ITam INSTITUTO TECNOLÓGICO AUTÓNOMO DE MÉXICO

CENTRO DE INVESTIGACIÓN ECONÓMICA

Discussion Paper Series

Transformations of the State Variable and Learning Dynamics

Shurojit Chatterji
Instituto Tecnológico Autónomo de México
Warwick University

and

Ignacio N. Lobato
Instituto Tecnológico Autónomo de México

Transformations of the State Variable and Learning Dynamics[1]

Shurojit Chatterji

Centro de Investigación Económica, ITAM and

Department of Economics, Warwick University

Ignacio N. Lobato

Centro de Investigación Económica, ITAM

Ave Camino Santa Teresa 930, Mexico D.F. 10700

December 2007

Abstract

This article studies dynamics in a model where agents forecast a one dimensional state variable via ordinary least squares regressions on the lagged values of the state variable. We study the stability properties of alternative transformations of the state variable that the agent can endogenously set forth. We study the consequences on the economy's stability of the typical transformations that an econometrician would attempt, such as differencing, detrending, or taking instantaneous concave transformations, such as logarithms. Surprisingly, for the considered class of economies, we found that these transformations are destabilizing, whereas alternative transformations, which an econometrician would never consider, such as convex transformations, are stabilizing. Therefore, we ironically find that in our set-up, an active agent, who is concerned about learning the economy's dynamics and transforms the state variable, in an attempt to improve forecasting, is more likely to deviate from the steady state than a passive agent.

KEYWORDS: Temporary Equilibrium, Ordinary Least Squares Learning, Globally Stable Formulations

1   Introduction

A long held view about the axiom of perfect foresight is that while it is an important conceptual tool for understanding those aspects of the formal content of economic models that do not rely on agents making forecasting errors, it is an exceedingly strong assumption[2] that can at best be justified in a stationary environment. Whether an economy ends up in a stationary environment in turn depends on how agents forecast and learn about the dynamics. One way of modelling this adjustment process is via "bounded rationality"[3], as exemplified by the temporary equilibrium approach, where forecasts are allowed to be based on a given statistical procedure, with agents estimating some structural parameters from past data.

The simplest stationary environment is a deterministic steady state. Recent studies, such as Grandmont (1998), have shown that a steady state is locally unstable under learning dynamics whenever an agent's forecasting rule extrapolates a large enough set of trends in deviations from the steady state. For example, the learning dynamics generated by ordinary least squares (OLS, henceforth) learning, where agents forecast the endogenous state variable via regressions on its lagged values, extrapolates all trends in past data and consequently produces local instability. In the face of such instability, it is sensible to consider that a Walrasian agent would modify the state variable to improve the forecasting capability of her model. This practice is habitual in time-series econometrics, where series are routinely differenced, detrended or subjected to instantaneous transformations, such as logarithms. This article asks whether such modifications might make the learning dynamics more stable and lead to improved forecasts in the long run.

The formal model is one with a unique deterministic steady state. Agents will be assumed to generate a point forecast of the future value of the state variable via OLS on its lagged values.[4] With OLS on endogenous variables, the asymptotic shape of the underlying

---

[2]Radner (1982) writes "Although it is capable of describing a richer set of institutions and behaviour than is the Arrow-Debreu model, the perfect foresight approach is contrary to the spirit of much of competitive market theory" and goes on to state that "this approach seems to require of the traders a capacity for imagination and computation far beyond what is realistic".

[3]In support of the bounded rationality approach, Radner (1982) writes "In a theory of adjustment towards rational expectations equilibrium, what are the appropriate assumptions about the agents' rationality during the "learning process"? As agents revise their market models, the true market model changes in a way that, in principle, depends on the revision rules of all agents. Thus, a theory of thorough -going rationality would seem to point to a treatment of the learning and adjustment process as a sequential game with incomplete and imperfect information. In my opinion, such an approach would be unrealistic and contrary to the spirit of a process of adjustment and learning. A more realistic alternative would envisage some form of "bounded rationality" during the adjustment process, which if stable would converge to a fully rational equilibrium".

[4]We will focus on the dynamics in a deterministic setting. The results are robust to small independent and identically distributed shocks in the market clearing process. Throughout we restrict agents to point

equilibrium map of the economy (which summarizes the dependence of the current value of the state variable on its forecasted value) determines what rates of growth the learning dynamics can sustain in the long run and drives the stability of the learning dynamics. For instance, a linear equilibrium map is known to generate instability for initial parameter estimates that are large enough. We define an *Unstable Formulation* (UF, henceforth) as one which essentially works like a linear equilibrium map asymptotically and hence leads to divergent paths whenever the initial parameter estimate is large enough. We also consider a bound on the asymptotic growth rate of the equilibrium map that is shown to preclude the divergence to infinity of the state variable and the parameter estimates. An equilibrium map with this property is said to induce a *Globally Stable Formulation* (GSF, henceforth) of the learning dynamics[5]. The article recognizes that, in a framework where an agent modifies certain structural features of her linear regression model in the face of instability, whether or not a GSF is induced is an endogenous aspect of the learning dynamics, and examines the implications of this endogeneity.

We examine the case of instantaneous transformations induced via a change of variable on the state variable forecasted initially via the linear regression model and find, counter intuitively, that the standard transformations used in time series analysis to stabilize the data, namely concave transformations, tend to have a *destabilizing* effect on the learning dynamics. In contrast, sufficiently convex transformations, that appear to exaggerate the divergent trends in the data, are shown to have a stabilizing effect, in that they may induce a GSF when the dynamics in the original variable are unstable. We also examine the case where an agent starting with a linear regression model, reformulates the model on the first differences of the (logged) state variable or on the detrended state variable. In both cases, we show that if the original formulation is unstable, then so is the new model, that is, neither differencing nor detrending are stability-enhancing.

In modelling learning in a decentralized Walrasian setting where agents are truly uncertain about the dynamics of the system, a natural requirement to impose on the learning scheme is that agents are willing to interact with market data and extrapolate a wide set of trends, in particular, divergent trends from the steady state. Earlier studies (Chatterji (1995), Grandmont (1998)) have pointed out that a greater willingness to learn in the above sense, somewhat paradoxically, makes it less likely that agents will end up in the steady state. This instability principle is a local phenomenon. While there exist formulations where the learning dynamics may remain bounded or even globally convergent to the steady state in

---

forecasts. The issues examined here continue to be relevant, but are considerably more complicated when expectations are stochastic.

[5]The notion of stability is a weak one. It merely ensures that the state variable does not diverge to infinity. Under an additional contracting condition on the equilibrium map, this notion of stability leads to global convergence to the steady state.

the long run (in spite of the local instability; see 2.1 below), there are arbitrarily long periods where the dynamics appear to be divergent. Our results imply, once again somewhat paradoxically, that these stable configurations require that agents have to systematically ignore the fact that their model appears to be an inadequate model for the (arbitrarily long) "unstable" phase of the learning process where the dynamics appear to be exploding, and desist from making the natural transformations, as these would eventually destabilize the global dynamics via the effect they induce on the asymptotic behaviour of the equilibrium map. As a practical conclusion, this raises doubts on the usefulness of applying the standard econometric transformations that attempt to stabilize the series before forecasting. An agent who disregards the local instabilities and stick to her initial model would end up in the perfect foresight steady state. On the other hand, an agent attempting to actively learn the true dynamics of the economy, will end up destabilizing the economy and will never arrive at a perfect foresight equilibrium.

The article is organized as follows. Section 2 provides an overview and establishes the basic model and concepts. Sections 3 and 4 consider the relation between stability and transformations for the cases of instantaneous transformations and dynamic transformations respectively. Section 5 concludes and offers some suggestions for future research. All proofs are gathered in Section 6. An Appendix specifies a standard contracting condition under which the dynamics are globally convergent to the steady state and provides some extensions.

2 Preliminaries

This section provides an overview of the paper, specifies the model and the two main concepts related to stability, and then establishes a preliminary result.

2.1 An Overview

We present here a motivating example together with an informal exposition of the issues studied in the article and discuss related literature.

Example 1. Consider a model where the state variable is a ratio of two prices p_c and p_m, with a unique steady state value. Agents can either forecast the relative price $X_t \equiv p_c/p_m$, or the relative price $Y \equiv p_m/p_c$. OLS learning can be very sensitive to whether agents forecast X or Y. Assuming agents know the steady state, agents forecast the deviation $x = X - X^*$(respectively, $y = Y - Y^*$). Agents generate the forecast x_{t+1}^e (resp, y_{t+1}^e) via OLS on lagged values of x (resp, y). Assuming that at date t, when the forecast x_{t+1}^e (resp, y_{t+1}^e) is made, information up to only $t-1$ (a standard assumption in the literature to avoid simultaneous determination of x_{t+1}^e and x_t) is available, one has $x_{t+1}^e = \beta_{t-1}^2 x_{t-1}$(resp,

$y_{t+1}^e = \gamma_{t-1}^2 y_{t-1}$). There is an obvious qualitative difference between the two forecasting rules as x (and consequently X) go to infinity. As $x_{t-1} \to \infty$, β_{t-1} can be arbitrarily large depending on the rate at which x_{t-1} diverges. On the other hand, as $x_{t-1} \to \infty$, $Y_{t-1} \to 0$, and consequently $y_{t-1} \to -Y^*$. Since the parameter estimate γ_{t-1} is a convex combination of past ratios $y_{t-j}/y_{t-1-j}, j \geq 1$, as $y_{t-1} \to -Y^*$, we have $\gamma_{t-1} \to 1$ and the forecasts in the limit are $y_{t+1}^e \simeq y_{t-1}$, independently of the rate at which x_{t-1} diverges. Therefore, it is not surprising that the stability results are qualitatively different in some cases depending on whether agents forecast X or Y. Indeed one may have scenarios where infinity is attracting when agents forecast X, but repelling when agents forecast Y, i.e. $y_t \nrightarrow 0$. The Appendix provides details of an economic context where the contrast is yet more stark; the learning dynamics are divergent for a significant set of initial conditions when agents forecast X, but globally convergent to the steady state Y^* when the agents forecast Y. ▲

The above example shows that the OLS learning dynamics are quite sensitive to transformations of the state variable that agents forecast. Next we briefly examine this dependence in a general set up. Let $x_t = f(x_{t+1}^e)$ be the Temporary Equilibrium Map (henceforth, TEM) that describes the dependence of the current value of the state variable on its forecasted value. Let $y = h(x)$ be a smooth change of variable. When agents forecast the representation y of the state variable, the TEM becomes $y_t = g(y_{t+1})$ where $h(f(h^{-1}(y_{t+1}^e))) \equiv g(y_{t+1}^e)$. The perfect foresight dynamics when agents forecast x are given by the map $x_t = f(x_{t+1})$ (where for simplicity, f is assumed to be invertible), and when agents forecast y are given by the map $y_t = g(y_{t+1})$. The two maps are topologically conjugate and therefore generate equivalent dynamics. Thus if one assumes at the very outset that agents have perfect foresight, it does not matter which variable the agents forecast.

The literature on learning assumes that agents forecast the future in terms of the past. So instead of the perfect foresight postulate $x_{t+1}^e \equiv x_{t+1}$ for all t, consider the case where x_{t+1}^e is predicted on the basis of its lagged values $x_j, j \leq t-1$. For the sake of illustration assume that the forecasts take the form $x_{t+1}^e \equiv \lambda x_{t-1}$ and $y_{t+1}^e \equiv \lambda y_{t-1}$. The dynamics are now given by $x_t = f(\lambda x_{t-1})$ and $y_t = g(\lambda y_{t-1})$ and in general they are not equivalent, unless one restricts $\lambda = 1$. (In this case, the dynamics are the reverse of the perfect foresight dynamics and are equivalent to each other). If one moves away from this very restrictive assumption and allows λ to be different from unity, as indeed would be the case if λ were a parameter that agents estimate using past data, one should expect the learning dynamics to be qualitatively different in general. In this article, we examine the case where λ may depend on t, that is $\lambda = \lambda_t$, and, in fact, λ_t is obtained via OLS on past data.

As mentioned in Example 1, there may exist important global stabilizing forces at play which stem from the variable that agents forecast. In that example, the stabilizing force acts via the restriction that as $x \to \infty$, y tends to a finite fixed point $-Y^*$, which in turn

ties down the parameter estimate γ_t to 1 in the limit; this causes a "turn around" of the learning dynamics around $-Y^*$. This stabilizing mechanism stems from the bounds on the domain of the problem. This article uncovers a more general version (Proposition 1, Section 2.3) of the stabilizing mechanism that works without any bounds either on the domain or the range of the TEM, or on the parameter estimates, and then studies the dependence of the stabilizing mechanism on the variable $y = h(x)$ that agents formulate their OLS model on. The analysis is driven by the non linearities induced via these changes of variables.

We examine the consequences on the economy's dynamics of the representative agent's use of the standard data transformations econometricians routinely employ to deal with non-stationarities. We first examine the stability implications of instantaneous transformations and find that the standard concave transformations are destabilizing. Depending on the agent's beliefs about the structure of the nonstationarity, there are two main approaches an econometrician may employ to achieve stationary data. The first approach is called "trend-stationarity" (TS) and it means that the agent believes that the data is stationary around a deterministic trend. In the simplest framework where the deterministic trend is linear, this model entails that the agent believes that the state variable has a constant growth rate, so that in order to achieve stationarity the agent should just linearly detrend the data. In case the agent believes the trend is a polynomial of higher order, stationarity is achieved by detrending the data, which requires that one substract from the data the polynomial trend estimated by least squares. The second approach is called "difference-stationarity" (DS) and it means that the agent believes that the state variable has an stochastic stationary growth rate. In this case it is said that the state variable has a unit-root. DS has been very popular since the 80's in econometrics, and in this case in order to achieve stationarity the agent should difference the (logged) data. In some cases (when the growth rate is stochastic but nonstationary) first differences are not enough and the series has to be differenced twice. In both, TS and DS, cases, we will see that for an economy with feedback, that is, where the actual motion of the economy depends on the agent's beliefs about the evolution of economy, neither differencing nor detrending will in general help the agent to learn properly the economy's dynamics. A similar conclusion holds for the case of log differencing.

Related Literature. The remainder of this section briefly describes the setting of the article and clarifies its relationship to the literature on learning dynamics. We analyse the temporary equilibrium dynamics generated by OLS in a framework where the state variable is one dimensional and where agents forecast one period ahead, and focuses on some global aspects of the dynamics in this set up.

By considering a model where agents forecasts are based on OLS learning, we follow a

well established tradition.[6] Specifically, this paper belongs to a stream of literature which recognizes the importance of including past realizations of the endogenous variables[7] in the data set; agents forecast by performing OLS on lagged values of the endogenous state variable. The resulting formulation of the learning dynamics is of particular interest since it justifies the forecasting procedure employed by agents on the basis of a standard econometric procedure, while capturing nicely an important feature of learning in decentralized or Walrasian set ups, whereby agents extract *all* linear trends from past data and extrapolate them into the future.[8] While the issues studied in this paper apply to any learning scheme that extrapolates trends that differ from unity from past data in deviations, it appears appropriate, for the reasons cited above, to examine in detail the case of OLS with endogenous variables.

The dynamics under OLS are discontinuous[9] at the steady state and are divergent for

---

[6]See, e.g., Bray (1982), Bray and Savin (1986), Frydman (1982), Lucas (1986), Marcet and Sargent (1989), Evans and Honkapohja (1998), Grandmont (1998), and Chatterji and Chattopadhyay (2000).

[7]See Marcet and Sargent (1989), Grandmont and Laroque (1991), Evans and Honkapohja (1998), Grandmont (1998), Chatterji and Chattopadhyay (2000). To see that this case is of relevance, it suffices to consider the nonstationary perfect foresight solutions to deterministic economies since by doing so one realizes that often the *only* exogenous variables in a dynamic model are time and the lagged values of the state variable. But then any useful regression model in such an environment *must* be based on detecting regularities in the time series of the endogenous variable. The argument sketched above does not depend on the consideration of a deterministic model. It is known that the inclusion of exogenous shocks which impinge on the fundamentals of a model need not always lead to solutions which can be written as functions of a finite number of past realizations of the exogenous shocks (see, e.g. Spear (1985)). In fact, for large classes of economies, the only solutions in the stochastic case are ones in which a joint distribution over current and past endogenous and exogenous variables is determined. i.e., it could be the case that *every* equilibrium displays "memory" (see, e.g. Spear and Srivastava (1986)).

[8]As is emphasized by Grandmont (1998) in his critique of learning and convergence to rational expectations, this ought to be an important part of any learning story with a competitive sector where agents are "small" and they cannot act collectively (by hypothesis) to either control or stabilize the economy and so it is appropriate for them to extrapolate trends that past data show. It is therefore of interest to examine learning schemes that do not arbitrarily restrict the set of trends competitive agents extrapolate into the future from past data.

[9]We note at the outset that in a framework where agents use differentiable forecasting rules, one expects the local stability results to depend on the specific parameterization chosen for the forecasting rule and the variables used; this is justified by the usual eigenvalue analysis with sufficiently rich parameterizations. An early example formalizing this idea is Saari and Williams (1986) and a more recent one is Van Zandt (2003). The analysis of this paper cannot be framed in these terms since the dynamical system one works with is *nondifferentiable* and precludes the usual eigenvalue analysis and necessitates a *global* argument (in the line of Chatterji and Chattopadhyay (2000) to understand the stability properties. The analysis of global convergence in Evans and Honkapohja (1998) Theorem 2 does not apply to our framework; their assumption D.2. rules out "feedback" in the dynamical system and in effect they analyse a model where regressions are performed on an exogenous variable.

an open set of initial conditions around the steady state (Grandmont and Laroque (1991), Grandmont (1998)). Marcet and Sargent's (1989) seminal article was the first to study OLS learning on endogenous variables in a general multidimensional set up. Their analysis brought to the fore the importance of the map that transforms beliefs into the actual growth rate of the economy. They presented local stability results under stochastic dynamics with a linear equilibrium map by constraining the parameter estimates to always lie in an appropriate bounded interval (via a *Projection Facility*) that effectively rules out the divergence subsequently uncovered by Grandmont and Laroque (1991). Evans and Honkapohja (1998) dispensed with the Projection Facility and obtained bounds on the probabilities of local convergence of the stochastic dynamics, showing thereby that it was possible to carry out local analysis without a Projection Facility. Chatterji and Chattopadhyay (2000) studied the deterministic global dynamics without imposing bounds on the parameter estimates and showed that certain configurations of bounds on the range of the equilibrium map may ensure that the dynamics are globally convergent to the steady state.[10] This paper studies the global dynamics without imposing bounds either on the parameter estimates, or on the domain or the range of the equilibrium map, and focusses on the divergence phenomenon. Earlier work by Bray and Savin (1986), analysed the case where agents question the validity of their OLS learning model using recursive least squares and the traditional Durbin-Watson serial correlation test. However, in their model OLS is performed on a well behaved exogenous stochastic process, and the instability problem examined here is consequently absent.[11]

2.2   The Model

We consider a framework where the state variable is a real number. In the deterministic version of the model, the economy will be assumed to have a unique steady state value. The state variable is denoted Y; the unique steady state value is denoted by Y^*. Agents will be assumed to know the steady state and one accordingly formulates the model in the deviation $y \equiv Y - Y^*$ of the state variable from the steady state value 0. The stochastic version of the model specifies at each date t, an independent identically distributed (henceforth, i.i.d.) random term ν_t, also a real number, drawn from the interval $[\underline{\nu}, \overline{\nu}]$.

Dynamic interactions in the economy are captured by the TEM $g : R \times [\underline{\nu}, \overline{\nu}] \to R$ that links the forecasted value of the state variable (and the current realization of random term)

---

[10]This seemingly contradictory finding on the co-existence of local instability and global stability stems from the non differentiability of the learning dynamics at the steady state. Chatterji (2002) extends the global stability finding to the case where the economy is subjected to stochastic shocks, and provides a more detailed exposition of the stability picture sketched here.

[11]This formulation restricts the set of trends agents extrapolate around the steady state; see Grandmont (1998) for details.

to the current equilibrium value of the state variable

$$y_t = g(y_{t+1}^e, \nu_t). \tag{1}$$

The map g is assumed to be well defined.

Agents predict deviations $y_t \ (= Y_t - Y^*)$ of the state variable from its steady state value. At date t, the agent is assumed to have a "belief" γ_{t-1} about the ratio of the deviations of the state variable across periods $t-1$ and t. This belief is extrapolated twice, a standard procedure in the literature, see Marcet and Sargent (1989), Grandmont (1998), Chatterji and Chattopadhyay (2000), to generate the point forecast

$$y_{t+1}^e = \gamma_{t-1}^2 y_{t-1}. \tag{2}$$

The belief at date t, γ_{t-1} is a convex combination of all past growth rates y_{j+1}/y_j, $j = -L, ., .., t-2$ in the state variable and is given by

$$\gamma_{t-1} = \sum_{-L}^{t-2} \eta_j \frac{y_{j+1}}{y_j}, \text{ where } \eta_j = \frac{y_j^2}{\sum\limits_{-L}^{t-2} y_j^2}, \text{ provided } \sum_{-L}^{t-2} y_j^2 \neq 0. \tag{3}$$

The forecasting rule this paper considers would, for instance, arise if agents estimate the best linear predictor of y_{t+1} given y_t, that is, a linear model such as

$$y_{t+1} = \gamma y_t + \epsilon_{t+1}, \tag{4}$$

where ϵ is white noise and γ is the slope parameter that agents estimate using the OLS estimator based on all past realizations of y. Since the steady state is assumed known to the agents, (4) does not contain an intercept. We work with the following recursive formulation of OLS:

$$\gamma_t = m(\omega_{t-1}y_{t-1})\gamma_{t-1} + [1 - m(\omega_{t-1}y_{t-1})]\frac{y_t}{y_{t-1}} \tag{5}$$

$$\omega_t^2 = m(\omega_{t-1}y_{t-1})\omega_{t-1}^2 \tag{6}$$

with $m(z) = (1 + z^2)^{-1}$, and subject to the initial conditions $\omega_0^2 = (\sum_{-L}^{-1} y_j^2)^{-1}$, and $\gamma_0 = (\sum_{-L}^{-1} y_j y_{j+1})/(\sum_{-L}^{-1} y_j^2)$.

The *OLS learning dynamics* are described by the TEM (1) in conjunction with the forecasting rule (2) and the updating rules (5) and (6).

The stochastic TEM g will be assumed to satisfy

Assumption 1: *The TEM $g : R \times [\underline{\nu}, \overline{\nu}] \to R$ satisfies (i) $E(g(0,\nu)) = 0$ and (ii) for every $\mathcal{L} > 0$, there exists $y(\mathcal{L}) > 0$ such that $|g(y,\nu)| < y(\mathcal{L})$ for all $y \in [-\mathcal{L}, \mathcal{L}]$ and all $\nu \in [\underline{\nu}, \overline{\nu}]$.*

Example 2. Suppose $\widehat{g} : R \rightarrow R$ satisfies $\widehat{g}(0) = 0$ and the Lipschitz condition $|\widehat{g}(y)| < a|y|$ for all y and some $a > 0$. (i) Now let $g(y_{t+1}^e, \nu_t) \equiv \widehat{g}(y_{t+1}^e) + \nu_t$, where $\nu_t \in [-\overline{\nu}, \overline{\nu}], \overline{\nu} > 0$ and $E(\nu) = 0$. Then g satisfies assumption 1. (ii) Let $g(y_{t+1}^e, \nu_t) \equiv \widehat{g}(y_{t+1}^e)\nu_t$, where $\nu_t \in [1 - \overline{\nu}, 1 + \overline{\nu}], \overline{\nu} > 0$ and $E(\nu) = 1$. Then g satisfies assumption 1.

A deterministic economy results when the random term ν_t is absent so that the TEM becomes $y_t = g(y_{t+1}^e)$. In this case the TEM will be assumed to satisfy

Assumption 2: *The deterministic TEM $g : R \rightarrow R$ has 0 as a fixed point, $g(0) = 0$, and satisfies the global Lipschitz condition $0 \leq b \leq |g(y)|/|y| \leq a$ for all y and some fixed a, b.*

2.3 Globally Stable Formulations and Learning Dynamics

An important determinant of the stability of the learning dynamics is the set of beliefs that agents hold about the rate of growth of the economy and extrapolate into the future. In particular, the set of beliefs that the formulation of the learning dynamics can sustain will be important. That is, if agents believe (or estimate) a growth rate γ and extrapolate that into the future, is it the case that the system grows at a rate which is at least as large as γ in the long run? The answer, not surprisingly, depends on the asymptotic rate of growth of the TEM g. If g is linear, it is known from earlier work (Marcet and Sargent (1989), Grandmont (1998)), that the learning dynamics can sustain all sufficiently high growth rates. The first result of this paper shows that if the TEM does not diverge too fast, then the learning dynamics cannot sustain arbitrarily high rates of growth. A bound on the rate of divergence of the TEM is specified in Definition 1 below and a TEM that evokes it is said to induce a Globally Stable Formulation (henceforth, *GSF*) of the learning dynamics. For a given function $f(y)$, the condition $f(y) = \mathcal{O}(\sqrt{y})$, will mean that there exist constants $\mathcal{M} > 0$ and $\mathcal{N} > 0$, arbitrarily large but fixed, such that $|y| \geq \mathcal{M} \Rightarrow |f(y)| \leq \mathcal{N}\sqrt{y}$.

Definition 1. *A TEM g that satisfies $g(y) = \mathcal{O}(\sqrt{y})$ is said to induce a Globally Stable Formulation.*[12]

By contrast, an Unstable Formulation (henceforth, *UF*) occurs whenever the TEM grows at least at a linear rate eventually.

Definition 2: *A TEM g is said to induce an Unstable Formulation of the learning dynamics if there exist constants $k > 0$ and $S > 0$ such that either the condition $|g(y)| > ky$ for $y > S$, or $|g(y)| > -ky$ for $y < -S$, or both hold.*[13]

---

[12] We are providing this definition for simplicity, since it suffices for the article. A more general definition, that allows for a random component with unbounded support, would state that "*a TEM g that satisfies $g(y, \nu) = \mathcal{O}_p(\sqrt{y})$ is said to induce a Globally Stable Formulation*" where the condition $f(y) = \mathcal{O}_p(\sqrt{y})$ means that for any arbitrarily large constant $\mathcal{M} > 0$ and for any $\varepsilon > 0$ there exists a constant C such that $|y| \geq \mathcal{M} \Rightarrow P(|f(y)| > C\sqrt{y}) < \varepsilon$. Given assumption 1, there is no loss of generality.

[13] Similarly, a more general definition that allow for stochastic component would state that the TEM g is

In a GSF, one gets for large $|y|$ and $|\gamma|$ an inequality of the form

$$|\frac{g(\gamma^2 y, \nu)}{y}| < |\gamma|$$

uniformly in $\nu \in [-\overline{v}, \overline{\nu}]$, which indicates that the actual growth induced by the belief γ is smaller than γ in modulus. This is shown to disallow the learning dynamics to sustain arbitrarily high rates of growth in the long run. In an UF, however, the inequality goes the other way and the conclusion is the opposite. The Proposition below summarizes the dynamic implications of the two definitions and the subsequent analysis focusses on how these contrasting formulations may arise.

Proposition 1. *Let the TEM g satisfy assumption 1 or assumption 2. (i) If g induces an UF of the learning dynamics, then, along every sample path, there exist γ_n and y_n, such that $|\gamma_t|$ forms an increasing sequence for $t \geq n$, and $|y_n| \to \infty$. (ii) If g induces a GSF on the learning dynamics, then, along every sample path, $|y_n| \nrightarrow \infty$, $|\gamma_n| \nrightarrow \infty$, i.e., for every trajectory of the learning dynamics, there exists a positive constant $\overline{\mathcal{Q}}$ such that $|\gamma_n| < \overline{\mathcal{Q}}$ and $|y_n| < \overline{\mathcal{Q}}$ infinitely often.*

3   Representations and Globally Stable Formulations

A GSF does not allow the learning dynamics to sustain high growth rates. This section examines the dependence of this feature of the learning dynamics on the specific representation of the state variable, in the context of deterministic economies.[14]

A state variable can be given infinitely many representations. To fix ideas, we arbitrarily choose one representation x and call it the primitive. The TEM in this representation is

$$x_t = f(x_{t+1}^e).$$

It is referred to as the primitive TEM and is assumed to satisfy assumption 2. An alternate representation of the state variable arises via a change of variable and is given by $y = h(x)$ where h is a smooth, one to one change of variable satisfying $h(0) = 0$, so that 0 remains the steady state in the new representation as well. If the agent chooses to formulate the linear regression model on the variable y, the ensuing point forecast is given by (2) and induces the point forecast $x_{t+1}^e = h^{-1}(y_{t+1}^e)$, which yields $x_t = f(h^{-1}(y_{t+1}^e))$. Finally, one obtains the TEM g, which summarises the dependence of the representation y_t on its forecasted value y_{t+1}^e, as

$$y_t = h(x_t) = h(f(h^{-1}(y_{t+1}^e))) \equiv g(y_{t+1}^e)$$

said to induce an Unstable Formulation of the learning dynamics if for any $\varepsilon > 0$ there exist constants $k > 0$ and $S > 0$ such that for $|y| > S$, $P(|g(y, \nu)| > k|y|) > 1 - \varepsilon$.

[14] An extension to stochastic economies is provided in part 2 of the Appendix.

as is hypothesized in (1), (without the stochastic term). The TEM g will henceforth be referred to as the TEM *induced* by the representation h.[15]

Notice that since h and f are not necessarily linear, the maps f and g may have different rates of growth asymptotically and may therefore differ in the rates of growth that they can sustain asymptotically. It is through this channel that the representation affects whether or not the learning dynamics can sustain high rates of growth in the long run.

A representation $h(x)$ will be assumed to satisfy

Assumption 3: $h\colon R \to R$ *is a monotone, smooth change of variable that has 0 as a fixed point, $h(0) = 0$.*

This section specifies sufficient conditions on h which induce an UF and a GSF respectively. In particular, an agent can move from an UF to a GSF and vice versa. As a first step, we observe that the two scenarios considered in this paper, respectively an UF, and a GSF, are robust, in the following sense: if one starts from either of these formulations and applies a transformation h that satisfies the Lipschitz condition outlined in assumption 4 below, one preserves the formulation in the representation g that h induces. We state the assumption below and then the robustness property which is summarized as Proposition 2.

Assumption 4: *h satisfies the global Lipschitz condition $0 < c < |h(x)|/|x| < d$, for all x and some fixed c, d.*

Proposition 2. *(i) Suppose that f satisfies assumption 2 and that $f(x) = \mathcal{O}(\sqrt{x})$, so that the identity map induces a GSF, and that h satisfies assumption 3 and assumption 4. Then the induced TEM g satisfies $g(y) = \mathcal{O}(\sqrt{y})$ and h, accordingly, induces a GSF of the learning dynamics.*

(ii) Suppose that f satisfies assumption 2 and the condition of Definition 2 so that the identity map induces an UF, and h satisfies assumption 3 and assumption 4. Then the induced TEM g satisfies the condition of Definition 2 and, accordingly, *the representation h induces an UF of the learning dynamics.*

---

[15]The change of variable h may arise from a change of variable applied to the *levels* of the primitive state variable X. Agents first fix a representation $Y = H(X)$ of the state variable on which they formulate their forecasting model and the TEM in the representation Y accordingly becomes $Y_t = G(Y_{t+1}^e)$ where $G = H \circ F \circ H^{-1}$, which is well defined since H is assumed to be monotone. We assume throughout that the agents know the steady state value of the state variable. Thus, once H is fixed, agents know that the steady state of the system is $Y^* = H(X^*)$. It is appropriate under the hypothesis that the agents know the steady state $Y^* = H(X^*)$, to formulate the forecasting procedure directly on the deviation y of the state variable from the steady state. The change of variable H in absolute levels induces the change of variable $h(x)$ in the variable $x(= X - X^*)$ using the identity

$$Y = h(X - X^*) + Y^* = H(X).$$

We will work with representations expressed in deviations $y = h(x)$ and with the induced TEM $g(y_{t+1}^e)$.

We now examine the role of eventually concave and eventually convex functions, (which violate assumption 4), in reversing the qualitative features of the learning dynamics. To understand how the eventual concavity or convexity of h translates to non linearities in the induced TEM g, it is convenient to specialize the setting to an f, which in addition to satisfying assumption 2, is eventually increasing. This ensures that any representation h satisfying assumption 3 will induce a TEM g that is eventually increasing. It will also be useful to assume that h is increasing. These assumptions allow us to work with a formula that relates the risk aversion of h with the risk aversion of the induced TEM g. Letting $\rho(y) \equiv h^{-1}(y)$, one gets $g = h \circ f \circ \rho$. For a given increasing function $r(z)$ that is defined for $z > 0$, let $Rr(z)$ denote its relative risk aversion. The following formula relates the relative risk aversion of $g(y)$ (restricting the discussion to $y > 0$) to the relative risk aversion of $h(x)$ and will be useful in the sequel,

$$Rg(y) = \frac{y\rho'(y)}{\rho(y)}[Rf(\rho(y)) + Rh(f(\rho(y)))\frac{f'(\rho(y))\rho(y)}{f(\rho(y))} - Rh(\rho(y))]. \tag{7}$$

It is instructive to consider the effect of the non-linearities of h on the induced TEM g when the primitive TEM f is linear. Observe therefore that if $f(x) = ax$, the formula (7) simplifies to

$$Rg(y) = \frac{y\rho'(y)}{\rho(y)}[Rh(a\rho(y)) - Rh(\rho(y))]. \tag{8}$$

One sees that a linear h, or indeed one with constant relative risk aversion does not cause the induced TEM g to be concave. In fact, if $0 < a < 1$, the concavity of h is inversely related to the concavity of g. This also makes apparent the potential "stabilizing" effect that a "very convex" h may have starting from a linear TEM f where $0 < a < 1$. Indeed if the relative risk aversion of h is a sufficiently rapidly increasing function, one might achieve that the relative risk aversion of g is positive and large enough so as to induce a GSF. Concave transformations on the other hand, have a "destabilizing" effect, that is, they can transform an f that induces a GSF (under the identity transformation) into a g that induces an UF. The following sections elaborate on these features.

3.1   Representations that induce Stability

In this subsection we transform, via an h, an UF into a GSF. Consider a linear primitive TEM f given by $f(x) = \bar{a}x$ where $\bar{a}$ is a positive constant. Under the additional assumption that f is a contraction, i.e $0 < \bar{a} < 1$, we identify a condition on h that causes the induced TEM g to satisfy $g(y) = \mathcal{O}(\sqrt{y})$.

Furthermore, an identical conclusion holds even if the primitive f is non linear but it

is eventually a contraction, i.e $|f(x)|/|x| < \bar{a} < 1$ for $|x|$ large. This follows from the next inequality, which holds, for large $|y|$, for any f satisfying assumption 2 and h satisfying assumption 3:

$$\frac{|h(f(h^{-1}(y)))|}{|y|} \leq \frac{|h(\bar{a}h^{-1}(y))|}{|y|}.$$

If h is a sufficiently convex function of $|x|$ eventually, one gets $|g(y)|$ to be a sufficiently concave function of $|y|$ so as to satisfy the condition $g(y) = \mathcal{O}(\sqrt{y})$ eventually and thereby induce a GSF. Noting that the Relative Risk aversion of the map $\bar{g}(y) = \sqrt{y}$ is $1/2$, it is straightforward to verify that the condition $Rg(y) \geq 1/2$ suffices to establish that $g(y) = \mathcal{O}(\sqrt{y})$. Since we are working with a linear TEM, the formula for $Rg(y)$ is given by (8). These observations are summarized below as a proposition without proof.

Proposition 3. *Let f satisfy assumption 2 and in addition the following contracting property: $|f(x)| \leq \bar{a}|x|$ for some $0 < \bar{a} < 1$ and $|x|$ large. Let h be an increasing (respectively, decreasing) transformation, satisfying assumption 3, such that $h_+(x) = h_1(x)$ (resp, $-h_1(x)$) and $h_-(x) = -h_2(-x)$ (resp, $h_2(-x)$) where $h_1(z)$ and $h_2(z)$ map R_+ to R_+, and are eventually sufficiently convex functions so that $\frac{y\rho'(y)}{\rho(y)}[Rh_j(\bar{a}\rho(y)) - Rh_j(\rho(y))]$ exceeds $1/2$ eventually for $j=1,2$. Then the induced TEM g satisfies $g(y) = \mathcal{O}(\sqrt{y})$ and h consequently induces a GSF.*

The example below gives an explicitly parameterized functional form for $h(x)$ to show that enough convexity of h leads to a GSF being induced whenever the underlying TEM f is eventually a contraction. As argued earlier, it suffices to establish this for a TEM f that is linear and a contraction.

Example 3. Let $f(x) = \bar{a}x$ with $0 < \bar{a} < 1$. We now verify the existence of maps h that lead $Rg(y)$ to eventually exceed $1/2$. Consider the following family of increasing transformations, indexed by n, where for $x \geq 0, h_+(x) = e^{x^n} - 1$ and $h_-(x) = -h_+(-x)$ for $x < 0$. The formula for $Rg(y)$ reduces to

$$Rg(y) = \frac{y}{y+1}[1 - \bar{a}^n]$$

Since $\bar{a} < 1$, for n large enough, the above eventually exceeds $\frac{1}{2}$ as required for $g(y) = \mathcal{O}(\sqrt{y})$. Figure 1 represents this transformation for the case $n = 2$, and Figure 2 plots the initial TEM with $\bar{a} = 0.5$ and the induced TEM.
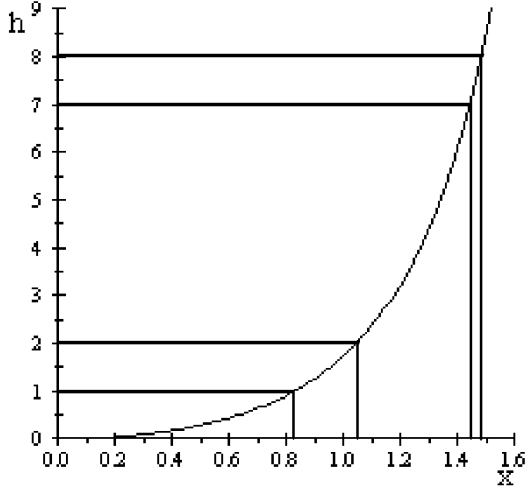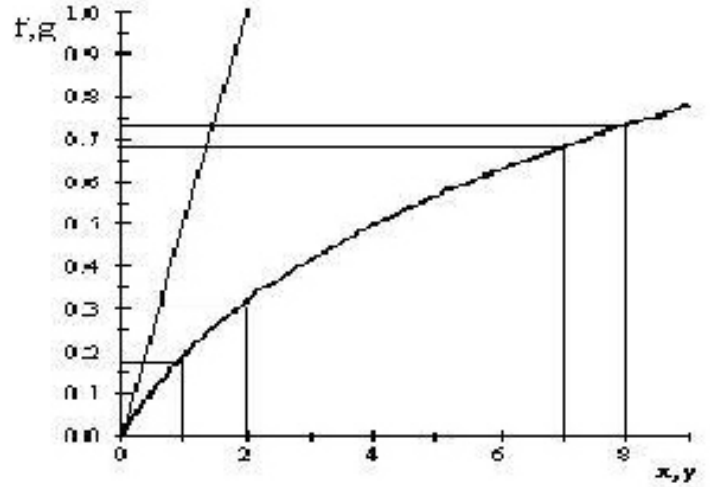
Figure 1. Convex transformation.

Figure 2. Initial TEM (thin) and induced TEM (thick).

3.2 Representations that induce an UF

Here we provide a set of conditions on the primitive f and the representation h that ensures that the induced TEM g describes an UF of the learning dynamics. It suffices to ensure that g is eventually monotone, diverges in modulus and its first derivative is bounded away from zero eventually. Indeed these conditions ensure that g eventually lies, for the case where the derivative is bounded away from zero by a positive number, above a line τy, $\tau > 0$, for y large enough. An analogous conclusion holds if the derivative is bounded away from zero by a negative number. These are summarized in the proposition below, a formal proof of which is omitted.

Proposition 4. *Assume f satisfies assumption 2 and h satisfies assumption 3. (i) Suppose that f is eventually increasing and either $f(y) > 0$ for $y > \mathcal{C}$, or $f(y) < 0$ for $y < -\mathcal{C}$, or both, for some $\mathcal{C} > 0$. Then the induced TEM g is also increasing and satisfies either $g(y) > 0$ for $y > \mathcal{C}$, or $g(y) < -\mathcal{C}$, or both. If instead, f is eventually decreasing and either $f(y) < 0$ for $y > \mathcal{C}$, or $f(y) > 0$ for $y < -\mathcal{C}$, for some $\mathcal{C} > 0$, or both hold, then the induced TEM g is decreasing and satisfies either $g(y) < 0$ for $y > \mathcal{C}$, or $g(y) > 0$ for $y < -\mathcal{C}$, or both.*

(ii) If in addition the induced TEM g is increasing (respectively, decreasing), satisfies $g(y) > 0$ for $y > C$(resp, $y < -C$), and if $g(y) \geq \overline{g}(y)$ where $\overline{g}\prime(y) > c > 0$ (resp, $\overline{g}\prime(y) < -c < 0$) for some constant c and y large enough, then g satisfies $g(y) > \tau y$ (resp, $-\tau y$) for y large enough, for some positive constant τ, and accordingly induces an UF of the

15

learning dynamics. If g is decreasing (resp, increasing), satisfies $g(y) < 0$ for $y > C$(resp, $y < -C$), and if $g(y) \leq \bar{g}(y)$ where $\bar{g}\prime(y) < c < 0$ (resp, $\bar{g}\prime(y) > c > 0$) for some constant c and y large enough, then g satisfies $g(y) < -\tau y$ (τy) for y large enough, for some positive constant τ, and accordingly h induces an UF of the learning dynamics.

It turns out that increasing and eventually concave representations h induce an UF starting from a primitive TEM f that eventually grows at some some minimal rate, so as to satisfy the condition that its graph lie above the graph of a function of the form of $Kx^\alpha, \alpha > 0$. The Corollary below states this possibility.

Corollary 1. *Let f satisfy assumption 2 and grow eventually at least as fast as $\bar{f}(x) = Kx^\alpha, \alpha > 0$, that is $f(x) \geq \bar{f}(x) = Kx^\alpha$ for $x > M$, where M and K are positive constants. Then, there exist increasing representations h satisfying assumption 3 and eventually concave, that induces an UF. Specifically, suppose $h(x) = ln(1 + x)$ for $x \geq 0$. Then, the induced TEM g satisfies $g(y) \geq \bar{g}(y) \equiv \ln(1 + K(e^y - 1)^\alpha)$ for y large with $\bar{g}\prime(y) > c > 0$ eventually, and induces an UF of the learning dynamics.*

In Figure 3 we have plotted the transformation $ln(1 + x)$, and in Figure 4 we have plotted the initial TEM with $\alpha = 0.4$ and $K=1$, and the TEM induced by the logarithmic transformation.
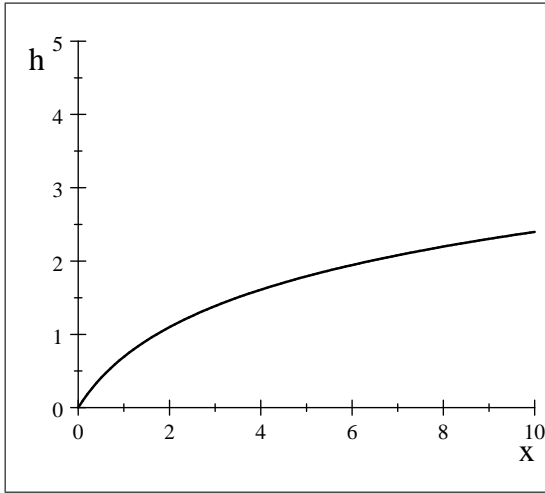

Figure 3. Concave transformation.
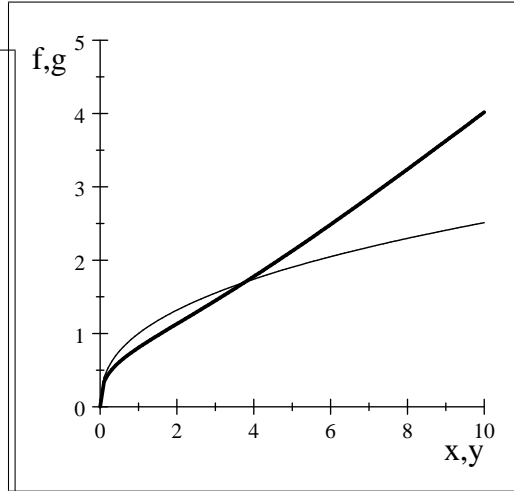
Figure 4. Initial TEM (thin) and induced TEM (thick).

Remark 1. The transformation $h(x) = ln(1 + x)$ for $x \geq 0$ has its relative risk aversion increasing and equal to 1 asymptotically. It can be verified using (7) that it suffices to work with an h that has constant relative risk aversion that is no less than 1. For example, if $h(x) = K \ln x + c$, for $x \geq M > 0$, $K > 0$, then $Rh = 1$ for x large and h accordingly induces an UF of the OLS dynamics. (One can verify that if Rh is of the form $1 - \beta, 0 < \beta < 1$, as

would arise for instance via $h(x) = Kx^\beta$, for $x \geq M > 0$, one is left with a concave g which does not ensure an UF). One could sharpen the result further by assuming that $\overline{f}(x)$ grow slower than assumed and be given, for instance, by $\overline{f}(x) = K \ln x$. One then needs to assume a greater degree of concavity of h to reach an identical conclusion. Analogous statements hold for a primitive TEM that is decreasing.

3.3   Destabilization via Concave Representations

A standard econometric practice to stabilize a time series sequence consists of applying the logarithm to the original series of the state variable. The previous result shows that in a learning model, such as the one we consider, applying this transformation does not attenuate the instability. Therefore, in the presence of local instability, an agent following this practice, termed an active agent, would never stabilize the system, whereas a passive agent who would not attempt any transformation would do. Figure 5 shows this case. In particular, we have plotted for an initial linear TEM $f(x) = ax, a = 1/2$, the estimated "beliefs" about the growth rates of the state variable for an active and a passive agent. Note the steady convergence for the passive agent and the exponential increase for the active agent.

Let γ_t^x and γ_t^y refer to the parameter estimate using the initial data $x_{j+1}/x_j, j = -L, ., .., 2$ and $h(x_{j+1})/h(x_j)$, $j = -L, ., .., 2$ respectively.

We emphasize the destabilizing effect of concave transformations via the following corollary.

Corollary 2. *Let f satisfy assumption 1, the global contraction condition $|f(x)| < a|x|$ for all x and, in addition, that it eventually satisfies $\overline{f}_1(x) \leq f(x) \leq \overline{f}_2(x)$ for $x > S > 0$ where the bounds grow eventually at least as fast as $\overline{f}_j(x) = Kx^{\alpha_j}, \alpha_j > 0$ with $0 < \alpha_1 < \alpha_2 < 1/2$. Then, OLS formulated on x is globally stable, i.e., $(x_t, \gamma_t^x, \omega_t) \rightarrow (0, \overline{\gamma}^x, \overline{\omega})$ with $|\overline{\gamma}^x| \leq 1$. There exist increasing concave transformations $h(x)$ s.t OLS formulated on $y = h(x)$ leads to an UF with $y_t \rightarrow \infty$ whenever the initial parameter estimate γ_0^y is large enough. In particular, one obtains instability for initial conditions that satisfy $\gamma_0^x > \gamma_0^y$.*

The condition $\alpha_2 < 1/2$ above ensures that a GSF is induced by the TEM f. In the Appendix we show that this condition along with the contracting condition that is imposed in the Corollary above, ensures that learning dynamics formulated on x are globally convergent to the steady state. On the other hand, if agents forecast $y = h(x)$, then under the conditions on h specified in the Corollary 2, Corollary 1 applies and one gets an UF with the dynamics of the state variable diverging to infinity for large enough γ_0^y. In particular, the parameter estimate γ_0^y obtained after applying a concave transform may be less than the original parameter estimate γ_0^x, but nevertheless leads to divergent dynamics. In conclusion, the

agent´s efforts to transform the data, so as to *dampen* the divergent trends in the original data, may be *destabilizing* for the system.
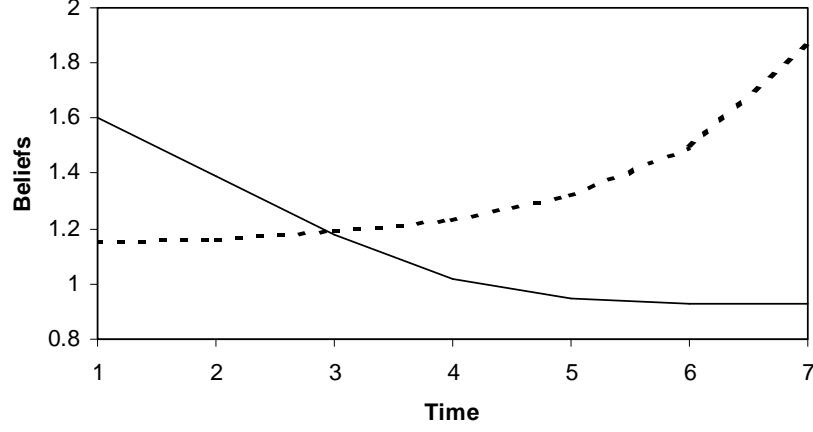

Figure 5. Estimated beliefs for passive agent (solid) and active agent (dotted).

4   Differencing and Detrending

In this section we consider the case where instead of applying an instantaneous transformation $h(x)$ to the data, the agent either differences or detrends the state variable. These two are the most common transformations that time series econometricians apply to nonstationary data to induce stationarity, according to whether the series are supposed to be DS or TS.

First, consider that the agent reformulates the forecasting model in terms of the differences of the current value of the state variable from its lagged value. Thus, the agents model is of the form (4) with $y_t = x_t - x_{t-1}$ and forecasts and parameter estimates continue to be given by (2) and (3) respectively. The forecast y_{t+1}^e induces the forecast $x_{t+1}^e = y_{t+1}^e + x_t^e$ with $x_t^e = y_t^e + x_{t-1}$. We retain the assumption that information up to $t - 1$ is available at the time of forecasting x_{t+1}^e, and this gives

$$x_{t+1}^e = [\gamma_{t-1}^2 + \gamma_{t-1}]y_{t-1} + x_{t-1}.$$

Assuming for exposition that the primitive TEM is linear and given by $x_t = ax_{t+1}^e$, one obtains

$$y_t = a\gamma_{t-1}^2 y_{t-1} + R(\gamma_{t-1}, y_{t-1}, x_{t-1}),$$

where $R(\gamma_{t-1}, y_{t-1}, x_{t-1}) = a\gamma_{t-1}y_{t-1} + [a-1]x_{t-1}$. As could be expected, differencing leads to stability in some cases. For instance, for a linear TEM in which the beliefs converge

18

monotonically from above to $1/a$. However, differencing is not, in general stability enhancing. One can show that if the initial parameter estimate γ_0 is sufficiently large, one gets the inequality $y_t > a\gamma_{t-1}^2 y_{t-1}$ which is an UF and accordingly the instability problem shows up again. Thus, differencing the data is not stability enhancing in a meaningful sense. The Proposition below summarizes this fact for any primitive TEM that induces an UF.

Proposition 5. *Assume the primitive TEM f satisfies assumption 1 and induces an UF of the learning dynamics. Suppose now agents formulate OLS learning on first differences $y_t = x_t - x_{t-1}$. Assume that the initial conditions $x_0, ..., x_{-L}$ are such that x_0 and x_{-1} satisfy $|(x_0/x_{-1})| > 1 + \delta$, where $\delta > 0$, and the parameter estimate γ_0 (given by (3) with $y_t = x_t - x_{t-1}$) is large enough. Then, under the dynamics with OLS learning formulated on y, $|\gamma_t|$ forms a strictly increasing sequence and the induced $|x_t|$ sequence diverges to infinity.*

Next, consider that the agent detrends the state variable using a polynomial and formulates the forecasting model in terms of the resulting residuals. For simplicity, we assume that there is no uncertainty associated to the estimation of the parameters of the polynomial. Thus, the agents model is of the form (4) with $y_t = x_t - P_n(t)$ where $P_n(t) = \sum_{j=0}^{n} \alpha_j t^j$, and forecasts and parameter estimates continue to be given by (2) and (3) respectively. The forecast y_{t+1}^e induces the forecast $x_{t+1}^e = y_{t+1}^e + P_n(t+1)$. We retain the assumption that information up to $t-1$ is available at the time of forecasting x_{t+1}^e. Assume for illustration that the primitive TEM is linear and given by $x_t = ax_{t+1}^e$. Then,

$$y_t = a\gamma_{t-1}^2 y_{t-1} + R_n(t), \tag{9}$$

where $R_n(t) = [aP_n(t+1) - P_n(t)]$. Assume that the agent starts to detrend the data at time t_0, then similarly to above, there exists an initial value γ_{t_0} large enough such that the system $y_t = a\gamma_{t-1}^2 y_{t-1}$ would lead to y_t growing exponentially, so that $R_n(t)$, which is just a polynomial, does not affect its eventual behavior. Thus, polynomial detrending is not stability enhancing. The Proposition below summarizes this fact.

Proposition 6. *Assume the primitive TEM f satisfies assumption 1 and induces an UF of the learning dynamics. Suppose at some date t_0 the representative agent decides to detrend the data and formulate OLS learning on the detrended data $y_t = x_t - P_n(t)$. Then, if*

$$\gamma_{t_0} > \max \left\{ \frac{1}{2a} + \frac{1}{2a}\sqrt{1 - \frac{4aR_n(t_0+1)}{y_{t_0}}}, \frac{|R_n(t_0+2)|}{|R_n(t_0+1)|} \right\}$$

$$\alpha_j \geq 0 \text{ for all } j$$

and

$$a < (1 + \frac{1}{t_0})^{-n} \tag{10}$$

then y_t diverges exponentially to infinity.

Note that the role of condition (10) is to assure that $R_n(t) < 0$ for all $t > t_0$. The previous proposition identifies a particular set of conditions that induce instability. Instability is a general outcome, in fact, we have performed a variety of simulations for alternative initial specifications and we have not been able to find a single case in which the series y_t would not diverge. Note, from equation (9), that divergence either takes an exponential form, when the leading term is $a\gamma_{t-1}^2 y_{t-1}$, or it takes a polynomial form, when the leading term is $R_n(t)$.

In practice, the most common transformation in time series takes first differences on the logarithm of the original data. We assume that the primitive variable satisfies a positivity constraint in levels, that is, $X \geq 0$, and has a unique steady state value X^*. In the next two examples we examine this combined transformation for a linear TEM and for a concave TEM.

Example 4. Assume that the initial TEM when expressed in levels X eventually behaves like

$$X_t = b + aX_{t+1}^e$$

where a and b are positive constants. Notice that the steady state X^* is no longer 0 but $b/(1-a)$. Now assume that the agent considers the standard transformation in econometrics, where the logarithms are applied to the data in levels rather than in deviations,

$$Y_t = \log(X_t/X_{t-1}). \tag{11}$$

In this case the unique steady state for the new Y_t variable is 0. Notice that the law of motion for Y_t satisfies

$$Y_t = \log\left(\frac{b}{X_{t-1}} + a\exp(Y_{t+1}^e + Y_t^e)\right).$$

When $X_{t-1} = b/(1-a)$, the steady state value, then

$$Y_t = \log(1 - a + a\exp(Y_{t+1}^e + Y_t^e))$$

so that the steady state corresponds to $Y = 0$. When X_{t-1} is very large compared to b, the map will approximately be $Y_t \approx \log a + (Y_{t+1}^e + Y_t^e)$. In fact, for $X_{t-1} > 0$, one has $Y_t > \log a + (Y_{t+1}^e + Y_t^e)$. In this case, assuming agents apply the forecasting rule (2) to Y,

$$Y_{t+1}^e = \gamma_{t-1}^2 Y_{t-1} \text{ and } Y_t^e = \gamma_{t-1}Y_{t-1}, \tag{12}$$

so that

$$Y_t > \log a + (\gamma_{t-1}^2 + \gamma_{t-1})Y_{t-1}.$$

The above law of motion is akin to an UF. Indeed for $Y_{t-1} > 0$ and fixed, if γ_{t-1} is large enough, one gets the configuration

$$\frac{Y_{t+j}}{Y_{t-1+j}} > \gamma_{t-1+j} > 1, \text{ for } j \geq 1$$

so that Y_t diverges and γ_t forms an increasing sequence. ▲

The previous example shows that applying the most typical transformation to a time series, namely taking differences in the logged data, is not in general stability enhancing. In fact, the next example conveys a stronger message: even with an initial TEM that evokes a GSF, this combined transformation leads to the possibility of divergence to infinity of the transformed state variable.

Example 5. Consider again a state variable in levels, $X \geq 0$, with a unique steady state value X^*. Assume that the TEM in this representation eventually behaves like

$$X_t = K \left(X_{t+1}^e \right)^{\alpha}$$

where $0 < \alpha < 1/2$.

Suppose first that agents formulate their forecasting model on X. Following the procedure employed in this article, we express the linear forecasting model (2) in deviations x from the known unique steady state X^*, $x_t = X_t - X^*$. It follows then that the TEM, when expressed in deviations, becomes eventually

$$x_t = K \left(x_{t+1}^e + X^* \right)^{\alpha} - X^*.$$

Since we have assumed $\alpha < 1/2$, this TEM satisfies the requirement of a GSF. The dynamics under OLS formulated on x therefore do not allow x_t to diverge to infinity.

Now suppose the agent considers the transformation (11) and the agent's beliefs are given by applying (2) to Y to obtain (12). Then, if at some point t_0 the beliefs satisfy $\gamma_{t_0} > -1/2 + 1/2\sqrt{1 + 4/\alpha}$, then Y_t constitute a nonstationary process. This result is immediate by noticing that

$$X_{t-1} = K^{\frac{1}{1-\alpha}} \frac{\exp(Y_{t+1}^e + Y_t^e)^{\frac{\alpha}{1-\alpha}}}{\exp(Y_t)^{\frac{1}{1-\alpha}}}$$

so that Y_t follows a linear second order difference equation

$$Y_{t+1} = \alpha(1 + \gamma_t + \gamma_t^2)Y_t - \alpha(\gamma_{t-1} + \gamma_{t-1}^2)Y_{t-1}$$

and a sufficient condition for the explosiveness of this process is that $\left| \alpha(\gamma_{t-1} + \gamma_{t-1}^2) \right| > 1$, that is, $\gamma_{t-1} > -1/2 + 1/2\sqrt{1 + 4/\alpha}.$▲

Therefore, taking differences in the logged data can be destabilizing even in the configuration of Example 5. Similarly to Section 3.3, the results from the previous example can be extended to more general settings where the initial TEM is eventually bounded below and above by power functions with power coefficients bounded above by 0.5, and in addition, it satisfies an appropriate contracting condition. For these general cases, learning dynamics

formulated on the original variable leads to global convergence to the steady state, whereas the learning dynamics formulated after taking differences in logged data leads to divergence to infinity. The situation is equivalent to the one graphed in Figure 5: an active agent would fare worse in terms of convergence to the steady state than a passive agent.

5   Concluding Remarks

In this article, we have modelled learning in a Walrasian set up in the sense that agents do not recognize the consequences of their actions on the market clearing prices or the dependence of the prices on the actions and forecasts of other agents. The only information available to the agents are past values of the only endogenous state variable, the price. Even in the simple deterministic set up examined here, this information structure is, in principle, compatible with the agents learning and converging to the steady state.[16] Agents generate a point forecast of future prices using past prices via OLS and we examine the resulting dynamics. Earlier studies have shown that there is a tendency for the resulting dynamics to be divergent. This article has attempted to introduce some sophistication in the learning process, in that, a representative agent attempts actively to improve the forecasting ability of her model by considering several transformations of the state variable within the framework of OLS learning. We find that for a class of economies, standard transformations used in econometrics, like differencing, detrending or concave transformations like taking logarithms, are not stability enhancing. As a practical conclusion, an active agent, who employs the standard econometric transformations to induce stability, may eventually fare worse than one that passively sticks to a fixed formulation in spite of bad forecasts in the short run. More work is needed before we obtain a better understanding of what assumptions on individual behaviour may actually underlie the perfect foresight approach to dynamic economics. Our somewhat negative findings however cast doubts on whether perfect foresight is at all compatible with decentralized Walrasian behaviour[17] in the learning phase of an economy.

We finally mention some possibilities for further work. We have considered the standard

---

[16]This need not be the case in more general models. Hellwig (1982) points out that Markovian processes with rational expectations need not exist unless the state space of the economy is expanded to include past endowments and expectations. A sensible formulation of a Walrasian agents learning problem in this set up is an important challenge.

[17]One alternative approach would be to model the learning phase as a dynamic strategic market game where agents explicitly take into account the price formation rules (as in Chatterji and Ghosal (2004)) but are uncertain about the strategies employed by others, or more generally, model learning in accordance with the literature on learning in games (Fudenberg and Levine 1998). A full blown strategic approach would be subject to the Radner critique (footnote 3).

econometric practice in which the agent, by inspecting the evolution of the time series of interest, decides to stabilize the data by the use of some transformation. Alternatively, an agent could base her decisions on the outcome of specification tests. These tests can be carried out in a variety of forms. In this article we have considered that the agent's beliefs are based on a simple linear autoregression of order one. Using this model as the null hypothesis, there are a variety of specification tests the agent could attempt. For instance, she could consider testing that the order of the autoregression is one against a higher order. In this case she could perform a simple Wald or Lagrange Multiplier (hencefoth, LM) test in the augmented model. Alternatively, she could question the linearity of the autoregression and employ general omnibus specification tests, such as those in Lobato (2003). In case the agent had some other alternative specific nonlinear model in mind, she could employ likelihood ratio tests. Another possibility is testing whether the autoregression of order one has constant or time-varying parameters, in this case LM tests are the simplest to implement. Finally, following the spirit of this article, the agent could stick to a linear autoregresion of order one with time constant parameters and test the necessity of the alternative transformations considered. In this respect, the simplest approach would again be the construction of LM tests such as those in Coulson and Robins (1987) and Wooldridge (1994), see also Atkinson and Lawrance (1989) for a comparison between competing tests.

6 Proofs

As a first step, we note the fact summarized in Lemma 1. A verification of this can be found in the proof of Proposition 1 of Chatterji and Chattopadhyay (2000) and is accordingly omitted.

Lemma 1. *Along any trajectory, the inequality*

$$|\gamma_t| \leq 1 + \frac{y_t^2}{2\sum_{-L}^{t-1} y_j^2}$$

is satisfied.

Proof of Proposition 1. The proof of divergence to infinity of the state variable under an UF is a standard replication of known arguments (Grandmont 1998) and is omitted. We prove that if g satisfies Assumption 2, then the dynamics of the state variable and the parameter estimate do not diverge to infinity. An identical proof holds for the case when g satisfies Assumption 1.

Step *(i)* Consider a sample path such that $|\gamma_n^2 y_n| < \mathcal{L}$ infinitely often for some constant $\mathcal{L} > 0$. By assumption 1, $|\gamma_n^2 y_n| < \mathcal{L}$ implies that $|y_{n+1}| = |g(\gamma_n^2 y_n)| < y(\mathcal{L})$ and so $|y_{n+1}| <$

$y(\mathcal{L})$ infinitely often where $y(\mathcal{L})$ is a positive constant. Next, by Lemma 1, $|\gamma_n| < 1 + \delta$, where $\delta > 0$, infinitely often. (To verify this, notice from the right hand side of the inequality stated in Lemma 1, that if $\sum\limits_{-L}^{n-1} y_j^2$ converges, then it must be the case that $y_n \to 0$ and the conclusion follows, while if $\sum\limits_{-L}^{n-1} y_j^2$ diverges, then for for any $\delta > 0$, if n is sufficiently large, one gets infinitely often the inequality $y_t^2 \left[2\sum\limits_{-L}^{n-1} y_j^2 \right]^{-1} < y(\mathcal{L})^2 \left[2\sum\limits_{-L}^{n-1} y_j^2 \right]^{-1} < \delta$ as required.) If one sets $\overline{\mathcal{Q}} = \max\{y(\mathcal{L}), 1 + \delta\}$, then $|\gamma_n| < \overline{\mathcal{Q}}$ and $|y_n| < \overline{\mathcal{Q}}$ infinitely often and the proof is complete.

Step (ii): Suppose now that $|\gamma_n^2 y_n|$ diverges. If $|y_{n+1}| < y(\mathcal{L})$ infinitely often for some positive constant $y(\mathcal{L})$, then the analysis of case (i) above applies; one has therefore $|\gamma_n| < \overline{\mathcal{Q}}$ and $|y_n| < \overline{\mathcal{Q}}$ infinitely often, so that $|\gamma_n^2 y_n|$ cannot diverge. So assume $|y_{n+1}|$ diverges as well. Since $g(y) = \mathcal{O}(\sqrt{y})$, there exist constants $\mathcal{M} > 0$ and $\mathcal{N} > 0$, arbitrarily large but fixed, such that $|y| \geq \mathcal{M} \Rightarrow |g(y)| \leq \mathcal{N}\sqrt{y}$. Since $|\gamma_n^2 y_n|$ and $|y_n|$ diverge, there exists T such that $|\gamma_{T+t}^2 y_{T+t}| > \mathcal{M}$ and $|y_{T+t}| > \max\{\mathcal{M}, 4\mathcal{N}^2\}$ for all $t \geq 0$. It follows that for all $t \geq 0$

$$\frac{|g(\gamma_{T+t}^2 y_{T+t})|}{|y_{T+t}|} \leq \mathcal{N}\frac{|\gamma_{T+t}|}{|y_{T+t}|^{\frac{1}{2}}} \leq \frac{1}{2}|\gamma_{T+t}|.$$

Since γ_{t+1} is a convex combination (with positive weights) of γ_t and $\frac{g(\gamma_t^2 y_t)}{y_t}$, one gets $|\gamma_{T+t+1}| \leq |\gamma_{T+t}|$ for all $t \geq 0$ and so $|\gamma_{T+t}|$ forms a decreasing sequence for $t \geq 1$. This implies in particular that along any sequence, the sequence of parameter estimates $|\gamma_t|$ remains bounded by a quantity $\widehat{\gamma} > 0$. This in turn leads to the inequality

$$\frac{|y_{T+t+1}|}{|y_{T+t}|} = \frac{|g(\gamma_{T+t}^2 y_{T+t})|}{|y_{T+t}|} \leq \mathcal{N}\frac{\widehat{\gamma}}{|y_{T+t}|^{\frac{1}{2}}}.$$

Now consider δ such that $0 < \delta < 1$. For some T', $|y_{T'+t}| \geq \mathcal{N}\widehat{\gamma}/\delta$ for all $t \geq 0$. It follows that, for all $t \geq 0$, $\frac{|y_{T'+t+1}|}{|y_{T'+t}|} < \delta$ and we must conclude that $y_t \to 0$. This contradicts the hypothesis that $|y_n|$ is unbounded. Thus this case cannot occur and we must always have $|y_{n+1}| < y(\mathcal{L})$ infinitely often for some positive constant $y(\mathcal{L})$ and the proof is complete.▲

Proof of Proposition 2. We first prove part (ii) of the Proposition. Assume h is increasing. (a) Suppose $|f(x)| > kx$ for $x > S > 0$. Since f is continuous, either $f(x) > kx$ for $x > S$ or $f(x) < -kx$ for $x > S$. Suppose $f(x) > kx$ for $x > S$ and let $y > h(S)$. Then $f(h^{-1}(y)) > 0$. Furthermore, $f(h^{-1}(y)) > kh^{-1}(y) > 0$. One then obtains $g(y) \equiv h(f(h^{-1}(y))) > h(kh^{-1}(y)) > h(\frac{k}{d}y) > \frac{ck}{d}y$. One thus has $\frac{g(y)}{y} > \overline{k} \equiv \frac{ck}{d}$ for $y > h(S)$. Suppose next that $f(x) < -kx$ for $x > S$. Then $f(h^{-1}(y)) < -bh^{-1}(y) < 0$ for all $y > h(S)$ and one gets $g(y) \equiv h(f(h^{-1}(y))) < h(-kh^{-1}(y)) < h(-\frac{k}{d}y) < -\frac{ck}{d}y = -\overline{k}y$. (b) Suppose

24

$|f(x)| > kx$ for $x < -S < 0$. If $f(x) > -kx$ for $x < -S < 0$, by an identical argument as in case (a), for $y < h(-S)$, one obtains $g(y) \equiv h(f(h^{-1}(y))) > h(-kh^{-1}(y)) > h(-\frac{k}{d}y) > -\frac{ck}{d}y = -\overline{k}y$. Finally, suppose $f(x) < kx$ for all $x < -S < 0$. One gets here $g(y) \equiv h(f(h^{-1}(y))) < h(kh^{-1}(y)) < h(\frac{k}{d}y) < \frac{ck}{d}y = \overline{k}y$ for $y < h(-S)$. Thus if h is increasing, the induced TEM g satisfies the requirement of Definition 1. Analogous arguments apply for an h that is decreasing. This completes the proof of *(i)* of the proposition.

We now prove part *(i)* of the Proposition. Assume h is increasing. By hypothesis, $f(x) = \mathcal{O}(\sqrt{x})$, which implies that for large $|x|, |f(x)| < \mathcal{L}\sqrt{x}$, where $\mathcal{L} > 0$. Observe that $f(h^{-1}(y))$ satisfies $-\mathcal{L}\sqrt{|h^{-1}(y)|} < f(h^{-1}(y)) < \mathcal{L}\sqrt{|h^{-1}(y)|}$. Since $|h^{-1}(y)| < \frac{1}{c}|y|$, the one gets $-\mathcal{L}\sqrt{\frac{|y|}{c}} < f(h^{-1}(y)) < \mathcal{L}\sqrt{\frac{|y|}{c}}$, which since h is increasing leads to $h(-\mathcal{L}\sqrt{\frac{|y|}{c}}) < h(f(h^{-1}(y))) < h(\mathcal{L}\sqrt{\frac{|y|}{c}})$. Finally since $|h(x)| < d|x|$, one gets $-d\mathcal{L}\sqrt{\frac{|y|}{c}} < h(f(h^{-1}(y))) < d\mathcal{L}\sqrt{\frac{|y|}{c}}$ which proves that $h(f(h^{-1}(y))) = g(y) = \mathcal{O}(\sqrt{y})$. Analogous arguments apply for an h that is decreasing. ▲

Proof of Proposition 5. For simplicity, we consider the case where the primitive TEM satisfies $f(x) > ax$ for x large with $a > 0$ (the other possibilities for an UF are treated analogously) and that the initial conditions $x_0, ..., x_{-L}$ are such that there exists $\delta > 0$ such that $|(x_0/x_{-1})| > 1 + \delta$, and the parameter estimate γ_0, given by (5), is large enough so that it satisfies that $\gamma_0 > -\frac{1}{2} + \frac{1}{\sqrt{a}}\sqrt{\frac{5}{4} + \frac{1}{\delta}}$. We have $x_2^e = [\gamma_0^2 + \gamma_0]y_0 + x_0$, which for large x_0 gives $x_1 > ax_2^e$ (since the TEM satisfies $f(x) > ax$ for x large) and accordingly $\frac{y_1}{y_0} > a\gamma_0^2 + a\gamma_0 + \frac{[a-1]x_0}{y_0}$. We first establish that $\frac{y_1}{y_0} > a\gamma_0^2$ for a γ_0 that satisfies the previous conditions. This is equivalent to showing that $a\gamma_0 + \frac{[a-1]x_0}{y_0} > 0$. Noting that $y_0 = x_0 - x_{-1}$ and rearranging terms, the required inequality becomes $\frac{x_0}{x_{-1}} > \frac{a\gamma_0}{a\gamma_0 + a - 1}$. Since we assume that $\frac{x_0}{x_{-1}} > 1 + \delta$, with $\delta > 0$, the inequality holds. This ensures that γ_1 which is a convex combination of γ_0 and $\frac{y_1}{y_0}$ will exceed γ_0. We show next that the condition $\frac{x_0}{x_{-1}} > 1 + \delta$ and $\gamma_0 > -\frac{1}{2} + \frac{1}{\sqrt{a}}\sqrt{\frac{5}{4} + \frac{1}{\delta}}$ implies that $\frac{x_1}{x_0} > 1 + \delta$. We have $\frac{x_1}{x_0} > 1 + \delta$ iff $\frac{y_1}{x_0} > \delta$. Next, $\frac{y_1}{x_0} = a(\gamma_0^2 + \gamma_0)(1 - \frac{x_{-1}}{x_0}) + (a - 1)$ and we have the inequality $\frac{y_1}{x_0} > a(\gamma_0^2 + \gamma_0)(1 - \frac{1}{1+\delta}) + (a - 1) = a(\gamma_0^2 + \gamma_0)\delta + (a - 1)$, which exceeds δ as required when γ_0 satisfies the previous conditions. Therefore, under the stated condition, $\frac{x_1}{x_0} > 1 + \delta$ and $\gamma_1 > \gamma_0$, and, by induction, the sequence x_t diverges to infinity and γ_t forms an increasing sequence.

Proof of Proposition 6. We consider the case where the primitive TEM satisfies $f(x) > ax$ for x large and $a > 0$. The other possibilities for an UF are treated analogously. Assume without loss of generality that γ_{t_0} and y_{t_0} are both positive. We will show that γ_t for $t = t_0 + 1, ...,$ form an increasing sequence so that y_t diverges exponentially to infinity. First, note that the stated conditions guarantee that $R_n(t) < 0$ for all $t \geq t_0$ because $R_n(t) = \sum_{j=0}^{n} \alpha_j[a(1 + t^{-1})^j - 1]t^j$. Second, note that a necessary and sufficient condition for $\gamma_{t_0+1} > \gamma_{t_0}$ is that $y_{t_0+1} > \gamma_{t_0}y_{t_0}$ and using that $y_{t_0+1} > a\gamma_{t_0}^2 y_{t_0} + R_n(t_0 + 1)$, the

condition $y_{t_0+1} > \gamma_{t_0} y_{t_0}$ holds when $\gamma_{t_0} > \frac{1}{2a} + \frac{1}{2a}\sqrt{1 - \frac{4aR_n(t_0+1)}{y_{t_0}}}$. For the second period, we need to verify that $\gamma_{t_0+2} > \gamma_{t_0+1}$, that is, $y_{t_0+2} > \gamma_{t_0+1} y_{t_0+1}$. Similarly as above, note that this holds as long as $\gamma_{t_0+1} > \frac{1}{2a} + \frac{1}{2a}\sqrt{1 - \frac{4aR_n(t_0+2)}{y_{t_0+1}}}$, but this condition holds because we know from the first period that $\gamma_{t_0+1} > \gamma_{t_0} > \frac{1}{2a} + \frac{1}{2a}\sqrt{1 - \frac{4aR_n(t_0+1)}{y_{t_0}}}$, and using that $\frac{|R_n(t_0+2)|}{y_{t_0+1}} < \frac{|R_n(t_0+1)|}{y_{t_0}}$, which follows from the fact that $y_{t_0+1} > \gamma_{t_0} y_{t_0}$ and the condition that $\gamma_{t_0} > \frac{|R_n(t_0+2)|}{|R_n(t_0+1)|}$. Then, recursively note that a necessary and sufficient condition for $\gamma_{t+1} > \gamma_t$ is that $y_{t+1} > \gamma_t y_t$ and using that $y_{t+1} > a\gamma_t^2 y_t + R_n(t+1)$, the condition $y_{t+1} > \gamma_t y_t$ holds when $\gamma_t > \frac{1}{2a} + \frac{1}{2a}\sqrt{1 - \frac{4aR_n(t+1)}{y_t}}$. Finally, notice that the stated conditions guarantee that the sequence $|R_n(t+1)|/y_t$ forms a decreasing sequence. To see this, note that $|R_n(t)|/y_{t-1} > |R_n(t+1)|/y_t$ is equivalent to $y_t/y_{t-1} > |R_n(t+1)|/|R_n(t)|$ and, while y_t/y_{t-1} is bounded below by γ_{t-1}, the sequence $|R_n(t+1)|/|R_n(t)|$ decreases monotonically to 1 as t tends to infinity using the Cauchy-Schwartz inequality.

References

Atkinson, A. C and A. J. Lawrance (1989), "A Comparison of Asymptotically Equivalent Test Statistics for Regression Transformation," *Biometrika*, 76, 223-229.

Bray, M. (1982), "Learning, Estimation, and the Stability of Rational Expectations," *Journal of Economic Theory*, 26,318-339.

Bray, M. and N. E. Savin (1986), "Rational Expectations Equilibria, Learning, and Model Specification," *Econometrica*, 54, 1129-1160.

Chatterji, S. (1995), "Temporary Equilibrium Dynamics with Bayesian Learning", *Journal of Economic Theory*, 67, 590-598.

Chatterji, S. (2002), "The Convergence of Least Squares Learning in Stochastic Temporary Equilibrium Models," *Economic Theory*, 20, 837-847.

Chatterji, S. and S. Chattopadhyay (2000), "Global Stability in spite of "Local Instability" with Learning," *Journal of Mathematical Economics*, 33, 155-165.

Chatterji, S. and S. Chattopadhyay (2002), "Learning, Global Dynamics and Normalization Rules" working paper.

Chatterji, S. and S. Ghosal (2004), "Local Coordination and Market Equilibria", *Journal of Economic Theory*, *114,* 255-279.

Coulson, N.E. and R. P. Robins (1987), "A Test of the First Difference Transformation in Time Series Models," *The Review of Economics and Statistics*, 69, 723-726.

Evans, G.W. and S. Honkapoja (1998), "Economic Dynamics with Learning: New Stability Results," *Review of Economic Studies*, 65, 23-44.

Frydman, R. (1982), "Toward an Understanding of Market Processes," *American Economic Review*, 72, 652-658.

Fudenberg, D and D. Levine (1998), Learning and the Theory of Games, MIT press, Cambridge, MA.

Grandmont, J.M. (1998), "Expectations Formation and Stability of Large Socio-economic Systems," *Econometrica*, 66, 741-781.

Grandmont, J.M., and Laroque, G. (1991), "Economic dynamics with learning: some instability examples", In Barnett, W.A. et al. (Eds), Equilibrium Theory and Applications. Cambridge University Press.

Hellwig, M.F. (1982) "Rational Expectations and the Markov Property of Temporary Equilibrium Processes", *Journal of Mathematical Economics* 9, 135-144.

Lobato, I.N. (2003), "Testing for nonlinear autoregression," *Journal of Business and Economic Statistics*, 21, 164-173.

Lucas, R. J. (1986), "Adaptive Behaviour and Economic Theory," *Journal of Business*, 59, 401-426.

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

Radner, R (1982) "Equilibrium under Uncertainty", Chapter 20, Handbook of Mathematical Economics, Volume 11, Editors Arrow, Kenneth J., and Intriligator, Michael, D., Elsevier Science Publishers.

Spear, S. (1985), "Rational Expectations in the Overlapping Generations Model," *Journal of Economic Theory*, 35,251-275.

Spear, S. and S. Srivastava (1986), "Markov Rational Expectations Equilibria in an Overlapping Generations Model," *Journal of Economic Theory*, 38, 35-62.

Saari, D. G., and S. R. Williams (1986), "On the Local Convergence of Economic Mechanisms," *Journal of Economic Theory*, 40, 152-167.

Van Zandt, T. (2003), "Robustness of Adaptive Expectations as an Equilibrium Selection Device," *Macroeconomic Dynamics, 7,* 89-118.

Wooldridge, J.M. (1994), "A Simple Specification Test for the Predictive Ability of Transformation Models," *The Review of Economics and Statistics*, 76, 59-65.

Appendix

Part 1 of this appendix provides conditions for the global convergence of the learning dynamics to the steady state. Part 2 extends the analysis of Section 3 of the paper to a stochastic formulation. Part 3 examines a set up where the domain of the problem is bounded. Part 4 gathers some proofs.

1. Global Stability.

A formulation that induces a GSF does not allow the state variable y_t and the parameter estimate γ_t to diverge to infinity. This of course does not imply convergence to the steady state 0. If one now imposes conditions that guarantee that the induced TEM g is a contraction, in addition to imposing a GSF, then one obtains global stability with all sequences y_t converging to the steady state and the parameter estimates converging to a finite limit. We assume in this section that the underlying TEM f, in addition to satisfying assumption 1, is a contraction, i.e., $|f(x)| < a|x|, 0 < a < 1$.

Consider now the TEM g induced by an h that induces a GSF. The TEM's f and g are related via the topological conjugacy $g = h \circ f \circ h^{-1}$ where h is smooth monotone change of variable satisfying $h(0) = 0$. By topological conjugacy, the iterated dynamics of the maps f and g must be qualitatively similar. Thus, if f generates trajectories $x_t \to 0$, then trajectories y_t generated by g must also satisfy $y_t \to 0$. For the OLS learning dynamics to have convergent dynamics, it is important to work with maps that satisfy a contracting condition.[18] We next specify a simple sufficient condition on h, namely the symmetry of h around the steady state 0, that will ensure that the induced TEM g will inherit a contracting property from the primitive TEM f.

Assumption 5. *Let h be symmetric around 0 in the sense that $h(x) = -h(-x)$ for all x.*

The symmetry condition specified above (which is by no means necessary) along with the following regularity condition guarantees that the induced TEM g has a nice contracting property which is described in Lemma 2 below.

Assumption 6. *The pair h and f satisfy satisfy the following regularity condition*

$$|g\prime(0)| \equiv |lim_{y \to 0} \frac{h(f(h^{-1}(y)))}{y}| \neq 1.$$

Lemma 2. *Assume f satisfies assumption 2 and $a < 1$, so that f is a global contraction. Let h satisfy assumption 3 and assumption 5. Then the induced TEM $g = h \circ f \circ h^{-1}$*

---

[18]It does not however follow that the contracting property $|f(x) < a|x|, 0 < a < 1$, that f is assumed to possess, is necessarily inherited by g. It may for instance, be the case that g does not satisfy a contracting condition $|g(y)| < a\prime|y|$, for any value of $0 < a\prime < 1$, but still generates trajectories y_t that converge to zero by virtue of the nth iterate of g, g^n satisfying the contracting condition $|g^n(y)| < a\prime|y|$ for $0 < a\prime < 1$.

satisfies $|g(y)| < |y|$ *for all* $|y| > 0$. *If one assumes in addition that the pair* h *and* f *satisfy assumption 6, one obtains the existence, for any compact interval* $[-\mathcal{Q},\mathcal{Q}]$, *of a constant* $a(\mathcal{Q})$ *satisfying* $0 < a(\mathcal{Q}) < 1$, *such that* $|g(y)| < a(\mathcal{Q})|y|$ *for all* $y \in [-\mathcal{Q},\mathcal{Q}]$.

One can now establish the following global convergence result.

Proposition 7 *Assume* f *satisfies assumption 1 and* $a < 1$, *so that* f *is a global contraction. Let* h, *satisfying assumption 3 (i) induce a GSF, and (ii) induce the condition that for any compact interval* $[-\mathcal{Q}, \mathcal{Q}]$, *there exist a constant* $a(\mathcal{Q})$ *satisfying* $0 < a(\mathcal{Q}) < 1$, *such that* $|g(y)| < a(\mathcal{Q})|y|$ *for all* $y \in [-\mathcal{Q}, \mathcal{Q}]$. *Then the OLS learning dynamics formulated on* $y = h(x)$ *are globally convergent to the steady state, i.e., for all initial conditions, one has* $(y_t, \gamma_t, \omega_t) \rightarrow (0, \overline{\gamma}, \overline{\omega})$ *with* $|\overline{\gamma}| \leq 1$.

Remark 2. This proposition generalizes the global stability proposition 2 of Chatterji and Chattopadhyay (2000). The latter required bounds on the range of the TEM which have been dispensed with in Proposition 7; indeed global stability can obtain for a TEM g satisfying $g(y) = \mathcal{O}(\sqrt{y})$, whose range consequently does not diverges "too fast". Importantly, as Proposition 3 makes clear, one may start with a TEM whose range diverges to infinity at a rate faster than a linear rate, and yet obtain global stability if agents forecast a suitably transformed variable $y = h(x)$ that induces a GSF and satisfies in addition assumption 5 and assumption 6. Conversely, there exist TEM's g which satisfy $g(y) = \mathcal{O}(\sqrt{y})$ and the contracting property, but which can be destabilized by an appropriate representation as specified in the previous section. However, the global stability finding of Proposition 2(CC) being driven by bounds on the range of the TEM survives for any representation h that preserves the contracting property. ▲

2. An Extension to Stochastic TEM's.

The main ideas of section 3 extend to stochastic economies. We focus on the case where the primitive TEM has, for simplicity, an additive stochastic component and is accordingly given by

$$x_t = f(x_{t+1}^e) + \nu_t$$

where f satisfies assumption 1 as before and $\{\nu_t\}$ is a sequence of mean zero (non-degenerate) i.i.d random variables whose support lies in the compact interval $[-\overline{\nu}, \overline{\nu}]$. We restrict attention for convenience to a representation h that is increasing and satisfies assumption 3. As observed earlier, f satisfies assumption 1. Assume in addition the following contracting property: $|f(x)| \leq a|x|$ for some $0 < a < 1$ and $|x|$ large. The induced TEM here is given by the relation

$$y_t = h(x_t) = h[f(h^{-1}(y_{t+1}^e)) + \nu_t] \equiv g(y_{t+1}^e, \nu_t)$$

and consequently satisfies the inequality

$$|y_t| = |h[f(h^{-1}(y_{t+1}^e)) + \nu_t]| \le |h[a(h^{-1}(y_{t+1}^e)) + \overline{\nu}]|$$

To show that the induced TEM $g(y, \nu_t) = \mathcal{O}(\sqrt{y})$, we focus on the positive part of the mapping as before and so it suffices to work with the map $\widehat{g}(y) \equiv h[a(h^{-1}(y)) + \overline{\nu}]$ and establish that it satisfies $\widehat{g}(y) = \mathcal{O}(\sqrt{y})$. The relative risk aversion of this map is given by the formula

$$R\widehat{g}(y) = \frac{y\rho'(y)}{\rho(y)}[Rh(a\rho(y) + \overline{\nu})\frac{a\rho(y)}{a\rho(y) + \overline{\nu}} - Rh(\rho(y))]$$

where as before $\rho(y) \equiv h^{-1}(y)$. This formula resembles (8) and essentially the same stabilizing mechanism is at play here. Therefore, an h that is sufficiently convex in the sense of Proposition 3 leads to a GSF as well. Example 6 below verifies these conditions where the primitive TEM is linear.

Example 6. Let the TEM be given by the stochastic relation $f(x) = ax + \nu_t$ with $0 < a < 1/2$. The TEM now violates the requirement $f(x) = \mathcal{O}(\sqrt{x})$. One has a UF of the learning dynamics if agents forecast x. We now verify the existence of a representation h that induces a TEM that satisfies the condition $\widehat{g}(y) = \mathcal{O}(\sqrt{y})$. Let $h(x)$ for x large be given by $h(x) = e^x - K$. For $x < 0$, the map $h(x)$ is defined by symmetry as $h(x) = -h(-x)$. The formula for $R\widehat{g}(y)$ for y large reduces to

$$R\widehat{g}(y) = \frac{y}{y + K}[1 - a]$$

Since $a < \frac{1}{2}$, the above eventually exceeds $\frac{1}{2}$ as required for $\widehat{g}(y) = \mathcal{O}(\sqrt{y})$.

Remark 3. Under additional assumptions, one can obtain a global stability result for stochastic TEMs as well. For instance, consider the formulation of Example 6. Assume that h is linear, for some $\overline{\lambda} > 0$, $h(x) = \overline{\lambda}x$ for $0 \le |x| \le Z$, then the TEM in some interval around 0 takes the conveniently additive form

$$y_t = \overline{g}(y_{t+1}^e) + \eta_t$$

where $\overline{g}(y_{t+1}^e) = h[f(h^{-1}(y_{t+1}^e)]$ and $\eta_t = \overline{\lambda}\nu_t$. If in addition to inducing a GSF, one has the contracting condition that there exist positive constants W and $\overline{a} < 1$ such that $|h[ah^{-1}(y)]| < \overline{a}|y|$ for all y and $|\widehat{g}(y)| < \overline{a}|y|$ for all $|y| > W$, along with the condition that h be linear in a large enough interval, i.e. Z be large enough, the stochastic dynamics eventually get trapped in a region where the above additive formulation applies. (Indeed in the formulation of Example 6 above, one needs to choose $Z > max\{(1-a)^{-1}\overline{\nu}, (1-a)^{-1}\overline{\nu}\overline{\lambda}\}$ and let $h(x) = \overline{\lambda}x$ for $0 \le x \le Z$, be weakly convex for all $z > Z$ and finally, as imposed in the example, let $h(x)$ for x large be given by $h(x) = e^x - K$ and for $x < 0$, the map $h(x)$ be defined by symmetry as $h(x) = -h(-x)$.) One can then adapt the method of Chatterji

(2002) to prove a global stability result wherein $\gamma_t \to 0$ almost surely and y_t in the limit is white noise around the steady state.▲

3. An Extension to Bounded Domains. [19]

In economic models, it is often the case that the state variable is bounded. These bounds may stem from positivity conditions of feasibility conditions etc. To incorporate such situations into the analysis, we examine the case where the range of h is bounded. Qualitatively similar results appear here under a mild regularity condition. A version of the GSF arises here by virtue of the fact that as agents perform regressions on a bounded variable, the parameter estimate is tied down in the limit, along sequences that approach the boundary of the bounded set.

We restate the basic assumptions in this setting as follows. The primitive TEM f now is not necessarily defined on the entire real line.

Assumption 7. *The TEM $f : D = (-d, \infty) \to R$, where $-\infty \leq -d < 0$, is continuously differentiable, has 0 as a fixed point, $f(0) = 0$ and satisfies the global Lipschitz condition $0 \leq b \leq \frac{|f(x)|}{|x|} \leq a$ for all x.*

Assumption 8. *h: $D \to V \equiv (-\underline{v}, \overline{v})$, where $\underline{v}, \overline{v} > 0$, is a monotone, smooth change of variable that has 0 as a fixed point, $h(0) = 0$.*

The forecasting model when formulated on the variable $y = h(x)$, now admits the possibility that the forecast generated by equation (2) falls outside the interval V, the admissible domain for the forecasts given that the agents forecast y. To prevent this one has to specify bounds on the forecasts.

Let the forecast ensuing from (2) be denoted as

$$\widetilde{y}^e_{t+1} = \gamma^2_{t-1} y_{t-1} \tag{13}$$

Unrestricted use of (13) leads to the possibility that values of y lying outside the permissible interval V are forecasted, a possibility which can arise in the model when V is a bounded interval and not the entire real line. In case $\underline{v}$ is finite, to ensure that forecasts of y do not fall below $-\underline{v}$, the lower bound for V, we truncate (13) at $-\underline{k}_t$ where $0 < \underline{k}_t < \underline{v}$ and $y_j \geq -\underline{k}_t$ for $j = t-1, t-2, ., 0, .., -L$, so that the lower bound $-\underline{k}_t$ is less than all past realizations of the state variable up to period $t-1$. Furthermore, if $y_t \geq -\underline{k}_t$ then $\underline{k}_{t+1} = \underline{k}_t$, while if $y_t < -\underline{k}_t$ then $-\underline{k}_{t+1} \in (-\underline{v}, y_t)$ so that the lower bound for the subsequent period, $-\underline{k}_{t+1}$, is updated if and only if the subsequent observation y_t falls below $-\underline{k}_t$. So $\underline{k}_t$ is a parameter at date t and depends upon information available up to date $t-1$. A similar procedure is adopted to ensure that the forecast for Y does not exceed $\overline{v}$ in case $\overline{v}$ is finite.

---

[19]The material presented here is based on Chatterji and Chattopadhyay (2002).

Here we truncate (13) at $\bar{k}_t$ where $0 < \bar{k}_t < \bar{v}$ and $y_j \leq \bar{k}_t$ for $j = t-1, t-2, ., 0, .., -L$. Here too $\bar{k}_t$ is a parameter at date t which is updated analogously to some quantity in $(y_t, \bar{v})$ if and only if y_t exceeds $\bar{k}_t$. Formally

$$y_{t+1}^e = \widetilde{y}_{t+1}^e \quad if \quad \widetilde{y}_{t+1}^e \in [-\underline{k}_t, \bar{k}_t]$$
$$y_{t+1}^e = -\underline{k}_t(\bar{k}_t) \quad if \quad \widetilde{y}_{t+1}^e < -\underline{k}_t(\widetilde{y}_{t+1}^e > \bar{k}_t)$$

The following is a heuristic description of how a GSF arises in this set up and implies that $y_t \nrightarrow -\underline{v}(\bar{v})$. The hypothesis $y_t \rightarrow -\underline{v}$ implies that the sequence γ_t tends to 1 and the sequence y_{t+1}^e tends to $-\underline{v}$. The case $y_t \rightarrow -\underline{v}$ can only arise if $\lim_{y^e \rightarrow -\underline{v}} g(y^e) = -\underline{v}$. It is convenient to extend the map g to $-\underline{v}$ by continuity so as to obtain the limiting fixed point $g(-\underline{v}) = -\underline{v}$. The derivative of the map g evaluated at this limiting fixed point cannot be lesser than 1 in modulus since the primitive TEM f is assumed to be contracting for large x. (Infinity is therefore repelling in the dynamics induced by f and since f and g are topological conjugates, $-\underline{v}$, the image of infinity under the change of variable h, cannot be attracting under g.) This observation, in conjunction with the mild that regularity condition $\lim_{y \rightarrow -\underline{v}(\bar{v})} g'(y)$ exists and is not equal to 1, implies that $-\underline{v}$ is a locally repelling fixed point of the dynamics defined by g, and this rules out convergence to $-\underline{v}$. We emphasize that the formal proof of this result is independent of the specific choice of the bounds on the forecasting procedure.

Proposition 8. *Assume f satisfies assumption 7 and in addition the following contracting property for large $|x|$: $|f(x)| \leq \bar{a}|x|$ for some $0 < \bar{a} < 1$ and $|x|$ large. If h satisfies assumption 8 with $h(x) \rightarrow -\underline{v}(\bar{v})$ as $x \rightarrow \infty$, and the regularity condition $[g(y) \rightarrow -\underline{v}$ or $g(y) \rightarrow \bar{v}$ as $y \rightarrow -\underline{v}(\bar{v})] \Rightarrow \lim_{y \rightarrow -\underline{v}(\bar{v})} g'(y)$ exists and is not equal to 1 holds, then the OLS learning dynamics formulated on $y(= h(x)) \in V$ satisfies the property that for every sequence generated by the learning dynamics, $y_t \nrightarrow -\underline{v}(\bar{v})$, or equivalently, $x_t(= h^{-1}(y_t)) < \widetilde{Q}$ infinitely often.*

A GSF as defined in Section 2, where there are no bounds on the domain or the range of the TEM, works via non linearities that do not allow the dynamics to sustain arbitrarily high parameter estimates. In this setting with bounds, the regularity condition causes a fixed point of the form $-\underline{v}$ or $\bar{v}$ to be repelling and therefore does not allow the dynamics to sustain the parameter estimate of 1 in the limit. While it might appear that a GSF appears here under a mild regularity condition, this added generality is what one gains for making the more restrictive hypothesis that regressions are performed on a bounded state variable, which effectively ties down parameter estimates as the state variable approaches the boundary of the set.

We finally sketch the details of Example 1 to illustrate the workings of a model where the primitive state variable has a positivity constraint.

Example 1 continued. *An Overlapping Generations Economy.*

Consider a standard version of the Overlapping Generations Model model in which one good is available in every period, there is a fixed stock M of money and one agent is born every period and lives for two periods. We consider the learning dynamics when the utility function of a young agent is given by $u(c_1, c_2) = \log c_1 c_2$. Let p_c and p_m be the prices of the consumption good and money respectively. In order to decide how much to consume in the first period of life, a young agent has to forecast the relative price for the next period. Assume first that agents forecast the nominal price level $X_t \equiv p_c/p_m$. The TEM in the nominal price level turns out to be affine and is expressed as

$$X_t = \bar{a} X_{t+1}^e + \bar{c} \tag{14}$$

where $X_t > 0$ and $0 < \bar{a} < 1$ and $\bar{c} > 0$, and the Golden Rule steady state value is given by $X^* = \bar{c}/(1 - \bar{a})$. Since this steady state will be assumed known to agents, it is appropriate to represent the TEM in deviations $x (\equiv X - X^*)$ from the steady state as

$$x_t = f(x_{t+1}^e) \equiv \bar{a} x_{t+1}^e \tag{15}$$

where $x_{t+1}^e \in D \equiv (-X^*, \infty)$.

Now assume that agents forecast the real balance $Y \equiv p_m/p_c$ instead of the nominal price. The TEM when expressed in the variable Y is

$$Y_t = \frac{Y_{t+1}^e}{\bar{a} + \bar{c} Y_{t+1}^e} \tag{16}$$

with $Y^* = X^{*-1} = \bar{c}^{-1}(1 - \bar{a})$. Note that now 0 is a fixed point of (16) even though its inverse image under H, namely $+\infty$, is not a fixed point of (14); denote this additional fixed point as $\overline{Y}$. Here too it is appropriate to represent the TEM in deviations $y (\equiv Y - Y^*)$ from the steady state as

$$y_t = g(y_{t+1}^e)$$

where $y_{t+1}^e \in D \equiv (-Y^*, \infty)$.

Notice that x and y are linked via the change of variable $y = h(x)$ where h is defined via the identity $Y = h(X - X^*) + Y^* = H(X)$ with $Y = H(X) = 1/X$. The two TEMs $f(x_{t+1}^e)$ and $g(y_{t+1}^e)$ are accordingly linked via the identity $h(f(h^{-1}(y_{t+1}^e))) \equiv g(y_{t+1}^e)$. Note that as $x \to \infty, h(x) \to -Y^*$. One extends the map g to $-Y^*$ by continuity, wherein $-Y^*$ becomes a fixed point of g even though $+\infty$, is not a fixed point of (15); also $g\prime(-Y^*) = 1/\bar{a} > 1$.

Now suppose agents generate the forecast x_{t+1}^e via OLS on lagged values of x. Then one has $x_{t+1}^e = \beta_{t-1}^2 x_{t-1}$ and so $x_t = \bar{a}\beta_{t-1}^2 x_{t-1}$. For $x > 0$ and β_{t-1}^2 sufficiently large, one has $x_t/x_{t-1} > \bar{a}\beta_{t-1}^2$. Consequently, (see Marcet and Sargent (1989) and Grandmont (1998)), β_{t-1}^2 forms an increasing sequence while $x_t \to \infty$. Thus the formulation of the learning

dynamics in this case is compatible with arbitrarily large parameter estimates along with the state variable diverging to infinity, which here implies that consumption goes to autarky. When agents forecast the variable y instead via OLS so that $y_{t+1}^e = \gamma_{t-1}^2 y_{t-1}$, Proposition 8 applies and the fixed point $-Y^*$ that corresponds to autarky turns out to be *unstable* under the learning dynamics, so that $y_t \nrightarrow -Y^*$. In fact, one can show that the learning dynamics converge to steady state $y = 0$ for all initial conditions so that one has in this case global convergence to the golden rule steady state.[20] ▲

4. Proofs.

Proof of Lemma 2. Assume without loss of generality that h is increasing and suppose, for the moment, that $y > 0$. Then, using that $h(0) = 0$, one gets $h^{-1}(y) > 0$. The contracting condition $|f(x)| < a|x|$ for $0 < a < 1$ implies that $|f(h^{-1}(y))| < ah^{-1}(y)$. Since h is increasing, one obtains $h(|f(h^{-1}(y))|) < h(ah^{-1}(y)) < y$, since, using that $a < 1$, $h(ah^{-1}(y)) < h(h^{-1}(y)) = y$. Then, $\frac{|g(y)|}{y} < 1$ using that $|g(y)| < h(|f(h^{-1}(y))|)$. By symmetry, an identical argument applies for $y < 0$, and one obtains thereby the inequality $\frac{|g(y)|}{|y|} < 1$ as required. The case where h is decreasing is treated analogously. One thus has in all cases $|g(y)| < |y|$ for all $y \neq 0$. Since f and g are topologically conjugate, the dynamics around 0 are equivalent. Since f is a contraction, $|f\prime(0)| < 1$, this implies that $|g\prime(0)| \ngtr 1$. The regularity condition then implies that $|g\prime(0)| < 1$. This in conjunction with the fact that $|g(y)| < |y|$ for all y implies the existence, for each $[-\mathcal{Q}, \mathcal{Q}]$, of a constant $a(\mathcal{Q})$ satisfying $0 < a(\mathcal{Q}) < 1$, such that $|g(y)| < a(\mathcal{Q})|y|$ for all $y \in [-\mathcal{Q}, \mathcal{Q}]$. ▲

Proof of Proposition 7. As a first step, we note the fact summarized in Lemma 3, the proof of which is a straightforward adaptation of the proof of Proposition 1 of Chatterji and Chattopadhyay (2000) and is accordingly omitted.

Lemma 3. *Consider a TEM $g : R \rightarrow R$, with $g(0) = 0$ that satisfies the following Lipschitz condition: for every positive number L, there exists a positive constant $a(L)$ such that $|g(y)| \leq a(L)|y|$ for all $|y| < L$. Suppose that for some $n, |\gamma_n|$ and $|y_n|$ are such that $|\gamma_n| < \frac{1}{a(Q)}$ and $|\gamma_n|^2 < \frac{1}{a(Q)}$ where $Q > \max\{|\gamma_n^2 y_n|, |y_n|\}$. Then, for $t \geq n$, the sequence $|\gamma_n|$ is decreasing and $(y_t, \gamma_t, \omega_t) \rightarrow (0, \overline{\gamma}, \overline{\omega})$ with $|\overline{\gamma}| \leq 1$.*

By Proposition 1, we know that there exists a constant $\overline{\mathcal{Q}}$, one has that $|y_n| < \overline{\mathcal{Q}}$ infinitely often. Pick $\mathcal{Q} > \overline{\mathcal{Q}}$. By assumption, $a(\mathcal{Q}) < a(\mathcal{Q})^{\frac{1}{2}} < 1$ and therefore $1 < \frac{1}{a(\mathcal{Q})^{\frac{1}{2}}} < \frac{1}{a(\mathcal{Q})}$. Select $\delta > 0$ such that $\mathcal{Q} > (1 + \delta)^2 \overline{\mathcal{Q}}$ and $(1 + \delta) < \frac{1}{a(\mathcal{Q})^{\frac{1}{2}}}$. Here too, as in step *(i)* of the proof of Proposition 1, Lemma 1 implies that $|\gamma_n| < 1 + \delta$ infinitely often . So eventually

---

[20]This result is valid for a class of preferences in a set up where (i) the configuration of endowments and preferences is such that, at constant prices, the agent wishes to transfer income from youth to old age, i.e., the Samuelson case; this ensures that equilibria in which fiat money has a positive exchange value exist; and, (ii) consumption when old is a normal good. The discussion in section 2 applies, and there is always local instability for an open set of initial conditions. For details see Chatterji and Chattopadhyay (2002).

for some n, one gets $|\gamma_n| < \frac{1}{a(\mathcal{Q})^{\frac{1}{2}}} < \frac{1}{a(\mathcal{Q})}$ for $\mathcal{Q} > \max\{|\gamma_n{}^2 y_n|, |y_n|\}$, and by Lemma 3, $(y_t, \gamma_t, \omega_t) \to (0, \overline{\gamma}, \overline{\omega})$ *with* $|\overline{\gamma}| \leq 1$ *for* $t \geq n$. ▲

Proof of Proposition 8. The regularity conditions imply that $g\prime(-\underline{v})(\equiv \lim_{y \to -\underline{v}} g\prime(y)) > 1$. This implies that $-\underline{v}$ is a locally repelling fixed point of the iterated dynamics defined by the map g. The hypothesis $y_t \to -\underline{v}$ implies that the sequence $\gamma_t \to 1$ and the sequence $y_{t+1}^e \to -\underline{v}$. Since y_t^e converges to $-\underline{v}$, one has (from the fact that $-\underline{v}$ is a locally repelling fixed point of the dynamics defined by g) that $y_t > \alpha' y_{t+1}^e > y_{t+1}^e$ for $0 < \alpha' < 1$ eventually. So, if $y_{t+1}^e \geq -\underline{k}_t$, then $y_t \geq -\underline{k}_t$ also. It follows that $\underline{k}_{t+1} = \underline{k}_t$ and since γ_t converges to 1, one also has $\tilde{y}_{t+2}^e = \gamma_t^2 y_t > y_{t+1}^e \geq -\underline{k}_t$ for t sufficiently large. Hence $y_{t+2}^e = \tilde{y}_{t+2}^e$, so that $y_{t+2}^e \geq -\underline{k}_t$. An induction argument lets us contradict the hypothesis that y_{t+1}^e converges to $-\underline{v}$ and this rules out $y_t \to -\underline{v}$. A similar argument rules out case $y_t \to \overline{v}$. ▲