No. 2003/03

Learning and Equilibrium Selection in a Monetary Overlapping Generations Model with Sticky Prices

Klaus Adam

Learning and Equilibrium Selection in a Monetary Overlapping Generations Model with Sticky Prices

Klaus Adam*

November 2002

Abstract:
We study adaptive learning in a monetary overlapping generations model with sticky prices and monopolistic competition for the case where learning agents observe current endogenous variables. Observability of current variables is essential for informational consistency of the learning setup with the model set up but generates multiple temporary equilibria when prices are flexible and prevents a straightforward construction of the learning dynamics. Sticky prices overcome this problem by avoiding simultaneity between prices and price expectations. Adaptive learning then robustly selects the determinate (monetary) steady state independent from the degree of imperfect competition. The indeterminate (non-monetary) steady state and non-stationary equilibria are never stable. Stability in a deterministic version of the model may differ because perfect foresight equilibria can be the limit of restricted perceptions equilibria of the stochastic economy with vanishing noise and thereby inherit different stability properties. This discontinuity at the zero variance of shocks suggests to analyze learning in stochastic models.

*Klaus Adam, University of Frankfurt, Mertonstr.17, PF 94, 60054 Frankfurt, Germany, and Research Fellow at the Center for Financial Studies, E-mail: kladam@wiwi.uni-frankfurt.de

1   Introduction

Monetary macroeconomic models often suffer from multiplicity of rational expectations equilibria. Overlapping generations (OLG) models with money are perhaps the best-known example where rational expectations do not lead to a unique equilibrium prediction. Typically OLG models possess both a determinate monetary steady state and a continuum of paths converging to the indeterminate non-monetary steady state.[1]

To sharpen the predictions of rational expectations it is - by now - common practice in the literature to study the stability of equilibria under adaptive learning schemes (Marcet and Sargent (1989a), Evans and Honkapohja (2001)). Consistent with experimental evidence, e.g. Marimon and Sunder (1993), equilibria whose expectations can be acquired via simple learning rules are widely believed to constitute more plausible model predictions than equilibria that would require more sophisticated coordination devices.

Surprisingly, the stability properties of rational expectations equilibria in monetary overlapping generations models is still an open issue. More precisely, the literature suggests that the stability properties of rational expectations equilibria depend on the information set available to agents and in particular on whether agents use current or lagged prices to update their inflation expectations.

The literature on learning in OLG models has for the most part assumed that agents use only lagged prices to update inflation expectations. With this assumption it seems to be a robust finding that the determinate monetary steady state is the only stable equilibrium under adaptive, least-squares, or related learning schemes. This was first argued by Marcet and Sargent (1989b). Arifovic (1995) showed the result for genetic algorithm learning, and Evans, Honkapohja, and Marimon (2001) for heterogenous learning rules.

However, a number of recent papers find that stability properties can be reversed once agents use current period variables to update expectations (Duffy (1994), Evans and Honkapohja (2001, section 3.4), Bullard and Mitra (2002)). In particular, Van Zandt and Lettau (2002) have shown that use of current prices in an OLG model can result in instability of the determinate and stability of the indeterminate steady state.

---

[1]There may also exist equilibrium cycles and sunspot equilibria, which have been analyzed by Grandmont (1985) and Woodford (1990), respectively.

1

It is important to allow agents to use current prices to update expectations to insure that the information structure for the learning setup is consistent with that of the model. Since agents in OLG models choose the *real* value of money balances, they necessarily know current prices at the time they make their money demand decisions. Real money demand, however, depends on expected inflation and it seems unlikely that agents would not use the same price information to update their inflation expectations. This has not been recognized before.

These arguments suggest that stability of the determinate monetary steady state, as found for the case of lagged price information, may rest upon an inconsistent informational setup. However, the results of Van Zandt and Lettau (2002) for the case of current price information are difficult to interpret. Agents' use of current prices introduces simultaneity between prices and expectations and thereby typically generates multiple market clearing price-expectations pairs. This multiplicity requires auxiliary selection assumptions to be able to construct the learning dynamics, as explained further in section 2, and causes difficulties in understanding the precise source of the stability reversal.

The main objective of this paper is to resolve the issue of stability of rational expectations equilibria under adaptive learning in monetary overlapping generations models when agents can condition expectations on current prices. To do so it is proposed to consider the competitive limit of a model with sticky prices and monopolistic competition instead of the customary competitive model with flexible prices.

Sticky prices are essential to improve upon the shortcomings of previous selection attempts. When prices are sticky agents can use theses prices to update their expectations without creating simultaneity between these variables. Sticky prices thereby allow for an informationally consistent setup but avoid multiplicity of temporary equilibria.

The introduction of sticky prices is also of interest in its own because sticky prices are widely believed to be an important ingredient of empirically plausible macroeconomic models (e.g. Galí and Gertler (1999)). Yet, their implications for equilibrium selection have not been analyzed so far.

Besides extending the stability analysis to the case of sticky prices, the paper also analyzes stability of (non-stationary) equilibria near the steady states, which have not been studied previously, and clarifies the relationship between learning in nonstochastic and stochastic versions of the model.

In summary, the results are the following. Most importantly, adaptive learning

schemes robustly select the same equilibrium independent of the degree of imperfect competition. Furthermore, the determinate monetary steady state is the only learnable equilibrium.[2] This confirms the early findings of Marcet and Sargent (1989b) and suggests that equilibrium selection in monetary OLG models does not depend on the agents' information set.

The results are derived in several steps. The paper first analyzes a deterministic model and studies the learnability of the monetary and non-monetary steady states under constant and decreasing gain learning rules. While the indeterminate non-monetary steady state is always unstable, the determinate monetary steady state is locally stable if the elasticity of labor supply is larger than one-half but unstable otherwise.

The paper then analyzes least squares learning in a stochastic environment and shows that the (stochastic) low inflation steady state is always locally stable independently of the labor supply elasticity. The indeterminate high inflation steady state and the non-stationary equilibria are always unstable under least-squares learning.

Although the deterministic model suggests that the elasticity of labor supply has an important impact on the stability of the monetary steady state, the stochastic model shows that its role arises only due to a singularity in the transition from a stochastic to a deterministic environment.

In particular, I find that the perfect foresight solution of the deterministic economy is of a simpler structure than the rational expectations solution in the stochastic economy even as the support of the noise becomes arbitrarily small.[3] This implies that agents in a stochastic environment must learn a different equilibrium law of motion to acquire rational expectations, which leads to different stability conditions.

The elasticity condition of the deterministic setup is found to correspond to the stability condition of a 'restricted perceptions equilibrium' in the stochastic economy where agents underparameterize the economy's law of motion. Thus, the perfect foresight equilibrium of the deterministic model can be interpreted as the limit of a restricted perceptions equilibrium of the stochastic model with vanishing noise.

This suggests to analyze learning in stochastic models and complements the argument in favor of stochastic models made by Evans and Honkapohja (1998) who showed

---

[2]Some qualifications apply as described later in the introduction.

[3]The equilibrium laws of motion differ because in a deterministic equilibrium some variables settle down to constant values while they keep on fluctuating in the stochastic equilibrium, which requires that agents condition their forecasts permanently on these variables.

that discontinuities can arise because some of the serial correlation properties of the data may disappear when shocks have zero variance.

The paper is organized as follows. Section 2 briefly shows how learning leads to multiple market clearing prices in flexible price models when agents' expectations depend on current prices and how sticky prices help to avoid this problem. Section 3 then introduces the OLG model with sticky prices and imperfect competition. The model's perfect foresight equilibria are determined in section 4 where it is also shown that these equilibria approach the ones of the flexible price model as the degree of imperfect competition vanishes. Section 5 analyzes learning in the deterministic model and section 6 considers a stochastic version of the model. Most technical details and proofs can be found in the appendix.

2 The Multiplicity Problem

As argued in the introduction, an informationally consistent setup requires that learning agents can use current prices to update their inflation expectations. This section shows that use of current prices easily leads to multiple market clearing prices and prevents a straightforward construction of the learning dynamics. It is also shown how sticky prices can be used to overcome the problem.

To illustrate the potential for multiplicity I assume simple money demand and supply functions and an equally simple learning rule. Let *real* money demand be given by

$$m_t^d = a - b \cdot \Pi_{t+1}^e \qquad a > b > 0 \tag{1}$$

where Π_{t+1}^e denotes current expectations of the future inflation factor. Money demand functions of this form can be derived from OLG models with money and also feature prominently in monetary models (Cagan (1956), Sargent and Wallace (1987)).

Furthermore, let *real* money supply be given by

$$m_t^s = \frac{\lambda m_{t-1}}{\Pi_t} + g \tag{2}$$

where $\lambda > 0$ denotes the nominal money growth factor, $g \geq 0$ the amount of real seignorage, m_{t-1} the previous period's stock of real balances, and Π_t the inflation factor. The supply rule nests nominal money growth rules and real seignorage rules and is a (hyperbolically) decreasing function of current inflation.

In models with learning, the expectations in the demand function (1) are determined through an explicit updating mechanism. Consider the following rule which has been widely used in the literature:

$$\Pi_{t+1}^e = \Pi_t^e + \gamma(\Pi_t - \Pi_t^e) \qquad (3)$$

The parameter $\gamma > 0$ determines how fast expectations are updated in response to past forecast errors. Note that agents use current inflation to update, as required for an informationally consistent setup. As a result, current inflation affects current expectations and money demand will decrease (linearly) with current inflation.

Figure 1 depicts money demand and supply when expectations are given by learning rule (3). The figure shows that market clearing prices are generically non-unique, given they exist.[4]

Multiplicities, as the one illustrated in figure 1, create problems for constructing the dynamics of prices and beliefs under learning since they require to select between different market clearing prices. Such a selection is not innocent because it affects current expectations and, via the learning rule, the whole path of future prices. It is then difficult to understand whether the learning rule or the temporary equilibrium selection ultimately drives the stability properties of equilibria.

To avoid the multiplicity problem Grandmont (1985) has added a restriction on the expectations functions in his analysis of the standard OLG model. However, this restriction does not hold for learning rule (3) and the remaining learning rules considered in this paper.

Instead of restricting the learning rule, I propose to introduce sticky prices in the form of monopolistically-competitive entrepreneurs who set prices one period in advance. If entrepreneurs do not know the current price level at the time they set their own prices, then their expectations are determined by history and the price setting problem has a unique solution. Once prices are predetermined, expectations entering the money demand function can safely depend on current prices without creating

---

[4]Existence is guaranteed for all $\gamma > 0$ when $g > 0$ is not too large, λ not too different from 1, and $1 < \Pi_t^e < \frac{a}{b}$. This follows from the fact that at $\Pi_t = \Pi_t^e$: $m_t^s \approx \frac{m_{t-1}}{\Pi_t^e} < m_t^d = m_{t-1}$ where the equality follows from the learning rule (3) and market clearing in t-1.
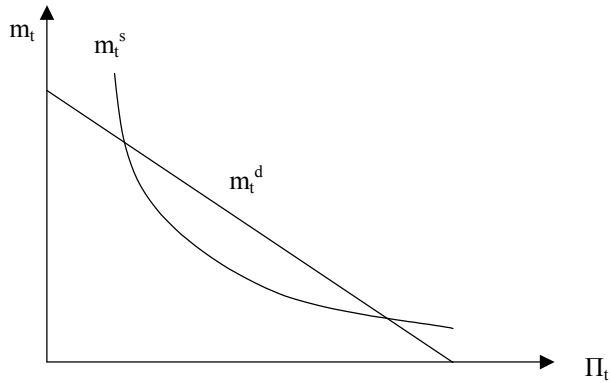
Figure 1: Money Market Clearing

multiplicity.[5] The construction of learning dynamics is then straightforward.

Although sticky prices insure uniqueness of temporary equilibria under learning they do not insure uniqueness when expectations are rational. In rational expectations equilibrium current prices are functions of entrepreneurs' information set and these functions are known. Therefore, sticky price models suffers from the same kind of indeterminacy of rational expectations equilibria as flexible price models. This has the advantage that results for the sticky price model remain easily comparable with the flexible price setup.

3 An OLG Model with Sticky Prices

I consider a simple overlapping generations model with production where each generation of agents lives for two periods - works when young and consumes when old- and may transfer wealth across time via fiat money. There is also an infinitely lived government that finances a constant real deficit through seignorage. This setup generates an environment closely related to the seignorage model of Sargent and Wallace (1987).

In each time period a new generation of agents is born. In contrast to standard models agents of a given generation are either born as workers or entrepreneurs with a

---

[5]Since prices are preset (money and labor) markets will be cleared by the real wage. Provided money demand is monotone in real wages, there is a unique market clearing real wage. Real wages do not show up in the money demand function of the flexible price economy because we assume a production technology that is linear in labor, which implies that real wages are constant.

unit mass of each of them in every new generation.[6]

Workers are homogeneous and offer their labor force at a competitive labor market in return for a wage income.

Entrepreneurs are in monopolistic competition with each entrepreneur $i \in [0, 1]$ producing a good q^i which is an imperfect substitute in the production of the aggregate consumption good c (Dixit and Stiglitz (1977)):

$$c = \left(\int_{i \in [0,1]} \left(q^i \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \quad \text{with } 1 > \sigma \geq 0$$

With $\sigma = 0$ goods are perfect substitutes and entrepreneurs are in perfect competition. When $\sigma > 0$ goods are imperfectly substitutable and entrepreneurs will earn monopolistic rents.

The timing of events is as follows. At the beginning of each period, young entrepreneurs commit to a price at which they are willing to sell the product. Then old agents, i.e. old workers and old entrepreneurs, spend all their money holdings to order goods. At the same time, the government orders goods for government consumption. Firms accept any amount of orders at the price they posted and then hire the work force that is necessary to produce the ordered quantities. The labor market clears and production takes place. Young workers are paid their wage, young entrepreneurs retain their profits, and the produced goods are delivered for consumption to the old agents and the government. Then a new period starts.

Let P_t^i denote the price posted for good i at the beginning of period t and P_t the Dixit-Stiglitz price index.[7] Furthermore, let M_t denote the stock of nominal money at the end of period t. The real value of outstanding money balances $m_t = \frac{M_t}{P_t}$ evolves according to

$$m_t = \frac{m_{t-1}}{\Pi_t} + g \tag{4}$$

where $\Pi_t = \frac{P_t}{P_{t-1}}$ is the inflation factor from $t-1$ to t, and $g \geq 0$ denotes real seignorage revenue raised for government consumption.[8]

---

[6] Any numbers could be chosen as long as they are constant through time.

[7] We have $P_t = \left(\int_{[0,1]} \left(P_t^i \right)^{\frac{\sigma-1}{\sigma}} di \right)^{\frac{\sigma}{\sigma-1}}$.

[8] Government consumption is assumed to be either wasteful or to generate utility that is separable from that of private consumption and leisure.

3.1 Workers

The representative worker who is born in period t maximizes

$$\max_{n_t, c_{t+1}^w} E_t^* \left[-v(n_t) + u(c_{t+1}^w) \right]$$

subject to

$$m_t^w = n_t w_t \tag{5}$$

$$c_{t+1}^w \leq \frac{m_t^w}{\Pi_{t+1}}$$

where n_t denotes hours of work, c_{t+1}^w consumption in $t+1$, w_t the real wage (in terms of period t consumption), and m_t^w the worker's end of period t real money holdings. Utility functions are assumed strictly concave and twice continuously differentiable.

Note that workers maximize utility with respect to some (potentially) subjective expectations operator E_t^* that is based on the information set

$$H_t = \sigma(P_t, P_{t-1}, \ldots; m_t, m_{t-1}, \ldots)$$

which contains past and current values of prices and real balances. Workers' expectations therefore depend on current prices, as required for an informationally consistent setup.

The first order conditions of the utility maximization problem implicitly define the workers' labor supply as a function of the current real wage and expected inflation:

$$n(w_t, E_t^*(\Pi_{t+1})) \tag{6}$$

Alternatively, the first order conditions define a real wage function

$$w(n_t, E_t^*(\Pi_{t+1})) \tag{7}$$

that determines the real wage that has to be paid to induce the representative worker to supply n_t units of labor when her inflation expectations are given by $E_t^*(\Pi_{t+1})$. The real wage function is continuously differentiable for $E_t^*(\Pi_{t+1}) > 0$ at all feasible levels n_t.

The labor supply function (6), or the real wage function (7), summarizes workers' optimal economic decisions. To insure that real money demand is downward sloping in expected inflation and to insure the existence of a monetary and a non-monetary steady state for small enough levels of government seignorage, one can impose the following conditions on the labor supply function (6):

Condition 1   *1. Labor supply increases in w_t and decreases in Π^e_{t+1}*

2. Labor supply is positive when both $w_t = 1 - \sigma$ and $\Pi^e_{t+1} = 1$

3. Labor supply is zero when $w_t = 1 - \sigma$ for some $\Pi^e_{t+1} < \infty$, where $\Pi^* < \infty$ denotes the lowest inflation rate (the infimum) for which this is the case.

A sufficient condition for the first requirement to hold is that the coefficient of relative risk aversion of the utility function $u(\cdot)$ is larger than one such that the substitution effect dominates the income effect. The second and third requirement hold whenever $v'(0) < u'(0) < \infty$, i.e. whenever the marginal utility of consumption is sufficiently high but bounded.

3.2   Entrepreneurs

At the beginning of each time period young entrepreneurs simultaneously decide about their prices. Profit maximization by the competitive sector implies that the demand curve faced by entrepreneur j is given by

$$q^j_t = m_t \left(\frac{P_t}{P^j_t} \right)^{\frac{1}{\sigma}} \tag{8}$$

With the production technology being linear in labor, each entrepreneur j maximizes

$$\max_{P^j_t} E^*_{t-1} \left[m_t \left(\frac{P_t}{P^j_t} \right)^{\frac{1}{\sigma}} \left[P^j_t - P_t w_t \right] \right] \tag{9}$$

Note that entrepreneurs maximize with respect to some (potentially) subjective expectations operator E^*_{t-1} which is based on an information set that does not contain time t variables.

The first order condition of (9) determines the profit maximizing price:

$$P^j_t = \frac{1}{1 - \sigma} P_{t-1} E^*_{t-1} (\Pi_t w_t) \tag{10}$$

As usual, optimal prices are a mark-up over expected marginal costs where the mark-up factor depends on the degree of imperfect competition σ. Equation (10) summarizes profit maximization behavior of entrepreneurs.

3.3 The Role of Subjective Expectations

The behavior of workers, entrepreneurs, and the government can be summarized by equations (7), (10), and (4), respectively. When expectations are potentially non-rational, as in the setup of this model, the nature of the non-rationality may determine how these equations can be combined with each other.

For example, when entrepreneurs' wage expectations in equation (10) are inconsistent with their inflation expectations, then it would not be possible to use the real wage function (7) to express wage expectations as a function of the expected inflation path.[9] To insure that the model equations can be used as it is the case when expectations are objectively rational, I have to impose restrictions on agent's subjective expectations. These restrictions are in several ways quite strong, though of course the standard 'rational expectations' makes even stronger assumptions.

Firstly, it is assumed that entrepreneurs and workers hold the same inflation expectations whenever they are given the same information set.[10] Secondly, entrepreneurs' forecasts of future wages and inflation rates are assumed to be always consistent with the model structure, in particular with workers' labor supply function and the economy's production function. Thirdly, expectations are supposed to obey the law of iterated expectations, i.e. $E_{t-1}^* [E_t^* [\cdot]] = E_{t-1}^* [\cdot]$, which says that agents expect their forecasts to be unbiased as they do not expect forecasts to move in predictable ways as time proceeds. Finally, expectations are point expectations.[11]

Using the previous assumptions one can express entrepreneurs' wage expectations as a function of expected future inflation rates and known variables, even when expectations are potentially non-rational. With equation (7) entrepreneurs' wage forecasts are then given by

$$E_{t-1}^* [w_t] = w(E_{t-1}^* [n_t], E_{t-1}^* [\Pi_{t+1}]) \tag{11}$$

---

[9]Inconsistent means that entrepreneurs expect a different wage than the one implied by the model if their inflation expectations were to become true.

[10]Heterogeneous expectations in a monetary hyperinflation model have been studied in Evans, Honkapohja and Marimon (2001).

[11]Expectations can also be interpreted as mean expectations once I consider the linearized version of the model.

From the linearity of the production function and equation (8) follows that labor demand n_t can be expressed as

$$n_t = \int q_t^i \mathrm{di} = m_t \left(P_t\right)^{\frac{1}{\sigma}} \left[\int \left(P_t^i\right)^{-\frac{1}{\sigma}} \mathrm{di}\right] = m_t \tag{12}$$

where the last equality follows from the fact that entrepreneurs set identical prices because they hold identical expectations. Combining this result with equations (4) and (11) delivers

$$E_{t-1}^* \left[w_t\right] = w \left(\frac{m_{t-1}}{E_{t-1}^* \left[\Pi_t\right]} + g, E_{t-1}^* \left[\Pi_{t+1}\right]\right) \tag{13}$$

Expected real wages are now a function of known variables and expected future inflation rates. By equation (10) the same holds for profit maximizing prices.

Using result (13) all relevant equations, i.e. (4), (7), and (10), now involve only inflation expectations. Therefore, as is the case with the flexible price economy, a single learning rule for inflation expectations is sufficient to analyze the model's learning dynamics.

4 Perfect Foresight Equilibria

This section characterizes the model's perfect foresight equilibria (PFE) and shows that the equilibria of the sticky price economy approach the ones of the flexible price economy as the degree of imperfect competition σ vanishes.

Equation (10) together with the fact that all firms charge the same price implies that current inflation can be expressed as

$$\Pi_t = \frac{1}{1 - \sigma} E_{t-1}[\Pi_t w_t] \tag{14}$$

Under perfect foresight Π_t is part of the $t - 1$ information set. Therefore, the price setting equation only determines the equilibrium real wage but not the current price level, i.e.

$$w_t = 1 - \sigma \tag{15}$$

When combining this with equations (4) and (12) one obtains a single equation characterizing PFE:

$$n(1 - \sigma, \Pi_{t+1}) = \frac{n(1 - \sigma, \Pi_t)}{\Pi_t} + g \tag{16}$$
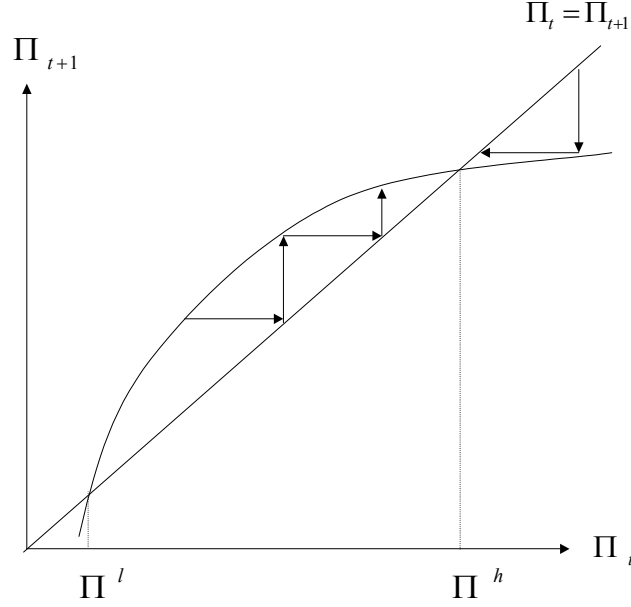
11

Figure 2: Perfect Foresight Paths

PFE can be generated by choosing an initial inflation rate Π_0 and iterating equation (16) forward. Figure 2 depicts the map from Π_t to Π_{t+1} implied by equation (16) when the level of government seignorage is given by $g = 0$.

Condition 1 guarantees existence of two steady states. There is a low inflation steady state (Π^l, m^l) where money is valued and inflation is equal to zero

$$\Pi^l = 1 \quad and \quad m^l = n(1 - \sigma, \Pi^l) > 0$$

and a high inflation steady state (Π^h, m^h) where money is not valued:[12]

$$\Pi^h = \Pi^* \quad and \quad m^h = n(1 - \sigma, \Pi^h) = 0$$

Since $\frac{\partial \Pi^h(g)}{\partial g} < 0$ at $g = 0$ money starts to become valued in the high inflation steady state for small positive levels of seignorage g.

When applying the implicit function theorem to equation (16), one finds that $\frac{d\Pi_{t+1}}{d\Pi_t} > 1$ at the low inflation steady state and $\frac{d\Pi_{t+1}}{d\Pi_t} < 1$ at the high inflation steady state, which generates the graph depicted in figure 2. It implies that there exist initially non-stationary PFE where inflation asymptotically approaches Π^h, as illustrated in the figure.

---

[12]The 'inflation rate' Π^* reported for this steady state is the limit of the inflation rates as $g \to 0$.

Now consider the relation to the PFE in the flexible price economy. In a flexible price economy PFE are characterized by an equation similar to (16) but with the equilibrium real wage given by $w_t = 1$, which follows from the linearity of the production function. Since $n(\cdot, \cdot)$ is continuously differentiable and since the inflation rates in a PFE must lie in the bounded interval $[0, \Pi^h]$, the PFE of the sticky price economy approach the ones of the flexible price economy as the degree of imperfect competition σ approaches zero.[13]

5 Stability of Steady States under Adaptive Learning

This section analyzes the stability of the steady state equilibria under adaptive learning schemes. The analysis of non-stationary equilibria is deferred to section 6.

Assume agents forecast future inflation rates according to the following adaptive rule:

$$_t\Pi^e_{t+1} = {}_{t-1}\Pi^e_t + \gamma_t(\Pi_t - {}_{t-1}\Pi^e_t) \tag{17}$$

Left-hand side subscripts denote the time at which expectations are formed and right-hand side subscripts the date for which the indexed variable is forecasted. Rule (17) says that the new inflation forecast is equal to the previous forecast plus γ_t times the latest forecast error, which is given by the term in the brackets. Note that the current inflation rate Π_t enters into the forecast made at time t, as required.

The gain parameter $\gamma_t \in (0, 1)$ determines how fast expectations adapt in response to forecast errors. Two kinds of gain sequences will be considered: constant gain learning rules where $\gamma_t = \gamma$ and decreasing gain learning rules where $\gamma_t \to 0$ with $\sum \gamma_t = \infty$. The motivation for considering constant and decreasing gain rules is that Van Zandt and Lettau (2002) have reported that stability may depend on which kind of gain sequence is chosen.

An alternative way to express learning and forecasting rule (17) is to say that agents

---

[13] More precisely: for any initial value Π_0 let Π^σ_t denote the path of inflation generated by iterating (16) forward with $w_t = 1 - \sigma$, then for any $\varepsilon > 0$ there exists a $\sigma > 0$ such that for all Π_0 : $\sup_t |\Pi^\sigma_t - \Pi^0_t| < \varepsilon$.

perceive the inflation process to be given by

$$\Pi_t = a + \varepsilon_t \tag{18}$$

where a is the (unknown) steady state inflation rate and ε_t a white noise shock. Equation (18) is usually referred to as agents' Perceived Law of Motion (PLM). Equation (17) then implies that agents estimate the parameter a according to

$$\widehat{a}_t = \sum_{j=0}^{t-1} \left(\left(\prod_{i=0}^{j-1}(1 - \gamma_{t-i}) \right) \gamma_{t-j} \Pi_{t-j} \right) \tag{19}$$

and predict by

$$_t\Pi_{t+1}^e = \widehat{a}_t \tag{20}$$

When $\gamma_t = \frac{1}{t}$ then $\widehat{a}_t$ is the least squares estimate of a, i.e. the average of past inflation rates. For $\gamma_t = \gamma < 1$, $\widehat{a}_t$ is an exponentially distributed weighted average of past inflation rates.

Entrepreneurs also require two period forecasts of inflation. These forecasts can be obtained by applying the law of iterated expectations to equation (17).[14] Writing equation (17) for $_{t+1}\Pi_{t+2}^e$ and taking time t expectations on both sides results in

$$_t\Pi_{t+2}^e = {}_t\Pi_{t+1}^e \tag{21}$$

which says that forecasts for the far future are identical to forecasts for the near future. This follows form the fact that agents perceive inflation to fluctuate around a steady state and believe that their forecast for the near future incorporates all available information.

5.1   The General (In-)Stability Result

This section characterizes the conditions for which the economy converges to a steady state when agents learn as outlined in the previous section.

To derive the result I must impose some regularity conditions, which are summarized as condition 2 in appendix 8.1. Condition 2 insures that various matrices are of full rank and that their eigenvectors do not lie right on the unit circle.

The following proposition states the main result of this section:

---

[14]The same result is obtained when using equations (18) and (20) instead.

Proposition 1 *Consider a steady state* (Π, m) *and assume condition 1 and regularity condition 2 hold. Let* w_i *denote the derivative of the real wage function (7) with respect to its i-th argument evaluated at* (Π, m). *If at the steady state*

$$\frac{w_2}{1-\sigma}(\Pi - 1) < \frac{w_1}{1-\sigma}\frac{m}{\Pi} < \Pi + 1 \tag{22}$$

then there exists a $\overline{\gamma} > 0$ *such that the steady state is locally asymptotically stable*

1. *for all constant gain learning rules with adaptation rates* $0 < \gamma < \overline{\gamma}$.

2. *for all decreasing gain learning rules.*

   If (22) does not hold, then there exists a $\overline{\gamma} > 0$ *such that the steady state is unstable*

1. *for all constant gain learning rules with adaptation rates* $0 < \gamma < \overline{\gamma}$.

2. *for all decreasing gain learning rules.*

The proof can be found in appendix 8.2. It is straightforward for the constant gain learning rule but more involved for the decreasing gain rule where one has to consider a non-autonomous difference equation. The proof is further complicated by the fact that one of the eigenvalues converges to 1 as γ_t converges to zero.

Condition (22), which determines the stability properties, has an immediate economic interpretation in terms of the stability of the steady state with respect to deviations of real money balances and inflation expectations from their steady state values.

First, consider demand deviations. The term in the middle of equation (22) is the elasticity of the inflation rate with respect to the $t-1$ money stock m_{t-1}, see equations (13) and (14). Since $w_1 > 0$ it follows that inflation rises in response to positive demand shocks, which should stabilize the economy because increased inflation helps devaluating excessive money balances and thereby pushes the economy back towards the steady state. Yet, if the inflation reaction is too strong then a deviation of money in one direction is followed by an even larger deviation in the opposite direction and the system starts to oscillate with increasing amplitude around the equilibrium. The term on the very right of (22) is the bound on the elasticity that prevents this from happening.

Next, consider deviations of inflation expectations from steady state inflation and the following inequality

$$w_2 < w_1 \frac{m}{\Pi^2} \qquad (23)$$

which is a sufficient condition for the inequality on the left-hand side of (22) to hold. A positive shock to agents' inflation expectations has two opposing effects on inflation which are captured by the two sides of inequality (23).

Firstly, firms anticipate lower product demand because inflation devaluates old agents' real money balances. This causes a fall in expected labor demand and expected wages and thereby puts downward pressure on inflation. This move down the labor supply function is captured by the term on the right of (23), which is the derivative of the real wage with respect to labor demand times the derivative of real money m_t (which is identical to labor demand n_t) with respect to Π_t.

Secondly, higher expected inflation taxes move the labor supply schedule upwards because workers have to be compensated with a higher real wage to offer any given amount of labor. This effect puts upward pressure on inflation and is captured by the term on the left of (23).

If the first effect dominates, then real wages will decrease in response to an increase in inflation expectations. The price setting equation (14) then implies that realized inflation will be lower than expected inflation. From the learning rule it then follows that inflation expectations will return over time to the steady state value.[15]

5.2   Learning the High Inflation Steady State

This section applies proposition 1 to study the stability of the high inflation steady state under adaptive learning. The main result is

Corollary 1 *For government expenditures g close enough to zero, condition (22) never holds for the high-inflation steady state.*

Proof: By contradiction suppose that the left-hand side of (22) holds. This requires

$$w_2 \Pi^2 - w_2 \Pi - w_1 m < 0$$

---

[15]When equation (23) does not hold then the steady state is not necessarily unstable, see equation (22), since current expectations influence also future expectations and the expected future money stocks. This channel has been ignored in the previous argument and is captured by the additional term on the left of equation (22).

With $w_2 > 0$, a necessary condition for this is

$$\Pi < \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{w_1}{w_2}m}$$

As $g \to 0$, $\Pi \to \Pi^h$ and $m \to 0$. Since $w_2 > 0$ and $w_1 > 0$ at $(\Pi, m) = (\Pi^h, 0)$, this condition boils down to $\Pi < 1$, which contradicts $\Pi^h > 1$.■

As in Marcet and Sargent (1989a), where agents could use only lagged prices to update expectations, the high inflation steady state is unstable. This holds independent of the degree of imperfect competition σ.

It is rather surprising that sticky prices do not affect the stability properties because even for $\sigma \approx 0$ the learning dynamics in a flexible price model and a sticky price model differ considerably: with flexible prices inflation always adjusts to bring the money market into equilibrium; with sticky prices inflation is chosen by entrepreneurs. Their choice brings the money market into equilibrium only if their expectations are correct. Outside the rational expectations equilibrium the real wage adjusts to insure the (temporary) market equilibrium. Therefore, inflation dynamics outside equilibrium differ even though the equilibrium dynamics of the two models are (almost) identical for small σ.

5.3 Learning the Low Inflation Steady State

Applying proposition 1 to the low inflation steady state delivers:

Corollary 2 *For government expenditures g close enough to zero, the stability condition (22) holds at the low inflation steady state if and only if*

$$\varepsilon_{n,w} > \frac{1}{2} \tag{24}$$

where $\varepsilon_{n,w}$ is the real wage elasticity of labor supply at the steady state.

Proof: All the terms in (22) are continuous in g. Therefore, when (22) holds for $g = 0$, it will also hold for sufficiently small but positive g. At $g = 0$, one has $\Pi = 1$ and $m = n(1-\sigma, 1) > 0$. For these values $\frac{1+\Pi}{\Pi} = 2$ and $\frac{w_2}{1-\sigma}\frac{\Pi-1}{\Pi} = 0$. Since $w_1 > 0$ and $m > 0$, the inequality in the left of (22) holds. Since $w = 1-\sigma$ in equilibrium, the term in the middle of (22) is equal to $\varepsilon_{w,n} = \frac{1}{\varepsilon_{n,w}}$, which establishes the claim.■

For a sufficiently elastic labor supply the low inflation steady state is stable. Again this holds independent of the degree of imperfect competition. If labor supply is too inelastic $(\varepsilon_{n,w} < \frac{1}{2})$ then inflation reacts strongly to demand deviations because firms predict a strong reaction in real wages. The strong inflation reaction leads to demand deviations of the opposite sign in the next period and generates the oscillating behavior described in section 5.1.

The result of corollary 2 differs notably from that in Marcet and Sargent (1989b) who found the low inflation steady state to be stable independent of the labor supply elasticity. One might be tempted to conclude that it is the different learning dynamics of the sticky price and flexible price economy that ultimately matter for this difference in stability: in a flexible price model real production costs remain unaffected by nominal demand conditions.

Yet, it is also possible that the role of the labor supply elasticity in the sticky price model arises only due to a singularity in the transition from a stochastic to a deterministic model: In a deterministic economy steady state money balances are constant which implies that a simple perceived law of motion of the form (18) is sufficient to acquire rational expectations. If, however, the money creation process contained a random component, then money balances would permanently fluctuate and, due to the price stickiness, inflation would react with some lag. Consequently, agents would have to condition forecasts on lagged money to learn to predict inflation rationally. This may well alter the stability property of the equilibrium.

Therefore, the next section studies an economy with stochastic demand shocks.

6   A stochastic model

This section considers an augmented version of the model with government seignorage shocks. The stochastic setup facilitates the stability analysis of non-stationary rational expectations equilibria and helps to put into perspective the role for the elasticity of labor supply found in the previous section.[16]

Suppose government seignorage is composed of a fixed and a random component

$$g_t = g + v_t$$

---

[16]Analysis of non-stationary equilibria in deterministic models can be difficult because variables may become asymptotically collinear as they settle down to constant values.

where v_t is a white noise shock with zero mean and small bounded support. Real money then evolves according to the following stochastic law of motion

$$m_t = \frac{m_{t-1}}{\Pi_t} + g + v_t \tag{25}$$

For $g > 0$ not too large and $v_t \equiv 0$ there are again the two deterministic steady states (Π^n, m^n) with $n = l, h$ around which equations (14) and (25) can be linearized:[17]

$$\begin{pmatrix} \Pi_t \\ m_t \end{pmatrix} = \alpha^n + \beta_0^n E_{t-1} \begin{bmatrix} \Pi_t \\ m_t \end{bmatrix} + \beta_1^n E_{t-1} \begin{bmatrix} \Pi_{t+1} \\ m_{t+1} \end{bmatrix} + \delta^n \begin{pmatrix} \Pi_{t-1} \\ m_{t-1} \end{pmatrix} + \begin{pmatrix} 0 \\ v_t \end{pmatrix} \tag{26}$$

The linearization coefficients $(\alpha^n, \beta_0^n, \beta_1^n, \delta^n)$ for $n = l, h$ are reported in appendix 8.3, where it is also shown that the rational expectations solutions of equation (26) have a minimum state variable representation as an AR(1) process

$$\begin{pmatrix} \Pi_t \\ m_t \end{pmatrix} = a + B \begin{pmatrix} \Pi_{t-1} \\ m_{t-1} \end{pmatrix} + \begin{pmatrix} 0 \\ v_t \end{pmatrix} \tag{27}$$

where a is a 2x1 vector and $B = (b_{i,j})$ a 2x2 matrix with $b_{1,2} \neq 0$.

In a rational expectations equilibrium the coefficient on lagged money in the inflation equation of (27) is never equal to zero, which implies that agents must condition their inflation expectations on lagged money balances if they wish to forecast rationally.

Demand shocks are important for forecasting future inflation either because they are eliminated by an appropriate reaction of inflation in the subsequent period or because they persist into the future where they affect future production costs. With money balances being permanently shocked, their importance for predicting inflation does not vanish over time.

I now describe in somewhat greater detail the rational expectations equilibria (REE) for the linearized stochastic model. The claims below are made precise in appendix 8.3.2.

In the neighborhood of the deterministic low inflation steady state there exist two stochastic REE. There is a stochastic steady state denoted by $(a^{l,1}, B^{l,1})$ and an explosive REE denoted by $(a^{l,2}, B^{l,2})$ where the matrix $B^{l,2}$ has an eigenvalue that is larger than one.

---

[17] The linearization uses equation (13) to express the wage expectations in equation (14).

In the neighborhood of the deterministic high inflation steady state there also exist two stochastic REE. There is again a stochastic steady state denoted by $(a^{h,1}, B^{h,1})$ and an asymptotically stationary REE, denoted by $(a^{h,2}, B^{h,2})$. Since the eigenvalues of $B^{h,2}$ are inside the unit circle the mean of this solution converges to the deterministic high inflation steady state.

Therefore, the situation in the stochastic model looks quite similar to that in the deterministic model. The main difference is that the equilibrium law of motion for inflation now depends on lagged money balances.

6.1   Expectational Stability of REE

This section analyzes the stability of the REE when agents use least squares estimation to learn about the parameters (a, B) of the minimum state variable solution (27). Equation (27) is the simplest equation agents could estimate with the hope of acquiring rational beliefs in the long run.

The considered least squares learning dynamics can be described as follows. Let (a_{t-1}, B_{t-1}) denote the least squares estimates of (a, B) based on information up to time $t-1$. By the law of iterated expectations, agents' forecasts will be given by

$$E_{t-1} \begin{bmatrix} \Pi_t \\ m_t \end{bmatrix} = a_{t-1} + B_{t-1} \begin{pmatrix} \Pi_{t-1} \\ m_{t-1} \end{pmatrix} \tag{28}$$

$$E_{t-1} \begin{bmatrix} \Pi_{t+1} \\ m_{t+1} \end{bmatrix} = a_{t-1} + B_{t-1} \left(a_{t-1} + B_{t-1} \begin{pmatrix} \Pi_{t-1} \\ m_{t-1} \end{pmatrix} \right) \tag{29}$$

Inserting these expectations into equation (26) delivers a new data point, which is the outcome of optimal price setting decisions when expectations are given by (28) and (29). Using the new data point agents calculate a new least squares estimate (a_t, B_t) and the process repeats itself.

For a wide range of economic models it has been shown that expectational stability (E-Stability) of a REE determines whether the least squares estimates (a_t, B_t) locally converge to their REE-values under least squares learning (Evans and Honkapohja (1994, 2001)).

E-Stability is defined in terms of the mapping from the Perceived Law of Motion (PLM), parameterized by (a, B) in (27), to the implied parameters $T(a, B)$ of the

Actual Law of Motion (ALM). The ALM parameters are obtained by inserting the expectations (28) and (29) into the model (26), which yields

$$T(a, B) = \left((\alpha + \beta_0 a + \beta_1 a + \beta_1 B a), (\beta_0 B + \beta_1 B^2 + \delta) \right) \qquad (30)$$

E-Stability of a REE is then determined by the local stability properties of the ordinary differential equation

$$\frac{d}{d\tau}(a, B) = T(a, B) - (a, B) \qquad (31)$$

at the REE-values of (a, B), where τ denotes notional or virtual time.

The following proposition reports the E-Stability properties for the REE of the model. The proof can be found in appendix 8.3.3.

Proposition 2 *There exists a level of government expenditures $\overline{g} > 0$ such that for all levels $0 \leq g < \overline{g}$ the following holds.*

For the linearization at the deterministic low inflation steady state

- *the stochastic steady state $(a^{l,1}, B^{l,1})$ is E-stable,*

- *the explosive solution $(a^{l,2}, B^{l,2})$ is E-unstable,*

For the linearization at the deterministic high inflation steady state

- *the stochastic steady state $(a^{h,1}, B^{h,1})$ is E-unstable,*

- *the asymptotically stationary solution $(a^{h,2}, B^{h,2})$ is E-unstable.*

The proposition shows that the stochastic low inflation steady state is the only E-stable REE. Again, this holds independent of the degree of imperfect competition. Furthermore, the low inflation stochastic steady state is now E-stable independent of the value of the labor supply elasticity. The latter feature emerges because the coefficient on lagged money in the inflation equation is given by $\frac{1}{m^l}$ at this equilibrium, which implies a unit-elasticity of inflation with respect to demand shocks. Consequently, demand shocks are expected to be devaluated by a corresponding amount of inflation in the subsequent period and the labor supply elasticity becomes irrelevant for entrepreneur's price setting behavior.

Given this difference to the deterministic model the next subsection shows how the stability conditions for the deterministic and the stochastic model are related to each other.

6.2 Relation with the Deterministic Model

Suppose for a moment agents used least squares to estimate the PLM of the deterministic model, i.e.

$$\Pi_t = a + \varepsilon_t \tag{32}$$

Clearly, forecasts based on this PLM will never be rational in the presence of demand shocks. However, the economy can still converge to a 'restricted perceptions equilibrium'. In such an equilibrium expectations are restricted to be of the form (32) and minimize the mean-squared forecast error for the ALM (26) generated by the expectations (32).

Appendix 8.3.4 shows that convergence to a restricted perceptions equilibrium occurs for the linearization at the deterministic low inflation steady state whenever $\varepsilon_{n,w} > \frac{1}{2}$, which is precisely the stability condition of corollary 2. Moreover, the resulting restricted perceptions equilibrium converges to the perfect foresight steady state as the support for the shock v_t vanishes. Furthermore, and this is corollary 2, the restricted perceptions equilibrium is the perfect foresight steady state if the support is literally zero.

Thus, the PFE analyzed in section 5 is *not* the limit of the minimum state variable REE for a vanishing support of the shocks. Instead it is the limit of a restricted perceptions equilibrium. The stability conditions for restricted perceptions equilibria need not be identical to those of the REE, as has been noted by Evans and Honkapohja (2001,chapter 13) in a more simple setting.

The previous findings also help to interpret the results of Van Zandt and Lettau (2002) whose stability reversals are obtained (mainly) for constant gain learning rules in a deterministic model.[18] A constant gain rule with a gain that is bounded from zero would never converge to a rational expectations equilibrium in a stochastic environment. This suggests that their stability conditions do not capture the stability properties of a rational expectations equilibrium in a stochastic version of their model but rather the stability properties to some other equilibrium, as is the case with the perfect foresight equilibria in this paper.

---

[18]For the cases where reversals are reported for decreasing gain learning rules re-interpretation as constant gain rules is possible as also argued by these authors.

7　Conclusions

This paper has shown that price stickiness helps to obtain a unique temporary equilibrium in models where informational consistency requires that agents can use current prices to update their expectations. Price stickiness thereby permits to study the stability of rational expectations equilibria under learning dynamics in a rigorous way.

　　The main result for the overlapping generations model is that, independent of the degree of imperfect competition, the determinate monetary steady state is the unique stable equilibrium. However, the paper has also shown that stability properties might differ between deterministic and stochastic models since there might be a discontinuity at the zero variance for shocks. In the present case this discontinuity emerged because the equilibrium law of motion for inflation for the deterministic economy was considerably simpler than the equilibrium law for a stochastic economy with vanishing noise.

　　This latter finding is interesting because it illustrates that agents who do not start out by estimating a forecasting equation whose structure is consistent with the rational expectations solution might either learn a different equilibrium, e.g. a restricted perceptions equilibrium, or affect the stability of the economy. After all, it is a rather strong assumption to postulate that agents learn about a model that is consistent with the law of motion of the economy once learning is complete.

　　A natural step to take is to study models where agents must learn not only about the parameterization of a given model but also about which model to use for learning. In Adam (2002) I have studied some of the implications of learning about forecast models but further research in this area would be most welcome.

8 Appendix

8.1 Regularity Condition

Condition 2 *At a stationary rational expectations equilibrium (Π, m)*

$$\frac{w_1}{1-\sigma}\frac{m}{\Pi^2} \neq \frac{w_2}{1-\sigma}\frac{\Pi-1}{\Pi}$$

$$\frac{w_1}{1-\sigma}\frac{m}{\Pi^2} \neq \frac{\Pi+1}{\Pi}$$

$$\gamma_t \neq \frac{m}{\Pi^2}\frac{w_1}{1-\sigma} - \frac{1}{\Pi} \quad \forall \ t$$

where

$$w_1 = \frac{\partial w}{\delta n}(m, \Pi) \quad and \quad w_2 = \frac{\partial w}{\partial_{t-1}\Pi_{t+1}^e}(m, \Pi)$$

8.2 Proof of Proposition 1

Substituting (14) into (17) and using (13) and (21) delivers:

$$_t\Pi_{t+1}^e = {_{t-1}}\Pi_t^e + \gamma_t \left(M(_{t-1}\Pi_t^e, m_{t-1}) - {_{t-1}}\Pi_t^e \right) \quad \text{with} \tag{33}$$

$$M(_{t-1}\Pi_t^e, m_{t-1}) \equiv {_{t-1}}\Pi_t^e \frac{w\left(\frac{m_{t-1}}{t-1\Pi_t^e} + g, \ {_{t-1}}\Pi_t^e\right)}{1-\sigma} \tag{34}$$

The equation above describes the new inflation expectations as a function of past expectations and past real money holdings. Real money evolves according to

$$m_t = \frac{m_{t-1}}{M(_{t-1}\Pi_t^e, m_{t-1})} + g \tag{35}$$

Linearizing (33) and (35) around a steady state (Π, m) yields

$$\begin{pmatrix} \theta_{1,t} \\ \theta_{2,t} \end{pmatrix} = A(\gamma_t) \begin{pmatrix} \theta_{1,t-1} \\ \theta_{2,t-1} \end{pmatrix} + \begin{pmatrix} \gamma_t r_{1,t} \\ r_{2,t} \end{pmatrix} \tag{36}$$

where

$$\theta_{1,t} = {_t}\Pi_{t+1}^e - \Pi$$

$$\theta_{2,t} = m_t - m$$

are the deviations from the equilibrium values, the $r_{i,t}$ are second order approximation errors, and $A(\gamma_t)$ is a 2x2 matrix given by

$$A(\gamma_t) = \begin{pmatrix} 1 + \gamma_t(M_1(\Pi, m) - 1) & \gamma_t M_2(\Pi, m) \\ -\frac{m}{\Pi^2}M_1(\Pi, m) & \frac{1}{\Pi} - \frac{m}{\Pi^2}M_2(\Pi, m) \end{pmatrix}$$

24

where M_i is the partial derivative of M with respect to the i-th argument.

The eigenvalues of $A(0)$ are given by

$$\lambda_1 = 1$$
$$\lambda_2 = \frac{1}{\Pi} - \frac{m}{\Pi^2} \frac{\omega_1}{(1-\sigma)}$$

and the eigenvalues of $A(\gamma_t)$ by

$$\lambda_{1,t} = \lambda_1 + \frac{\partial \lambda_1}{\partial \gamma} \gamma_t + O(\gamma_t^2) \tag{37}$$

$$\lambda_{2,t} = \lambda_2 + \frac{\partial \lambda_2}{\partial \gamma} \gamma_t + O(\gamma_t^2) \tag{38}$$

where the last terms are second order approximation errors.

If condition (22) holds, then $|\lambda_2| < 1$. The regularity condition 2 implies $\left| \frac{\partial \lambda_2}{\partial \gamma} \right| < \infty$ at $\gamma = 0$. Thus, $|\lambda_{2,t}| < 1$ for small enough γ_t. Next, consider the eigenvalue $\lambda_{1,t}$. Use the characteristic polynomial of $A(\gamma)$ given by

$$P(A(\gamma), \lambda) = (A_{11}(\gamma) - \lambda)(A_{22}(\gamma) - \lambda) - A_{12}(\gamma)A_{21}(\gamma)$$

where the $A_{ij}(\gamma)$ denote the matrix entries of $A(\gamma)$. Then apply the implicit function theorem to obtain $\frac{\partial \lambda_1}{\partial \gamma}$ at $\gamma = 0, \lambda = \lambda_1 = 1$:

$$\frac{\partial \lambda_1}{\partial \gamma} = -\frac{\partial P(A(\gamma), \lambda)/\partial \gamma}{\partial P(A(\gamma), \lambda)/\partial \lambda}$$
$$= \frac{\frac{1}{1-\sigma} \left(w_2(\Pi - 1) - w_1 \frac{m}{\Pi} \right)}{1 - \lambda_2}$$

Condition (22) implies that $\frac{\partial \lambda_1}{\partial \gamma} < 0$ and thereby $|\lambda_{1,t}| < 1$ for small enough γ_t. Thus, (22) implies that both eigenvalues of $A(\gamma_t)$ are within the unit circle for γ_t sufficiently small. Otherwise, at least one eigenvalue lies outside the unit circle. This establishes the stability and instability claims for the constant gain learning rules.

Below I consider the case of decreasing gain learning rules. The proof for the decreasing gain rules is complicated by the fact that the difference equation (36) is non-autonomous and that one of the eigenvalues of $A(\gamma_t)$ converges to 1 as $\gamma_t \to 0$. The proof uses similar techniques as Evans and Honkapohja (2000) who treat a simpler general setup. The main difference to Evans and Honkapohja is that the current setting contains a state variable ($\theta_{2,t}$) that is not a belief and therefore does not contain a gain sequence converging to zero.

The function $M(\cdot, \cdot)$ as defined in (34) is continuously differentiable in both arguments in a neighborhood of $\theta = (0,0)$. Therefore, for all $K_1 > 0$ and $K_2 > 0$ with K_1 and K_2 arbitrarily small there exists a neighborhood to $(0,0)$ where the absolute values of the approximation errors in (36) are bounded by

$$|r_{1,t}| \leq K_1 \left(|\theta_{1,t-1}| + |\theta_{2,t-1}| \right) \tag{39a}$$

$$|r_{2,t}| \leq K_2 \left(|\theta_{1,t-1}| + |\theta_{2,t-1}| \right) \tag{39b}$$

Next, consider the eigenvectors e_1 and e_2 of $A(0)$ corresponding to the eigenvectors λ_1 and λ_2, respectively:

$$e_1 = \begin{pmatrix} e_{11} \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1 - \frac{1}{\Pi} + \frac{m}{\Pi^2} M_2}{-\frac{m}{\Pi^2} M_1} \\ 1 \end{pmatrix}$$

$$e_2 = \begin{pmatrix} e_{21} \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

The eigenvectors $e_{1,t}$ and $e_{2,t}$ of $A(\gamma_t)$ corresponding to the eigenvalues $\lambda_{1,t}$ and $\lambda_{2,t}$, respectively, are given by

$$e_{1,t} = \begin{pmatrix} e_{11,t} \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\lambda_{1,t} - \frac{1}{\Pi} + \frac{m}{\Pi^2} M_2}{-\frac{m}{\Pi^2} M_1} \\ 1 \end{pmatrix}$$

$$e_{2,t} = \begin{pmatrix} e_{21,t} \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{\lambda_{2,t} - \frac{1}{\Pi} + \frac{m}{\Pi^2} M_2}{-\frac{m}{\Pi^2} M_1} \\ 1 \end{pmatrix}$$

Now consider the vector base consisting of the eigenvectors $(e_{1,t}, e_{2,t})$ of $A(\gamma_t)$. Let the vector (θ_1, θ_2) have representation $(\rho_1, \rho_2)_{\gamma_t}$ with this new base, i.e.

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = (e_{1,t}, e_{2,t}) \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}_{\gamma_t} \tag{40}$$

where the subscript γ_t indicates the base to which the coordinates refer. Then from (36)

$$\begin{pmatrix} \rho_{1,t} \\ \rho_{2,t} \end{pmatrix}_{\gamma_t} = \begin{pmatrix} \lambda_{1,t} & 0 \\ 0 & \lambda_{2,t} \end{pmatrix} \begin{pmatrix} \rho_{1,t-1} \\ \rho_{2,t-1} \end{pmatrix}_{\gamma_t} + \begin{pmatrix} s_{1,t} \\ s_{2,t} \end{pmatrix}_{\gamma_t} \tag{41}$$

where the approximation errors are given by

$$\begin{pmatrix} s_{1,t} \\ s_{2,t} \end{pmatrix}_{\gamma_t} = \begin{pmatrix} \frac{1}{e_{11,t} - e_{21,t}} (\gamma_t r_{1,t} - e_{21,t} r_{2,t}) \\ \frac{1}{e_{11,t} - e_{21,t}} (-\gamma_t r_{1,t} + e_{11,t} r_{2,t}) \end{pmatrix}_{\gamma_t}$$

26

Using (39a), (39b), and (40) the approximation errors can be bounded as follows:

$$|s_{1,t}| \leq \frac{1}{|e_{11,t} - e_{21,t}|}(\gamma_t |r_{1,t}| + |e_{21,t}| |r_{2,t}|)$$

$$\leq \frac{1}{|e_{11,t} - e_{21,t}|}(\gamma_t K_1 + |e_{21,t}| K_2)((|\theta_{1,t-1}| + |\theta_{2,t-1}|))$$

$$= \frac{1}{|e_{11,t} - e_{21,t}|}(\gamma_t K_1 + |e_{21,t}| K_2)(|\rho_{1,t-1}e_{11,t} + \rho_{2,t-1}e_{21,t}| + |\rho_{1,t-1} + \rho_{2,t-1}|)$$

$$\leq \frac{1}{|e_{11,t} - e_{21,t}|}(\gamma_t K_1 + |e_{21,t}| K_2)(|\rho_{1,t-1}| (|e_{11,t}| + 1) + |\rho_{2,t-1}| (|e_{21,t}| + 1))$$

Since $\lim_{t\to\infty} |e_{11,t} - e_{21,t}| > 0$, $|e_{21,t}| \sim O(\gamma_t)$, and since K_1 and K_2 can be made arbitrarily small it follows that

$$|s_{1,t}| \leq \gamma_t K_1'(|\rho_{1,t-1}| + |\rho_{2,t-1}|) \tag{42}$$

for some $K_1' > 0$ that can also be made arbitrarily small by considering a sufficiently small neighborhood around the steady state. Similarly,

$$|s_{2,t}| \leq \frac{1}{|e_{11,t} - e_{21,t}|}(\gamma_t |r_{1,t}| + |e_{11,t}| |r_{2,t}|)$$

$$\leq \frac{1}{|e_{11,t} - e_{21,t}|}(\gamma_t K_1 + |e_{11,t}| K_2)(|\rho_{1,t-1}| (|e_{11,t}| + 1) + |\rho_{2,t-1}| (|e_{21,t}| + 1))$$

$$\leq K_2'(|\rho_{1,t-1}| + |\rho_{2,t-1}|) \tag{43}$$

for some $K_2' > 0$ arbitrarily small.

An inconvenient feature of (41) is that the coordinates are expressed in terms of a different vector base for each γ_t. Therefore, I rewrite (41) with coordinates from the vector base (e_1, e_2). This base is almost identical to the base $(e_{1,t}, e_{2,t})$ for small γ_t.

Let $(\rho_1, \rho_2)_{\gamma_t}$ have representation $(\alpha_1, \alpha_2)_0$ with base (e_1, e_1), i.e.

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}_0 = \begin{pmatrix} a_{11,t} & a_{12,t} \\ a_{21,t} & a_{22}, t \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}_{\gamma_t}$$

$$= (e_1, e_2)^{-1} (e_{1,t}, e_{2,t}) \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}_{\gamma_t}$$

$$= \begin{pmatrix} \frac{e_{11,t}}{e_{11}} & \frac{e_{21,t}}{e_{11}} \\ 1 - \frac{e_{11,t}}{e_{11}} & 1 - \frac{e_{21,t}}{e_{11}} \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}_{\gamma_t} \tag{44}$$

or conversely

$$
\begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}_{\gamma_t} = \begin{pmatrix} b_{11,t} & b_{12,t} \\ b_{21,t} & b_{22,t} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}_0
$$

$$
= \begin{pmatrix} \frac{e_{11}-e_{21,t}}{e_{11,t}-e_{21,t}} & -\frac{e_{21,t}}{e_{11,t}-e_{21,t}} \\ \frac{e_{11,t}-e_{11}}{e_{11,t}-e_{21,t}} & \frac{e_{11,t}}{e_{11,t}-e_{21,t}} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}_0 \tag{45}
$$

One can express the bound on the approximation error in (42) in new coordinates

$$
\begin{aligned}
|s_{1,t}| &\leq \gamma_t K_1'(|b_{11,t}\alpha_{1,t-1} + b_{12,t}\alpha_{2,t-1}| + |b_{21,t}\alpha_{1,t-1} + b_{22,t}\alpha_{2,t-1}|) \\
&\leq \gamma_t K_1' \left((|b_{11,t}| + |b_{21,t}|) |\alpha_{1,t-1}| + (|b_{12,t}| + |b_{22,t}|) |\alpha_{2,t-1}| \right) \\
&\leq \gamma_t K_1'' \left(|\alpha_{1,t-1}| + |\alpha_{2,t-1}| \right) \tag{46}
\end{aligned}
$$

for $K_1'' > 0$ and arbitrarily small for a sufficiently small neighborhood. Similarly for the bound in (43)

$$
|s_{2,t}| \leq K_2''(|\alpha_{1,t-1}| + |\alpha_{2,t-1}|) \tag{47}
$$

with $K_2'' > 0$, arbitrarily small. From (41), (44), and (45)

$$
\begin{aligned}
\alpha_{1,t} &= a_{11,t}\rho_{1,t} + a_{12,t}\rho_{2,t} \\
&= a_{11,t} \left(\lambda_{1,t}\rho_{1,t-1} + s_{1,t} \right) + a_{12,t}((\lambda_2\rho_{2,t-1} + s_{2,t}) \\
&= a_{11,t} \left(\lambda_{1,t} \left(b_{11,t}\alpha_{1,t-1} + b_{12,t}\alpha_{2,t-1} \right) + s_{1,t} \right) \\
&\quad + a_{12,t} \left(\lambda_{2,t} \left(b_{21,t}\alpha_{1,t-1} + b_{22,t}\alpha_{2,t-1} \right) + s_{2,t} \right) \\
&= \left(a_{11,t}\lambda_{1,t}b_{11,t} + a_{12,t}\lambda_{2,t}b_{21,t} \right) \alpha_{1,t-1} \\
&\quad + \left(a_{11,t}\lambda_{1,t}b_{12,t} + a_{12,t}\lambda_{2,t}b_{22,t} \right) \alpha_{2,t-1} + a_{11,t}s_{1,t} + a_{12,t}s_{2,t} \tag{48}
\end{aligned}
$$

and similarly

$$
\begin{aligned}
\alpha_{2,t} &= \left(a_{21,t}\lambda_{1,t}b_{11,t} + a_{22,t}\lambda_{2,t}b_{21,t} \right) \alpha_{1,t-1} \\
&\quad + \left(a_{21,t}\lambda_{1,t}b_{12,t} + a_{22,t}\lambda_{2,t}b_{22,t} \right) \alpha_{2,t-1} + a_{21,t}s_{1,t} + a_{22,t}s_{2,t} \tag{49}
\end{aligned}
$$

Using (46), (47), and (48) one can construct upper and lower bounds for $|\alpha_{1,t}|$:

$$
\begin{aligned}
|\alpha_{1,t}| &\leq |(a_{11,t}\lambda_{1,t}b_{11,t} + a_{12,t}\lambda_{2,t}b_{21,t})| \, |\alpha_{1,t-1}| \\
&\quad + |(a_{11,t}\lambda_{1,t}b_{12,t} + a_{12,t}\lambda_{2,t}b_{22,t})| \, |\alpha_{2,t-1}| + |a_{11,t}| \, |s_{1,t}| + |a_{12,t}| \, |s_{2,t}| \\
&\leq (|(a_{11,t}\lambda_{1,t}b_{11,t} + a_{12,t}\lambda_{2,t}b_{21,t})| + \gamma_t |a_{11,t}| \, K_1'' + |a_{12,t}| \, K_2'') |\alpha_{1,t-1}| \\
&\quad + (|(a_{11,t}\lambda_{1,t}b_{12,t} + a_{12,t}\lambda_{2,t}b_{22,t})| + \gamma_t |a_{11,t}| \, K_1'' + |a_{12,t}| \, K_2'') |\alpha_{2,t-1}|
\end{aligned}
$$

28

Note that the terms $a_{12,t}$, $b_{12,t}$, and $b_{21,t}$ are of order $O(\gamma_t)$. Moreover, using a Taylor series expansion

$$
\begin{aligned}
a_{11,t}\lambda_{1,t}b_{11,t} &= \left(1 + C\frac{\partial\lambda_1}{\partial\gamma}\gamma_t + O(\gamma_t^2)\right)\left(1 + \frac{\partial\lambda_1}{\partial\gamma}\gamma_t + O(\gamma_t^2)\right)\left(1 - C\frac{\partial\lambda_1}{\partial\gamma}\gamma_t + O(\gamma_t^2)\right) \\
&= 1 + \frac{\partial\lambda_1}{\partial\gamma}\gamma_t + O(\gamma_t^2)
\end{aligned}
$$

Therefore

$$
|\alpha_{1,t}| \leq (1 + \gamma_t V_{11})\,|\alpha_{1,t-1}| + \gamma_t V_{12}\,|\alpha_{2,t-1}|
$$

where $V_{12} > 0$ and V_{11} of the same sign as $\frac{\partial\lambda_1}{\partial\gamma}$. Also V_{11} and can be made arbitrarily close to $\frac{\partial\lambda_1}{\partial\gamma}$ by choosing a sufficiently small neighborhood.

A lower bound for $|\alpha_{1,t}|$ is given by

$$
\begin{aligned}
|\alpha_{1,t}| &\geq |(a_{11,t}\lambda_{1,t}b_{11,t} + a_{12,t}\lambda_{2,t}b_{21,t})|\,|\alpha_{1,t-1}| \\
&\quad - |(a_{11,t}\lambda_{1,t}b_{12,t} + a_{12,t}\lambda_{2,t}b_{22,t})|\,|\alpha_{2,t-1}| - |a_{11,t}|\,|s_{1,t}| - |a_{12,t}|\,|s_{2,t}| \\
&\geq (|(a_{11,t}\lambda_{1,t}b_{11,t} + a_{12,t}\lambda_{2,t}b_{21,t})| - \gamma_t\,|a_{11,t}|\,K_1'' - |a_{12,t}|\,K_2''\,)\,|\alpha_{1,t-1}| \\
&\quad - (|(a_{11,t}\lambda_{1,t}b_{12,t} + a_{12,t}\lambda_{2,t}b_{22,t})| + \gamma_t\,|a_{11,t}|\,K_1'' + |a_{12,t}|\,K_2''\,)\,|\alpha_{2,t-1}| \\
&\geq (1 + \gamma_t W_{11})\,|\alpha_{1,t-1}| - \gamma_t W_{12}\,|\alpha_{2,t-1}|
\end{aligned}
$$

with $W_{12} > 0$ and W_{11} of the same sign and arbitrarily close to $\frac{\partial\lambda_1}{\partial\gamma}$, by the same arguments as above.

Next use (46), (47), and (49) to get bounds for $|\alpha_{2,t}|$:

$$
\begin{aligned}
|\alpha_{2,t}| &\leq (|(a_{21,t}\lambda_{1,t}b_{11,t} + a_{22,t}\lambda_{2,t}b_{21,t})| + \gamma_t\,|a_{21,t}|\,K_1'' + |a_{22,t}|\,K_2''\,)\,|\alpha_{1,t-1}| \\
&\quad + (|(a_{21}\lambda_{1,t}b_{12,t} + a_{22,t}\lambda_{2,t}b_{22,t})| + \gamma_t\,|a_{21,t}|\,K_1'' + |a_{22,t}|\,K_2''\,)\,|\alpha_{2,t-1}|
\end{aligned}
$$

Since $a_{21,t}$, $b_{12,t}$, and $b_{21,t}$ are of order $O(\gamma_t)$

$$
|\alpha_{2,t}| \leq V_{21}\,|\alpha_{1,t-1}| + V_{22}\,|\alpha_{2,t-1}|
$$

with $V_{21} > 0$, $V_{22} > 0$. Moreover, by choosing a sufficiently small neighborhood and a t large enough one can choose V_{21} arbitrarily close to zero. Also, since

$$
\lim a_{22,t} = \lim b_{22,t} = 1
$$

one can choose $V_{22} < 1$ when $|\lambda_2| < 1$ and $V_{22} > 1$ when $|\lambda_2| > 1$ for all t sufficiently large and all sufficiently small neighborhoods.

29

A lower bound for $|\alpha_{2,t}|$ is given by

$$
\begin{aligned}
|\alpha_{2,t}| \geq {} & - \left(|(a_{21,t}\lambda_{1,t}b_{11,t} + a_{22,t}\lambda_{2,t}b_{21,t})| + \gamma_t |a_{21,t}| K_1'' + |a_{22,t}| K_2'' \right) |\alpha_{1,t-1}| \\
& + \left(|(a_{21,t}\lambda_{1,t}b_{12,t} + a_{22,t}\lambda_{2,t}b_{22,t})| - \gamma_t |a_{21,t}| K_1'' - |a_{22,t}| K_2'' \right) |\alpha_{2,t-1}| \\
\geq {} & -W_{21} |\alpha_{1,t-1}| + W_{22} |\alpha_{1,t-1}|
\end{aligned}
$$

with $W_{21} > 0$, $W_{22} > 0$. By the same arguments as above, for t sufficiently large and a sufficiently small neighborhood $W_{22} < 1$ if $|\lambda_2| < 1$ and $W_{22} > 1$ if $|\lambda_2| > 1$.

Collecting the previous bounds we have

$$
W_t \begin{pmatrix} |\alpha_{1,t-1}| \\ |\alpha_{2,t-1}| \end{pmatrix} \leq \begin{pmatrix} |\alpha_{1,t}| \\ |\alpha_{2,t}| \end{pmatrix} \leq V_t \begin{pmatrix} |\alpha_{1,t-1}| \\ |\alpha_{2,t-1}| \end{pmatrix} \tag{50}
$$

$$
where \quad W_t = \begin{pmatrix} 1 + \gamma_t W_{11} & -\gamma_t W_{12} \\ -W_{21} & W_{22} \end{pmatrix}
$$

$$
V_t = \begin{pmatrix} 1 + \gamma_t V_{11} & \gamma_t V_{12} \\ V_{21} & V_{22} \end{pmatrix}
$$

where the inequalities should be interpreted component-wise. Now take a time t^* and a neighborhood U such that one can choose $W_{22} < 1$ ($W_{22} > 1$), and $V_{22} < 1$ ($V_{22} > 1$) if $|\lambda_2| < 1$ ($|\lambda_2| > 1$).

I now assume (22) holds and will prove the stability part for decreasing gain learning rules. First, construct a matrix norm $\|\cdot\|_h$ and a compatible vector norm $|\cdot|_h$ such that

$$
|Sx|_h \leq \|S\|_h |x|_h \tag{51}
$$

for all 2×2 matrices S and 2×1 vectors x. Define the matrix norm as follows

$$
\|S\|_h = \left\| D_h S D_h^{-1} \right\|_{\max} \quad \text{with}
$$

$$
D_h = \begin{pmatrix} h^2 & 0 \\ 0 & h \end{pmatrix}
$$

where $\|\cdot\|_{\max}$ is the maximum absolute norm defined by $\|M\|_{\max} = \max_{i,j} |M_{i,j}|$. A compatible vector norm is given by (see Horn and Johnson (1985), p.297)

$$
|x|_h = \|(x, x)\|_h
$$

where (x, x) is the matrix whose columns consist of the vectors x.

With these definitions we have

$$\|V_t\|_h = \left\| \begin{pmatrix} 1 + \gamma_t V_{11} & \gamma_t h V_{12} \\ h^{-1} V_{21} & V_{22} \end{pmatrix} \right\|_{\text{max}}$$

Now choose h large enough such that

$$h^{-1} V_{21} < V_{22}$$

and a time $t^{**} \geq t^*$ large enough such that for all $t \geq t^{**}$

$$\gamma_t h V_{12} < V_{22} < 1 + \gamma_t V_{11}$$

Then for $t \geq t^{**}$

$$\|V_t\|_h = 1 + \gamma_t V_{11}$$

and by (51)

$$\left\| V_t \cdot \begin{pmatrix} |\alpha_{1,t-1}| \\ |\alpha_{2,t-1}| \end{pmatrix} \right\|_h \leq (1 + \gamma_t V_{11}) \left| \begin{pmatrix} |\alpha_{1,t-1}| \\ |\alpha_{2,t-1}| \end{pmatrix} \right|_h \tag{52}$$

Since the vector norm $|\cdot|_h$ is absolute, i.e.

$$|x|_h = |\,|x|\,|_h$$

it follows (from Horn and Johnson (1985), p.285) that it is monotone. From (50) and (52) we therefore have that

$$\left| \begin{pmatrix} |\alpha_{1,t}| \\ |\alpha_{2,t}| \end{pmatrix} \right|_h - \gamma_t V_{11} \left| \begin{pmatrix} |\alpha_{1,t-1}| \\ |\alpha_{2,t-1}| \end{pmatrix} \right|_h \leq \left| \begin{pmatrix} |\alpha_{1,t-1}| \\ |\alpha_{2,t-1}| \end{pmatrix} \right|_h \tag{53}$$

Since $V_{11} < 0$ when (22) holds, $|\alpha_t|$ is a strictly decreasing positive sequence. This implies that it has a limit $\alpha^* \geq 0$. I now show that $\alpha^* = 0$. Summing the left- and right-hand side of equation (53) for t to $t + s$ yields

$$\left| \begin{pmatrix} |\alpha_{1,t+s}| \\ |\alpha_{2,t+s}| \end{pmatrix} \right|_h - V_{11} \sum_{i=0}^{s} \gamma_{t+i} \left| \begin{pmatrix} |\alpha_{1,t-1+i}| \\ |\alpha_{2,t-1+i}| \end{pmatrix} \right|_h \leq \left| \begin{pmatrix} |\alpha_{1,t-1}| \\ |\alpha_{2,t-1}| \end{pmatrix} \right|_h$$

Now assume $\alpha^* > 0$. Then we can divide the previous expression by the norm of $|\alpha_{t+s-1}|_h$ which together with the fact that $|\alpha_t|_h$ is decreasing yields

$$\frac{\left| \begin{pmatrix} |\alpha_{1,t+s}| \\ |\alpha_{2,t+s}| \end{pmatrix} \right|_h}{\left| \begin{pmatrix} |\alpha_{1,t-1+s}| \\ |\alpha_{2,t-1+s}| \end{pmatrix} \right|_h} - V_{11} \sum_{i=0}^{s} \gamma_{t+i} \leq \frac{\left| \begin{pmatrix} |\alpha_{1,t-1}| \\ |\alpha_{2,t-1}| \end{pmatrix} \right|_h}{\left| \begin{pmatrix} |\alpha_{1,t-1+s}| \\ |\alpha_{2,t-1+s}| \end{pmatrix} \right|_h}$$

31

Since $\sum_t \gamma_t = \infty$ the left-hand side will increase without bound as s increases. But then $|a_{t-1+s}|_h$ must converge to zero, a contradiction. Therefore, $\lim_{t\to\infty} |\alpha_t|_h = \alpha^* = 0$.

This establishes that there exists a neighborhood U of $\alpha = (0,0)$ such that if $\alpha_t \in U$ at a time $t \geq t^{**}$, then $\alpha_t \to (0,0)$. By continuity of $M_i(\cdot, \cdot)$ and the fact that (36) has a fixed point at $(0,0)$ for all γ_t, α_t remains in U for $t \leq t^{**}$ if the initial values α_0 are chosen from another sufficiently small neighborhood $U' \subset U$. This establishes the asymptotic stability result for decreasing gain learning rules.

I now proceed with the instability part of the proposition. When (22) does not hold then $\frac{\partial \lambda_1}{\partial \gamma} > 0$, or $|\lambda_2| > 1$, or both.

First suppose $|\lambda_2| > 1$. Then $W_{22} > 1$ for $t \geq t^*$. Consider the cone $C_\beta = \{(\alpha_1, \alpha_2) | \, |\alpha_2| \geq \beta |\alpha_1|\}$. I will show that there exists a finite time $t^{**} \geq t^*$ and a neighborhood $U' \subset U$ such that if $\alpha_t \in C_\beta \cap U'$ at a time $t \geq t^{**}$, it follows that $\alpha_{t+1} \in C_\beta$. In other words, α_t must leave U' before it can leave C_β.

From (50) we have that for $\alpha_t \in C_\beta$ and β large enough

$$|\alpha_{2,t+1}| - |\alpha_{2,t}| \geq (-\frac{W_{21}}{\beta} + W_{22} - 1) |\alpha_{2,t}| = Z |\alpha_{2,t}| \tag{54}$$

For β large enough $Z > 0$. Also from (50)

$$|\alpha_{1,t+1}| - |\alpha_{1,t}| \leq \gamma_{t+1} (V_{12} |\alpha_{1,t}| + V_{22} |\alpha_{2,t}|)$$
$$\leq \gamma_{t+1}(\frac{V_{12}}{\beta} + V_{22}) |\alpha_{2,t}|$$

Choosing t^{**} large enough such that for all $t \geq t^{**}$

$$\beta \gamma_{t+1}(\frac{V_{12}}{\beta} + V_{22}) < Z$$

we have

$$\beta (|\alpha_{1,t+1}| - |\alpha_{1,t}|) \leq |\alpha_{2,t+1}| - |\alpha_{2,t}|$$

From $\alpha_t \in C_\beta$ we have

$$\beta |\alpha_{1,t}| \leq |\alpha_{2,t}|$$

Adding up the last two equations implies $\alpha_{t+1} \in C_\beta$.

Now note that from (54) it follows that for $t \geq t^{**}$ and any $\alpha_t \in C_\beta \cap U'$ with $\alpha_{2,t} \neq 0$ the sequence $\{\alpha_{t+i}\}_{i=1}^\infty$ will leave U' in finite time.

32

It remains to show that for any small neighborhood $U'' \subset U'$ there is a point $\alpha_0 \in U''$ that is mapped in a fixed number of steps t^{**} into a non-zero point $\alpha_{t^{**}} \in C_\beta \cap U'$. The mapping $M(\Pi, m)$ as defined in (34) is continuously differentiable for $\Pi > 0$. Furthermore, $M(\Pi, m) > 0$ for $\Pi > 0$. Since in any stationary rational expectations equilibrium (Π, m) with $g \geq 0$ we have $\Pi \geq 1$, (36) is continuously differentiable in a neighborhood of $\alpha = (0, 0)$. By the regularity assumptions, the matrices $A(\gamma_t)$ are non-singular. The mapping (36), therefore, fulfills the assumptions of the inverse function theorem (see e.g. Hirsch and Smale (1974), p.337). Moreover, they have a fixed point at $(0, 0)$ for all t. Therefore, the t^{**}-iterative map also fulfills the assumptions of the inverse function theorem and has a fixed point at $(0, 0)$. Now fix an arbitrary $U'' \subset U'$ and choose a $\alpha_{t^{**}} \in C_\beta \cap U''$ with $|\alpha_{1,t^*}|$ and $|\alpha_{2,t^*}|$ sufficiently small. Then by the continuous differentiability of the t^{**}-iterative map and the fixed point property, the pre-image (θ_0, ρ_0) must be in U''. But I have shown that from t^{**} onwards one obtains a divergent trajectory.

Next, suppose $\frac{\partial \lambda_1}{\partial \gamma} > 1$ and without loss of generality $|\lambda_2| < 1$. Then $W_{11} > 0$ and $V_{22} < 1$ for $t \geq t^*$. Define the cone $C'_\beta = \{(\alpha_1, \alpha_2)| \, |\alpha_1| \geq \beta \, |\alpha_2|\}$. With $\alpha_t \in C_\beta \cap U$ equation (50) implies for $t \geq t^*$

$$|\alpha_{1,t+1}| - |\alpha_{1,t}| \geq (\gamma_{t+1} W_{11} \, |\alpha_{1,t}| - \gamma_{t+1} W_{12} \, |\alpha_{2,t}| \,)$$
$$\geq \gamma_{t+1} \left(W_{11} - \frac{W_{12}}{\beta} \right) |\alpha_{1,t}| \tag{55}$$

Similarly, (50) implies

$$\beta \left(|\alpha_{2,t+1}| - |\alpha_{2,t}| \right) \leq \beta \left(V_{21} \, |\alpha_{1,t}| + (V_{22} - 1) \right) |\alpha_{2,t}|$$
$$\leq \left(\beta V_{21} + (V_{22} - 1) \right) |\alpha_{1,t}| \tag{56}$$

Now choose a β such that

$$\left(W_{11} - \frac{W_{12}}{\beta} \right) > 0 \tag{57}$$

This implies that $|\alpha_{1,t}|$ is increasing in $C'_\beta \cap U$.

Restricting consideration to a sufficiently small neighborhood $U' \subset U$ one can choose V_{21} arbitrarily close to zero and V_{22} arbitrarily close to $|\lambda_2| < 1$. With a sufficiently small U' it holds that

$$(\beta V_{21} + (V_{22} - 1)) < 0$$

33

and $|\alpha_{2,t}|$ is decreasing in $C'_\beta \cap U'$. This implies that $\alpha_{t+1} \in C'_\beta$ whenever $\alpha_t \in C'_\beta \cap U'$ for $t \geq t^*$. As before, α_t must leave U' before it can leave C'_β. At the same time (55),(57), and the fact that $\sum \gamma_t = \infty$ imply that α_t will leave U' in finite time. Then choosing an $\alpha_{t^*} \in C'_\beta \cap U''$ sufficiently close to zero will insure that the pre-image α_0 of α_{t^*} will be from any arbitrarily small neighborhood $U'' \subset U'$. But from t^{**} onwards one gets a divergent trajectory.

8.3 Appendix to Section 6

8.3.1 Linearization Coefficients

For $g = 0$, the coefficient matrices around the low inflation equilibrium (Π^l, m^l) are given by

$$\alpha^l = \begin{pmatrix} -\frac{w_2^l}{1-\sigma} \\ m^l(1 + \frac{w_2^l}{1-\sigma}) \end{pmatrix} \tag{58a}$$

$$\beta_0^l = \begin{pmatrix} 1 - \frac{m^l w_1^l}{1-\sigma} & 0 \\ -m^l + (m^l)^2 \frac{w_1^l}{1-\sigma} & 0 \end{pmatrix} \tag{58b}$$

$$\beta_1^l = \begin{pmatrix} \frac{w_2^l}{1-\sigma} & 0 \\ -m^l \frac{w_2^l}{1-\sigma} & 0 \end{pmatrix} \tag{58c}$$

$$\delta^l = \begin{pmatrix} 0 & \frac{w_1^l}{1-\sigma} \\ 0 & 1 - m^l \frac{w_1^l}{1-\sigma} \end{pmatrix} \tag{58d}$$

where

$$w_1^l = \frac{\partial w}{\partial n_t}(m^l, \Pi^l) \text{ and } w_2^l = \frac{\partial w}{\partial E_{t-1}[\Pi_t]}(m^l, \Pi^l)$$

Similarly, the coefficients for the high inflation equilibrium (Π^h, m^h) at $g = 0$ are given by

$$\alpha^h = \begin{pmatrix} -(\Pi^h)^2 \frac{w_2^h}{1-\sigma} \\ 0 \end{pmatrix} \tag{59a}$$

$$\beta_0^h = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{59b}$$

$$\beta_1^h = \begin{pmatrix} \Pi^h \frac{w_2^h}{1-\sigma} & 0 \\ 0 & 0 \end{pmatrix} \tag{59c}$$

$$\delta^h = \begin{pmatrix} 0 & \frac{w_1^h}{1-\sigma} \\ 0 & \frac{1}{\Pi^h} \end{pmatrix} \tag{59d}$$

34

where w_1^h and w_2^h are now evaluated at (m^h, Π^h).

8.3.2 Minimum State Variable Solutions

Consider a stochastic linear expectational difference equation of the form

$$x_t = k + B_0 E_{t-1}\left[x_t\right] + B_1 E_{t-1}\left[x_{t+1}\right] + Dx_{t-1} + u_t \tag{60}$$

with $x_t, u_t, k \in R^n$, $B_0, B_1, D \in R^{n \times n}$, and $B_1 \neq 0$, $D \neq 0$. The minimum state variable solutions of (60) take the form

$$x_t = a + Bx_{t-1} + u_t$$

provided there exists a real solution to the matrix quadratic equation

$$B_1 B^2 - (B_0 - I)B + D = 0 \tag{61}$$

see chapter 10 in Evans and Honkapohja (2001). Then a is given by

$$(I - B_0 - B_1(1 + B))a - k = 0 \tag{62}$$

The minimum state variable rational expectations solutions can be calculated by solving the matrix equations (61) for B and then using (62) to calculate a where k, B_1, B_2, and D are given by the linearization coefficients in appendix 8.3.1. Some lengthy algebra shows that around the low inflation steady state (Π^l, m^l) there are two AR(1) rational expectations solutions given by

$$\begin{pmatrix} \Pi_t \\ m_t \end{pmatrix} = a^{l,1} + B^{l,1} \begin{pmatrix} \Pi_{t-1} \\ m_{t-1} \end{pmatrix} + \begin{pmatrix} 0 \\ v_t \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ m^l \end{pmatrix} + \begin{pmatrix} 0 & \frac{1}{m^l} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \Pi_{t-1} \\ m_{t-1} \end{pmatrix} + \begin{pmatrix} 0 \\ v_t \end{pmatrix}$$

and

$$\begin{pmatrix} \Pi_t \\ m_t \end{pmatrix} = a^{l,2} + B^{l,2} \begin{pmatrix} \Pi_{t-1} \\ m_{t-1} \end{pmatrix} + \begin{pmatrix} 0 \\ v_t \end{pmatrix}$$

$$= \begin{pmatrix} 1 + m^l \frac{w_1^l}{w_2^l} \\ -\left(m^l\right)^2 \frac{w_1^l}{w_2^l} \end{pmatrix} + \begin{pmatrix} 0 & -\frac{w_1^l}{w_2^l} \\ 0 & 1 + m^l \frac{w_1^l}{w_2^l} \end{pmatrix} \begin{pmatrix} \Pi_{t-1} \\ m_{t-1} \end{pmatrix} + \begin{pmatrix} 0 \\ v_t \end{pmatrix}$$

With the eigenvalues of $B^{l,1}$ being equal to zero, this solution is stationary. However, since $m^l \frac{w_1^l}{w_2^l} > 0$ the solution $\left(a^{l,2}, B^{l,2}\right)$ has an eigenvalue large than one, which implies that it is explosive.

Around the high inflation steady state (Π^h, m^h) there is a single AR(1) rational expectations solution given by

$$
\begin{pmatrix} \Pi_t \\ m_t \end{pmatrix} = a^{h,2} + B^{h,2} \begin{pmatrix} \Pi_t \\ m_t \end{pmatrix} + \begin{pmatrix} 0 \\ v_t \end{pmatrix}
$$

$$
= \begin{pmatrix} \Pi^h \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & -\frac{w_1^h}{w_2^h} \\ 0 & \frac{1}{\Pi^h} \end{pmatrix} \begin{pmatrix} \Pi_{t-1} \\ m_{t-1} \end{pmatrix} + \begin{pmatrix} 0 \\ v_t \end{pmatrix}
$$

Since $\Pi^h > 1$ the eigenvalues of $B^{h,2}$ are inside the unit circle, which implies that the solution is asymptotically stationary. For $g >$ the linearization coefficients around $(\Pi^h(g), m^h(g))$ are given by

$$
\alpha^h = \begin{pmatrix} -\left(\Pi^h\right)^2 \frac{w_2^h}{1-\sigma} \\ m^h + \frac{m^h w_2}{(1-\sigma)} \end{pmatrix} \tag{63a}
$$

$$
\beta_0^h = \begin{pmatrix} 1 - \frac{m^h w_1^h}{\Pi^h(1-\sigma)} & 0 \\ -\frac{m^h}{\left(\Pi^h\right)^2} + \frac{\left(m^h\right)^2 w_1^h}{\left(\Pi^h\right)^3 (1-\sigma)} & 0 \end{pmatrix} \tag{63b}
$$

$$
\beta_1^h = \begin{pmatrix} \Pi^h \frac{w_2^h}{1-\sigma} & 0 \\ -\frac{m^h w_2^h}{\Pi^h(1-\sigma)} & 0 \end{pmatrix} \tag{63c}
$$

$$
\delta^h = \begin{pmatrix} 0 & \frac{w_1^h}{1-\sigma} \\ 0 & \frac{1}{\Pi^h} - \frac{m^h w_1}{\left(\Pi^h\right)^2 (1-\sigma)} \end{pmatrix} \tag{63d}
$$

and there exists another rational expectations solution, which ceases to exist at $g = 0$:

$$
\begin{pmatrix} \Pi_t \\ m_t \end{pmatrix} = a^{h,1} + B^{h,1} \begin{pmatrix} \Pi_t \\ m_t \end{pmatrix} + \begin{pmatrix} 0 \\ v_t \end{pmatrix}
$$

$$
= \begin{pmatrix} 0 \\ m^h \end{pmatrix} + \begin{pmatrix} 0 & \frac{\Pi^h}{m^h} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \Pi_{t-1} \\ m_{t-1} \end{pmatrix} + \begin{pmatrix} 0 \\ v_t \end{pmatrix}
$$

Since both eigenvalues of $B^{h,1}$ are equal to zero this solution is stationary.

8.3.3 Proof of Proposition 2

The differential equation (31) is stable (unstable) at some REE (a^*, B^*) if all eigenvalues (some eigenvalue) of

$$
\frac{\partial vec\left(T(a, B)\right)}{\partial vec(a, B)}
$$

at $(a, B) = (a^*, B^*)$ are smaller than one (is larger than one), where $vec(\cdot)$ denotes the (columnwise) vectorization operator. Let

$$T(a, B) = (T^a(a, B), T^B(a, B))$$

where T^a denotes the implied ALM for a, and T^B the ALM for B. Since $T^B(a, B)$ is independent of a, see (30), one has to consider only the eigenvalues of $\frac{\partial \, vec\, T^B(B)}{\partial \, vec\, B}$ and $\frac{\partial T^a(a,B)}{\partial a}$. The following table lists the eigenvalues of these two matrices for the respective rational expectations solutions at $g = 0$:[19]

RE-Solution	EV's of $\frac{\partial \, vec\, T^B(B)}{\partial \, vec\, B}$	EV's of $\frac{\partial T^a(a,B)}{\partial a}$
$a^{l,1}, B^{l,1}$	$\lambda_1 = \lambda_2 = 0, \lambda_3 = \lambda_4 = 1 - \frac{w_2^l}{1-\sigma} - m^l \frac{w_1^l}{1-\sigma}$	$\lambda_5 = 0, \lambda_6 = 1 - m^l \frac{w_1^l}{1-\sigma}$
$a^{l,2}, B^{l,2}$	$\lambda_1 = \lambda_2 = 0, \lambda_3 = 1, \lambda_4 = 1 + \frac{w_2^l}{1-\sigma} + m^l \frac{w_1^l}{1-\sigma}$	$\lambda_5 = 0, \lambda_6 = 1 + w_2^l$
$a^{h,1}, B^{h,1}$	$\lambda_1 = \lambda_2 = 0, \lambda_3 = \lambda_4 = 1$	$\lambda_5 = 0, \lambda_6 = 1 + (\Pi^h - 1)\frac{w_2^h}{1-\sigma}$
$a^{h,2}, B^{h,2}$	$\lambda_1 = \lambda_2 = 0, \lambda_3 = 1, \lambda_4 = 1 + \frac{w_2^h}{1-\sigma}$	$\lambda_5 = 0, \lambda_6 = 1 + \Pi^h \frac{w_2^h}{1-\sigma}$

Since $w_1^n > 0$, $w_2^n > 0$ for $n = l, h$ and $\Pi^h > 1$, it follows that the rational expectations solution $(a^{l,1}, B^{l,1})$ is E-stable and the solutions $(a^{l,2}, B^{l,2})$, $(a^{h,1}, B^{h,1})$ and $(a^{h,2}, B^{h,2})$ are E-unstable.

8.3.4   The Restricted Perceptions Equilibrium

Suppose agents forecast using the PLM (32), which implies $E_{t-1}(\Pi_t) = E_{t-1}(\Pi_{t+1}) = a_{t-1}$, where a_{t-1} is the $t-1$ estimate of a. For given $a_{t-1} = a$ the implied ALM (26) is stationary for the linearization coefficients (58) at the low inflation steady state if and only if $\left| 1 - m^l \frac{w_1^l}{1-\sigma} \right| < 1$. Since $m^l > 0$, $w_1^l > 0$, and $\sigma \in (0, 1)$ this is the case if and only if $\varepsilon_{n,w} > \frac{1}{2}$. Provided the elasticity condition holds the mapping from the PLM to the ALM, $T(a)$, is given by the projection of the ALM (26) with expectations given by a onto the PLM (32), i.e. $T(a) = E(\Pi(a))$. From (26) it follows that $T(a) = 1$. Therefore the stochastic differential equation determining E-stability of the restricted perceptions equilibrium

$$\frac{da}{d\tau} = T(a) - a = 1 - a$$

has a unique fixed point at $a = 1$ (the low steady state inflation rate) which is always E-stable (provided $\varepsilon_{n,w} > \frac{1}{2}$). It follows that the ALM (26) with $E_{t-1}(\Pi_t) = E_{t-1}(\Pi_{t+1}) = 1$ fluctuates around the perfect foresight steady state with deviations from it converging to zero as the support of the shock v_t vanishes.

---

[19] For the solution $(a^{h,1}, B^{h,1})$ it lists the limits of the eigenvalues for $g \to 0$.

References

Adam, Klaus, "Adaptive Learning and Cyclical Behavior of Output and Inflation," *University of Frankfurt Mimeo, www.wiwi.uni-frankfurt.de/professoren/wieland/adam,* 2002.

Arifovic, Jasmina, "Genetic Algorithms and Inflationary Economies," *Journal of Monetary Economics*, 1995, *36*, 219–243.

Bullard, James B. and Kaushik Mitra, "Learning About Monetary Policy Rules," *Journal of Monetary Economics (forthcoming)*, 2002.

Cagan, P., *The Monetary Dynamics of Hyper-Inflation*, in Milton Friedman (ed.), "Studies in the Quantity Theory of Money", University of Chicago Press, 1956.

Dixit, Avinash K. and Joseph E. Stiglitz, "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 1977, *67*, 297–308.

Duffy, John, "On Learning and the Nonuniqueness of Equilibrium in an Overlapping Generations Model with Fiat Money," *Journal of Economic Theory*, 1994, *64*, 541–553.

Evans, George W. and Seppo Honkapohja, "Convergence of Least Squares Learning to a Non-Stationary Equilibrium," *Economics Letters*, 1994, *46*, 131–136.

___ and ___ , "Economic Dynamics with Learning: New Stability Results," *Review of Economic Studies*, 1998, *65*, 23–44.

___ and ___ , "Convergence for Difference Equations with Vanishing Time Dependence, With Applications to Adaptive Learning," *Economic Theory*, 2000, *15*, 717–725.

___ and ___ , *Learning and Expectations in Macroeconomics*, Princeton: Princeton University Press, 2001.

___ , ___ , and **Ramon Marimon**, "Convergence in Monetary Inflation Models with Heterogeneous Learning Rules," *Macroeconomic Dynamics*, 2001, *5*, 1–31.

Galí, Jordi and Mark Gertler, "Inflation Dynamics: A Structural Econometric Analysis," *Journal of Monetary Economics*, 1999, *44*, 195–222.

Grandmont, Jean-Michel, "On Endogenous Competitive Business Cycles," *Econometrica*, 1985, *53*, 995–1046.

Hirsch, Morris and Stephen Smale, *Differential Equations, Dynamical Systems, and Linear Algebra*, New York: Academic Press, 1974.

Horn, Roger and Charles Johnson, *Matrix Analysis*, Cambridge: Cambridge University Press, 1985.

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

⎯⎯ **and Thomas Sargent**, "Least Squares Learning and the Dynamics of Hyperinflation," in W. Barnett, J. Geweke, and K. Shell, eds., *Economic Complexity: Chaos, Sunspots, Bubbles, and Nonlinearity*, Cambridge University Press 1989, pp. 119–137.

Marimon, Ramon and Shyam Sunder, "Indeterminacy of Equilibria in a Hyperinflationary World: Experimental Evidence," *Econometrica*, 1993, *61(5)*, 1073–1107.

Sargent, Thomas J. and Neill Wallace, "Inflation and the Government Budget Constraint," in Assaf Razin and Efraim Sadka, eds., *Economic Policy in Theory and Practice*, Macmillan 1987.

Van-Zandt, Timothy and Martin Lettau, "Robustness of Adaptive Expectations as an Equilibrium Selection Device," *Macroeconomic Dynamics (forthcoming)*, 2002.

Woodford, Michael, "Learning to Believe in Sunspots," *Econometrica*, 1990, *58*, 277–307.

CFS Working Paper Series:

No.	Author(s)	Title
2002/08	**Andreas Bascha** **Uwe Walz**	Financing Practices in the German Venture Capital Industry An Empirical Assessment
2002/09	**Marco Hoeberichts**	The Credibility of Central Bank Announcements
2002/10	**Markus Haas** **Stefan Mittnik** **Marc S. Paolella**	Mixed Normal Conditional Heteroskedasticity
2002/11	**Elke Hahn**	Core Inflation in the Euro Area: An Application of the Generalized Dynamic Factor Model
2002/12	**Yunus Aksoy** **Hanno Lustig**	On the Short and Long Term Real Effects of Nominal Exchange Rates
2002/13	**Andreas A. Jobst**	Collateralised Loan Obligations (CLOs) - A Primer
2002/14	**Andreas A. Jobst**	The Pricing Puzzle: The Default Term Structure of Collateralised Loan Obligations
2002/15	**Athanasios Orphanides**	Activist Stabilization Policy and Inflation: The Taylor Rule in the 1970s
2003/01	**Klaus Adam**	Learning to Forecast and Cyclical Behavior of Output and Inflation
2003/02	**Klaus Adam**	On the Relation between Robust and Bayesian Decision Making
2003/03	**Klaus Adam**	Learning and Equilibrium Selection in a Monetary Overlapping Generations Model with Sticky Prices

Copies of working papers are available at the Center for Financial Studies or can be downloaded (http://www.ifk-cfs.de).