{u'(c_{t+1})}{u'(c_t)}$ in (4.17) has an interpretation of a one-period stochastic discount factor. The multiplicative adjustment factor $\delta_t^r = \frac{d\mathbb{Q}_t^{t+1}}{d\mathbb{M}_t^{t+1}}$ is due to the "model" or "endogenous" uncertainty. It augments the usual measure of intertemporal marginal rate of substitution, $\delta_t^f = \frac{u'(c_{t+1})}{u'(c_t)}$. Intuitively, δ_t^r creates a distortion in the consumption decision. This distortion arises due to the model uncertainty, or, ultimately, from the use of rational beliefs. Therefore, this factor increases the variability of the usual stochastic discount factor and may help better explain relations between intertemporal profiles of consumption and asset prices, i.e explain the risk premium.

From the one-period stochastic discount factor, one can easily compute the *market price of risk* by[25]:

$$p_t = \frac{\sqrt{Var_{\mathbb{M}_t} \delta_t}}{E_{\mathbb{M}_t} \delta_t}. \tag{4.18}$$

Indeed, as an approximation, one can view one period payoff of a security as a bundle of the two attributes: conditional mean and conditional standard deviation. Hansen and Jagannathan (1991) suggest the way how these attributes can be valued. To see this, consider variance decomposition for the one-period security price q_t:

$$q_t = E_{\mathbb{M}_t^{t+1}}(\delta_t) E_{\mathbb{M}_t^{t+1}}(\beta R_{jt+1}) - Cov_{\mathbb{M}_t^{t+1}}(\delta_t, \beta R_{jt+1}) \tag{4.19}$$

which further yields the price bound:

$$q_t \geq E_{\mathbb{M}_t^{t+1}}(\delta_t) E_{\mathbb{M}_t^{t+1}}(\beta R_{jt+1}) - Std_{\mathbb{M}_t^{t+1}}(\delta_t) Std_{\mathbb{M}_t^{t+1}}(\beta R_{jt+1}), \tag{4.20}$$

where Std denotes standard deviation operator.

[23]Long-lived securities are treated similarly

[24]Note that the existence of Radon-Nikodym derivative of conditional belief w.r.t. conditional stationary measure (i.e. $\frac{d\mathbb{Q}_t^{t+k}}{d\mathbb{M}_t^{t+k}}$), does not imply that $\frac{d\mathbb{Q}}{d\mathbb{M}}$ exists. See Kurz (1997a).

[25]$E_{\mathbb{M}_t}\delta_t = 1$, if risk-free bond is available.

Along the efficient frontier, the 'price of risk', relative to expected return, is given by (4.18). This ratio is one way to portray risk-aversion of representative agent at the equilibrium consumption process.

In turn, $Var_{\mathbb{M}_t^{t+1}}(\delta_t) \approx Var_{\mathbb{M}_t^{t+1}}(\delta_t^r)$ if $Var_{\mathbb{M}_t^{t+1}}(\delta_t^f)$ is relatively small and vice versa. Comparison of the two represents quantitative decomposition of the market price of risk into components: one that is due to model or endogenous uncertainty and the other that is due to variability of the marginal rate of substitution in consumption.

4.3 Implementation

The example of implementation will be of a simple kind. This is done so that the methods offered above are easily understood.

For the purposes of the illustration of methodology we assume the preferences are of the constant relative risk aversion type:

$$u(c_t) = \frac{(c_t)^\gamma}{\gamma}, \quad \gamma < 1$$

In this case, the marginal utility is given by:

$$u'(c_t) = (c_t)^\alpha, \quad \alpha \equiv \gamma - 1.$$

If we consider the example of the section above and confine out attention to stocks, we have the Euler equations:

$$\begin{aligned} & E_{\mathbb{Q}_t^{t+K}}\Big(\beta(x_{1t+k})^\alpha R_{jt+k}\Big) = 1, \\ & (\text{ for } j = 1, \ldots, N) \\ & (\text{ for } k = 1, 2, \ldots, K) \end{aligned} \tag{4.21}$$

where x_{1t+k} is the ratio of consumption in time period $t+k$ to consumption in time period t, the k-period real return, R_{jt+k}, is given by $(q_{jt+k} + D_{jt+k})/q_{jt}$, and D_{jt+k} denotes the dividend gain.

Appendix shows, according to the rationality conditions, that the preference parameters $\theta = (\alpha, \beta)$ can be estimated by method of moments of Hansen and Singleton, by considering the sample moment conditions:

$$\frac{1}{T_e} \sum_{t=H-T_e+1}^{H} \Big((\beta(x_{1t+k})^\alpha R_{jt+k} - \mathbf{1}) z(I_t) \Big) = \mathbf{0} \tag{4.22}$$

for a fixed (or growing collection of functions) z of I_t, i.e. "instruments".

$\hat{\theta}$ is the minimizer of the quadratic metric of the sample moment conditions (4.22), i.e. letting $g_{T_e}(\theta) = \frac{1}{T_e}\sum_{t=H-T_e+1}^{H}\Big((\beta(x_{1t+k})^{\alpha}R_{jt+k} - 1)z(I_t)\Big)$:

$$\hat{\theta} = arg\min_{\theta} Q_{1T_e}(\theta) = arg\min_{\theta} g_{T_e}(\theta)'W_1 g_{T_e}(\theta) \tag{4.23}$$

where W_1 is a p.d. matrix. The estimator $\hat{\theta}$ is consistent and asymptotically normal under regularity conditions stated in the appendix (In particular, as $T_e \to \infty$, $H \to \infty$, $T_e = O(H - T_e)$), a result that is not seemingly surprising, but requires us to impose certain *regularity conditions* on the *learning mechanism the agent uses*. Therefore, the regularity conditions underlying the consistency and the limit distribution are an additional burden, in particular the seemingly odd condition that the length of history since the beginning of time, H, should grow large and the interval during which the agent is observed, T_e, should also grow and that the rates of its growth should be curbed relative to H.

Next, define the stochastic process $x_t = (x_{1t}, R_{2t}, \ldots, R_{Nt})$. Suppose that $\ln x_t^K = (\ln x_{t+k}, k = 1, ..., K)$ is conditionally normally distributed according to the stationary measure and the beliefs. That is the conditional stationary measure is normal, and the beliefs are conditionally normal too. Reader may refer to the example of the previous section. Further we suppose that the structure of the agent's belief is such that the sentiments are generated only about the mean growth rate of cum dividend asset price, whereas the rest of the required parameters of beliefs is given purely by his assessment of the stationary measure, with no sentiments added. This means that $\mathbb{Q}_t^{t+k}$ is defined by:

$$N\left(\mu_t^K + m^K(I_t, \lambda_t^K), S^K(I_t, \lambda_t^K)\right),$$

In what follows, for the sake of simplicity, we only look at one asset and $K = 1$. Then in such case $\mu_t^1 = (0, \mu_{2t})'$. Next, the equations (4.21) with unknown sentiments α_t^1 can be reduced to a simple set of equations as follows. Since

$$\begin{aligned} &E_{\mathbb{Q}_t^{t+1}} \ln\Big[\beta(x_{1t+1})^{\alpha}R_{1t+1}\Big] \\ &\qquad = \ln\beta + \mu_{1t}^1 + \alpha\Big[m_1^1(I_t, \lambda_t^1)\Big] + \Big[m_2^1(I_t, \lambda_t^1)\Big], \text{ and} \end{aligned}$$

$$\begin{aligned} &Var_{\mathbb{Q}_t^{t+1}} \ln\Big[\beta(x_{1t+1})^{\alpha}R_{1t+1}\Big] \\ &\qquad = \alpha^2\Sigma_{11}(I_t, \lambda_t^1) + \Sigma_{22}(I_t, \lambda_t^1) + 2\alpha\Sigma_{12}(I_t, \lambda_t^1). \end{aligned}$$

Equation (4.21) is then equivalent in this case to:

$$\begin{aligned} &\ln\beta + \mu_{2t}^1 + \alpha\Big[m_1^1(I_t, \lambda_t^1)\Big] + \Big[m_2^1(I_t, \lambda_t^1)\Big] \\ &\qquad + \frac{\alpha^2\Sigma_{11}(I_t, \lambda_t^1) + \Sigma_{22}(I_t, \lambda_t^1) + 2\alpha\Sigma_{12}(I_t, \lambda_t^1)}{2} = 0 \end{aligned} \tag{4.24}$$

Similarly, the analogous construct defines the system where all sentiments are generated only about the conditional mean of consumption growth rate. We shall refer to the first specification as to the system of beliefs R, and the second as the system C.

Next step is clear: by substituting for λ_t^1, and (α, β), our estimates, $\hat{\lambda}^1$, and $(\hat{\alpha}, \hat{\beta})$, we obtain the estimates $\hat{\alpha}_t^1$. Further, test statistics can be computed as outlined in the appendix.

The following empirical example covers the U.S. monthly economy data for the post-war period. This includes the total of 460 observations on aggregate consumption, SP500 prices and dividend rates (from Shiller and other sources).

The model implements the analysis of two sub-families of beliefs:

1 sub-family C, when the conditional beliefs are given almost entirely by the conditional stationary measure, and the sentiments are generated only about the unknown true conditional mean of consumption;

2 sub-family R, when the conditional beliefs are given almost entirely by the conditional stationary measure, and the sentiments are generated only about the unknown true conditional mean of returns;

Such systems, if the sentiments were all zero, by construction would be equal to conditional stationary measure and therefore would be rational. We shall examine whether the sentiments introduce non-rationality.

The estimate of the conditional stationary measure for real log-consumption per capita growth and real log-return (deflated by cpi) was constructed by linear non-recursive least squares, using the whole sample. Such non-recursivity is permitted by the regularity conditions in the appendix. The conditional variances were also taken of linear form and estimated non-recursively. In the *canonical model*, this estimate is then taken as the true conditional stationary measure.

According to the regularity conditions, the estimates of the sentiments can only be constructed based on the sample with the size comparable to the length of history size or smaller. T_e, therefore has been set at H/7, the last seventh part of the whole history.

The sequence of figures (1) - (6) summarizes succinctly our findings. We discuss each figure in turn.

Figure 1 plots the estimated mean of the real log-consumption growth according to the conditional stationary measure (solid line with triangular vertices) and the sentiments α_t^1 for consumption growth rate (dotted line with diamonds). Oftentimes, the two quantities are moving in the opposite directions, inducing the relatively smooth forecast (by the conditional belief) of the future consumption growth.

Figure 2 displays the actual real log-consumption growth series (solid), the estimates of the mean of the conditional stationary measure, and the (sentiment-corrected) mean of the conditional belief, that is the sum of the estimated mean of the conditional stationary measure and the estimated sentiment. Periods of pessimism and optimism, relative to the conditional stationary measure, are present. This makes for interesting dynamics of the conditional log-consumption growth forecast.

Turning to the family R of conditional beliefs, we report the following observations.

According to the figure 3, which depicts the estimated conditional mean of the real log-return to SP500 (cum dividend) under the conditional stationary measure, and the estimated sentiments, variability of sentiments adds significantly to the volatility of the estimated mean of the conditional stationary measure. On the other hand, the sentiment tends to move inversely to the movements in the mean of the conditional stationary measure, so the overall effect is slight reduction in the volatility of the stock return forecasts.

Indeed, figure 4 shows that (sentiment-corrected) mean of the conditional belief is relatively smooth, and tends to lie below the actual series. Overall, its dynamics exhibits both periods of clear pessimism and optimism.

Notably, the interpretation of the mean of the conditional belief of the family C as a forecast of the future log-consumption ($\ln x_{1t+1}$) allows to achieve good forecasting performance relative to what is attained by using the conditional stationary measure. Indeed, the mean squared error (MSE) of conditional belief based forecast of future consumption ($\ln x_{1t+1}$) is 0.003776, while MSE of the forecast based on the conditional stationary measure is 0.004111. On the other hand, forecasting of returns by the conditional belief in the family R (MSE = 0.02567) is slightly worse compared to the results attained by the conditional stationary measure based forecast (MSE = 0.02241).

In both systems of beliefs we were also able to conclude that *endogenous* or *model* uncertainty arising from the inability to know the true law is the predominant component of volatility[26].

Finally, we get to the rationality.

It can be shown that the earlier posed statistics, under the assumption of the correct specification of the conditional stationary measure, can be reduced under the further assumptions posed in section 4.2 (in the finite class $\mathcal{G}$ example) to the comparisons between:

(1) time average of the conditional means of the variables (real log-consumption growth and log-return) implied by the conditional beliefs, and

(2) observed historical means.

This simplification is due to the nature of sub-families C and R, that ascribe no sentiments about volatility. This implies that if the model of the conditional stationary measure is correctly specified, then by construction the conditional volatility should integrate to

[26]Simulation studies of Kurz (1997a) offer similar conclusions.

the unconditional one.

What are the rationality evaluation results?

The historical mean of log-consumption growth is computed as 0.001121. The sample mean implied by the system of conditional beliefs C (i.e. time average of means of conditional belief), computed from the last T_e observations, is 0.0003579, which is fairly different. This suggests some pessimism "on the average" implied by the beliefs. Is this a rational deviate in the statistical sense?

Similarly, for the system R the beliefs imply the average of 0.003474, while the true historical average is 0.0056. The numbers perhaps reflect slight pessimism about the returns. Again, is the system of beliefs R a rational sub-family?

The answer to both of these questions is given in figures 5 and 6. These figures plot estimates of the bootstrap distribution (densities[27]) of the estimated long-run mean implied by the beliefs and the historical means. The bootstrap distributions are valid asymptotic approximations of the true asymptotic distributions of the estimators. The plotted densities characterize thus the estimation uncertainty.

It should be noted that in the canonical model, the historical mean is known and its value is fixed and therefore the distribution of the estimator is degenerate. This is represented by Dirac density at the historical mean. This is shown in the figures as a large shaded rhomb. If the historical mean would lie outside of the "reasonable" confidence intervals for the estimated long-run mean from the conditional belief, the researcher would be lead to conclude that implied sentiments about future stock returns are not rational.

The empiricist model differs by putting the band around the rhomb, thus accounting for the present estimation uncertainty of the long-run mean [28].

In that respect, under the strict canonical interpretation of rationality, the result in figure 5 offers strong evidence in favor of *rationality* of the system of beliefs C. This is so since the historical mean lies within confidence intervals of the usual significance levels. On the other hand, the system of beliefs R is not rational under strict canonical interpretation of rationality (see figure 6). The historical mean lies well outside the $.98$ confidence interval for the estimated mean return generated by the beliefs. In the model of rationality, that accounts for finite history, in contrast, *both* of the implied belief systems *are* rational (figures 5 and 6). Note also that figures allow to construct confidence intervals of *any* prespecified level. Accounting for the additional estimation uncertainty in the empiricist

[27]The plotted are the Nadaraya- Watson kernel estimate (with a MISE-optimal bandwidth) of the densities of the estimators (of mean generated by the beliefs and the limit of the historical mean) based on the bootstrap scheme of Politis and Romano (1994). Validity of the scheme of Politis and Romano in the non-stationary settings was recently establish by White et al.

[28]Another part where finiteness of history matters is in the specification of the mean implied by the beliefs, since it contains *estimates* of the conditional stationary measure. This second difference is ignored in the current implementation.

model makes the interpretation more involved, though, using Bonferroni bounds, we can conclude that the significance levels at which hypothesis for systems C and R can not be rejected include all commonly used levels. Such an interpretation is valid under the assumption that $T_e/H \to 0$, as $Te, H \to \infty$. In our calculations, we made choice $T_e = H/7$. If such a choice is deemed not credible, the conclusions above are still valid, but the p-values and the significance levels should be interpreted as in the test of rationality *combined* with the contiguity of long-run learning (by the agent) of the true stationary measure relative to the econometrician's estimate of it. That is, if the agent consistently learns the truth, she asymptotically gets to the rationality (see discussion in the first section), however this rate of learning can be so slow that approximating her estimates of the conditional stationary measure with the econometrician's estimate may invalidate the inference approach. If her learning is comparable in rate to the rate of convergence of the econometrician's estimator (i.e. "contiguous"), then the p-values would be interpretable as both the degree of adherence to asymptotic rationality condition as well as their contiguity to econometrician's estimate. The rest is up to the subjective interpretation of the reader, who should, however, realize the implications of his or her discriminating decisions.

4.4 Figures

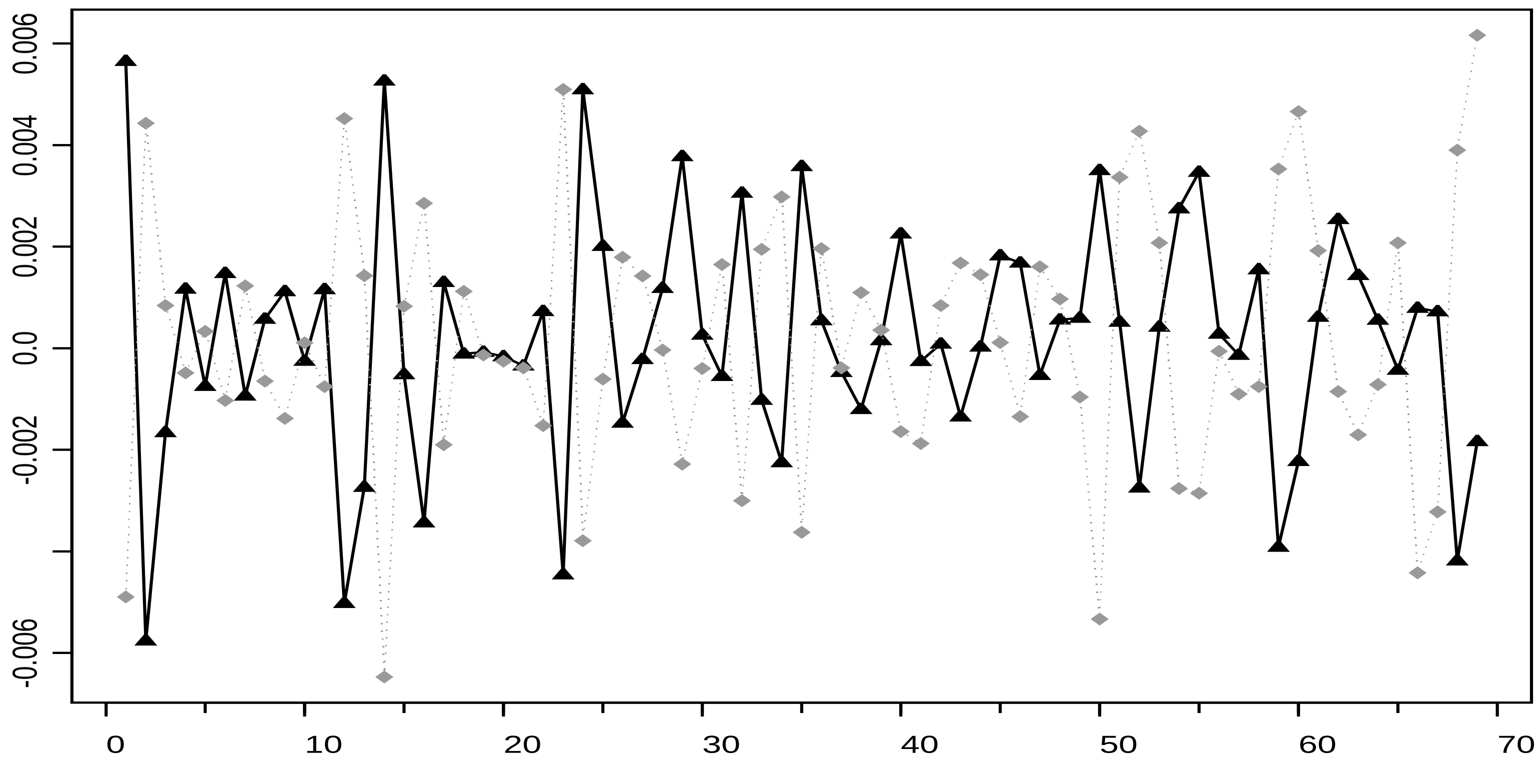

Figure 1: The Estimated Mean of the Conditional Stationary Measure ($-\blacktriangle-$) and the Sentiments α_t^1 ($\cdot\cdot\blacklozenge\cdot\cdot$) for the Growth Rate of Consumption (dotted).

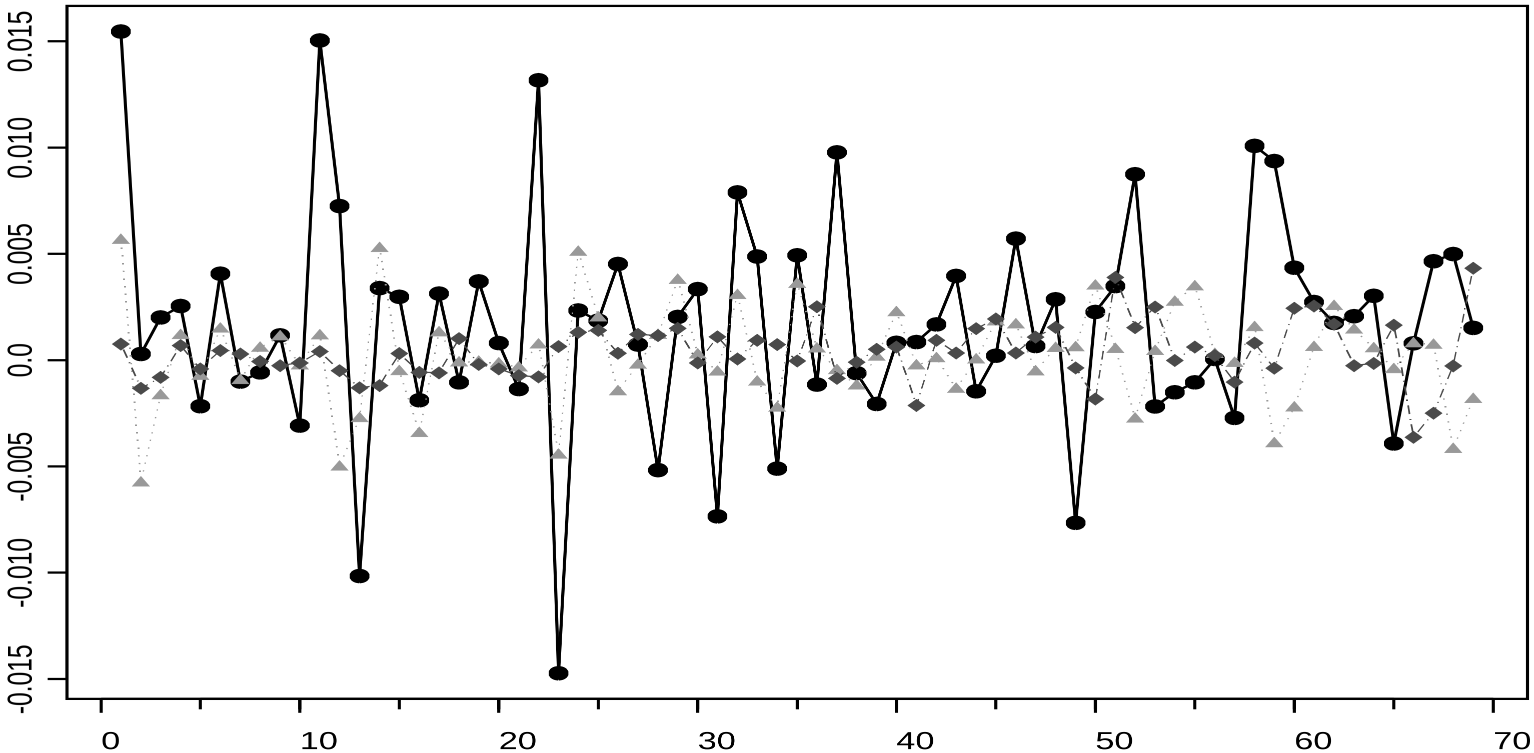

Figure 2: The Actual Consumption Growth (− • −), the Mean of Conditional Stationary Measure (·▲·) and the Mean of Conditional Belief (−. − ♦ − .−)

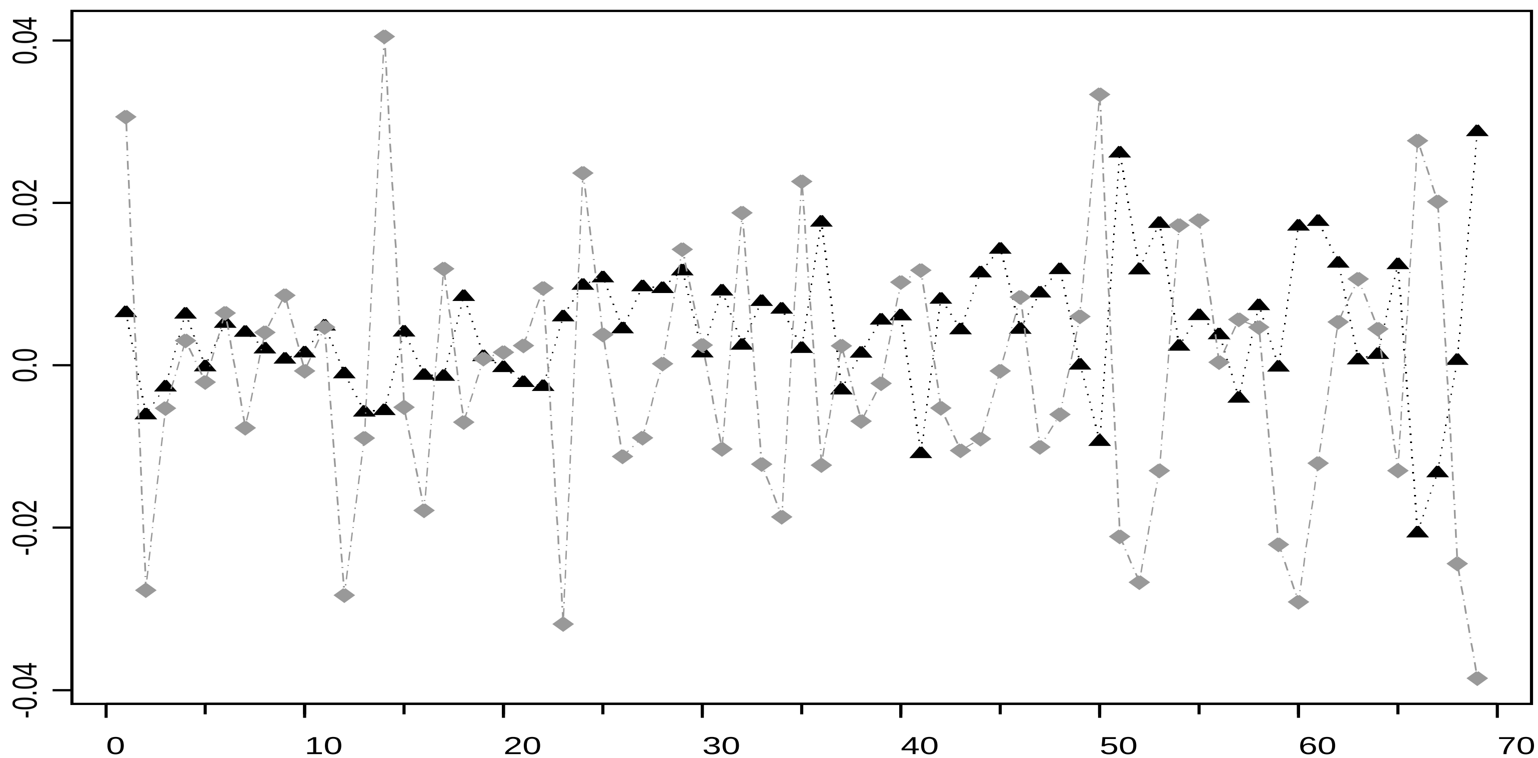

Figure 3: The Estimated Mean of the Conditional Stationary Measure ($-\blacktriangle-$) and the Sentiments α_t^1 for the Stock Return ($\cdot\cdot\blacklozenge\cdot\cdot$).

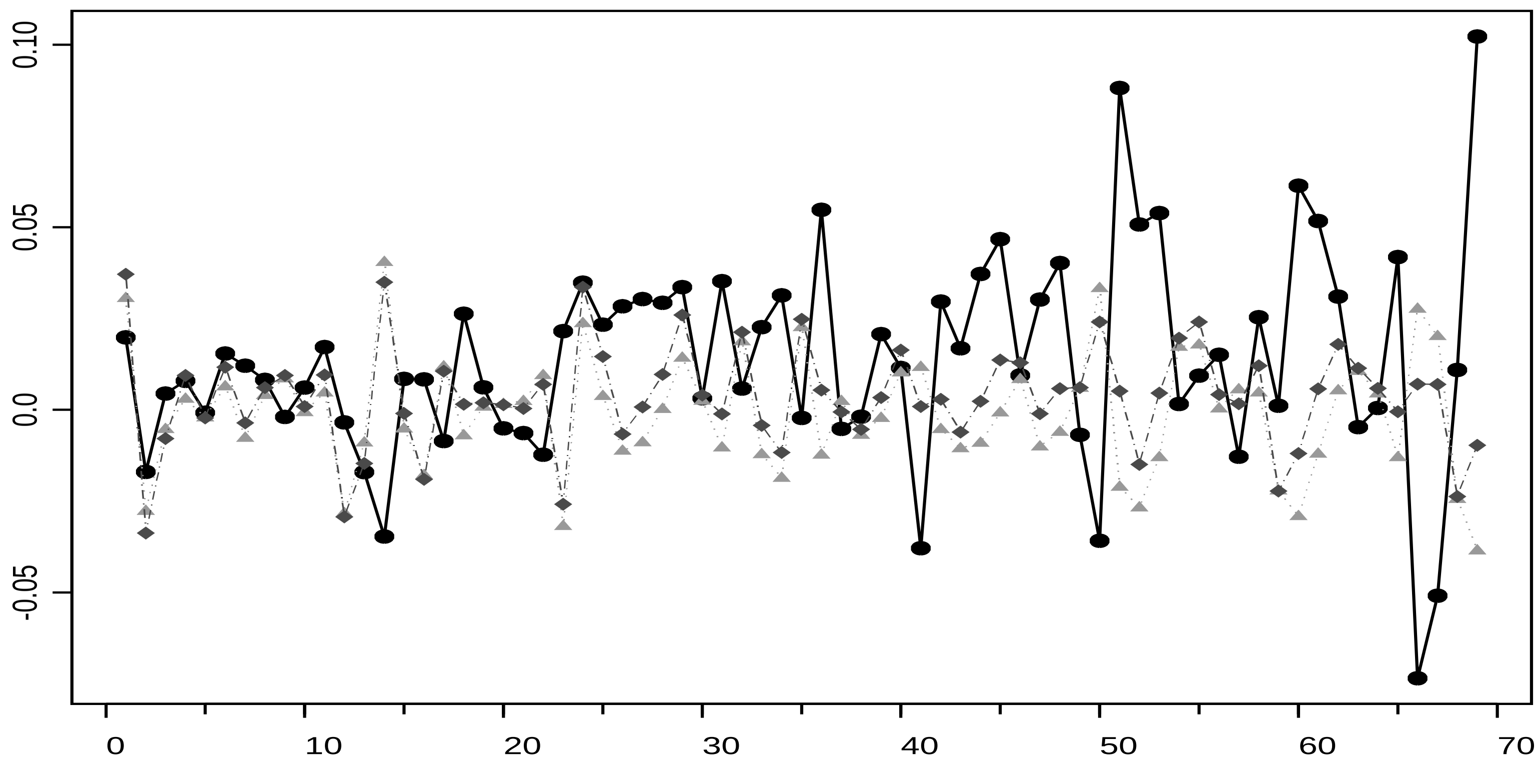

Figure 4: The Actual Return (−•−), the Mean of Conditional Stationary Measure (··▲··) and the Mean of Conditional Belief (−. − ⧫ − .−).

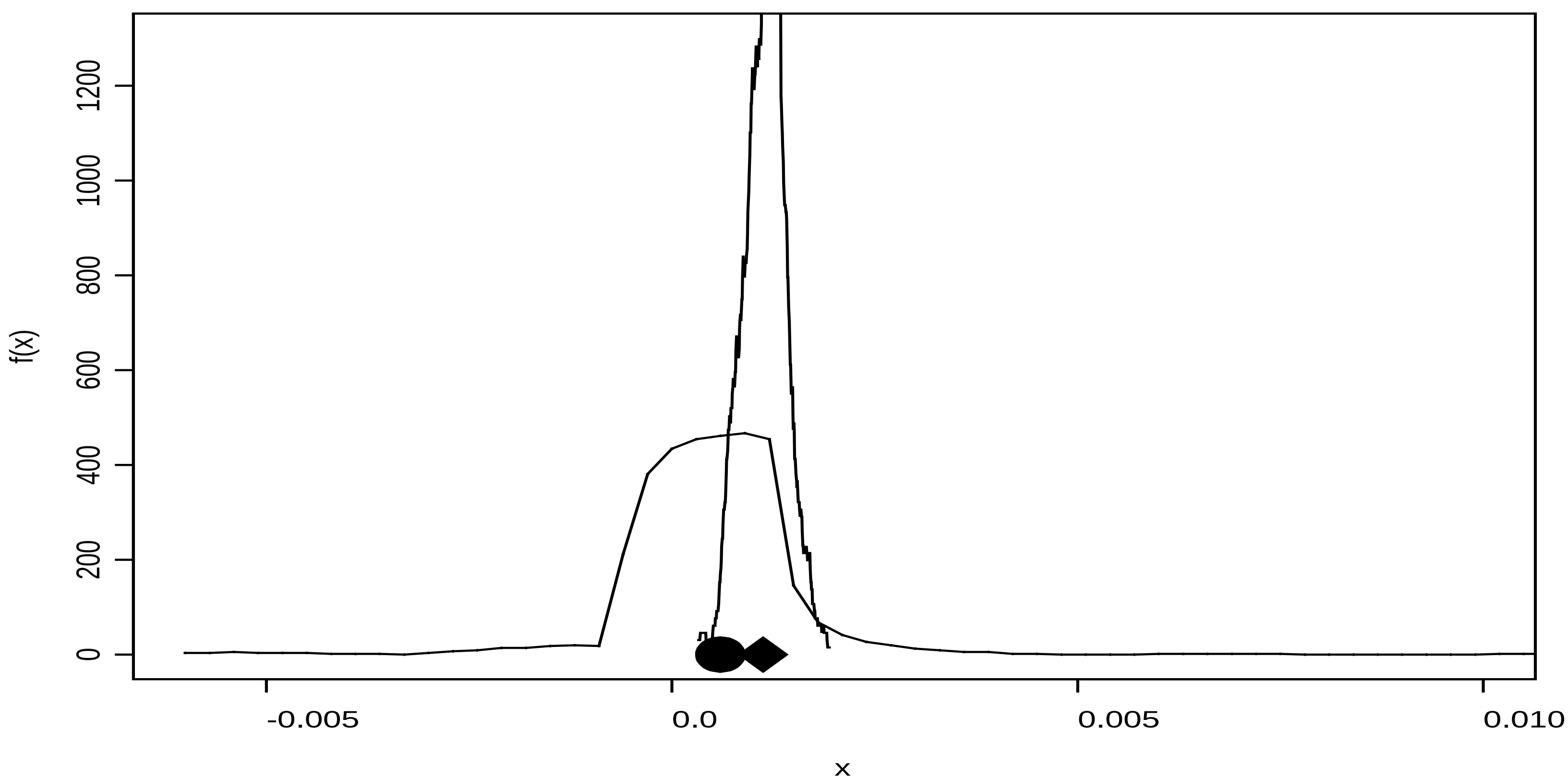

Figure 5: The Bootstrap (Kernel-Smoothed) Densities for the Historical Mean (higher curve) and for the Implied Long-Run Mean (lower curve) from System of Beliefs C, (Consumption Growth Rate Case).

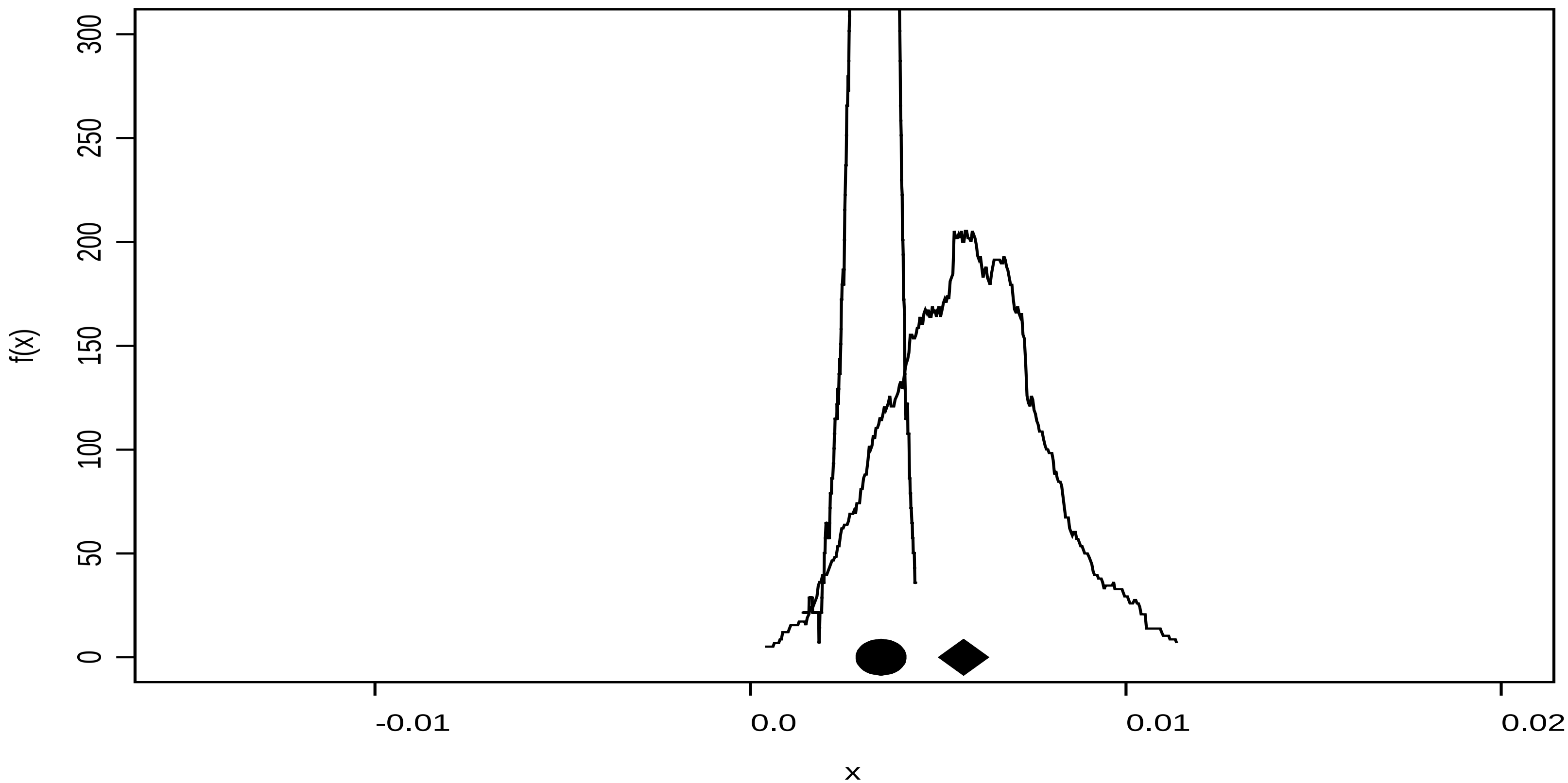

Figure 6: Bootstrap (Kernel-Smoothed) Densities for the Estimated Historical Mean (lower curve) and the Estimated Implied Long-Run Mean (higher curve) from System of Beliefs R, (Return Case).

5 Conclusion

Since Kurz (1997a) introduced his rationality concept, the RBE modeling has evolved as a unified paradigm capable of explaining virtually all ”puzzles” of the real macroeconomy that the RE models could not successfully capture. A series of studies has been conducted to assess plausibility and validity of RB models. Simulation was the main tool to complement intuitively appealing theoretical argument put forward to apprehend that part of world that has to do with anticipating the uncertain, in the presence of long history.

The present work has attempted to find and then establish the validity of the econometric tools to explore such a real world paradigm. This apparatus, apart from some technical issues, provides methods for evaluation of the Kurz canonical model of Rational Beliefs. At the same time, we seek to offer variations to the RB theory from statistical and economic point of view. In particular, we attempted to evaluate the implications of finiteness of large history for modeling perspective of both the economic agents, who inhabit the model, and the outside observer (econometrician), both of whom perform statistical inference with the same finite data. This novel frame of reference has been termed the *empiricist* one. Implications for tools of inference has been demonstrated, notably in that the conditions have been established to distinguish when the interpretation of results is equivalent or very similar to the canonical model, and when it is not.

With the general econometric framework in place, we have developed a simple but suggestive model of consumption and aggregate stock market behavior. Remarkably, it predicts that the primary source of asset price volatility is the variation in the market participants' beliefs, and is in this sense *endogenous* to the model. Furthermore, it suggests a way to build sentiment-based forecasts, and assess the rationality of the beliefs in the presence of finite history.

Departing from this natural benchmark, numerous extensions can be pursued. For example, specification of preferences, driving stochastic process for returns and learning structure may be relaxed in non-trivial directions while econometric procedures can be further refined. Dynamic asset valuation in finite history economies populated by heterogeneous consumers with imperfect knowledge of the economic environment can be exceedingly hard, but it can also be very fruitful.

A Some Preliminaries of Convergence of Stochastic Maps (Empirical Processes)

The discussion of weak convergence and related concepts we use here can be found in texts Hoffmann-Jørgensen (1994), Hoffmann-Jørgensen (1991), Pollard (1990),van der Vaart and Wellner (1996). We only provide necessary minimum of definitions for our and reader's convenience. Let $L^{\infty}(\mathcal{M})$ be the space of bounded functions (for $k \geq 1$) mapping $\mathcal{M}$ to $\mathbb{R}^k$. Equip $L^{\infty}(\mathcal{M})$ with the sup norm. $\mathbb{E}$ denotes the expectation w.r.t $\mathbb{P}$. $\mathbb{P}^*, \mathbb{P}_*, \mathbb{E}^*, \mathbb{E}_*$ denote outer and inner probabilities and expectations (cf. van der Vaart and Wellner (1996)). Notation of this section is independent of the all other sections.

A sequence of arbitrary maps $v_T(\cdot) : \Omega \to L^{\infty}(\mathcal{M})$ converges weakly to a process, $v_0(\cdot)$, if $\mathbb{E}^* v_T(\cdot) \to \mathbb{E} v_0(\cdot)$ for all continuous maps $g : L^{\infty}(\mathcal{M}) \to \mathbb{R}$ (cf. van der Vaart and Wellner (1996)). If $\mathbb{P}^{(T)}$ is the same for all T, then denote such convergence as $\Rightarrow$. If $\mathbb{P}^{(T)}$ is a sequence of distinct probabilities, then we still write $\Rightarrow$. Using outer/inner probability measure above is a measure-theoretic subtlety. If variables under consideration are measurable, then outer/inner probability can be replaced by probability measure.

Process $\{v_T(\cdot)\}$ is *stochastically equicontinuous* ($\{v_T(\cdot)\}$ is $\mathbb{P}$-s.e.) if $\forall \epsilon > 0$ and $\eta > 0, \exists \delta > 0$ s.t.

$$\limsup_{T\to\infty} \mathbb{P}^* \left(\sup_{\rho(\tau_1,\tau_2)<\delta} |v_T(\tau_1) - v_T(\tau_2)| > \eta \right) < \epsilon$$

for some semi-metric ρ on $\mathcal{M}$, s.t. $(\mathcal{M}, \rho)$ is totally bounded.

It is also convenient to define the Gaussian Process or generalized $\mathbb{P}$-*Brownian Bridge* for fixed $\mathbb{P}$, as follows (cf. van der Vaart and Wellner (1996)). The centered Gaussian process is "random" function $\mathbf{G}_{\mathbb{P}} g$, if its distribution is *defined* by its finite dimensional distributions:

$$(\mathbf{G}_{\mathbb{P}} g_1, ...\mathbf{G}_{\mathbb{P}} g_k) \stackrel{\mathcal{D}}{=} N \left(0, \left\{ \mathbb{CV} \mathbf{G}_{\mathbb{P}} g_i, \mathbf{G}_{\mathbb{P}} g_j \right\}_{ij} \right) \tag{A.25}$$

and covariance operator $\mathbb{CV}$:

$$\mathbb{CV} \mathbf{G}_{\mathbb{P}} g_i, \mathbf{G}_{\mathbb{P}} g_j \equiv E_{\mathbb{P}} \left(\mathbf{G}_{\mathbb{P}} g_i \mathbf{G}_{\mathbb{P}} g_j \right) - E_{\mathbb{P}} \left(\mathbf{G}_{\mathbb{P}} g_i \right) E_{\mathbb{P}} \left(\mathbf{G}_{\mathbb{P}} g_j \right)$$

Distribution of process $\mathbb{G}_{\mathbb{P}} f\,(W, \cdot)$ is fully characterized by (A.25) if the process is *tight* (cf. van der Vaart and Wellner (1996)). In what follows, tightness of the limiting process that we obtain will be a consequence of *asymptotic tightness*, a concept that is equivalent to *stochastic equicontinuity* (cf. van der Vaart and Wellner (1996)).

B Fundamental conditions

Existence of the stationary measure $\mathbb{M}$ is assumed, in the sense given in Kurz (1994b). Further, suppose

(BC1) (POINTWISE LLN) Assume that for any random variables $\{w_t^K\} \equiv (x_{t+1}, \ldots, x_{t+K})$ on$(\Omega, \mathcal{F}, \mathbb{P})$ that

$$\nu_N^1(g) = \frac{1}{N}\sum_{t=O+1}^{O+N} g(w_t^K) \xrightarrow{\mathbb{P}^*} \lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N} E_{\mathbb{P}}g(w_t^K) = E_{\mathbb{M}}g(w_t^K),\ \forall g \in \mathcal{M} \supseteq \mathcal{G} \tag{B.26}$$

for any fixed starting summation time O (e.g. $H - T_e$), given functional classes $\mathcal{M}$ and $\mathcal{G}$. Class $\mathcal{G}$ has the natural interpretation of the "reference" class to which agents coordinate their beliefs.

(BC2) STOCHASTIC EQUICONTINUITY

$\sqrt{N}\nu_N^1(g)$ is $\mathbb{P}$-s.e. on $(\mathcal{M}, \rho_1)$ for some semi-metric ρ_1

The next subsection gives a catalogue of sufficient conditions that imply (BC1) and (BC2) in the general frameworks including the non-stationarity and long-range dependence.

Note that above conditions assert that functional classes $\mathcal{M}$ and $\mathcal{G}$, ($\mathcal{G} \subseteq \mathcal{M}$) are Donsker and Glivenko-Cantelli.

(BC3) CLT (FIDI[29] WEAK LIMITS)

$$\left\{\sqrt{N}\nu_N^1(g_1), \ldots, \sqrt{N}\nu_N^1(g_s)\right\} \xrightarrow{D} N(0, \{\Omega_{ij}\}),$$

$$\Omega_{ij} \equiv \lim_{N\to\infty}\frac{1}{N}\sum_{t=O+1}^{N+O}\sum_{l=O+1}^{N+O} Cov_{\mathbb{P}}(g_i(w_t^K), g_j(w_l^K)) + 1_{(l\neq t)}Cov_{\mathbb{P}}(g_i(w_l^K), g_j(w_t^K))$$

B.1 Sufficient Conditions for Fundamental Conditions

The statement of the above results requires the limit results (LLN, CLT, s.e) for *triangular* arrays of random variables $\{w_{tN}^K\}$. In many instances the available limit theorems are stated for non-array settings. However, in most such instances the extensions to the arrays may be made straightforwardly. In the cases of mixing, the results are extended by imposing the required conditions not only through the array, but also uniformly in N.

Besides the mixing options above, there are many other more general or different options that imply BC1, BC3. These include strong, absolutely regular, uniform, ρ-mixing, other types of mixing, as well as more general conditions such as near-epoch dependence and approximability conditions, $m(n)$-decomposability and many others (see Davidson (1994), Doukhan (1994), Dehling, Denker, and Philipp (1986), Gallant and White (1988), Herrndorf (1984a), Herrndorf (1984b), Herrndorf (1983b), McLeish (1974), Corollary 2.11, Berkes and Philipp (1998), Philipp (1986) and sources therein and many others). For sufficient conditions implying BC2, see Andrews and Pollard (1994), Andrews (1994), and references therein.

We do not restate all of these conditions here for brevity.

[29] fidi = finite-dimensional.

C Definitions of Statistics and Estimators

C.1 Estimators

$\hat{\theta}$ *(Estimator of Preference Parameters)*

$$\hat{\theta} \equiv \arg\min_{\theta\in\Theta} Q_{1T_e}(\theta) = \arg\min_{\theta\in\Theta} f_{1T_e}(\theta)' W_{1T_e} f_{1T_e}(\theta),$$

$$f_{1T_e}(\theta) \equiv \frac{1}{T_e}\sum_{t=H-T_e+1}^{H} h(x_t,\theta)\otimes z_t \equiv \frac{1}{T_e}\sum_{t=H-T_e+1}^{H} m_1(w_t^K, I_t, \theta),$$

$$z_t \equiv m(I_t), \text{ for a Borel function } m,$$

W_{1T_e} is a matrix that converges to a positive definite matrix W_1 as $T\to\infty$

$\hat{\lambda}^K$ *(Parameters of the Conditional Stationarity Measure)*

$$\hat{\lambda}^K \equiv \arg\min_{\theta\in\Theta} Q_{2H}(\theta) = \arg\min_{\theta\in\Theta} f_{2H}(\theta)' W_{2H} f_{2H}(\theta),$$

$$f_{2H}(\theta) \equiv \frac{1}{H}\sum_{t=1}^{H}\left(g(w_t^K) - E_{\mathbb{M}_t^{t+K}(\lambda^K)} g(w_t^K)\right) \equiv \frac{1}{H}\sum_{t=1}^{H} m_2(w_t^K, I_t, \lambda),$$

where we abuse notation by using $\mathbb{M}_t^{t+K}(\lambda^K)$ to indicate parametric dependence of the measure $\mathbb{M}_t^{t+K}(\cdot)$ on λ^K. W_{2T} is a matrix that converges to a positive definite matrix W_2 as $T\to\infty$.

COMMENT: Other Periods could be used as well.

α_t^K *(Sentiment Parameter, when λ^K, θ is known – in the Canonical model)*

$$\alpha_t^K \equiv \arg\min_{\alpha\in\mathcal{A}} \left[f_1(w_t^K, \theta, \lambda^K, \alpha)\right]^2$$

$\hat{\alpha}_t^K$ *(Estimator of Sentiment Parameter – for both the Canonical and Empiricist Model)*

$$\hat{\alpha}_t^K \equiv \arg\min_{\alpha\in\mathcal{A}} \left[f_1(w_t^K, \hat{\theta}, \hat{\lambda}^K, \alpha)\right]^2$$

$\tilde{\alpha}_t^K$ *(Agent's Sentiment)*

$$\tilde{\alpha}_t^K \equiv \arg\min_{\alpha\in\mathcal{A}} \left[f_1(w_t^K, \theta, \lambda_t^K, \alpha)\right]^2$$

COMMENT: $\tilde{\alpha}_t^K = \alpha_t^K$ in the canonical model.

The sentiment parameters are assumed to be unique solutions of the above equations, i.e. deducible form them. This may require additional restrictions in practice – see discussion in the text. See also there the definition of notation that is not defined here.

C.2 Rationality Score Processes

Agent's Rationality Score Process (a functional on $\mathcal{G}$) in the Canonical Model

$$S_M^{a1}(g) \equiv \frac{1}{M} \sum_{t=H-M+1}^{H} E_{\mathbb{Q}_t^{t+k}(\lambda^K)} g(w_t^K) - \frac{1}{H} \sum_{t=1}^{H} g(w_t^K), \quad H \geq M \geq 0$$

where $\mathbb{Q}_t^{t+k}(\lambda^K) = \mathbb{B}(\cdot, \alpha_t^K + l^K(I_t, \lambda^K))$.

Agent's Rationality Score Processes (a functional on $\mathcal{G}$) in the Empiricist Model

$$S_M^{a2}(g) \equiv \frac{1}{M} \sum_{t=H-M+1}^{H} E_{\mathbb{Q}_t^{t+k}(\lambda_t^K)} g(w_t^K) - \frac{1}{H} \sum_{t=1}^{H} g(w_t^K), \quad H \geq M \geq 0$$

where $\mathbb{Q}_t^{t+k}(\lambda_t^K) = \mathbb{B}(\cdot, \tilde{\alpha}_t^K + l^K(I_t, \lambda_t^K))$.

$$S_M^{a3}(g) \equiv \frac{1}{M} \sum_{t=H-M+1}^{H} E_{\mathbb{Q}_t^{t+k}(\lambda_{H-M}^K)} g(w_t^K) - \frac{1}{H} \sum_{t=1}^{H} g(w_t^K),$$

$$S_M^{a4}(g) \equiv \frac{1}{M} \sum_{t=H-M+1}^{H} E_{\mathbb{Q}_t^{t+k}(\lambda_H^K)} g(w_t^K) - \frac{1}{H} \sum_{t=1}^{H} g(w_t^K).$$

COMMENT: Under proposed progressivity of learning, the difference between the rationality score processes will converge to zero uniformly in outer probability.

Econometrician's Rationality Score Process (a functional on $\mathcal{G}$)

$$S_T^{e}(g) \equiv \frac{1}{T_e} \sum_{t=H-T_e+1}^{H} E_{\hat{\mathbb{Q}}_t^{t+k}} g(w_t^K) - \frac{1}{H} \sum_{t=1}^{H} g(w_t^K), \quad H \geq M \geq 0$$

$\hat{\mathbb{Q}}_t^{t+k}$ is formed as $\mathbb{B}(\cdot, \hat{\alpha}_t^K + l(I_t, \hat{\lambda}^K))$ as defined in the text.

The score parameter take arbitrary $g \in \mathcal{G}$ and maps it to $L^{\infty}(\mathcal{G})$. The score processes serve as bases for the formulation of rationality test statistics.

It is important now to introduce a new piece of notation. This will latter simplify the notational complexity of some statements.

$$\int g(w_t^K) \mathbb{B}\left(dz, \alpha^K + \lambda_t^K\left(I_t, \lambda\right)\right) \equiv m_{3g}\left(w_t^K, I_t, \lambda^K, \theta, \alpha^K\right)$$

Hence

$$E_{\hat{\mathbb{Q}}_t^{t+k}} g(w_t^K) \equiv m_{3g}\left(w_t^K, I_t, \hat{\lambda}^K, \hat{\theta}, \hat{\alpha}_t^K\right),$$
$$E_{\mathbb{Q}_t^{t+k}} g(w_t^K) \equiv m_{3g}\left(w_t^K, I_t, \lambda_t^K, \theta, \tilde{\alpha}_t^K\right)$$

C.3 Test Hypotheses (Asymptotic Form)

H_0: (Rationality):

$$S_M^{ai}(g) \xrightarrow{\mathbb{P}^*} 0, \text{ as } T, H \to \infty, \text{ uniformly in } \mathcal{G}$$

H_A: (Non-Rationality):

$$S_M^{ai}(g) \xrightarrow{\mathbb{P}^*} c \neq 0, \text{ as } T, H \to \infty, \text{ for some } g \in \mathcal{G}$$

The test hypotheses are in the asymptotic form.

COMMENT: Under proposed progressivity of learning, the difference between the agent's rationality score processes will converge to zero uniformly in outer probability, making the null equivalent for all i, as well as the alternative.

C.4 Rationality Test Statistics

Take a continuous functional $l : L^\infty(\mathcal{G}) \to \mathbb{R}_+$ and form a test statistics:

$$R_{T_e}^e = l(S_{T_e}^e(g)) \tag{C.27}$$

Therefore the test statistics are "random" composition maps from Ω to $\mathbb{R}_+$. They map the "random" score functionals to $\mathbb{R}_+$.

In particular, two forms are considered:

$$U_{T_e} \equiv \sup_{g \in \mathcal{G}} \left| S_{T_e}^e(g) \right| \quad \mathcal{G}, \text{ finite or infinite}$$

$$W_{T_e} \equiv \left(S_{T_e}^e(\mathbf{g}) \right)' W_{3T} \left(S_{T_e}^e(\mathbf{g}) \right), \ \mathbf{g} \in \mathcal{G}, \ \mathcal{G} \text{ is finite}$$

where W_{3T} is positive definite matrix that converges to W_3, also positive definite, $\mathbb{P}$-a.s, $S_{T_e}^e(\mathbf{g}) = \{S_{T_e}^e(g_1), \ldots, S_{T_e}^e(g_l)\}'$.

C.5 Test Decision

The statistical test decision of the econometrician is of the binary form:

$$d \equiv 1_{R_{T_e}^e > \tau},$$

where c is the critical value. 1 classifies the beliefs as non-rational, and 0 classifies the beliefs as rational.

C.6 Consistency of Econometric Test

Decision $d = 1_{R^e > \tau}$ is asymptotically consistent if

$$\mathbb{P}_*(d = 1|H_A) \longrightarrow 1, \text{ as } T_e, H \to \infty$$

$$\mathbb{P}_*(d = 0|H_O) \longrightarrow 1, \text{ as } T_e, H \to \infty.$$

In finite sample the above statement holds only approximately, if sample size is large enough.

D Assumptions

D.1 Basic Assumptions

(BA1) Progressivity of Agent's Learning (Long-Run)

$$\left|\left|\lambda_t^K - \lambda^K\right|\right| = O_{\mathbb{P}^*}\left(\left|\left|\lambda_{H-M} - \lambda^K\right|\right|\right) \tag{D.28}$$

uniformly in $t \in [M+1, \dots, H]$ as $H \to \infty, T_e \to \infty$.

(BA2) Allowing History to "kick-in"

$$T_e = O(H - T_e)$$
$$M = O(H - M) \text{ (empiricist model only)}$$

(BA3) Allowing Econometrician to View History as Formed

$$T_e = o(H) \tag{D.29}$$

COMMENT: (D.29) is not always employed.

(BA4) $\left\{\frac{\partial^i}{\partial(\theta,\lambda)} h(\cdot, \theta)\right\} \subseteq \mathcal{M},\ \forall i = 0, 1, 2$

COMMENT: h and its derivatives are thus Glivenko-Cantelli and Donsker.

(BA5) $g \in \mathcal{G} \subseteq \mathcal{M}$

COMMENT: g is thus Glivenko-Cantelli and Donsker.

(BA6)

$$\left\{\nabla_\theta^i m_1(\cdot, \cdot, \theta), \theta \in \Theta\right\} \subseteq \mathcal{M},\ i = 0, 1$$

$$\left\{\nabla_\lambda^i m_2(\cdot, \cdot, \lambda), \lambda \in \Lambda\right\} \subseteq \mathcal{M},\ i = 0, 1$$

$$\left\{\nabla_\lambda^i m_{3g}(\cdot, \theta, \lambda, \alpha), \lambda \in \Lambda, \theta \in \Theta, \alpha \in \mathcal{A}\right\} \subseteq \mathcal{M},\ i = 0, 1$$

$$\left\{\nabla_{\theta,\lambda,\alpha}^i (f_1(w, \theta, \lambda), \alpha)^j, w \in \mathbf{W}, \lambda \in \Lambda, \theta \in \Theta, \alpha \in \mathcal{A}\right\} \subseteq \mathcal{M},\ i = 0, 1, j = 1, -1$$

Assumptions (BA4)–(BA6) are of more technical nature.

D.2 Technical Assumptions

(TA1) $\nabla_\lambda f_1(w,\theta,\lambda,\alpha)$ $\nabla_{\lambda\lambda} f_1(w,\theta,\lambda,\alpha)$ are bounded and continuous in all parameters. Eigenvalues of $\nabla_\alpha f_1(w,\theta,\lambda,\alpha)$ are bounded away form zero, uniformly in all parameters.

(TA2) Given (w,θ,λ), $f_1(w,\theta,\lambda,\alpha)=0$ has unique solution for α. This statement is uniform: $\exists$ ϵ-ball $N_\epsilon(w,\theta,\lambda,\alpha)$ around α such that:

$$\inf_{\theta,\lambda,w\in\Theta\times\Lambda\times\mathbf{W},\tilde{\alpha}\notin N(\epsilon,w,\theta,\lambda,\alpha)} \left(f_{1m}(w,\theta,\lambda,\tilde{\alpha})\right)^2 \geq \epsilon \tag{D.30}$$

(TA3) $(\lambda_t^K,\tilde{\alpha}_t^K,\alpha_t^K,\tilde{\alpha}_t^K,\theta)\in\Lambda\times\mathcal{A}^3\times\Theta$ - a compact, convex subset of a Euclidean space.

(TA4) There exists positive definite matrix A such that:

$$E_{\mathbb{M}} A m_1(w_t^K, I_t, \tilde{\theta}) \neq 0,\ \forall\tilde{\theta}\neq\theta$$

$$E_{\mathbb{M}} A m_2(w_t^K, I_t, \tilde{\lambda}) \neq 0,\ \forall\tilde{\lambda}\neq\lambda^K$$

E Main Theorems

E.1 The Key Results on Asymptotic Limits

Proposition 1 (Consistency Theorems) *Under assumptions BC, BA, TA, the following hold true:*

$$\hat{\theta} \xrightarrow{\mathbb{P}} \theta,$$

$$\hat{\lambda}^K \xrightarrow{\mathbb{P}} \lambda^K,$$

$$\sup_{H \geq t \geq H - T_e} \left\| \hat{\alpha}_t^K - \tilde{\alpha}_t^K \right\| \xrightarrow{\mathbb{P}^*} 0,$$

$$\sup_{H \geq t \geq H - T_e} \left\| \hat{\alpha}_t^K - \alpha_t^K \right\| \xrightarrow{\mathbb{P}^*} 0.$$

□

Lemma 1 (Convergence in Outer Probability of Agent's Rationality Score Process) *The following statement is true under H_O and H_A:*

$$\sup_{g \in \mathcal{G}} \left| S_{T_e}^{ai}(g) - \lim_{T} E_{\mathbb{P}}(S_{T_e}^{1a}(g)) \right| \xrightarrow{\mathbb{P}^*} 0, \; i = 1, 2, 3, 4. \tag{E.31}$$

COMMENT: $\lim_{M,H \longrightarrow \infty} E_{\mathbb{P}}(S_M^{a1}(g)) \equiv \begin{cases} 0, & \text{under } H_O \\ c \neq 0, & \text{under } H_A \end{cases}$. The lemma shows that first order asymptotics is the same for all score processes, i.e. the difference between all of them converges to zero (by triangle inequality). This means that null hypothesis of rationality is the same for all score processes, and therefore no ambiguity arises.

□

Lemma 2 (Convergence in Outer Probability of Econometrician's Rationality Score Process)

$$\sup_{g \in \mathcal{G}} \left| S_M^{a}(g) - \lim_{T} E_{\mathbb{P}}(S_{T_e}^{1a}(g)) \right| \xrightarrow{\mathbb{P}^*} 0.$$

COMMENT: $\lim_{M,H \longrightarrow \infty} E_{\mathbb{P}}(S_M^{a1}(g)) \equiv \begin{cases} 0, & \text{under } H_O \\ c \neq 0, & \text{under } H_A \end{cases}$. The above shows that the econometrician's approximation of the rationality score process is valid for all agent's score processes by the virtue of the previous lemma.

□

Proposition 2 (Rationality Test Consistency) *Under assumptions BC, BA, TA, the decision $d = 1(R^e > \tau)$ is asymptotically consistent, for τ fixed and small enough.*

□

Before we state and prove the next theorem the following notation must be defined:

$$\hat{J}_1(\theta) \equiv \frac{1}{T_e} \sum_{t=H-T_e+1}^{H} \nabla_\theta m_1(w_t^K, I_t, \theta),$$

$$\hat{J}_2(\lambda) \equiv \frac{1}{H} \sum_{t=1}^{H} \nabla_\lambda m_2(w_t^K, I_t, \lambda^K),$$

$$J_1(\theta) \equiv E_{\mathbb{M}} \nabla_\theta m_1(w_t^K, I_t, \theta)$$
$$J_2(\theta) \equiv E_{\mathbb{M}} \nabla_\theta m_2(w_t^K, I_t, \lambda^K).$$

$$\Omega_1 \equiv \lim_{T_e, H \longrightarrow \infty} \Bigg\{ \frac{1}{T_e} \sum_{t=H-T_e+1}^{H} E_{\mathbb{P}} m_1(w_t^K, I_t, \theta) m_1(w_t^K, I_t, \theta)' + \frac{1}{T_e} \sum_{t=H-T_e+1}^{H} \Bigg[\sum_{\substack{k=H-T_e+1 \\ k \neq t}}^{H} E_{\mathbb{P}} m_1(w_t^K, I_t, \theta) m_1(w_k^K, I_k, \theta)' + E_{\mathbb{P}} m_1(w_k^K, I_k, \theta) m_1(w_t^K, I_t, \theta)' \Bigg] \Bigg\}$$

$$\Omega_2 \equiv \lim_{H \longrightarrow \infty} \Bigg\{ \frac{1}{H} \sum_{t=1}^{H} E_{\mathbb{P}} m_2(w_t^K, I_t, \lambda^K) m_2(w_t^K, I_t, \lambda^K)' + \frac{1}{H} \sum_{t=1}^{H} \Bigg[\sum_{\substack{k=1 \\ k \neq t}}^{H} E_{\mathbb{P}} m_2(w_t^K, I_t, \lambda^K) m_2(w_k^K, I_k, \lambda^K)' + E_{\mathbb{P}} m_2(w_k^K, I_k, \lambda^K) m_2(w_t^K, I_t, \lambda^K)' \Bigg] \Bigg\}$$

Consistency of estimators $\hat{\lambda}^K$, $\hat{\theta}$ has been established in Theorem 1. The following theorem 2 derives the asymptotic distributions of the estimators, only using the pointwise convergence and stochastic equicontinuity results. Surely, this is not a "novel" result. However, we need it exactly in the form given, since it is referred too and used numerously in the subsequent proof. Since the proofs for specialized cases are widely available in the books, we only give a sketch. We shall be glad to cite the result if its available somewhere else precisely in the form stated below.

Proposition 3 (Asymptotic Representation of $\hat{\lambda}^K$, $\hat{\theta}$) *Under assumptions BC, BA, TA*

$$\sqrt{T_e}\Big(\hat{\theta} - \theta\Big) = -\Big[J_1(\theta) W_1 J_1(\theta)'\Big]^{-1} J(\theta) W_1 \Big(\frac{1}{\sqrt{T_e}} \sum_{t=H-T_e+1}^{H} m_1\left(w_t^K, I_t, \theta\right) + o_{\mathbb{P}}(1) \Big)$$

$$\sqrt{H}\Big(\hat{\lambda}-\lambda\Big) = -\Big[J_2(\lambda)W_2J_2(\lambda)'\Big]^{-1}J_2(\lambda)W_2\Big(\frac{1}{\sqrt{H}}\sum_{t=1}^{H}m_2\left(w_t^K, I_t, \lambda^K\right)+o_{\mathbb{P}}(1)\Big)$$

and the weak limits are given by:

$$\sqrt{T_e}\Big(\hat{\theta}-\theta\Big) \Rightarrow N\Big(0, \Big[J_1(\theta)W_1J_1(\theta)'\Big]^{-1}J_1(\theta)\Omega_1J_2(\theta)'\Big[J_1(\theta)W_1J_1(\theta)'\Big]^{-1}\Big),$$

$$\sqrt{H}\Big(\hat{\lambda}-\lambda\Big) \Rightarrow N\Big(0, \Big[J_2(\lambda)W_2J_2(\lambda)'\Big]^{-1}J_2(\lambda)\Omega_2J_2(\lambda)'\Big[J_2(\lambda)W_2J_2(\lambda)'\Big]^{-1}\Big)$$

□

Proposition 4 (Weak Convergence of the Econometrician's Rationality Score Processes) *The following weak limit results are true for convergence in $L^\infty(\mathcal{G})$, under H_O:*

$$\sqrt{T_e}S^e_{T_e}(g) \Rightarrow \mathbf{G}^e_{\mathbb{P}}(g)$$

and under H_A,

$$\sqrt{T_e}\Big(S^e_{T_e} - \mathbb{P}\lim S^e_{T_e}(g)\Big) \Rightarrow \mathbf{G}^e_{\mathbb{P}}(g)$$

where $\mathbf{G}^e_{\mathbb{P}}(g)$ a centered $\mathbb{P}$-Brownian Bridge, whose distribution is defined by the sample covariance operator:

$$\mathbb{CV}\left(\mathbf{G}^e_{\mathbb{P}}(g_1), \mathbf{G}^e_{\mathbb{P}}(g_2)\right) \equiv \mathbf{CV}_1(g_1, g_2)$$

where operator $\mathbf{CV}_1(g_1, g_2)$ is defined in the proof.

A dramatic simplification of the covariance operator is obtained under assumption BA3, i.e. by allowing the history be formed relative to the econometrician's sample.

Corollary 1 (Simplification of the Covariance Operator under Assumption BA3)

$$\mathbb{CV}\left(\mathbf{G}^e_{\mathbb{P}}(g_1), \mathbf{G}^e_{\mathbb{P}}(g_2)\right) \equiv \mathbf{CV}_1(g_1, g_2)$$

where operator $\mathbf{CV}_2(g_1, g_2)$ is defined in the proof.

Before we proceed it is useful to define the following test statistics. These are extremely useful when we later use the bootstrap to obtain the asymptotic approximations of the null distribution (and hence p-values) regardless of whether the null is true or not.

$$W^C_e \equiv \Big\{\ S^e_{T_e}(\mathbf{g}_i) - \mathbb{P}\lim S^e_{T_e}(g)\ \Big\}' W_{3T}\Big\{\ S^e_{T_e}(\mathbf{g}_i) - \mathbb{P}\lim S^e_{T_e}(g)\ \Big\}$$

where boldfacing denotes the usual vectorized notation. This statistics is formed when $\mathcal{G}$ is finite, it is a "centered" counterpart to W_e. The following statistic, U^C_e, is formed when $\mathcal{G}$ is infinite, and is a "centered" counterpart to U_e.

$$U^C_e \equiv \sup_{g\in\mathcal{G}}\ \Big|\ S^e_{T_e}(g) - \mathbb{P}\lim S^e_{T_e}(g)\ \Big|$$

Proposition 5 (Weak Convergence of the Econometrician's Test Statistics) *(i) The following weak limit results are true under H_O*

$$\sqrt{T_e} U_e \;\Rightarrow\; \sup_{g\in\mathcal{G}} \Big| \mathbf{G}^e_{\mathbb{P}}(g) \Big|$$

and under both H_O and H_A,

$$\sqrt{T_e} U^C_e \;\Rightarrow \sup_{g\in\mathcal{G}} \Big| \mathbf{G}^e_{\mathbb{P}}(g) \Big|$$

(i) The following weak limit results are true under H_O

$$T_e W_e \;\Rightarrow \Big\{ \mathbf{G}^e_{\mathbb{P}}(\mathbf{g}_i) \Big\} W_3 \Big\{ \mathbf{G}^e_{\mathbb{P}}(\mathbf{g}_i) \Big\}'$$

and under H_A,

$$T_e W^C_e \Rightarrow \Big\{ \mathbf{G}^e_{\mathbb{P}}(\mathbf{g}_i) \Big\} W_3 \Big\{ \mathbf{G}^e_{\mathbb{P}}(\mathbf{g}_i) \Big\}'$$

□

COMMENT: The above limit results could be restated in terms of generalized Bessel processes and their finite dimensional projections.

E.2 On Inference with the Sup Statistic U_e

A closed form asymptotic distribution for the test statistic is not known. However, the simulation and resampling methods can be used to construct valid approximations of its asymptotic distribution. An important question is how to approximate the distribution of continuous functionals of the coefficient and test processes. Of course, the finite dimensional distribution for any finite collection of g is already obtained in the statement of the theorems. Description of how to estimate the relevant quantities, needed for inference, is given in a next section. As a consequence of tightness (s.e.) of the process, its distribution can be approximated easily by considering, for example, the approximation of the processes by piecewise constants. Indeed, we can always choose a fine enough partition $\mathcal{G} = \cup_{i\in\mathcal{I}} \mathcal{G}_i$, so that to make the distance between the true and the approximate statistic, which is constant on each $\mathcal{G}_i$, sufficiently small. Indeed take δ as the size of largest cell in the partition. By definitions of s.e. there is T^0 large enough so that for $T > T^0$:

$$\mathbb{P}^* \left(\sup_{i\in\mathcal{I}} \sup_{g\in\mathcal{G}_i} \Big| Z_T(g) - Z_T(g_i) \Big| \geq \epsilon \right) \leq \epsilon$$

for any ϵ small. Denote the piecewise approximation by $\widetilde{Z}_T(\cdot)$. This therefore implies that

$$\mathbb{P}^* \left(\sup_{i\in\mathcal{I}} \Big| Z_T - \widetilde{Z}_T \Big| \geq \epsilon \right) \leq \epsilon$$

Exactly the same reasoning applies to the limit process Z. Denote its piecewise approximation by $\widetilde{Z}$, so that for a fine enough partition $\mathcal{G} = \cup_{i\in\mathcal{I}}\mathcal{G}_i$:

$$\mathbb{P}^*\left(\sup_{i\in\mathcal{I}}\left|Z - \widetilde{Z}\right| \geq \epsilon\right) \leq \epsilon.$$

From this we can validate the usage of distribution of $\sup_g \widetilde{Z}(g)$, which is finite-dimensional, for asymptotic inference about $\sup_g Z_T(g)$. This enables a variety of inference techniques: bootstrap, sub-sampling, inference using the limit distributions (substituting for estimated quantities). That is, any inference method applies as long as it can provide valid inference for finite-dimensional distribution. This also means that to obtain a desired precision, we can partition $\mathcal{G}$ in sufficiently small subsets and then approximate the distribution of the functional by a functional of the approximation.

Another important question is how to make the qualitative assessment of approximation. In practice, one may design a formal decision rule that considers successive approximations and measures the difference in improvements, stopping when the difference is small. When designing the stopping criteria, it helps to know the rate of convergence of successive approximation. The measures of rate depend both on the speed of convergence of the processes to the limit and on the 'quality' of stochastic equicontinuity, both of which could be assessed under given scenarios. There is a wealth of fairly recent but growing literature on the approximations. To cite only few, see Dehling, Denker, and Philipp (1986), Herrndorf (1983a), Philipp (1994), Koltchinskii (1994), Ledoux and Talagrand (1991), sources cited in there, and others. The settings that we consider here could be adopted by using the asymptotic representation obtained in metric space $L^\infty(\mathcal{M})$ and noting that these could also be treated as processes in various Banach, particularly Hilbert spaces, and so much of the above literature applies [30]. The test statistic would have to be redefined then and will be consistent only against a particular class of alternatives.

F Proofs

F.1 Proof of Theorem 1

Unless otherwise stated the limits here are taken as as $T_e \to \infty, H \to \infty$.

Uniform convergence of objective functions $Q_{1T}(\theta)$, $Q_{2T}(\lambda)$ follows form pointwise convergence (by BA4, BC1,) combined with s.e. (BA4, BC2). The textbook argument, Amemiya (1985), applies along with (TC4) implying the consistency of $\hat{\theta}$, $\hat{\lambda}$.

[30] Note, however, that we may get weaker notion of convergence – e.g. by switching to Hilbert space L^2, the set of continuous functionals on the processes will exclude many functionals that are continuous with respect the sup norm but not L^2 norm.

Assumption BA1 implies along with TA1 that

$$\begin{aligned}
&\left| f_1(w_t^K, \theta, \lambda_t^K, \alpha) - f_1(w_t^K, \theta, \lambda^K, \alpha) \right| \\
&\leq \left\| \nabla_\lambda f_1(w_t^K, \theta, \lambda_t^K, \alpha) \right\| \left\| \lambda_t^K - \lambda^{*K} \right\| \\
&\leq O_{\mathbb{P}^*}(1) o_{\mathbb{P}^*}(1), \quad \text{uniformly in } \alpha \in \mathcal{A}, \theta \in \Theta, t \in [H - T_e + 1, \dots, H] \\
&\text{as } T_e \to \infty, H \to \infty
\end{aligned}$$

Next we claim Assumption TA2 implies:

$$\left| \tilde{\alpha}_t^K - \alpha_t^K \right| = o_{\mathbb{P}^*}(1), \text{ uniformly in } t \in [H - T_e + 1, \dots, H] \text{ , as } T_e \to \infty, H \to \infty$$

Indeed, we have $\underline{\lim} \mathbb{P}_*(\mathbf{A}) = 1$, where

$$\mathbf{A} \equiv \left\{ f_1^2 \left(w_t^K, \theta, \lambda^K, \alpha_t^K \right) + \epsilon > f_1^2 \left(w_t^K, \theta, \lambda^K, \alpha_t^K \right) \right\}$$

by construction and since $\underline{\lim} \mathbb{P}_*(\mathbf{B}) = 1$ (by above uniform convergence) , where

$$\begin{aligned}
\mathbf{B} \equiv \Big\{ & f_1^2 \left(w_t^K, \theta, \lambda_t^K, \alpha_t^K \right) + \frac{\epsilon}{4} > f_1^2 \left(w_t^K, \theta, \lambda_t^K, \tilde{\alpha}_t^K \right), \\
& f_1^2 \left(w_t^K, \theta, \lambda^K, \alpha_t^K \right) + \frac{\epsilon}{4} > f_1^2 \left(w_t^K, \theta, \lambda_t^K, \tilde{\alpha}_t^K \right), \\
& f_1^2 \left(w_t^K, \theta, \lambda_t^K, \tilde{\alpha}_t^K \right) + \frac{\epsilon}{4} > f_1^2 \left(w_t^K, \theta, \lambda^K, \tilde{\alpha}_t^K \right) \Big\}
\end{aligned}$$

Next assumption TA2 yields that for ν small, there are balls $B(\nu, \alpha_t)$, $\forall t$ as $H \to \infty, T_e \to \infty$

$$\inf_{t \in [H - T_e + 1, \dots, H]} \left(\inf_{\alpha \in N(\nu, \alpha_t)} f_1^2 \left(w_t^K, \theta, \lambda^K, \alpha \right) - f_1^2 \left(W_t^K, \theta, \lambda^K, \alpha_t^K \right) \right) \geq \epsilon > 0$$

Then it follows that $\mathbf{A} \subset \mathbf{C}$ where

$$\mathbf{C} \equiv \left\{ \inf_{t \in [H - T_e + 1, \dots, H]} \left(\inf_{\alpha \in B(\nu, \alpha_t^K)} f_1^2 \left(w_t^K, \theta, \lambda^K, \alpha \right) - f_1^2 \left(w_t^K, \theta, \lambda^K, \tilde{\alpha}_t^K \right) \right) \geq \epsilon \right\}$$

and thus $\underline{\lim} \mathbb{P}_* \left(\tilde{\alpha} \in B(\nu, \alpha_t^K, \forall t \in [H - T_e + 1, \dots, H] \right) = 1$, for any ν fixed and small. Or equivalently, $\overline{\lim} \mathbb{P}^* \left(\tilde{\alpha} \notin B(\nu, \alpha_t^K, \forall t \in [H - T_e + 1, \dots, H] \right) = 0$, for any ν fixed and small, confirming the claim.

The same when reaplcing λ_t^K, $\hat{\lambda}^K$, hence

$$\left| \hat{\alpha}_t^K - \alpha_t^K \right| = o_{\mathbb{P}^*}(1), \text{ uniformly in } t \in [H - T_e + 1, \dots, H] \text{ , as } T_e \to \infty, H \to \infty$$

Minkowski yields:

$$\left| \hat{\alpha}_t^K - \tilde{\alpha}_t^K \right| = o_{\mathbb{P}^*}(1), \text{ uniformly in } t \in [H - T_e + 1, \dots, H] \text{ , as } T_e \to \infty, H \to \infty$$

This completes the proof of the last two statements.

■

F.2 Proof of Lemma 1

In this proof it is useful to write

$$\alpha_t(I_t, \lambda_t^K) \equiv \tilde{\alpha}_t^K$$

stressing its definition as an implicit function.

$$\begin{aligned}
A(g) &\equiv \frac{1}{M} \sum_{t=H-M+1}^{M} m_{3g}\left(w_t^K, I_t, \theta, \lambda_t^K, \alpha_t(\lambda_t^K)\right) \\
&= \frac{1}{M} \sum_{t=H-M+1}^{M} m_{3g}\left(w_t^K, I_t, \theta, \lambda_{H-M+1}^K, \alpha_t(\lambda_t^K)\right) \\
&+ \frac{1}{M} \sum_{t=H-M+1}^{M} \Big(\nabla_\lambda m_{3g}\left(w_t^K, I_t, \theta, \lambda_t^{*K}, \alpha_t(\lambda_t^K)\right) \\
&+ \nabla_\alpha m_{3g}\left(w_t^K, I_t, \theta, \lambda_t^K, \alpha_t(\lambda_t^{*K})\right) \left[\nabla_\alpha f_1(W_t^K, \theta, \lambda_t^{*K}, \alpha_t(\lambda_t^{*K}))\right]^{-1} \nabla_\lambda f_1(W_t^K, \theta, \lambda_t^{*K}, \alpha_t(\lambda_t^{*K}))\Big) \\
&\times \left(\lambda_t^K - \lambda_{H-M+1}^K\right)
\end{aligned}$$

where stars denote intermediate values, as in standard expansion,

$$\begin{aligned}
&= \frac{1}{M} \sum_{t=H-M+1}^{M} m_{3g}\left(w_t^K, I_t, \theta, \lambda_{H-M}^K, \alpha(\lambda_{H-M+1}^K)\right) + o_{\mathbb{P}}\left(\left\|\lambda^K - \lambda_{H-M+1}^K\right\|\right) \\
&= \frac{1}{M} \sum_{t=H-M+1}^{M} m_{3g}\left(w_t^K, I_t, \theta, \lambda^K, \alpha_t^K\right) + o_{\mathbb{P}}(1)
\end{aligned}$$

which uses BA1, TA, uniformly in $g \in \mathcal{G}$. Therefore, applying BC1 and BC2, we have uniformly in $g \in \mathcal{G}$

$$\begin{aligned}
S_M^{ai}(g) &\xrightarrow{\mathbb{P}^*} \lim_{M\to\infty, H\to\infty} \frac{1}{M} \sum_{t=H-M+1}^{M} E_{\mathbb{P}} m_{3g}\left(w_t^K, I_t, \theta, \lambda^K, \alpha_t^K\right) - E_{\mathbb{M}} g(w_t^K) \\
&\equiv E_{\mathbb{M}} m_{3g}\left(w_t^K, I_t, \theta, \lambda^K, \alpha_t^K\right) - E_{\mathbb{M}} g(w_t^K)
\end{aligned}$$

Under H_0 the right hand side is zero. The proof for $S^{ai}(g)$, $i = 1, 2, 4$ is similar and easier.

■

F.3 Proof of Lemma 2

It is very similar to the proof of lemma 1, and hence has been dropped for brevity.

■

F.4 Proof of Theorem 2

Follows immediately from lemma 1 and 2, using definition of outer probability.

■

F.5 Proof of Theorem 4

Consider the proof for H_O case.

Write

$$\begin{aligned}
\sqrt{T_e}A(g) = & \frac{1}{\sqrt{T_e}} \sum_{t=H-T_e+1}^{H} m_{3g}\left(w_t^K, I_t, \theta, \lambda^K, \alpha_t(\lambda_t^K)\right) \\
& + \left[\frac{1}{T_e} \sum_{t=H-T_e+1}^{H} \left(\nabla_\lambda m_{3g}\left(w_t^K, I_t, \hat{\theta}, \lambda^{*K}, \alpha_t(\hat{\lambda}^K)\right) \right.\right. \\
& + \frac{1}{T_e} \sum_{t=H-T_e+1}^{H} \left\{ \nabla_\alpha m_{3g}\left(w_t^K, I_t, \hat{\theta}, \hat{\lambda}^K, \alpha_t(\lambda^{*K})\right) \right. \\
& \left.\left.\left. \left[\nabla_\alpha f_1\left(w_t^K, \hat{\theta}, \lambda^{*K}, \alpha_t(\lambda^{*K})\right)\right]^{-1} \nabla_\lambda f_1\left(w_t^K, \hat{\theta}, \lambda^{*K}, \alpha_t(\lambda^{*K})\right) \right\}\right]\right. \\
& \times \sqrt{T_e}\left(\hat{\lambda}^K - \lambda^K\right) \\
& + \left[\frac{1}{T_e} \sum_{t=H-T_e+1}^{H} \nabla_\theta m_{3g}\left(w_t^K, I_t, \theta^*, \hat{\lambda}, \alpha_t(\hat{\lambda}^K)\right) \right] \\
& \times \sqrt{T_e}\left(\hat{\theta} - \theta\right)
\end{aligned}$$

Employing conditions BC, BA, and TA, the following representation is uniform in $g \in \mathcal{G}$:

$$\begin{aligned}
\sqrt{T_e}S_e(g) = & \left(\frac{1}{\sqrt{T_e}} \sum_{t=H-T_e+1}^{H} m_{3g}\left(w_t^K, I_t, \theta, \lambda^K, \alpha_t(\lambda_t^K)\right) - \frac{\sqrt{T_e}}{H} \sum_{t=1}^{H} g(w_t^K) \right) \\
& + \left[J_{3g}(\theta, \lambda^K) + J_{4g}(\theta, \lambda) + o_{\mathbb{P}^*}(1) \right] \times \sqrt{T_e}\left(\hat{\lambda}^K - \lambda^K\right) \\
& + \left[J_{5g}(\theta, \lambda^K) + o_{\mathbb{P}^*}(1) \right] \times \sqrt{T_e}\left(\hat{\theta} - \theta\right)
\end{aligned}$$

where J_3, J_4, J_5 are the uniform probability limits of the corresponding quantities enclosed in square brackets in the previous equation . These limits are defined in the next subsection, along with the covariance kernel.

Let $\frac{T_e}{H} \to \tau_e \in [0, 1)$, where τ_e is the asymptotic proportion of the sample the econometrician uses.

Further employing the representation stated in theorem 3, Slutsky implies:

$$\begin{aligned}
\sqrt{T_e} S_e(g) = & \left(\frac{1}{\sqrt{T_e}} \sum_{t=H-T_e+1}^{H} m_{3g}\left(w_t^K, I_t, \theta, \lambda^K, \alpha_t(\lambda_t^K)\right) - \sqrt{\tau_e} \frac{1}{\sqrt{H}} \sum_{t=1}^{H} g(w_t^K) \right) \\
& - \left[J_3(\theta, \lambda) + J_4(\theta, \lambda) \right] \\
& \times \left\{ \left[J_1(\theta) W_1 J_1(\theta)' \right]^{-1} J(\theta) W_1 \Big(\frac{1}{\sqrt{T_e}} \sum_{t=H-T_e+1}^{H} m_1\left(w_t^K, I_t, \theta\right) \Big) \right\} \\
& - \left[J_5(\theta, \lambda) \right] \\
& \times \left\{ \left[J_2(\lambda) W_2 J_2(\lambda)' \right]^{-1} J_2(\lambda) W_2 \Big(\frac{1}{\sqrt{H}} \sum_{t=1}^{H} m_2\left(w_t^K, I_t, \lambda\right) \Big) \right\} \\
& + o_{\mathbb{P}^*}(1)
\end{aligned}$$

Further, the fidi convergence in distribution is given by BC3 (under H_O):

$$\left\{ \sqrt{T_e} S_e(g_1), \ldots, \sqrt{T_e} S_e(g_K) \right\} \xrightarrow{\mathcal{D}} N(0, M)$$

where $M_{ij} = \mathbb{CV}^*(g_i, g_j)$. The operator $\mathbf{CV}$ is stated in the next section.

Employing BC2, conclude by Theorems 1.5.4 and 1.5.7 in van der Vaart and Wellner (1996):

$$\sqrt{T_e} S^e(g) \Rightarrow \mathbf{G}^e_{\mathbb{P}^*}(g)$$

with covariance kernel given by $\mathbf{CV}_1$, which is stated next and incorporates both H_O and H_A cases.

Note that the H_A case follows similarly, by additional centering around the asymptotic mean of the first term in the above representation.

■

F.6 Covariance Operators $\mathbf{CV}_1$, $\mathbf{CV}_2$

Let

$$J_3(\theta,\lambda) \equiv \lim E_{\mathbb{P}} \frac{1}{T_e} \sum_{t=H-T_e+1}^{H} \left(\nabla_\lambda m_{3g}\left(w_t^K, I_t, \theta, \lambda^K, \alpha_t(\lambda^K)\right) \right)$$

$$J_4(\theta,\lambda) \equiv \lim E_{\mathbb{P}} \frac{1}{T_e} \sum_{t=H-T_e+1}^{H} \Big\{ \nabla_\alpha m_{3g}\left(w_t^K, I_t, \theta, \lambda^K, \alpha_t(\lambda^K)\right)$$

$$\left[\nabla_\alpha f_1\left(w_t^K, \theta, \lambda^K, \alpha_t(\lambda^K)\right)\right]^{-1} \nabla_\lambda f_1\left(w_t^K, \theta, \lambda^K, \alpha_t(\lambda^K)\right) \Big\}$$

$$J_5(\theta,\lambda) \equiv \lim E_{\mathbb{P}} \frac{1}{T_e} \sum_{t=H-T_e+1}^{H} \left\{ \nabla_\theta m_{3g}\left(w_t^K, I_t, \theta, \lambda, \alpha_t(\lambda^K)\right) \right\}$$

Let

$$\begin{aligned}
\mu(g) \equiv & \left(\frac{1}{\sqrt{T_e}} \sum_{t=H-T_e+1}^{H} m_{3g}\left(w_t^K, I_t, \theta, \lambda^K, \alpha_t(\lambda_t^K)\right) - \sqrt{\tau_e} \frac{1}{\sqrt{H}} \sum_{t=1}^{H} g(w_t^K) - \sqrt{T_e} \mathbb{P} \lim S^e(g) \right) \\
& - \Big[J_3(\theta,\lambda) + J_4(\theta,\lambda) \Big] \\
& \times \left\{ \Big[J_1(\theta) W_1 J_1(\theta)' \Big]^{-1} J(\theta) W_1 \Big(\frac{1}{\sqrt{T_e}} \sum_{t=H-T_e+1}^{H} m_1\left(w_t^K, I_t, \theta\right) \Big) \right\} \\
& - \Big[J_5(\theta,\lambda) \Big] \\
& \times \left\{ \Big[J_2(\lambda) W_2 J_2(\lambda)' \Big]^{-1} J_2(\lambda) W_2 \Big(\frac{1}{\sqrt{H}} \sum_{t=1}^{H} m_2\left(w_t^K, I_t, \lambda\right) \Big) \right\}
\end{aligned}$$

Then under BA2,

$$\mathbf{CV}_1(g_1, g_2) \equiv \lim_{T_e, H \to \infty} Cov(\mu(g_1), \mu(g_2))$$

Consider then BA3, and let

$$\psi(g) \equiv \left(\frac{1}{\sqrt{T_e}} \sum_{t=H-T_e+1}^{H} m_{3g} \left(w_t^K, I_t, \theta, \lambda^K, \alpha_t(\lambda_t^K) \right) - \sqrt{T_e}(\mathbb{P} \lim S^e(g) - E_{\mathbb{M}}(g(w_t^K)) \right)$$

$$- \left[J_3(\theta, \lambda) + J_4(\theta, \lambda) \right]$$

$$\times \left\{ \left[J_1(\theta) W_1 J_1(\theta)' \right]^{-1} J(\theta) W_1 \Big(\frac{1}{\sqrt{T_e}} \sum_{t=H-T_e+1}^{H} m_1 \left(w_t^K, I_t, \theta \right) \Big) \right\}$$

Then under BA2,

$$\mathbf{CV}_2(g_1, g_2) \equiv \lim_{T_e, H \to \infty} Cov(\psi(g_1), \psi(g_2))$$

□

References

AMEMIYA, T. (1985): *Advanced Econometrics*. Harvard University Press.

ANDERSON, E., L. P. HANSEN, AND T. J. SARGENT (1999): "Risk and Robustness in Equilibrium," Manuscript.

ANDREWS, D., AND D. POLLARD (1994): "An Introduction to Functional Central Limit Theorems for Dependent Stochastic Processes," *International Statistical Review*, 62, 119–132.

ANDREWS, D. W. (1994): "An Introduction to Econometric Applications of Empirical Process Theory for Dependent Random Variables," *preprint*.

ANDREWS, D. W. K. (1989): "An Empirical Process Central Limit Theorem for Dependent Non-Identically Distributed Random Variable," Discussion paper, Yale University, Cowles Foundation, Discussion Paper 907.

——— (1993): "An Introduction to Econometric Applications of Empirical Process Theory for Dependent Random Variables," *Econometric Reviews*, 12(2), 183–216.

BARBERIS, N., A. SHLEIFER, AND R. VISHNY (1998): "The Model of Investor Sentiments," *Journal of Financial Economics*, 49, 307–341.

BERKES, I., AND W. PHILIPP (1998): "Limit Theorems for Mixing Sequences without Rate Assumptions," *Ann. Probab.*, 26(2), 805–831.

BLACK, S. (1997): "Forward Discount Puzzle in a Rational Beliefs Framework," Manuscript, Stanford University.

BREEDEN, D. T. (1979): "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics*, 7, 265–296.

BROCK, W. A., AND B. D. LEBARON (1996): "A Dynamic Structural Model for Stock Return Volatility and Trading Volume," *Review of Economics and Statistics*, 78(1), 94–110.

BURNSIDE, C. (1994): "Hansen-Jagannathan Bounds as Classical Tests of Asset Pricing Models," *Journal of Business and Economic Statistics*, 12, 57–79.

COCHRANE, J. H. (1989): "The Sensitivity Tests of the Intertemporal Allocation of Consumption to Near-Rational Alternatives," *American Economic Review*, 79, 319–337.

DAVIDSON, J. (1994): *Stochastic Limit Theory*. The Clarendon Press, Oxford University Press, New York.

DEHLING, H., M. DENKER, AND W. PHILIPP (1986): "Central Limit Theorems for Mixing Sequences of Random Variables under Minimal Conditions," *Ann. Probab.*, 14(4), 1359–1370.

DOUKHAN, P. (1994): *Mixing: Properties and Examples*. Springer-Verlag, New York.

EPSTEIN, L. G., AND T. WANG (1994): "Intertemporal Asset Pricing under Knightian Uncertainty," *Econometrica*, 62(3), 283–322.

FROOT, K. A., AND J. FRANKEL (1989): "Forward Discount Bias: Is it an Exchange Risk Premium?," *Quarterly Journal of Economics*, 104, 139–161.

GALLANT, A., AND H. WHITE (1988): *Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*. New York: Bassil Blackwell.

GEWEKE, J. (1999): "Computational Experiments and Reality," Manuscript, University of Minnesota.

GILBOA, I., AND D. SCHMEIDLER (1989): "Maxmin Expected Utility with Non-unique Prior," *Journal of Mathematical Economics*, 18, 141–153.

HANSEN, L. P., AND R. JAGANNATHAN (1991): "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy*, 99, 225–262.

HANSEN, L. P., AND T. J. SARGENT (1993): "Aggregation over Time and the Inverse Optimal Predictor Problem for Adaptive Expectations in Continuous Time," *International Economic Review*, 24, 1–20.

HANSEN, L. P., T. J. SARGENT, AND T. D. TALLARINI (1999): "Robust Permanent Income and Pricing," Manuscript.

HANSEN, L. P., AND K. SINGLETON (1982): "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectation Models," *Econometrica*, 50(5), 1269–1286.

HERRNDORF, N. (1983a): "Approximation of Vector-valued Random Variables by Constants," *J. Approx. Theory*, 37(2), 175–181.

HERRNDORF, N. (1983b): "The Invariance Principle for φ-mixing Sequences," *Z. Wahrsch. Verw. Gebiete*, 63(1), 97–108.

——— (1984a): "A Functional Central Limit Theorem for ρ-mixing Sequences," *J. Multivariate Anal.*, 15(1), 141–146.

——— (1984b): "A Functional Central Limit Theorem for Weakly Dependent Sequences of Random Variables," *Ann. Probab.*, 12(1), 141–153.

HOFFMANN-JØRGENSEN, J. (1991): *Stochastic Processes on Polish spaces*. Aarhus Universitet, Matematisk Institut, Aarhus.

——— (1994): *Probability with a View toward Statistics.* Chapman & Hall, New York.

KOENKER, R., AND J. A. F. MACHADO (1997): "GMM Inference When the Number of Moment Conditions is Large," Manuscript, University of Illinois at Urbana-Champaign.

KOLTCHINSKII, V. I. (1994): "Komlos-Major-Tusnady Approximation for the General Empirical Process and Haar Expansions of Classes of Functions," *J. Theoret. Probab.*, 7(1), 73–118.

KRUSELL, P., AND A. SMITH (1996): "Rules of Thumb in Macroeconomic Equilibrium: a Quantitative Analysis," *Journal of Economic Dynamics and Control*, 20, 527–558.

KURZ, M. (1974): "The Kesten-Stignum Model and the Treatment of Uncertainty in Equilibrium Theory," in *Essays on Economic Behavior under Uncertainty*, ed. by M. S. Balch, D. L. McFadden, and S. Y. Wu, pp. 389–399. North-Holland, Amsterdam.

——— (1994a): "On Rational Belief Equilibria," *Economic Theory*, 4, 859–876.

——— (1994b): "On the Structure and Diversity of Rational Beliefs," *Economic Theory*, 4, 877–900.

——— (ed.) (1997a): *Endogenous Economic Fluctuations*, vol. 6 of *Studies in Economic Theory*. Springer-Verlag.

——— (1997b): *On the Volatility of Foreign Exchange Rates,*vol. 6 of *Studies in Economic Theory*, chap. 12, pp. 317–352. Springer-Verlag, Amsterdam.

——— (1998): "Social States of Belief and the Determinant of Equity Risk Premium in a Rational Belief Equilibrium," in *Functional Analysis and Economic Theory*, ed. by Y. A. Abramovich, E. Avgerinos, and N. C. Yannelis, Springer Series in Applied Mathematics and Operations Research, pp. 171–220. Springer-Verlag.

KURZ, M., AND A. BELTRATTI (1997): *The Equity Premium is no Puzzle,*vol. 6 of *Studies in Economic Theory*, chap. 11, pp. 283–315. Springer-Verlag.

KURZ, M., AND M. MOTOLESE (1998): "Beliefs and Ergodic Properties of Dynamic Economies," Manuscript, Stanford University.

LEDOUX, M., AND M. TALAGRAND (1991): *Probability in Banach spaces*. Springer-Verlag, Berlin, Isoperimetry and processes.

LUCAS, R. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46, 1429–1446.

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.

MCLEISH, D. L. (1974): "Dependent Central Limit Theorems and Invariance Principles," *Ann. Probability*, 2, 620–628.

MEHRA, R., AND E. C. PRESCOTT (1985): "The Equity Premium; a Puzzle," *Journal of Monetary Economics*, 16(145-161).

MOTOLESE, M. (1998): "Dynamic Non-neutrality of Money under Rational Beliefs: the Role of Endogenous Uncertainty," Ph.D. thesis, Università Degli Studi di Bologna.

MUTH, J. F. (1961): "Rational Expectations and the Theory of Price Movements," *Econometrica*, 29, 315–335.

NIELSEN, C. K. (1997a): *On some Topological Properties of Stable Measures,*vol. 6 of *Studies in Economic Theory*, chap. 11, pp. 69–92. Springer-Verlag.

——— (1997b): *Rational Belief Structures and Rational Belief Equilibria,*vol. 6 of *Studies in Economic Theory*, chap. 11, pp. 121–144. Springer-Verlag.

PHILIPP, W. (1986): "Invariance Principles for Independent and Weakly Dependent Random Variables," in *Dependence in Probability and Statistics (Oberwolfach, 1985)*, pp. 225–268. Birkhäuser Boston, Boston, MA.

——— (1994): "Empirical Distribution Functions and Strong Approximation Theorems for Dependent Random Variables. A Problem of Baker in Probabilistic Number Theory," *Trans. Amer. Math. Soc.*, 345(2), 705–727.

POLITIS, D. N., AND J. P. ROMANO (1994): "The stationary bootstrap," *J. Amer. Statist. Assoc.*, 89(428), 1303–1313.

POLLARD, D. (1990): *Empirical Processes: Theory and Applications.* Hayward, CA: Institute of Mathematical Statistics.

PRESCOTT, E., AND R. MEHRA (1980): "Recursive Competitive Equilibrium: The case of Homogeneous Households," *Econometrica*, 48, 1365–1380.

VAN DER VAART, A. W., AND J. A. WELLNER (1996): *Weak Convergence and Empirical Processes*. Springer-Verlag, New York.