"Some properties of autocoherent models"

Gilles Saint Paul

Institut d'Economie Industrielle (IDEI) – Manufacture des Tabacs
Aile Jean-Jacques Laffont – 21, allée de Brienne – 31000 TOULOUSE – FRANCE
Tél. + 33(0)5 61 12 85 89 – Fax + 33(0)5 61 12 86 37 – **www.idei.fr** – contact@cict.fr

INSTITUT
D'ECONOMIE
INDUSTRIELLE

Some properties of autocoherent models

Gilles Saint-Paul
Toulouse School of Economics and CEPR

April 4, 2012

ABSTRACT

An autocoherent model is a model which is validated by the data if people use it to form their expectations. A structural model may be incorrect but autocoherent, thus supporting a self-confirming equilibrium. This paper explores some mathematical properties of autocoherent models. The first part clarifies the relationship between autocoherence and identification.It establishes sufficient conditions under which an expert constraints is compelled to reveal the true value of some parameter. These conditions are related to the traditional notion of identification, but it must be amended to reflect the performativity of the perceived model and the fact that identification is different depending on the econometrician's assumptions about the perceived model's validity. The second part clearly spells out the conditions for an autocoherent model equilibrium to arise in the linear/Gaussian case, and provides an equivalent characterization based on an "interpretation". That is, an autocoherent model equilibrium can be constructed on the basis of a linear transformation which maps the actual realization of the shocks to their "interpreted counterpart", defined as the value of the shocks consistent with the observed outcomes on the basis of the (incorrect) perceived model. If such a transformation exists then the perceived model can support a self-confirming equilibrium.

KEYWORDS: Rational expectations, self-confirming equilibrium, identification, learning, autocoherent models, performativity

JEL: A11, E6.

1 Introduction

While rational expectations theory typically assumes that people use the correct model to form their forecasts, more recent research has studied what happens when that is not the case. Authors like Sargent (2008) and Fudenberg and Levine (2007) have argued that the economy may settle at a point where incorrect beliefs are sustained in equilibrium, because to invalidate those beliefs the economy would have to engage in an off-equilibrium path. This is the essence of the self-confirming equilibrium concept (SCE) of Fudenberg and Levine (1993).

Imposing that the economy settles at an SCE rather than a rational expectations equilibrium with correct beliefs is clearly less restrictive. In macroeconomics, we may want to impose that agents form their beliefs using an explicit structural model. The fact that such a model supports an SCE then means that in an equilibrium where agents use it, the model matches the data, i.e. correctly predicts the moments of the observables. This is what I define as an "autocoherent model". In such a world, a theory of how expectations are formed is a theory of which model people use, and it is then reasonable to impose that such a model be autocoherent. Otherwise, the model would be "counter-performative" in that its adoption to form expectations would lead it to be eventually rejected when confronted with the data[1].

If there are enough observables relative to the number of parameters of the relevant model, then only the correct model is likely to be autocoherent. Otherwise, there will typically exists a continuum of autocoherent models. Autocoherence alone then does not suffice to predict which model will be used; a positive theory of which model is actually used is needed. In Saint-Paul (2011a,b) I have analyzed, for some specific examples, the case where

---

[1]See McKenzie, 2006.

the model is produced by intellectuals who pursue their own agenda, while facing the constraint that the model be autocoherent. This approach was a first step in understanding the political economy of models. It delivers some plausible predictions regarding how ideological biases may influence the parameter values in the model proposed by the expert (such as the Keynesian multiplier).

Instead, this paper is more analytical and explores some mathematical properties of autocoherent models. It consists of two parts. The first part clarifies the relationship between autocoherence and the traditional econometric notion of identification. In particular, it establishes sufficient conditions under which an expert subject to autocoherence constraints is compelled to reveal the true value of some parameter. These conditions are closely related, of course, to the parameter being identified in the econometric sense, but must be amended to reflect the performativity of the expert (i.e., if people adopt his model, the equilibrium and therefore the data generating process change) and the fact that identification is different depending on whether or the econometrician assumes that people use the correct model to form their expectations. The second part clearly spells out the conditions for an autocoherent model equilibrium to arise in the linear/Gaussian case, and provides an equivalent characterization based on an "interpretation". That is, an autocoherent model equilibrium can be constructed on the basis of a linear transformation which maps the actual realization of the shocks to their "interpreted counterpart", defined as the value of the shocks consistent with the observed outcomes on the basis of the (incorrect) perceived model. If such a transformation, which must be orthogonal for the scalar product defined by the variance-covariance matrix of the observables, exists, then the perceived model is autocoherent.

2 Autocoherence and identification

We start by tackling the following question: Suppose that the model used by people in forming their expectations is designed by an expert, whose theory must be compatible with the data. How much discipline does such a constraint impose on the expert? One way to tackle this issue is to ask how many parameters' true value will the expert be forced to reveal. This in turn brings the question of the relationship between autocoherence and identification.

Let $v \in \mathbb{R}^n$ be the vector of correct structural parameters and $\hat{v} \in \mathbb{R}^n$ the vector of perceived structural parameters. In general, economic outcomes will depend both on the actual and perceived parameters, the former affecting outcomes directly and the latter through private expectations and government policies. Let $M(v, \hat{v})$ be a vector representing the relevant empirical moments of the distribution of observables[2]. This vector is treated as a function of the actual and perceived model.

Definition 1 – A model $\hat{v}$ is *autocoherent* if and only if

$$M(v, \hat{v}) = M(\hat{v}, \hat{v}).$$

We can denote by $AC(v)$ the set of autocoherent models. This set clearly depends on the correct model. The correct model is always autocoherent:

$$v \in AC(v).$$

---

[2]If it is common knowledge that the distribution of observables is part of a family spanned by a few parameters, then $M()$ is the vector of these parameters. For example if the observables are a Gaussian vector, $M()$ consists of the mean vector and its variance-covariance matrix. If there is no such common knowledge, M would then be the entire distribution, or at least the vector of coordinates of that distribution in a base of the functional space from which the distribution is drawn.

Throughout the whole paper, sample moments are assumed for simplicity to be equal to the true moments of the underlying distribution, that is, everything takes place as if there were an infinite number of observations.

People may also have prior ideas (or information) about which model may be correct. Therefore I will assume that there is a set V of admissible models, and that any model formulated by the expert must be in $V \cap AC(v)$. Finally it is also reasonable to assume that $v \in V$.

Let us now consider the inferences about the correct model that an econometrician would make in this world where the perceived model might be incorrect. One possibility is that the econometrician wrongly believes that people use the true model, as do the people themselves, but does not observe the model used by the people.

Definition 2 – A model $\tilde{v}$ is acceptable with unknown beliefs (or u-acceptable) if

$$M(v, \hat{v}) = M(\tilde{v}, \tilde{v}).$$

The set of u-acceptable models is denoted as $UA(v, \hat{v})$. Note that it depends on both v and $\hat{v}$, contrary to AC. Two different models in AC deliver two different self-confirming equilibria and two different values of the moment vector $M(v, \hat{v})$. In contrast, when the econometrician considers all the possible models that may explain the empirical moments, the equilibrium and therefore those moments remain invariant. That is, the econometrician lives in a given equilibrium and cannot change it, in contrast to the expert who may influence behavior through the formation of expectations. It is immediate to prove the following:

Proposition 1 –$\forall v$, $(\hat{v} \in AC(v) \iff \hat{v} \in UA(v, \hat{v}))$.

Saying that a model is autocoherent is equivalent to saying that it is one of the models that are acceptable, in an equilibrium where it is the perceived model. Note that the correct model is not in general u-acceptable. If people believed in the correct model, the empirical moments would be different. The incorrect identifying assumption that people use the correct model generally prevents the econometrician from considering it as potentially correct.

Alternatively, the econometrician may observe the perceived model, but be uncertain about whether it is the correct one. This leads to the following definition

Definition 3 – A model $\tilde{v}$ is acceptable with known beliefs (or k-acceptable) if

$$M(v, \hat{v}) = M(\tilde{v}, \hat{v}).$$

I will denote the set of k-acceptable models by $KA(v, \hat{v})$.

Proposition 2 – (i) $\forall \hat{v}$, $\forall v$, $v \in KA(v, \hat{v})$
(ii) $\forall v, (\hat{v} \in AC(v) \iff \hat{v} \in KA(v, \hat{v}))$.

Now the correct model is k-acceptable, and so is the perceived one, otherwise it would not be autocoherent.

For the sake of completeness, we might also consider an econometrician who could not observe $\hat{v}$ and would not assume that the perceived model is correct. He would then have to consider all the pairs $(\tilde{v}, v^*)$ such that $M(v, \hat{v}) = M(\tilde{v}, v^*)$. However, there is nothing we will do with such a definition of acceptability, so there is no need to pursue it further.

With this apparatus in hand, we can ask whether identification of a parameter compels the expert to reveal its true value. This question is not totally obvious, because a change in the value of a perceived parameter leads to a different equilibrium, where this different value might be acceptable even though it was not in the original equilibrium.

To address this question I define identification as follows.

Definition 4 – Let $\pi : \mathbb{R}^n \to \mathbb{R}^p$ be a function, also referred to as a "parameter". Let $S \subset \mathbb{R}^n$. Then π is S-identified if and only if: $|\pi(S)| = 1$.

In definition 4, $\pi()$ is a composite vector of parameters derived from the structural ones, and S the set of acceptable models according to some definition. If all those models deliver the same value for π, then it is identified.

Proposition 3 – Let $\pi()$ be a parameter. Assume that for any $\hat{v} \in V \cap AC(v)$, π is $KA(v, \hat{v})$-identified. Then for any $\hat{v} \in V \cap AC(v), \pi(\hat{v}) = \pi(v)$.

Proposition 3 tells us that if π is always identified, at any equilibrium supported by an admissible autocoherent model, then autocoherence requires the expert to choose the perceived parameters so as to reveal the true value of π. The proof is straightforward: for any $\hat{v} \in V \cap AC(v)$, we have that $\hat{v} \in KA(v, \hat{v})$ by (ii) in prop. 2. By (i), we also have that $v \in KA(v, \hat{v})$. Since π is $KA(v, \hat{v})$-identified, it follows that $\pi(\hat{v}) = \pi(v)$.

It might be that for many autocoherent choices of $\hat{v}$, π is $KA(v, \hat{v})$-identified, but that for some other choices it is not. In that case only the first choices will compel the expert to reveal the true π, while the other choices allow the expert to pick a different π since it is no longer identified by the data generating process in the equilibrium associated with those beliefs.

Proposition 3's proof rests on the property that the true model is k-acceptable, and so is the perceived model if it is to be autocoherent. The unique identified value of the parameter $\pi()$ must therefore be the correct one.

Things are different if the econometrician is wrong and considers only u-acceptable possibilities. Then typically $v \notin UA(v, \hat{v})$. Suppose that π is $UA(v, \hat{v})$-identified for any admissible autocoherent model. Then it means that in the equilibrium delivered by the perceived model $\hat{v}$, there is a unique value for parameter π across all the possible models that would match the same moments if they were correct and believed. But this unique value may well differ across perceived models, because different perceived models imply different equilibria and thus different moments, and there does not a priori

exist a model v^* which would be common to different sets $UA(v, \hat{v})$ (while instead the correct model is common to all the $KA(v, \hat{v})$). Consequently, if a parameter is identified under the wrong identifying assumption that any candidate correct model is also the perceived one, then not only the expert is not compelled to reveal its true value, but there is no unique value of that parameter across autocoherent models.

3 A linear framework

I now consider a general linear framework and try to elicit some formal properties of autocoherent models.

We consider a linear model of the following form:

$$Z = MX + QZ^e. \tag{1}$$

In this formulation , X is the vector of exogenous variables. It is common knowledge that it is distributed normally with zero mean[3] and known variance-covariance matrix Ω. The assumption of known distributions for the exogenous variables is less special than it seems. For example, in a Gaussian setting, any exogenous variable x with unknown variance σ^2 can be treated as endogenous, with $x = \sigma \varepsilon$ and ε an exogenous standard normal random variable, and similarly for vectors of exogenous variables with unknown distributions.

Z is the vector of endogenous variables. Because people may not observe all exogenous variables, they may form expectations of them, which in turn may affect outcomes. For this reasons I will assume that the vector Z also

___

[3]This follows the tradition of the literature on stabilization and rational expectations where means typically do not matter and are usually normalized to zero. Empirically, however, it is much easier to find evidence of mean-matching autocoherence conditions, than variance-matching ones. See Saint-Paul (2011a).

contains all the exogenous variables[4], and accordingly that the corresponding sub-matrix of M is the identity matrix. The vector Z^e gives me the value of Z which is *expected* by the public. The matrix M depicts the direct effect of exogenous variables on outcomes, while Q depicts the effect of expectations. If $Q \neq 0$, expectations matter and affect outcomes. If m is the number of exogenous variables and p the number of truly endogenous variables, then $n = m + p$ is the dimension of vector Z. M is an (n, m) matrix, and by reducing the number of endogenous variables when they are redundant, we can always assume it is of rank p and therefore that $p \leq m$.

This model describes how an economy actually behaves, conditional on expectations. As such, it is clearly incomplete. To compute the equilibrium, I need to know how people form expectations. I will limit myself to expectations formation processes that have the following two properties:

A. Expectations are *intrinsic,* that is, for a given realization of X there is a unique value of Z^e in equilibrium. This rules out "sunspot" equilibria where one might have $Z = f(X, X')$, where X' is a random variable not included in X. However, one could always allow for such equilibria by making X' part of X, i.e. making it intrinsic, and add only a unit diagonal term to M to reflect the fact that X' does not affect any true endogenous variable. So this restriction is not binding as long as X includes all the variables on which society may index its expectations.

B. Expectations are *linear,* that is, there must exist an equilibrium relationship between Z^e and X of the following form:

$$Z^e = KX.$$

Given K, the behavior of the economy is determined by

$$Z = (M + QK)X. \tag{2}$$

---

[4]What is really needed, though, is to add to the true endogenous variables only a subset of the exogenous variables, i.e. those that are observed and/or whose expectation intervenes in one of the equations of the model.

Under this simple structure, an equilibrium is simply a matrix H such that $Z = HX$. Hence, $M + QK$ is an equilibrium. However, to know K, one must specify how expectations are formed.

I assume that the information available to the agents is given by two information sets I and J. Both are represented as subspaces of $\mathbb{R}^n$ such that $I \subset J$. I is the information set of the agents when they form their expectations. That is, I assume that people observe $T_I Z$, where T_I is the projection operator on subspace I (and we also use T_I to denote its matrix). For example, if $Z = (z_1 \; z_2 \; z_3)'$ and only z_1 is observed,

$$I = \mathbb{R}. \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

and

$$T_I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

J is the "ex-post" information set, that is, the information set used by the agents to validate their model. As long as the model is validated "ex-post", it is natural to assume that $I \subset J$. For example, we may observe z_2 in addition to z_1 once all outcomes are realized, and then we will have

$$J = \mathbb{R}. \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \oplus \mathbb{R}. \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

and

$$T_J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Agents form their expectations on the basis of what they observe, i.e. the vector $T_I Z$. We know that if $X \sim N(0, \Omega)$, and if A is any $k \times m$ matrix of full rank such that $k \leq m$, then $E(X \mid AX = Y) = h(A, \Omega)Y$, where

$$h(A, \Omega) = \Omega A'(A\Omega A')^{-1}. \tag{3}$$

I know derive a few properties of operator $h()$ that will play a key role in proving my analytical results. Let B be an invertible matrix. Since $(BAX = Y) \iff (AX = B^{-1}Y)$, we have that

$$h(BA, \Omega) = h(A, \Omega)B^{-1}. \tag{P1}$$

Since $BE(X \mid ABX = Y) = E(BX \mid ABX = Y)$ and $BX \sim N(0, B\Omega B')$, we have that

$$h(AB, \Omega) = B^{-1}h(A, B\Omega B'). \tag{P2}$$

In what follows I will drop the variable Ω from the argument of $h()$, since Ω will remain the same throughout.

Since it is observed that $Y = AX$, the operator

$$\phi(A) = h(A)A \tag{4}$$

maps the realization of X into its expectation conditional on observing Y. The preceding properties imply that if B is invertible,

$$\phi(BA) = \phi(A). \tag{P3}$$

As a corollary, if A itself is invertible, then $\phi(A) = I$, meaning that the actual value of X can be recovered from observing Y. Finally, if one were to forecast Y out of the inferred X, one would pick up Y again, implying that

$$A\phi(A) = A. \tag{P4}$$

These properties can be directly proved using (3).

A useful corollary of the above is

Lemma 1 $- \phi(A)^2 = \phi(A)$.
Proof $- \phi(A)\phi(A) = h(A)A\phi(A) = h(A)A = \phi(A)$.

The theory outlined below rests on the assumption that whenever observing AX, people behave as if they believe that X is equal to $\phi(\hat{A})X$, where

11

$\hat{A}$ is the perceived model and ϕ, h two operators related through (4) and satisfying (P1)-(P4). It does not really matter that $h()$ be the optimal filter defined by (3). Thus potentially people can use a wrong model, a wrong filter, or both. The set of autocoherent models analyzed below is conditional on a given filter[5].

Let us now specify how expectations are formed. People observe $T_I Z$. They must infer from it the conditional distribution of X. Given the linear structure of the problem, they only care about the mean of this distribution. Then, given this mean, they must make a forecast for Z. For both of these operations, they need a model which tells us how Z relates to X. I assume that if people *observe* Y and they *believe* that $Y = \hat{A}X$, then they forecast $X^e = h(\hat{A})Y$. They then also *believe* that $X^e = \phi(\hat{A})X$ for the unobserved true realization of X.

In what follows, an equilibrium is defined as a reduced form matrix H which relates the endogenous variables Z to the exogenous ones X. To any equilibrium is associated a matrix K which maps the realization of the exogenous variables X into the expectations Z^e. In a rational expectations equilibrium (Muth 1961), people use the correct model to form their expectations, which leads us to the following definition[6]:

Definition 5 – A rational expectations equilibrium (REE) is a matrix H such that there exists $K \in \mathcal{M}_{nm}(\mathbb{R})$,

---

[5] Conversely, one could develop a theory of autocoherent filters conditional on a perceived model, which may or may not be the correct one. Such a theory would be of limited interest: While the correct model may be impossible to know because it is econometrically underidentified, the correct filter is available off the shelf of statistical theory. However, this putative 'dual' theory might have some merit if supplemented with cognitive constraints on the complexity of the filters that may be used.

Relatedly, it is conceivable that people use the right model, but the wrong filter, perhaps because they make approximmations. The results that follow remain valid in this case as long as the inference operators h and ϕ that are involved satisfy (P1)-(P4).

[6] If $h(A) \neq \Omega A'(A\Omega A')^{-1}$ then peole do not use the optimal filter and this is not an REE in the Muth sense. One could then relabel it a "correct model equilibrium".

(i) $H = M + QK$

(ii) $K = H\phi(T_I H)$

The second condition means the following. We have (dropping Ω from the notations). $Z^e = E(Z \mid T_I Z) = E(HX \mid T_I HX) = HE(X \mid T_I HX) = Hh(T_I H)T_I HX = H\phi(T_I H)X$. Therefore it must be that $K = H\phi(T_I H)$.

4   Autocoherent model equilibria

We now discuss the equilibria that may arise when people indeed use a model to set their forecast Z^e, but this model may not be the correct model of the economy.

Let us therefore assume that people use the following model

$$\tilde{Z} = \hat{M}\tilde{X} + \hat{Q}Z^e$$
$$Z^e = \hat{K}\tilde{X}.$$

Here, $\tilde{X}$ is a random variable which has the same (known) distribution as X. A natural interpretation of $\tilde{X}$ is that it is the agent's "perceived" value of $\tilde{X}$. But people do not think that they observe $\tilde{X}$ and do not need to know its realization, they just need to formulate a forecast. Similarly, $\tilde{Z}$ is the "perceived" vector of endogenous variables, but people again do not think that they observe it. They do observe $T_I Z$ at the time of forming expectations and $T_J Z$ when validating their model, and therefore they interpret those values as being equal to $T_I \tilde{Z}$ and $T_J \tilde{Z}$ for $\tilde{Z}$ drawn from the model.

Note that the set of exogenous variables upon which the people's perceived model is based is the same as for the true model[7]. Indeed, I focus on the case where the structural model used by people has the same specification (M, Q) as the true model. That is, people use the same mental steps as an

---

[7]Confer the above remark about sunspots.

economist who would want to compute a rational expectations equilibrium in that economy, using a not necessarily correct model. One deep justification for such an approach is the idea that models and theories are public knowledge, and will therefore be used by the people to form their beliefs. If everybody in this economy believes that the model is of the (Q, M) form, then the agents will solve this model to optimally set their forecasts.

A different option would be to assme that people have a simpler representation of the world and use a reduced form $Z = \hat{H}X$ instead, without making a distinction between the direct contribution of the exogenous shocks and that of expectations. In the literature on learning (Marcet and Sargent (1989) and Evans and Honkapohja (2003)), the perceived law of motion is such a reduced form. In the Appendix, I briefly discuss the differences between the two approaches.

The matrices $\hat{Q}$ and $\hat{M}$ describe the model that people have in their minds, and it may differ from the correct model, which is described by Q and M. The matrix $\hat{K}$ describes the mapping assumed by people from the realization of (perceived) exogenous variables $\tilde{X}$ to the forecast Z^e.

4.1 Forecasts

How do people form expectations here? They observe $T_I Z$, which they believe is equal to $T_I \tilde{Z} = T_I(\hat{M} + \hat{Q}\hat{K})\tilde{X}$. Therefore, $X^e = h(T_I(\hat{M} + \hat{Q}\hat{K}))T_I Z$ and

$$Z_i^e = (\hat{M} + \hat{Q}\hat{K})h(T_I(\hat{M} + \hat{Q}\hat{K}))T_I Z, \tag{5}$$

where the "i" subscript means that this is the forecast of an individual.

This equation allows us to recover the forecast as a function of Z, given people's mental representations $\hat{M}, \hat{Q}$, and $\hat{K}$. Furthermore, we also have

that

$$Z_i^e = (\hat{M} + \hat{Q}\hat{K})h(T_I(\hat{M} + \hat{Q}\hat{K}))T_I\tilde{Z}$$
$$Z_i^e = (\hat{M} + \hat{Q}\hat{K})\phi(T_I(\hat{M} + \hat{Q}\hat{K}))\tilde{X}. \tag{6}$$

This means that people believe that their own forecast is related to the realization of the exogenous variables by this relationship.

Next, we assume that everybody uses the same model and that this is common knowledge. In particular, the realization of the forecast variable Z^e is the same in the perceived model as in the true model, since people know their own forecast. Consequently

$$Z_i^e = Z^e = \hat{K}\tilde{X}.$$

Therefore, by (6) the $\hat{K}$ matrix, which describes the perceived process of expectation formation, must satisfy

$$\hat{K} = (\hat{M} + \hat{Q}\hat{K})\phi(T_I(\hat{M} + \hat{Q}\hat{K})).$$

How does this relate to the real world process of expectation formation, described by K? Using (5), (2), and the fact that $Z_i^e = Z^e$, we have that

$$Z^e = (\hat{M} + \hat{Q}\hat{K})h(T_I(\hat{M} + \hat{Q}\hat{K}))T_I(M + QK)X,$$

and therefore

$$K = (\hat{M} + \hat{Q}\hat{K})h(T_I(\hat{M} + \hat{Q}\hat{K}))T_I(M + QK).$$

4.2 Model validation: the autocoherence property

We are now going to impose an additional restriction on the model that people use: it must be consistent with their observed data. The observed data are given by the vector $T_J Z$. In equilibrium, this vector is given by

$$T_J(M + QK)X.$$

This determines its distribution.

On the other hand, agents can also use their own model to predict the distribution of $T_J Z$, since they know the distribution of X and their model is based on an exogenous random variable $\tilde{X}$ which has the same distribution. Thus when observing $T_J Z$, they interpret it as $T_J \tilde{Z} = T_J(\hat{M} + \hat{Q}\hat{K})\tilde{X}$. Therefore, for the model used by the people to replicate the distribution of $T_J Z$, it must be that $T_J Z$ and $T_J \tilde{Z}$ have the same distribution, a relationship commonly denoted by "~".

This discussion leads to the following definition of an *Autocoherent Model Equilibrium.* This is an equilibrium supported by a model $\hat{M}, \hat{Q}$, such that people use the model to form their forecast in a way consistent with this model.

Definition 6 – H is an Autocoherent Model Equilibrium (AME) for the model $\hat{M}, \hat{Q}$ iff there exists K (called a forecast process) and $\hat{K}$ (called a perceived forecast process) $\in \mathcal{M}_{nm}(\mathbb{R})$ such that

(i) $H = M + QK$.

(ii) $\hat{K} = (\hat{M} + \hat{Q}\hat{K})\phi(T_I(\hat{M} + \hat{Q}\hat{K}))$.

(iii) $K = (\hat{M} + \hat{Q}\hat{K})h(T_I(\hat{M} + \hat{Q}\hat{K}))T_I H$.

(iv) There exists a random variable $\tilde{X}$ such that $\tilde{X} \sim X$ and $T_J H X \sim T_J(\hat{M} + \hat{Q}\hat{K})\tilde{X}$.

It will be useful, in the sequel, to use the following properties of an AME:

Lemma 2 – The perceived forecast process of an AME associated with $(\hat{M}, \hat{Q})$ has the following properties:

(i) $T_I(\hat{M} + \hat{Q}\hat{K}) = T_I \hat{K}$

(ii) $\hat{K}\phi(T_I\hat{K}) = \hat{K}$

(iii) $T_I\hat{M}\phi(T_I\hat{K}) = T_I\hat{M}$.

Proof – (i) is proved by applying projector T_I to both sides of (ii) in

16

Definition 6 and using (P4); (ii) is obtained by right-multiplying both sides of (ii) in Definition 6 and applying Lemma 1. Finally, to get (iii), right-multiply both sides of (i) by $\phi()$, getting, for the LHS $T_I(\hat{M} + \hat{Q}\hat{K})\phi(T_I\hat{K}) = T_I\hat{M}\phi(T_IK) + T_I\hat{Q}\hat{K}\phi(T_I\hat{K}) = $ (by (ii)) $T_I\hat{M}\phi(T_IK) + T_I\hat{Q}\hat{K}$; then the RHS is $T_I\hat{K}\phi(T_I\hat{K}) = T_I\hat{K} = T_I(\hat{M} + \hat{Q}\hat{K})$. The terms in $T_I\hat{Q}\hat{K}$ cancel and one gets (iii). QED.

Lemma 2 expresses in matrix form some intuitive properties of an AME. Condition (i), for example, is equivalent to $T_IZ^e = T_IZ$, meaning that the observable part of my forecasts must match the actual observables. Condition (ii) means that the forecast of my forecast, using the perceived forecast process, is equal to my forecast. Condition (iii) means that since I know the forecasts, I believe I can correctly infer the part of the observables that are accounted for by the exogenous variables, i.e. the MX vector (this belief would be correct if the model were correct).

One issue is: Given the perceived model, is there a unique AME? This issue is somehow the generalization of the uniqueness problem in an REE to the AME case. My main result (Proposition 6 below) does not rely on the equilibrium being unique. However that question is interesting in its own right and has two aspects. First, given the perceived model $(\hat{M}, \hat{Q})$ and the perceived forecast process $\hat{K}$, is there a unique equilibrium H satisfying definition 6? The answer is that this is generically true. Second, given the perceived model $(\hat{M}, \hat{Q})$, is there a unique perceived forecast process satisfying (ii)? This is equivalent to having a unique REE if $(\hat{M}, \hat{Q})$ were the correct model. Thus the required conditions are the same, however they are more subtle that just $I - Q$ being invertible, which looking at (1) alone would suggest. All these issues are discussed more precisely in the Appendix.

4.3   Interpretation of outcomes

Definition 6 is not very practical. But we will shortly show an equivalent set of conditions which is far more practical. Before doing so, it is interesting to introduce the notion of an *interpretation.*

Definition 7 – Assume there exists an AME. Let z be a vector of $\mathbb{R}^n$. Then $\tilde{x}$ is an $I-$interpretation (resp. $J-$interpretation) of z iff

$$T_I z = T_I(\hat{M} + \hat{Q}\hat{K})\tilde{x}$$

(resp $T_J z = T_J(\hat{M} + \hat{Q}\hat{K})\tilde{x}$).

In short, an interpretation is a realization of the perceived exogenous variables which is compatible with a given observation, ex-ante or ex-post.

The two following results are obvious but useful:

Proposition 4 – Let $z \in \mathbb{R}^n$. Then if $\tilde{x}$ is a $J-$interpretation of z, it is also an $I-$interpretation.

Proof – Since $I \subset J$, $T_I T_J = T_I$. Thus if $T_J z = T_J(\hat{M} + \hat{Q}\hat{K})\tilde{x}$, then $T_I z = T_I(\hat{M} + \hat{Q}\hat{K})\tilde{x}$.

Proposition 5 – If $\tilde{x}$ and $\tilde{x}'$ are $I-$interpretations of z, then

$$\hat{K}\tilde{x} = \hat{K}\tilde{x}'$$

Proof – $\hat{K}\tilde{x} = (\hat{M} + \hat{Q}\hat{K})\phi(T_I(\hat{M} + \hat{Q}\hat{K}))\tilde{x} = (\hat{M} + \hat{Q}\hat{K})h(T_I(\hat{M} + \hat{Q}\hat{K}))T_I(\hat{M} + \hat{Q}\hat{K})\tilde{x}$

$= (\hat{M} + \hat{Q}\hat{K})h(T_I(\hat{M} + \hat{Q}\hat{K}))T_I z$, and we get the same expression if we perform the computations with $\tilde{x}'$ instead.

Proposition 5 tells us that if I have an $I-$interpretation of z, nothing is lost by assuming that this is indeed the realization of $\tilde{X}$ in order to compute the forecasts. This is not surprising since the forecast only depends on the observable $T_I z$ and is therefore the same for all interpretations.

4.4 Autoherent model equilibria are interpretable

We now establish a characterization of AMEs which has the merit of being easier to handle than definition 6, and at the same time can naturally be understood as a representation of an AME based on the agents' interpretative activity.

Proposition 6 – H is an Autocoherent Model Equilibrium for the model $\hat{M}, \hat{Q}$ if and only if there exists K and $\hat{K} \in \mathcal{M}_{nm}(\mathbb{R})$ and matrix P (called an interpretation matrix of the AME) such that

$$P\Omega P' = \Omega$$

and

(i) $H = M + QK$.
(ii) $\hat{K} = (\hat{M} + \hat{Q}\hat{K})\phi(T_I(\hat{M} + \hat{Q}\hat{K}))$.
(iii) $K = \hat{K}P^{-1}$.
(iv) $T_J(\hat{M} + \hat{Q}\hat{K}) = T_J HP$.

Proof – Assume that the conditions in Definition 6 hold. Then clearly (i) and (ii) above hold. Furthermore, let $A = T_J(\hat{M} + \hat{Q}\hat{K})$ and $B = T_J(M + QK)$. By (iv) in Def. 3, $BX \sim A\tilde{X}$. In particular , $EBXX'B' = EA\tilde{X}\tilde{X}'A'$. Since $X \sim X'$, $EXX' = E\tilde{X}\tilde{X}' = \Omega$. Hence $B\Omega B' = A\Omega A'$. Since $\det \Omega \neq 0$, there exists a matrix P such that $A = BP$, which is orthogonal for the scalar product defined by Ω, i.e. $P\Omega P' = \Omega$.[8] This proves (iv). Furthermore, by

---

[8] A proof is provided in the Appendix.

(iii) in Def. 6, we have

$$
\begin{aligned}
K &= (\hat{M} + \hat{Q}\hat{K})h(T_I(\hat{M} + \hat{Q}\hat{K}))T_I H \\
&= (\hat{M} + \hat{Q}\hat{K})h(T_I(\hat{M} + \hat{Q}\hat{K}))T_I T_J H \\
&= (\hat{M} + \hat{Q}\hat{K})h(T_I(\hat{M} + \hat{Q}\hat{K}))T_I T_J(\hat{M} + \hat{Q}\hat{K})P^{-1} \\
&= (\hat{M} + \hat{Q}\hat{K})h(T_I(\hat{M} + \hat{Q}\hat{K}))T_I(\hat{M} + \hat{Q}\hat{K})P^{-1} \\
&= (\hat{M} + \hat{Q}\hat{K})\phi(T_I(\hat{M} + \hat{Q}\hat{K}))P^{-1} \\
&= \hat{K}P^{-1}.
\end{aligned}
$$

Let us now prove the converse. Assume that (i)-(iv) hold. Clearly, (i) and (ii) hold in Def. 6. Let $\tilde{X}$ be the random variable defined by $\tilde{X} = P^{-1}X$. Since $X \sim N(0, \Omega)$ and $E\tilde{X}\tilde{X}' = P^{-1}\Omega P'^{-1} = P^{-1}P\Omega P'P'^{-1} = \Omega$, indeed $\tilde{X} \sim X$. Furthermore $T_J HX = T_J HP\tilde{X} = T_J(\hat{M} + \hat{Q}\hat{K})\tilde{X}$. Since these two variables are equal, they clearly have the same distribution. This proves (iv) in Definition 2. Finally, we have that $T_I(\hat{M} + \hat{Q}\hat{K}) = T_I(M + QK)P$, hence

$$
\begin{aligned}
K &= \hat{K}P^{-1} \\
&= (\hat{M} + \hat{Q}\hat{K})\phi(T_I(\hat{M} + \hat{Q}\hat{K}))P^{-1} \\
&= (\hat{M} + \hat{Q}\hat{K})h(T_I(\hat{M} + \hat{Q}\hat{K}))T_I(\hat{M} + \hat{Q}\hat{K})P^{-1} \\
&= (\hat{M} + \hat{Q}\hat{K})h(T_I(\hat{M} + \hat{Q}\hat{K}))T_I HPP^{-1} \\
&= (\hat{M} + \hat{Q}\hat{K})h(T_I(\hat{M} + \hat{Q}\hat{K}))T_I H.
\end{aligned}
$$

This proves (iii) in Def. 6.
QED.

Proposition 6 tells us that in the autocoherent model used by the people, everything takes place as if, for any realization of X, people were using an interpretation $\tilde{X} = P^{-1}X$ instead. Then, clearly, they will forecast $\hat{K}\tilde{X} = \hat{K}P^{-1}X = KX$.

This is confirmed by Proposition 7.

Proposition 7 – Let H be an AME and P an interpretation matrix, then for any realization x of X and its associated endogenous vector $z = Hx$, $P^{-1}x$ is a $J-$interpretation of z.

Proof – We just have to compute $T_J(\hat{M} + \hat{Q}\hat{K})P^{-1}x = T_J Hx = T_J z$. QED.

A well known example of such an interpretation matrix arises in the literature on structural VARs (Sims (1980), Blanchard and Quah (1989), Gali (1999)). There, the variance-covariance matrix of disturbances is known, but there are degrees of freedom in mapping the structural shocks to the econometric disturbances. As a result, one formulates identifying assumptions – typically that some structural shocks have zero effect on some variables – which amounts to imposing one interpretation of the structural disturbances. Generally the variance-covariance matrix of the shocks is normalized to identity, so that the interpreted ones will be related to the correct ones by an orthogonal transformation, which will be the identity matrix if the identifying assumptions are correct. In other words, the authors of the VAR literature are explicitly undertaking the interpretation exercise that our agents are performing in Proposition 4.

Condition (iv) implies that given their incorrect interpretation of the realization of x people can predict the ex-post observables as well as somebody who would use the correct model H to make those predictions and accordingly interpret the data correctly[9].

---

[9]Another property is that $T_I K = T_I H$. Indeed,

$$
\begin{aligned}
T_I K &= T_I \hat{K} P^{-1} \\
&= T_I(\hat{M} + \hat{Q}\hat{K})\phi(T_I(\hat{M} + \hat{Q}\hat{K}))P^{-1} \\
&= T_I(\hat{M} + \hat{Q}\hat{K})P^{-1} \\
&= T_I T_J(\hat{M} + \hat{Q}\hat{K})P^{-1} \\
&= T_I T_J H = T_I H,
\end{aligned}
$$

This result means that despite that people use the wrong model, they make the same

21

5 Conclusion

A self-confirming equilibrium imposes fewer restrictions on outcomes than a rational expectations equilibrium. If the dimension of the ex-post observable space is not too large, there will be a large number of such equilibria which will differ according to which autocoherent model people use to form their expectations. The present paper has discussed some analytical properties of autocoherent models. In the interesting case where the set of autocoherent models is not reduced to the correct one, we need to supplement the model by a meta-theory of how the perceived model is determined. Such a meta-theory may be provided by Bayesian learning, as in the learning literature of Marcet and Sargent (1989) and Evans and Honkapohja (2003). Or it can be based on positive political economy as in Saint-Paul (2011a,b). Crossing these two approaches and understanding how ideological preferences affect the learning strategies of experts and intellectuals is likely to be a realistic and fruitful direction for further research.

---

forecast on observables as somebody who would use the right model. This is clear since the forecast on observables is the observables themselves.

REFERENCES

Blanchard, Olivier Jean & Kahn, Charles M, 1980. "The Solution of Linear Difference Models under Rational Expectations," *Econometrica,* vol. 48(5), pages 1305-11, July.

Blanchard, Olivier and Danny Quah (1989), "The Dynamic Effects of Aggregate Demand and Supply Disturbances", *American Economic Review* , Vol. 79, No. 4 (Sep., 1989), pp. 655-673

Evans, George and Seppo Honkapohja (2003), "Expectational Stability of Stationary Sunspot Equilibria in a Forward-looking Linear Model", *Journal of Economic Dynamics and Control,* 28(1), 171-81

Fudenberg. D. and D.K. Levine (1993) "Self-Confirming Equilibrium" *Econometrica*, 61 , 523-546.

———————————————— (2007), "Self-confirming equilibrium and the Lucas critique", mimeo, Harvard U.

Futia, Carl A. (1981), "Rational Expectations in Stationary Linear Models", *Econometrica* , Vol. 49, No. 1, 171-192

Gali, Jordi (1999), "Technology, Employment, and the Business Cycle: Do Technology Schocks Explain Aggregate Fluctuations?" *American Economic Review*, 249-271, March

Marcet, A. and Thomas J. Sargent (1989), "Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models", *Journal of Economic Theory*, vol. 48, no 2

McKenzie, Donald (2006), *An Engine, Not a Camera: How Financial Models Shape Markets* MIT Press

Muth, John F. (1961) "Rational Expectations and the Theory of Price Movements", *Econometrica* 29, p. 315-335.

Saint-Paul, Gilles (2011a), "Toward a political economy of macroeconomic thinking", NBER WP, presented at the International Seminar on Macroeconomics

——————————- (2011b), "The possibility of ideological bias in structural macroeconomic models", Working paper, Toulouse School of Economics and CEPR WP 8465.

Sargent, Thomas (2008), "Evolution and Intelligent Design", *American Economic Review,* 98(1), 5-37

Sims, Christopher (1980), "Macroeconomics and reality", *Econometrica*

6   APPENDIX

6.1   Completing the proof of Prop. 3

Lemma A1 – Assume that A and B are two $k \times m$ matrices such that $\text{Rank}(A) = \text{Rank}(B) = k$, and $k \leq m$. Assume that $A\Omega A' = B\Omega B'$ for Ω definite positive. Then there exists an $m \times m$ matrix P such that $A = BP$ and $P\Omega P' = \Omega$.

Proof – To prove this, note that by replacing Ω by its Choleski decomposition, $\Omega = CC'$, the condition is equivalent to $A_1 A_1' = B_1 B_1'$ for $A_1 = AC$ and $B_1 = BC$. If we can prove that $A_1 = B_1 P_1$ for P_1 such that $P_1 P_1' = I$, then $A = A_1 C^{-1} = B_1 P_1 C^{-1} = BCP_1 C^{-1} = BP$, where $P = CP_1 C^{-1}$ clearly satisfies $P\Omega P' = \Omega$.

Thus we just have to focus on the case where $\Omega = I$, which I now assume.

Consider the case where $AA' = BB' = I$. Then the k row vectors of A are an orthonormal family of R^m, and similarly for B. Each of those families can be completed into an orthonormal basis of R^m. By stacking the row of these two bases, we get two invertible $m \times m$ matrices $\bar{A} = \begin{pmatrix} A \\ A_1 \end{pmatrix}$ and $\bar{B} = \begin{pmatrix} B \\ B_1 \end{pmatrix}$ such that $\bar{A}\bar{A}' = \bar{B}\bar{B}' = I$. Let $P = \bar{B}'\bar{A}'^{-1}$. Then clearly $PP' = I$ and $A = BP$. Suppose now that $AA' = BB' = M$. Let again be $M = DD'$ be the Choleski decomposition of M, where D is $k \times k$ invertible and triangular. Let $\tilde{A} = D^{-1}A$ and $\tilde{B} = D^{-1}B$. Then since $\tilde{A}\tilde{A}' = \tilde{B}\tilde{B}'$, $\exists P \in M_{mm}(R)$, $\tilde{A} = \tilde{B}P$, $PP' = I$. Multiplying both sides by D we get the required condition $A = BP$.

6.2   Uniqueness conditional on the model

First consider the uniqueness of K and H conditional on $\hat{K}, \hat{M}, \hat{Q}$. Given Definition 2, we note that K must solve

$$K = (\hat{M} + \hat{Q}\hat{K})h(T_I(\hat{M} + \hat{Q}\hat{K}))T_I(M + QK).$$

This can be rewritten $K = A + BK$, where $A = (\hat{M} + \hat{Q}\hat{K})h(T_I(\hat{M} + \hat{Q}\hat{K}))T_I M$ and $B = (\hat{M}+\hat{Q}\hat{K})h(T_I(\hat{M}+\hat{Q}\hat{K}))T_I Q$. Since $I-B$ is generically invertible, this condition will generically be satisfied by a unique value of K.

Second, consider whether $\hat{K}$ is unique given the perceived model. This means that there is only one matrix $\hat{K}$ which satisfies (ii) in Proposition 2. The following result provides sufficient conditions for this to hold:

Proposition A1 – Given $\hat{M}$ and $\hat{Q}$, there is at most one matrix $\hat{K}$ which satisfies (ii) provided $\det(I_n - \hat{Q}) \neq 0$ and one of the conditions are satisfied:

(a) $T_I\hat{Q} = 0$.

(b) $T_I\hat{Q} = \hat{Q}T_I$.

(c) $\mathrm{Rank}(T_I\hat{M}) = \min(\dim I, m)$.

Proof – From Lemma 2 we see that equation (ii) in Definition 5 and Prop. 4 implies that $\hat{K}$ must satisfy $\hat{K} = \hat{M}\phi(T_I\hat{K}) + \hat{Q}\hat{K}$. If $T_I\hat{Q} = 0$, we have that $T_I\hat{K} = T_I\hat{M}\phi(T_I\hat{K}) = T_I\hat{M}$ by Lemma 2 again. Therefore, $\phi(T_I\hat{K}) = \phi(T_I\hat{M})$ and $\hat{K}$ is solution to the linear matrix equation $\hat{K} = \hat{M}\phi(T_I\hat{M})+\hat{Q}\hat{K}$, which has a unique solution given the invertibility of $I_n-\hat{Q}$. This proves (a). If $T_I\hat{Q} = \hat{Q}T_I$, then we have that $T_I\hat{K} = T_I\hat{M}\phi(T_I\hat{K}) + \hat{Q}T_I\hat{K} = T_I\hat{M} + \hat{Q}T_I\hat{K}$ (lemma 2), so that $T_I\hat{K} = (I_n - \hat{Q})^{-1}T_I\hat{M}$, implying again, since $(I_n - \hat{Q})^{-1}$ is invertible, $\phi(T_I\hat{K}) = \phi(T_I\hat{M})$ (proposition 1), and the rest follows as for (a).

Let us now turn to case (c). Let $k = \mathrm{Rank}(T_I) = \dim(I)$. Let $\tilde{T}_k$ be the nn matrix defined by

$$\tilde{T}_k = \begin{pmatrix} I_k & 0_{k,n-k} \\ 0_{n-k,k} & 0_{n-k,n-k} \end{pmatrix}.$$

Then there exists an invertible nn matrix U such that $T_I = U\tilde{T}_k U^{-1}$. By Lemma 2 $\hat{K}$ satisfies $T_I\hat{M}\phi(T_I\hat{K}) = T_I\hat{M}$. This can be rewritten $U\tilde{T}_k U^{-1}\hat{M}h(U\tilde{T}_k U^{-1}\hat{K})U\tilde{T}_K U^{-1}\hat{K}$

$U\tilde{T}_k U^{-1}\hat{M}$. Denoting $\bar{M} = U^{-1}\hat{M}$ and $\bar{K} = U^{-1}\hat{K}$, and applying proposition 1, we see that this is equivalent to

$$\tilde{T}_k \bar{M} h(\tilde{T}_k \bar{K})\tilde{T}_k \bar{K} = \tilde{T}_k \bar{M}. \tag{7}$$

Assume first that $k \leq m$. Clearly, $\mathrm{Rank}(\tilde{T}_k \bar{M}) = \mathrm{Rank}(T_I \hat{M})$. By assumption, this is equal to k. Thus, by (7) we also have that $\mathrm{Rank}(\tilde{T}_k \bar{M} h(\tilde{T}_k \bar{K})\tilde{T}_k \bar{K}) = k$. Since $\mathrm{Rank}(\tilde{T}_k \bar{K}) \leq k$, it must be that $\mathrm{Rank}(\tilde{T}_k \bar{K}) = \mathrm{Rank}(\tilde{T}_k \bar{M} h(\tilde{T}_k \bar{K})) = k$. Now, the matrix $\tilde{T}_k \bar{M} h(\tilde{T}_k \bar{K})$ has the following form:

$$\tilde{T}_k \bar{M} h(\tilde{T}_k \bar{K}) = \left(\begin{array}{cc} A & B \\ 0_{n-k,k} & 0_{n-k,n-k} \end{array} \right),$$

while $\tilde{T}_k \bar{K}$ can be written

$$\tilde{T}_k \bar{K} = \left(\begin{array}{c} \bar{K}_1 \\ 0_{n-k,m} \end{array} \right).$$

Here A is kk, B is $k, n-k$ and $\bar{K}_1$ is km. We can then see that (7) is equivalent to

$$\left(\begin{array}{c} A\bar{K}_1 \\ 0_{n-k,m} \end{array} \right) = \tilde{T}_k \bar{M},$$

implying that $\mathrm{Rank}\, A = k$ and therefore that A is invertible. Consider now the matrix defined by

$$C = \left(\begin{array}{cc} A & 0_{k,n-k} \\ 0_{n-k,k} & I_{n-k} \end{array} \right).$$

Clearly, $\det C = \det A \neq 0$. Furthermore, $C\tilde{T}_k \bar{K} = \tilde{T}_k \bar{M}$. It then follows from Proposition 1 that $\phi(\tilde{T}_k \bar{K}) = \phi(\tilde{T}_k \bar{M})$. Therefore $\phi(T_I \hat{K}) = \phi(U\tilde{T}_k U^{-1}\hat{K}) = \phi(\tilde{T}_k \bar{K}) = \phi(\tilde{T}_k \bar{M}) = \phi(U\tilde{T}_k U^{-1}\hat{M}) = \phi(T_I \hat{M})$. The rest of the proof is the same as for (a) and (b).

Next, assume $m \leq k$. By assumption, $\mathrm{Rank}(\tilde{T}_k \bar{M}) = \mathrm{Rank}(T_I \hat{M}) = m$. Multiplying both sides of (7) by $(\tilde{T}_k \bar{M})'$, we get $(\tilde{T}_k \bar{M})'\tilde{T}_k \bar{M} h(\tilde{T}_k \bar{K})\tilde{T}_k \bar{K} = (\tilde{T}_k \bar{M})'\tilde{T}_k \bar{M}$. Since $(\tilde{T}_k \bar{M})'\tilde{T}_k \bar{M}$ is an mm matrix of rank m, it is invertible.

27

It follows that $h(\tilde{T}_k\bar{K})\tilde{T}_k\bar{K} = \phi(\tilde{T}_k\bar{K}) = \phi(T_I\hat{K}) = I_m$. Therefore, the only solution is $\hat{K} = (I_n - \hat{Q})^{-1}\hat{M}$.

QED.

Note: Proposition A1 refers to the properties of the perceived model regardless of whether it is correct and therefore also applies to the case where it is correct, i.e. to the uniqueness of an intrinsic rational expectations equilibrium. It addresses an issue which, to the best of my knowledge, has been overlooked in the literature, i.e. that the filter $\phi()$ is a nonlinear function of K which opens up the possibility of multiple equilibria even though $I - Q$ might be invertible. The meaning of that multiplicity is that the information available to the agents for forming their expectations may itself depend on the equilibrium matrix K, ie on the way expectations are formed. That is, how much filtering can be done differs across equilibria and some equilibria may be more informative than others. The less informative equilibria would be broken if people could make as precise inferences as in the more informative ones, but in those less informative equilibria expectations are formed in such a way that information is lost. In the literature (e.g. Blanchard and Kahn (1980), Futia (1981)) one is mostly in a context where this is ruled out and uniqueness boils down to the invertibility of $I - Q$. Indeed, since condition (c) is generic, this source of multiplicity is somewhat a curiosity.

6.3 Reduced form vs. Structural models

Definition A1 – H is an Autocoherent Reduced Form Model Equilibrium (ARFME) for the model $\hat{H}$ and the unbiased LIO $h()$ iff there exists a forecast process $K \in \mathcal{M}_{nm}(\mathbb{R})$ such that

(i) $H = M + QK$.
(ii) $K = \hat{H}h(T_I\hat{H})T_I H$
(iii) There exists a random variable $\tilde{X}$ such that $\tilde{X} \sim X$ and $T_J HX \sim$

28

$T_J \hat{H} \tilde{X}$.

There is then a straightforward counterpart to Proposition 4:

Proposition A2 – H is an Autocoherent Reduced Form Model Equilibrium (ARFME) for the model $\hat{H}$ and the unbiased LIO $h()$ iff there exists a forecast process $K \in \mathcal{M}_{nm}(\mathbb{R})$ and an orthogonal matrix P such that
(i) $H = M + QK$.
(ii) $K = \hat{H}h(T_I\hat{H})T_I H$
(iii) $T_J H \sim T_J \hat{H} P$.

How do ARFME relate to AME? Clearly, all AMEs are also ARFMEs for the reduced form implied by the model that people use:

Proposition A2 – Let H an AME for model $(\hat{M}, \hat{Q})$, and perceived forecast process $\hat{K}$. Then it is an ARME for model $\hat{H} = \hat{M} + \hat{Q}\hat{K}$.

Proof – By definition 5, (i), (ii) and (iii) in definition 9 hold.

It is also true that any ARFME is an AME for some model:

Proposition A3 – Let H be an ARFME for model $\hat{H}$. Then it is an AME for model $\hat{M} = H$, $\hat{Q} = 0$.
Proof – Straightforward.

These two propositions tell us that if all I am looking for is an AME, I can actually restrict myself to looking for an ARFME, which is simpler to characterize. However, the theory is really useful in a context where not all models are acceptable. Thus, we want to restrict the choice for $(\hat{Q}, \hat{M})$ to a subset of "acceptable" models. In this case the search for an acceptable AME cannot be reduced to looking for an ARFME. For example it may just not be plausible to think that $\hat{Q} = 0$, since it would mean that expectations are completely irrelevant.