Working Paper Series F-68

Learnability of an Equilibrium with Private Information

Ryuichi Nakagawa

Faculty of Economics
Kansai University
rc-naka@kansai-u.ac.jp

June 4, 2014

Economic Society of Kansai University
Osaka, 564-8680 Japan

Learnability of an Equilibrium with Private Information[1]

Ryuichi Nakagawa[2]
Faculty of Economics
Kansai University

This version: June 4, 2014

[1]The author is grateful to George W. Evans, Bruce McGough, and participants at the University of Oregon Macro Group, the Econometric Society, the European Economic Association, and the Royal Economic Society for their helpful comments and suggestions. For financial support, the author thanks KAKENHI (No. 20730139, No. 23530234), Murata Science Foundation, Nomura Foundation for Academic Promotion, and Zengin Foundation for Studies on Economics and Finance.

[2]Address: Faculty of Economics, Kansai University, 3-3-35 Yamate Suita, Osaka 564-8680, Japan. Phone: +81-6-6368-0590. Fax: +81-6-6339-7704. E-mail: ryu-naka@kansai-u.ac.jp. URL: http://www2.itc.kansai-u.ac.jp/∼ryu-naka/.

Learnability of an Equilibrium with Private Information

Abstract

This paper investigates the learnability of an equilibrium with private information. Our model consists of different types of agents where agents of each type have their own private information about an exogenous variable and conduct adaptive learning with a heterogeneously misspecified perceived laws of motion (PLM) that includes only this variable. The analytical result is that the learnability of the equilibrium is weakly increased as PLMs of different types become heterogeneous and misspecified by private information. The numerical analysis shows that in a New Keynesian model with private information about fundamental shocks, the Taylor principle of monetary policy continues to be a sufficient condition for the learnability. These results are applicable to a broad class of equilibria attainable under heterogeneous and/or misspecified learning.

JEL classification: C62; D83; E52

Keywords: Adaptive learning; Private information; Heterogeneous misspecification; Learnability; Taylor principle

1 Introduction

Rational expectations are based on an unlikely assumption that agents have perfect knowledge about the structure of the economy. Recent macroeconomic research has incorporated the concept of *adaptive learning* as an alternative framework in which agents are assumed to formulate their forecasts by estimating econometric models (i.e., perceived laws of motion, PLMs) through least-squares techniques (see Bray, 1982; Evans and Honkapohja, 2001, 2008). In this framework, it has been investigated whether an equilibrium attainable under adaptive learning is stable, that is, *learnable*. If agents' beliefs and their forecasts converge around the equilibrium, the equilibrium is considered to be learnable. Learnability has been increasingly emphasized as a criterion for selecting a stable equilibrium among nonexplosive multiple equilibria (e.g., Honkapohja and Mitra, 2004a). The necessary conditions imposed on macroeconomic policies for learnability have been investigated (e.g., Bullard and Mitra, 2002; Evans and Honkapohja, 2006; Guse, 2008; Anufriev et al., 2013).

Learnability has been analyzed in different structures of information sets of economic variables held by agents. Evans and Honkapohja (2001) provide a benchmark framework where each agent has a full information set to form a correctly specified PLM including all relevant variables (hereafter, CS learning). Branch (2004), Guse (2008), and Hommes and Zhu (2014) focus on the *restricted perceptions equilibrium* (hereafter, RPE) where agents' information sets are limited so that they are constrained to form underparameterized PLMs excluding unobservable variables. Adam et al. (2006) and Berardi (2007) introduce heterogeneity in agents' information sets, where a fraction of agents have limited information sets while other agents have full information sets. Honkapohja and Mitra (2004b) and Muto (2011) assume that the private sector and the central bank have different information sets.

This paper investigates the learnability of an equilibrium in the presence of private information about economic variables. There exist economic variables that are observable by specific agents and unobservable by others. For example, a preference shock possessed by a household continues to be observable only for this household (see Allen and Gale, 2004). In financial markets, the profitability of a borrower tends to be observable only by this borrower (see Stiglitz and Weiss, 1981). Then, our model consists of different types of agents, and agents of each type have their own private information about an exogenous variable, which makes information sets held by different types of agents limited and heterogeneous.[1] Under adaptive learning, each type of agent is constrained to form a *heterogeneously misspecified* PLM that includes only this observable variable (hereafter, HM learning). In this framework, there exists a generalization of the RPE, which we call a *heterogeneous misspecification equilibrium* (hereafter, HME). By investigating the learnability of the HME in a general multivariate expectations model, this paper provides analytical results about whether and how the learnability is affected by the existence of private information. Next, in a basic New Keynesian (NK) macroeconomic model, this paper evaluates analytically and numerically the impact of private information on learnability conditions imposed on monetary policy with a contemporaneous data interest rate rule. In addition, this paper examines the robustness of the impact when agents include lagged endogenous variables into their PLMs as public information.

The result of this paper is that the learnability of an equilibrium is weakly increased by the existence of private information about exogenous variables. Specifically, the learnability is weakly increased as the degree of heterogeneity in PLMs caused by private information is increased. If exogenous variables have similar stochastic characteristics, the learnability is also weakly increased as the degree of misspecification in PLMs increases. Honkapohja and Mitra (2006) show that an equilibrium under several types of heterogeneous learning is largely more learnable than the equilibrium under CS learning.[2] Our results complement their results by indicating that the heterogeneity in PLMs also leads to strong learnability of an equilibrium.

Second, in the NK model with private information about fundamental shocks, this paper shows that the *Taylor principle*—to raise nominal interest rates more than one-for-one in re-

[1] Branch (2007) provides evidence that information sets are limited and heterogeneous using the Michigan survey of inflation expectations. Bovi (2013) uses the survey data of the European Commission to show that heterogeneous beliefs are persistent in UK citizens.

[2] Honkapohja and Mitra (2006) consider 1) heterogeneous initial beliefs on parameters in agents' homogeneous PLMs, 2) heterogeneous learning algorithms (e.g., updating functions or gain parameters), and 3) structural heterogeneity in which the forecasts of different agents have different effects on the dynamics of the economy.

sponse to an increase in inflation (see Taylor, 1993)—is a sufficient condition to ensure the learnability of an equilibrium. Bullard and Mitra (2002) obtain the same result under CS learning. Our result confirms that their result is robust to the existence of private information about fundamental shocks.

Finally, the numerical analysis finds that the impact of private information on learnability is economically significant. In particular, the impact becomes more significant as the degree of heterogeneity in PLMs is increased, while the impact is unchanged even if the degree of misspecification in PLMs is increased. Furthermore, if lagged economic variables are included in PLMs as public information, the impact of private information is reduced, but remains significant.

These results are applicable to models under adaptive learning with a variety of agents' information sets; for example, a full information set considered in the benchmark analysis of CS learning, and a limited information set of exogenous variables assumed in the analysis of the RPE. In our model, those information sets can be reproduced by accommodating the degree of limitation and heterogeneity of information sets of exogenous variables. Hence, the results of this paper are robust for a broad class of heterogeneous and/or misspecified learning.

This paper is closely related to the literature on adaptive learning in the presence of private information. Marcet and Sargent (1989a) establish learning schemes for an equilibrium under private information about economic variables. Branch and McGough (2011) study business cycle dynamics under adaptive learning and private information. Heinemann (2009) considers the existence of private noisy signals of economic variables and investigates their effects on learnability. It has not, however, been investigated whether/how the private information that continuously makes agents' PLMs heterogeneously underparameterized has an impact on learnability.

Our analysis is also related to the literature on heterogeneity and misspecification in learning. Adam et al. (2006) and Berardi (2007, 2009) consider heterogeneous and misspecified learning where a fraction of agents form underparameterized PLMs, while other agents form correctly specified PLMs. Their models can be interpreted as allowing the existence of private information in the sense that the latter agents have the full information that the former agents do not have, but their models do not include a plausible case where each type of agent has his/her own private information and forms a heterogeneously misspecified PLM. Meanwhile, Honkapohja and Mitra (2004b) and Muto (2011) consider a symmetric case where the private sector and the central bank have their own private information. However, their models include not only the heterogeneity in PLMs but also structural heterogeneity, and hence the impact of the former heterogeneity is not indicated independently.

Furthermore, heterogeneously misspecified PLMs in our paper are considered in the context of the literature introducing *dynamic predictor selection*, which was originally established by Brock and Hommes (1997). Branch and Evans (2006, 2007), for example, analyze an economy where agents are allowed to choose among a list of heterogeneously misspecified PLMs (whereas in the present paper, each type of agent is constrained to form a specific heterogeneously misspecified PLM).[3] Their model appears to allow the existence of private information, but the model premises homogeneity in agents' information sets so that all agents can choose among the same list of the PLMs. Actually, the existence of private information causes intrinsic heterogeneity in information sets and constrains each agent to hold a specific heterogeneously misspecified PLM. Thus, our paper focuses on the framework of heterogeneous information sets, which is different from that in dynamic predictor selection models.

The paper is structured as follows. The next section presents our model and provides a benchmark analysis about learnability without private information, that is, CS learning. Section 3 introduces HM learning with private information and investigates the dynamics of the HME. Section 4 examines the impact of private information on the learnability of an equilibrium. Section 5 evaluates numerically the impact in a basic NK model. Finally, we present our conclusions.

[3] Dynamic predictor selection is also employed in different ways by Branch and McGough (2010), Guse (2010), Berardi (2011), Anufriev et al. (2013), and Pfajfar (2013).

2 Model

2.1 Setup

We establish the general form of the multivariate linear expectations model. The economy is represented by two vector equations:

$$y_t = A + BE_t^* y_{t+1} + Cw_t, \quad (1)$$

$$w_t = \Phi w_{t-1} + v_t. \quad (2)$$

The equations represent the dynamics of the endogenous variables and the evolution of the exogenous variables. The economy has m endogenous variables and n exogenous variables. y_t is an $m \times 1$ vector of endogenous variables at time t. $w_t = (w_{1t}, \ldots, w_{nt})'$ is an $n \times 1$ vector of autoregressive exogenous variables. The standard deviation of w_{it} for each i is defined by $\sigma_{ii} > 0$, and the correlation matrix of w_t is defined by $\Gamma \equiv (\rho_{ij})_{1 \leq i,j \leq n}$, where $\rho_{ij} = \rho_{ji}$ and $\rho_{ij} \in [0, 1]$ denotes the correlation between w_i and w_j for each $i, j \in \{1, \ldots, n\}$. v_t is an $n \times 1$ vector of fundamental shocks with means of zero that drive the stochastic process of w_t.[4] Matrix A is an $m \times 1$ vector of constant terms. B is an $m \times m$ coefficient matrix of $E_t^* y_{t+1}$. C is an $m \times n$ coefficient matrix of w_t, and Φ is an $n \times n$ matrix of coefficients of w_t. E_t^* is the operator of the aggregate expectation of y_{t+1} at time t, which is not necessarily rational under adaptive learning. The model is purely forward-looking, but it can be applied to the analysis of models with lagged endogenous variables y_{t-1}.

For ease of calculation, we impose regularity assumptions on these parameters in Appendix A. In particular, Φ is assumed to be a diagonal and nonnegative matrix whose diagonal elements exist in the interval $[0, 1)$: $\Phi \equiv diag\,(\varphi_i)_{1 \leq i \leq n}$ where $0 \leq \varphi_i < 1$ for each i. In addition, Γ is assumed to be a nonnegative matrix, in which $0 \leq \rho_{ij} \leq 1$ for each $i, j \in \{1, ..., n\}$. These assumptions are not crucial for our analysis, as most stationary linear models in the literature can be transformed to satisfy these conditions.

[4]Note that exogenous variables with nonzero and heterogeneous means can be transformed to the form (2).

2.2 CS Learning

Before considering the existence of private information, let us review the benchmark of adaptive learning with a full information set to make it easier later to highlight the impact of private information. If agents could formulate rational expectations, the system (1)–(2) has a nonexplosive fundamental REE, which takes the form of a minimal state variable (hereafter, MSV) solution (see McCallum, 2004):

$$y_t = a + cw_t + \varepsilon_t, \tag{3}$$

where a is an $m \times 1$ vector of constant terms, c is an $m \times n$ coefficient matrix for w_t, and ε_t is an $m \times 1$ vector of error terms that are perceived to be white noise. Using the method of undetermined coefficients, the solution $(\bar{a}, \bar{c})$ is uniquely obtained as

$$\bar{a} = (I_m - B)^{-1} A, \tag{4}$$

$$\bar{c} = B\bar{c}\Phi + C, \tag{5}$$

where I_m is an m-dimensional identity matrix. Note that the constant terms vector $\bar{a}$ corresponds to the steady state of the fundamental REE.

If agents do not have enough knowledge to develop rational expectations, agents might adopt adaptive learning using all available data to formulate their forecast $E_t^* y_{t+1}$. If all economic variables up to time t, $\{y_s, w_s\}_{s=1}^t$, are observable for all agents, they can estimate a correctly specified PLM of the form of the MSV solution (3). Using the estimated parameters (a, c), agents formulate $E_t^* y_{t+1} = a + c\Phi w_t$, which is incorporated into Eq. (1) and yields the actual law of motion (hereafter, ALM) of the economy, $y_t = (A + Ba) + (Bc\Phi + C)\, w_t$.[5] Evans and Honkapohja (2001, chapter 2) show that the global convergence of (a, c) in real-time learning through least-squares techniques is governed by the ordinary differential equation (hereafter,

[5]Forming $E_t^* y_{t+1}$ using an information set including current variables is considered by Honkapohja and Mitra (2006), McCallum (2007), Ellison and Pearlman (2011), and Bullard and Eusepi (2014). Evans and Honkapohja (2001) indicate that excluding those variables from information sets might provide different conclusions about the learnability of an equilibrium, but the analysis of alternative frameworks is omitted here because of limited space.

ODE):

$$\frac{d}{d\tau}(a,c) = T(a,c) - (a,c), \tag{6}$$

where $T(a,c) \equiv (T_a(a), T_c(c)) = (A + Ba, Bc\Phi + C)$ is the mapping from the PLM (a,c) to the ALM $T(a,c)$, and τ denotes notional time. If the ODE is globally asymptotically stable, then (a,c) converges to the solution (4)–(5), meaning that the fundamental REE is found to be learnable under CS learning.

3 HM Learning

In what follows, we consider adaptive learning by agents with private information and find the dynamics of an equilibrium under HM learning.

3.1 Private Information

First, we introduce private information about exogenous variables w_t, such that each agent has access to information on only a subset of those variables (see Marcet and Sargent, 1989a).[6]

Assumption 1 *For any $i \in \{1, ..., n\}$, the evolution of an exogenous variable $\{w_{is}\}_{s=1}^{t}$ is observable for the proportion $\frac{1}{n}$ of agents (hereafter, type i) and unobservable for agents of other types.*

For analytical tractability, the population of each type is assumed to be the same at $\frac{1}{n}$, but the distribution of the populations of different types are not important for our analysis (see Section 4.3). Assumption 1 implies that agents of type i recognize the stochastic characteristics of w_{it} and do not recognize the correlations $\left\{\rho_{ij} = \rho_{ji}\right\}_{j\neq i}^{n}$ and the quantity n of exogenous variables. In this situation, the agent of type i has the set of public and private information $\{y_s, w_{is}\}_{s=1}^{t}$, which is limited and different from the information sets held by other types in

[6]Heinemann (2009) establishes another framework of private information in which agents receive different private noisy signals about a single economic variable.

terms of $\{w_{it}\}_{i=1}^{n}$.[7] This type of private information might describe a feature of idiosyncratic shocks held by individual agents in the economy; for example, a preference shock in an agent's utility (see Allen and Gale, 2004).

Under Assumption 1, the number n and the correlations $\left\{\rho_{ij}\right\}_{i,j=1}^{n}$ of exogenous variables may be treated as the measures of the *limitation* and *heterogeneity* in information sets, respectively. If $n = 1$, the information set of each agent is reduced to the full information set in Section 2.2. The larger n is, the more limited each information set is relative to the full one. Thus, n represents not only the quantity of privately observable variables but also the degree of limitation of each information set. Similarly, if $\rho_{ij} = \rho_{ji} = 1$ for some types i and j (and hence, $\varphi_i = \varphi_j$ and $\frac{w_{it}}{\sigma_{ii}} = \frac{w_{jt}}{\sigma_{jj}}$), the information sets of both types are perfectly homogeneous as if both types observed the same variable. The smaller ρ_{ij} is, the more heterogeneous are both information sets. Thus, the correlations represent the degree of heterogeneity in information sets of different types.

3.2 Heterogeneous Misspecification Equilibrium

Next, we consider adaptive learning with the limited and heterogeneous information sets (that is, HM learning). In contrast to agents with the full information set in Section 2.2, the agent of type i with the information set $\{y_s, w_{is}\}_{s=1}^{t}$ is constrained to estimate a heterogeneously misspecified PLM:

$$y_t = a_i + c_i w_{it} + \varepsilon_{it}, \tag{7}$$

which underparameterizes the MSV solution (3) and differs from the PLMs of other types. Parameters a_i and c_i are $m \times 1$ vectors of constant terms and coefficients, and ε_{it} is an $m \times 1$ vector of error terms that are perceived to be white noise.[8] The agent estimates the parameter

[7] Note that Marcet and Sargent (1989a) consider the private information of not only exogenous variables but also endogenous variables. They also consider the existence of hidden state variables that are unobservable by all agents. Our assumption is designed to focus on clarifying the impact of private information on the learnability of an equilibrium.

[8] Agents might include the vector of lagged endogenous variables y_{t-1} in the PLM as public information. This case will be analyzed in Sections 4.2.3 and 5.3.2. On the other hand, we do not assume that agents include sunspot variables in their PLMs, because under the information sets that include current endogenous variables

matrix $\phi_i' \equiv (a_i, c_i)$ using a recursive least-squares (hereafter, RLS) method, but does not recognize the misspecification in the PLM (7) as w_{it} and ε_{it} are orthogonalized. Finally, using the PLM (7) and the estimated ϕ_i, the agent formulates the forecast $E_{it}^* y_{t+1}$ as

$$E_{it}^* y_{t+1} = a_i + c_i \varphi_i w_{it}, \tag{8}$$

where E_{it}^* is the operator of expectations formed by type i at time t.

In this framework, the *misspecification* and *heterogeneity* in the PLMs are characterized as follows:

Lemma 1 *Given Assumption 1 and the PLMs of the form (7) for all $i \in \{1, \ldots, n\}$,*

1. *the PLMs become more misspecified as the degree of limitation of each information set, n, increases;*

2. *the PLMs become more heterogeneous as the degree of heterogeneity of information sets, ρ_{ij} for any i, j, decreases.*

Later, the impacts of private information will be found by examining the effects of n and $\{\rho_{ij}\}_{i,j=1}^n$ on the learnability conditions.[9]

The aggregate PLM and the aggregate forecast $E_t^* y_{t+1}$ in Eq. (1) are determined by Eqs. (7) and (8) of all types. For simplicity, let us assume that forecasts of different types $\{E_{it}^* y_{t+1}\}_{i=1}^n$ have equal contributions to the dynamics of the economy.[10] Then, the aggregate PLM is defined as the average of the PLMs of all types in the same form as the MSV solution (3):

$$y_t = a + c w_t + \frac{1}{n} \sum_{i=1}^{n} \varepsilon_{it}, \tag{9}$$

y_t, sunspot equilibria cannot be learnable (see Evans and Honkapohja, 2001, section 10.5.1). Although Evans and McGough (2005) find a specific situation in which sunspot equilibria are learnable under those information sets, such a situation is treated as a special case and will not be discussed in our paper.

[9] The literature measures the degree of heterogeneity in PLMs by the proportions of different agents who specify different forms of PLMs (e.g., Branch and Evans, 2006; Berardi, 2007), but their heterogeneity can be reproduced in our model by accommodating n and $\{\rho_{ij}\}_{i,j=1}^n$ (see Section 4.3).

[10] Note that the proportions of contributions of the different forecasts do not affect our results (see Section 4.3).

where $a \equiv \frac{1}{n}\sum_{i=1}^{n} a_i$ is the average of the constant term vectors for all types, and $c \equiv \frac{1}{n}(c_1, ..., c_n)$ is an $m \times n$ matrix that combines the coefficients $\{c_i\}_{i=1}^{n}$ of the PLMs of the form (7) for all types and multiplies them by $\frac{1}{n}$. With the aggregate PLM (9), the aggregate forecast $E_t^* y_{t+1}$ is formulated by

$$E_t^* y_{t+1} = a + c\Phi w_t, \tag{10}$$

where E_t^* is the operator of the average of heterogeneous forecasts.

The ALM of the economy depends upon $E_t^* y_{t+1}$ in Eq. (10). Substituting Eq. (10) into the system (1)–(2), the ALM is determined by

$$y_t = (A + Ba) + (Bc\Phi + C)\, w_t. \tag{11}$$

The stability of the equilibrium attainable under HM learning is subject to whether the aggregate parameters $\phi' \equiv (a, c)$ converge to bounded values; the dynamics of ϕ is established by the real-time learning processes of agents of all types for $\{\phi_i\}_{i=1}^{n}$. If the PLM (7) were correctly specified, the *E-stability principle* of Evans and Honkapohja (2001, chapter 2) would hold; that is, the convergence of parameters in real-time learning would be characterized by the ODE that could be made by a mapping from the PLM (9) to the ALM (11). Under HM learning, where the PLMs are underparameterized, the principle does not necessarily hold, and the convergence of $\{\phi_i\}_{i=1}^{n}$ is inferred from the stochastic recursive algorithms of $\{\phi_i\}_{i=1}^{n}$ formulated by the PLMs of the form (7) for all i and the ALM (11). As a result, we find that the global convergence of $\{\phi_i\}_{i=1}^{n}$ is governed by the following associated ODE for aggregate parameters (a, c):

$$\frac{d}{d\tau}(a, c) = (T_a(a), T_c(c)) - (a, c), \tag{12}$$

where τ denotes notional time and

$$\begin{aligned} T_a(a) &\equiv A + Ba, \\ T_c(c) &\equiv (Bc\Phi + C)\left(\frac{1}{n}\Psi\right), \\ \Psi &\equiv diag(\sigma_{ii})_{1 \le i \le n} \cdot \Gamma \cdot diag(\sigma_{ii})_{1 \le i \le n}^{-1}. \end{aligned}$$

The derivation of the ODE is in Appendix B. The mapping $(T_a(a), T_c(c))$ provides the coefficients of the forecasts of y_t updated by agents of all types using their limited and heterogeneous information sets. The ODE of a in Eq. (12) is equivalent to each ODE of $\{a_i\}_{i=1}^n$ as a is an arithmetic average of $\{a_i\}_{i=1}^n$ which have the same form of ODEs. The ODE of c in Eq. (12) is a combination of the ODEs of $\{c_i\}_{i=1}^n$

The effects of misspecification n and heterogeneity $\{\rho_{ij}\}_{i,j=1}^n$ on the convergence of the aggregate parameters (a, c) are realized through $\frac{1}{n}\Psi$ in $T_c(c)$. Thus, the ODE (12) includes the benchmark ODE (6) under CS learning as follows:

Proposition 1 *The ODE (12) under HM learning is equivalent to the ODE (6) under CS learning*

1. *if* $n = 1$ *(no limitation) or*
2. *if* $\rho_{ij} = 1$ *for all* $i, j \in \{1, ..., n\}$ *(no heterogeneity).*

The first part is trivial, and the second part is proved by the fact that if $\rho_{ij} = 1$ for all i, j (and hence $\varphi_i = \varphi_j$ and $\frac{w_{it}}{\sigma_{ii}} = \frac{w_{jt}}{\sigma_{jj}}$ for all i, j), then $\left(\frac{1}{n}\Psi\right) w_t = w_t$, which makes the mapping $T_c(c)$ under HM learning equivalent to the mapping for c under CS learning in Eq. (6). The second part means that if privately observable variables are perfectly correlated, the existence of private information has no impact on the dynamics of the economy under adaptive learning although the PLMs of the form (7) remain misspecified at an individual level. This is intuitive because in this situation, all information sets of different types are perfectly homogeneous as if there exists no private information.

An equilibrium under HM learning is attained if the ODE are globally asymptotically stable such that parameter estimates (a, c) converge to the fixed point $(\bar{a}, \bar{c})$ of the ODE:

$$\bar{a} = (I_m - B)^{-1} A, \tag{13}$$

$$\bar{c} = (B\bar{c}\Phi + C)\left(\frac{1}{n}\Psi\right), \tag{14}$$

where the fixed point $\bar{a}$ corresponds to the steady state of the equilibrium. Let us call the equilibrium with the fixed point (13)–(14) the *heterogeneous misspecification equilibrium* (*HME*):

Definition 1 *The HME is a stationary stochastic process for $\{y_t\}_{t=0}^{\infty}$ following the system (1)–(2) given that $\{E_t^* y_{t+1}\}_{t=0}^{\infty}$ is formed by the aggregation of the PLMs of the form (7) for all i with the parameters $\{\phi_i' = (a_i, c_i)\}_{i=1}^{n}$ determined at the fixed point (13)–(14) of the ODE (12).*

It is trivial that Eqs. (13)–(14) have a unique solution:

Proposition 2 *There exists a unique HME.*

3.3 Observable Steady State

We also provide the ODE when the steady state $\bar{a}$ of the HME is observable for all agents. An observable steady state is a popular assumption in recent macroeconomic studies. In particular, a nonlinear macroeconomic model tends to be log-linearized around a steady state (or a point close to the steady state) by assuming the steady state to be observable (e.g., Bullard and Mitra, 2002; Mitra et al., 2013). If the steady state $\bar{a}$ is observable for agents of all types, agents of type i can estimate a PLM with a_i fixed at $\bar{a}$:

$$y_t = \bar{a} + c_i w_{it} + \varepsilon_{it}, \tag{15}$$

which is equivalent to the PLM excluding the constant term: $\tilde{y}_t = c_i w_{it} + \varepsilon_{it}$ where $\tilde{y}_t \equiv y_t - \bar{a}$ (see Slobodyan and Wouters, 2012).[11]

In this situation, the global convergence of $\{\phi_i' = (\bar{a}, c_i)\}_{i=1}^{n}$ of all types is subject to only the c part of the ODE (12). If the ODE is globally asymptotically stable, matrix c converges to the fixed point given by Eq. (14). Regardless of whether the steady state $\bar{a}$ is observable or not, the HME (13)–(14) is a unique equilibrium under HM learning.

[11] Note that if the means of $\{w_{it}\}_{i=1}^{n}$ are nonzero and/or heterogeneous, the fixed points of the constant terms of the PLMs of all types should be different from the steady state of y_t and/or heterogeneous. However, this case never violates our analytical results about the learnability.

4 Learnability of HME

By examining the learnability of the HME, we investigate the impact of private information on the learnability of an equilibrium. First, we provide the learnability conditions of the HME. Next, we examine the impact of heterogeneity and misspecification of the PLMs on the learnability, separately. Finally, we discuss the connections between the informational structures of our model and the related literature. For simplicity of exposition, we introduce the notation $\lambda[X]$ as the largest value of the real parts of the eigenvalues of the matrix X.

4.1 Learnability Conditions

If the steady state $\bar{a}$ is unobservable, the learnability of the HME is ensured by the stability of the ODE (12). According to Evans and Honkapohja (2001, section 6.6), the ODE is globally asymptotically stable if and only if their Jacobians,

$$
\begin{aligned}
D\left(T_a(a)-a\right) &= B - I_m, && (16)\\
D\left(T_c(c)-c\right) &= \left(\Phi\left(\frac{1}{n}\Psi\right)\right)' \otimes B - I_{mn}, && (17)
\end{aligned}
$$

have all negative real parts of eigenvalues; that is, $\lambda[B] < 1$ and $\lambda\left[\left(\Phi\left(\frac{1}{n}\Psi\right)\right)' \otimes B\right] < 1$, respectively.[12] If the steady state $\bar{a}$ is observable, the stability is governed solely by the c's ODE of (12), which is globally asymptotically stable if and only if Eq. (17) has all negative real parts of eigenvalues; that is, $\lambda\left[\left(\Phi\left(\frac{1}{n}\Psi\right)\right)' \otimes B\right] < 1$.

Matrix $\Phi\left(\frac{1}{n}\Psi\right)$ in the stability conditions has the following characteristics:

Lemma 2 *For any $n \geq 1$ and $i, j \in \{1, \cdots, n\}$,*

1. *all eigenvalues of $\Phi\left(\frac{1}{n}\Psi\right)$ are real and exist in the interval $[0,1)$;*
2. $\frac{d\lambda\left[\Phi\left(\frac{1}{n}\Psi\right)\right]}{d\rho_{ij}} \geq 0$*;*

[12] Actually, the ODE can also be asymptotically stable if they have one or more zero real parts of eigenvalues, which are ruled out in this paper as nongeneric cases.

3. $\frac{d\lambda\left[\Phi\left(\frac{1}{n}\Psi\right)\right]}{d\varphi_i} \geq 0$.

The proof of Lemma 2 is in Appendix C. Lemma 2.1 means $0 \leq \lambda\left[\Phi\left(\frac{1}{n}\Psi\right)\right] < 1$ so that if $\lambda[B] < 1$, then $\lambda\left[\left(\Phi\left(\frac{1}{n}\Psi\right)\right)' \otimes B\right] < 1$.

Therefore, the learnability conditions of the HME are summarized as follows:

Proposition 3 *When the steady state $\bar{a}$ is unobservable (respectively, observable), the HME is learnable if and only if $\lambda[B] < 1$ $\left(\lambda\left[\left(\Phi\left(\frac{1}{n}\Psi\right)\right)' \otimes B\right] < 1\right)$.*

4.2 Impact of HM Learning

Next, we clarify the impact of HM learning on the learnability. The degree of HM learning is represented by heterogeneity $\left\{\rho_{ij}\right\}_{i,j=1}^{n}$ and misspecification n in the PLMs (Lemma 1). We show the impacts of heterogeneity and misspecification in HM learning, separately. Proposition 3 indicates that if $\lambda[B] \leq 0$, the HME is always learnable regardless of the observability of $\bar{a}$. Hereafter, we restrict our discussions to the case of $\lambda[B] > 0$ when we provide the learnability conditions.

4.2.1 Impact of Heterogeneity

Proposition 3 means that when the steady state $\bar{a}$ is unobservable, the learnability of the HME is independent of $\left\{\rho_{ij}\right\}_{i,j=1}^{n}$. When $\bar{a}$ is observable and $\lambda[B] > 0$, the learnability is reduced as $\lambda\left[\left(\Phi\left(\frac{1}{n}\Psi\right)\right)' \otimes B\right]$ increases. According to Lemma 2.2, $\lambda\left[\Phi\left(\frac{1}{n}\Psi\right)\right]$ is weakly increased as $\left\{\rho_{ij}\right\}_{i,j=1}^{n}$ increase. Thus:

Proposition 4 *For any $n \geq 1$, when the steady state $\bar{a}$ is unobservable (respectively, observable), the learnability condition of the HME is independent of (weakly strengthened by) an increase in the correlations $\left\{\rho_{ij}\right\}_{i,j=1}^{n}$ of exogenous variables.*

That is, the existence of heterogeneity in the PLMs never makes an equilibrium less learnable. Furthermore, an increase in the degree of heterogeneity in the PLMs weakly increases the learnability.

The strong learnability of an equilibrium under heterogeneous learning results from the sluggish updating of the aggregate forecast $E_t^* y_{t+1}$. Under CS learning, if an exogenous variable w_{it} evolves, all agents update their forecasts $\{E_{it}^* y_{t+1}\}_{i=1}^n$ equivalently so that $E_t^* y_{t+1}$ is updated to the same degree. Under heterogeneous learning, in contrast, $E_t^* y_{t+1}$ is not updated to the same degree as under CS learning, because only the agent of type i updates his/her forecast $E_{it}^* y_{t+1}$, while the other types of agents do not to the same degree. If PLMs are highly heterogeneous by low correlations of $\{w_{it}\}_{i=1}^n$, an evolution of w_{it} is not synchronized with the evolutions of $\{w_{jt}\}_{j\neq i}^n$ to similar degrees, and hence the forecasts of the other types $\left\{E_{jt}^* y_{t+1}\right\}_{j\neq i}^n$ are not updated compared with the forecast of type i, $E_{it}^* y_{t+1}$. This sluggishness of the updating of the aggregate forecast $E_t^* y_{t+1}$ leads to the strong learnability of an equilibrium.[13] Note that when the steady state is unobservable, such sluggishness under heterogeneous learning disappears because forecasts of all agents are always updated through the estimation of the constant terms $\{a_i\}_{i=1}^n$.

This result provides an important result on the impact of HM learning caused by private information as a whole. Although the proposition shows the impact of only heterogeneity in the PLMs on the learnability, this feature holds for any degree of misspecification ($n \geq 1$). In addition, the HME is the least learnable if the heterogeneity vanishes under the perfect correlations $\left\{\rho_{ij} = 1\right\}_{i,j=1}^n$, where HM learning is equivalent to CS learning in terms of the dynamics of the economy (Proposition 1).[14] Therefore:

Proposition 5 *When the steady state $\bar{a}$ is unobservable (respectively, observable), an equilib-*

[13] In a different framework of heterogeneous learning, Berardi (2007) finds numerically that heterogeneity in PLMs makes an equilibrium more learnable.

[14] Proposition 3 provides the learnability of an equilibrium under CS learning as follows: when the steady state $\bar{a}$ is unobservable (respectively, observable), an equilibrium under CS learning is globally asymptotically stable if and only if $\lambda[B] < 1$ ($\lambda[\Phi_n' \otimes B] < 1$). Note that if $\rho_{ij} = 1$ (and hence $\varphi_i = \varphi_j$) for all i, j, then $\lambda\left[\Phi_n\left(\frac{1}{n}\Psi_n\right)\right] = \lambda[\Phi_n]$. This condition corresponds to Proposition 10.3 in Evans and Honkapohja (2001).

rium under HM learning is as learnable as (not less learnable than) an equilibrium under CS learning.

That is, HM learning (that is, the existence of private information) never makes an equilibrium less learnable.

4.2.2 Impact of Misspecification

The impact of the existence of misspecification in the PLMs is automatically found by Proposition 5:

Corollary 1 *The existence of misspecification in the PLMs never makes an equilibrium less learnable.*

On the other hand, the degree of misspecification does not have a monotonic relationship with the learnability as the degree of heterogeneity has in Proposition 4. Proposition 3 indicates that if $\bar{a}$ is observable, the learnability is nonmonotonically related with n through $\lambda\left[\left(\Phi\left(\frac{1}{n}\Psi\right)\right)' \otimes B\right]$.

However, a monotonic relationship between the degree of misspecification and the learnability is found in a plausible case where exogenous variables have similar stochastic characteristics: $\sigma_{ii} = \sigma > 0$, $\rho_{ij} = \rho \in [0,1)$, and $\varphi_i = \varphi \in [0,1)$ for any i, j. These characteristics are typical of, for example, idiosyncratic shocks held by similar economic agents. In this case, $\lambda\left[\Phi\left(\frac{1}{n}\Psi\right)\right] = \frac{\varphi}{n}\left(1 + (n-1)\rho\right)$, and $\frac{d\lambda\left[\Phi\left(\frac{1}{n}\Psi\right)\right]}{dn} \leq 0$ with the equality iff $\varphi = 0$. Thus:

Proposition 6 *Given exogenous variables with the stochastic characteristics of $\sigma_{ii} = \sigma > 0$, $\rho_{ij} = \rho \in [0,1)$, and $\varphi_i = \varphi \in [0,1)$ for any $i, j \in \{1, \cdots, n\}$, if the steady state $\bar{a}$ is unobservable (respectively, observable), the learnability condition of the HME is independent of (weakly weakened by) an increase in the quantity n of exogenous variables.*

That is, in a situation where exogenous variables have similar stochastic characteristics, the learnability of the HME is weakly increased by an increase in the degree of misspecification in the PLMs.

The strong learnability of an equilibrium under misspecified PLMs also results from the sluggish updating of the aggregate forecast $E_t^* y_{t+1}$, as is found in the mechanism under heterogeneous PLMs. Under CS learning, agents update their forecasts in response to the evolution of any variable. Under misspecified learning, agents update their forecasts only if their observable variables evolve. This sluggishness of the updating of individual forecasts leads to the strong learnability of the HME.[15]

4.2.3 Implications

Our results are summarized as follows. The existence of private information weakly increases the learnability of an equilibrium. Private information causes the heterogeneity and misspecification in agents' PLMs, the degrees of which depend upon the stochastic characteristics of privately observable exogenous variables. The learnability is weakly increased as the degree of heterogeneity in the PLMs is increased, and if exogenous variables have similar stochastic characteristics, the learnability is also weakly increased as the degree of misspecification in the PLMs increases.

Regarding the policy implications, the results of the paper suggest that the existence of private information imposes no additional constraint on ensuring the learnability of an equilibrium. In the real economy, there exists a large amount of private information that tends to make an equilibrium more learnable. If a government aims to ensure the learnability of an equilibrium with private information, the government should follow stabilization policies under CS learning as if there exists no private information.

One might think that the impact of private information should disappear if agents include

[15] Note that even if there exist an infinitely large number of exogenous variables, individual forecasts continue to be updated as long as exogenous variables are correlated and remain informative about the evolution of unobservable variables.

the lagged endogenous variable y_{t-1} in their PLMs as public information. Agents of all types have the information $\{y_s\}_{s=1}^{t-1}$, which reflects past evolutions of unobservable exogenous variables, and hence including y_{t-1} into PLMs seems to cover the disadvantage of updating in the presence of private information. However, the evolutions of unobservable exogenous variables cannot be perfectly identified as long as the number of unobservable exogenous variables for each agent is greater than that of observable endogenous variables ($n-1 > m$). Hence, the impact of private information is likely to remain even if y_{t-1} is included in the PLMs. In Section 5.3.2, this case will be illustrated in a basic NK model.

4.3 Connections to Other Frameworks

Finally, our analytical results hold true in a variety of informational structures that are limited and/or heterogeneous in terms of $\{w_{it}\}_{i=1}^{n}$, because Assumption 1 is able to reproduce those structures by accommodating the characteristics of $\{w_{it}\}_{i=1}^{n}$; that is, n, $\left\{\rho_{ij}\right\}_{i,j=1}^{n}$, and $\{\varphi_i, \sigma_{ii}\}_{i=1}^{n}$.

As an example, let us reproduce a limited and homogeneous information set considered in the RPE where there exists an exogenous variable unobservable by any agent (e.g., Evans and Honkapohja, 2001, section 13.1.1). In our model, suppose that there are a large number of exogenous variables ($n \simeq \infty$) such that the population of each type is infinitesimal ($\frac{1}{n} \simeq 0$), and such that an exogenous variable w_{1t} is imperfectly correlated with the other variables $\{w_{it}\}_{i=2}^{n}$ ($\{\rho_{1i} = \rho_{i1} < 1\}_{i=2}^{n}$), while the other variables are perfectly correlated ($\left\{\rho_{ij} = 1\right\}_{i,j=2}^{n}$). The HME in this structure is asymptotically equivalent to an RPE where in the presence of two exogenous variables $\{w_{it}\}_{i=1}^{2}$, all agents have information on only w_{2t} and form a misspecified PLM: $y_t = a + c_2 w_{2t}$.[16] See Appendix D for other examples: 1) asymmetric information sets of a full information set and a limited information set; 2) information sets partly overlapped with each other; and 3) different populations of different types of agents, instead of the symmetric

[16] The learnability condition of the above RPE with the observable steady state is obtained using Proposition 3 as $\lim_{n\to\infty} \lambda \left[\left(\Phi\left(\frac{1}{n}\Psi\right)\right)' \otimes B\right] = \varphi_2 \lambda [B] < 1$, which corresponds to Evans and Honkapohja (2001, section 13.1.1) if the observable steady state and AR(1) exogenous variables are assumed in their model.

distribution at $\frac{1}{n}$ assumed here.[17]

Hence, the results of this paper suggest that the equilibria with limited and heterogeneous information sets in terms of exogenous variables are not less learnable than the equilibrium under CS learning.

This applicability of our model makes it possible to compare learnability conditions of equilibria with different informational structures. Let us compare the learnability of 1) the equilibrium with private information (HME), 2) the equilibrium with a limited information set (above RPE), and 3) the equilibrium with a full information set (benchmark). The learnability of an equilibrium is weakly increased as $\left\{\rho_{ij}\right\}_{i,j=1}^{n}$ decrease (Proposition 4). Given $n \simeq \infty$, the above RPE is reproduced by the smallest number of imperfect correlations, and the equilibrium under CS learning is reproduced by perfect correlations. Thus, the learnability order of these equilibria is described as

$$\text{HME} \geq \text{RPE} \geq \text{Benchmark};$$

that is, the RPE is not more learnable than the HME and not less learnable than the benchmark equilibrium.

On the other hand, it is also found that there exist common features of learnability that are independent of the informational structure in terms of $\{w_{it}\}_{i=1}^{n}$. Proposition 3 provides:

Corollary 2 *The learnability condition of the HME is more stringent when the steady state $\bar{a}$ is unobservable than observable.*

Corollary 3 *When the steady state $\bar{a}$ is unobservable (respectively, observable), the learnability condition of the HME is independent of (weakly strengthened by) an increase in the autocorrelations $\{\varphi_i\}_{i=1}^{n}$ of exogenous variables.*

[17] Note that information sets that are limited and/or heterogeneous in terms of endogenous variables y_t cannot be reproduced in our model. See Berardi (2007) to consider heterogeneous PLMs in terms of endogenous variables.

The former corollary is trivial because of the fact that if $\lambda[B] < 1$, then $\lambda\left[\left(\Phi\left(\frac{1}{n}\Psi\right)\right)' \otimes B\right] < 1$ (but not vice versa), and the latter is obtained from Lemma 2.3. Corollary 2 is consistent with Bullard and Mitra (2002)'s finding that if constant terms are excluded from the PLMs, the learnability conditions under CS learning are relaxed. Similar to Corollary 3, Marcet and Sargent (1989a) and Bullard and Mitra (2002) find that when constant terms are excluded from the PLMs, an increase in the autocorrelation of exogenous variable $\{\varphi_i\}_{i=1}^n$ for any i reduces the learnability of an equilibrium under CS learning.

5 Numerical Example

In this section, we examine the learnability of the HME in a basic NK macroeconomic model.[18] The NK model is a benchmark macroeconomic framework for establishing DSGE models and analyzing optimal monetary policy.[19] We obtain the constraints that must be imposed on monetary policy rules to ensure the learnability of the HME. Next, we demonstrate numerically the impact of private information on learnability and the effect of including y_{t-1} in the PLMs.

5.1 NK Model

We consider a basic NK model with an aggregate demand shock g_t:

$$x_t = -\alpha\left(i_t - E_t^*\pi_{t+1}\right) + E_t^* x_{t+1} + g_t, \tag{18}$$

$$\pi_t = \kappa x_t + \beta E_t^* \pi_{t+1}. \tag{19}$$

The model has three endogenous variables: output gap x_t, inflation rate π_t, and nominal interest rate i_t. Eq. (18) is a log-linearized intertemporal Euler equation that is derived from the households' optimal choice of consumption. Eq. (19) is a forward-looking Phillips curve that

[18] Learnability in this section means the *local* stability under adaptive learning, because a NK model is a log-linearization of a nonlinear model.

[19] See Bullard and Mitra (2002), Evans and Honkapohja (2003a,b, 2006), Honkapohja and Mitra (2004a, 2005), Adam (2005), Berardi (2009), Branch and McGough (2009), and Airaudo and Zanna (2010) for an analysis of optimal monetary policy that ensures the learnability of an equilibrium.

is derived from the optimizing behavior of monopolistically competitive firms with Calvo price setting. $\alpha > 0$, $\kappa > 0$, and $0 \leq \beta < 1$ are assumed. The central bank adopts a Taylor-type contemporaneous data interest rate rule:

$$i_t = \phi_\pi \pi_t + \phi_x x_t, \tag{20}$$

where ϕ_π and ϕ_x are the policy parameters controlled by the central bank, and they are assumed to be nonnegative.

To consider the impact of private information in the NK model, we assume that g_t is the aggregation of idiosyncratic demand shocks: $g_t \equiv \sum_{i=1}^n g_{it}$, which might be obtained from, for example, households' utility functions with idiosyncratic preference shocks. The shock g_{it} for each i follows an AR(1) process: $g_{it} = \varphi_i g_{i,t-1} + v_{it}$, where $0 \leq \varphi_i < 1$ and the disturbance term v_{it} has a zero mean. The correlation of g_{it} and g_{jt} is $\rho_{ij} \geq 0$ for each i, j. Under CS learning, g_{it} for all i is observable for all agents, whereas under HM learning, the shock is observable for $1/n$ of agents and unobservable for other agents. The aggregate forecasts $(E_t^* x_{t+1}, E_t^* \pi_{t+1})$ are the averages of the forecasts of all types $\{(E_{it}^* x_{t+1}, E_{it}^* \pi_{t+1})\}_{i=1}^n$.[20]

5.2 Learnability Conditions

Let us obtain learnability conditions imposed on monetary policy parameters (ϕ_π, ϕ_x) in the interest rate rule (20). The learnability conditions of equilibria under CS and HM learning are provided in Table 1, the derivation of which is shown in Appendix E. The conditions under the unobservable and observable steady states are given for each type of learning, respectively. For ease of comparison, Figure 1 illustrates the domains of (ϕ_π, ϕ_x) that satisfy the learnability conditions given in Table 1.

The table shows that if the steady state is unobservable, the learnability condition under HM learning is exactly the same as the condition under CS learning:

$$\kappa (\phi_\pi - 1) + \phi_x (1 - \beta) > 0, \tag{21}$$

[20] To focus on investigating the properties of the HME, we follow the assumptions given by Branch and McGough (2009) for incorporating heterogeneous forecasts at an individual level within the NK model. The original form of the NK model (18)–(19) is obtained by aggregating individual forecasts.

which is the *Taylor principle* provided by Bullard and Mitra (2002). The figure demonstrates that if the steady state is observable, the learnability conditions of equilibria under CS and HM learning are not more stringent than the Taylor principle. In addition, the learnability condition under HM learning is not more stringent than the condition under CS learning.

Proposition 7 *In the NK model (18)–(19) with a Taylor-type nominal interest rate rule (20), the Taylor principle (21) is the sufficient condition for the learnability of an equilibrium under adaptive learning, regardless of the existence of private information and regardless of the observability of the steady state.*

The proposition reinforces the importance of the Taylor principle in monetary policy. Bullard and Mitra (2002) and Honkapohja and Mitra (2004a) find that with the contemporaneous data rule (20), the Taylor principle is a sufficient condition for the learnability of the REE under CS learning. Guse (2008) confirms that their result is robust even if agents' learning is misspecified. Proposition 7 emphasizes that the sufficiency of the Taylor principle remains true even if agents' learning is heterogeneously misspecified because of private information.

5.3 Calibrations

Next, we numerically evaluate the impact of private information on the learnability of an equilibrium. Furthermore, we examine whether the strong learnability of the HME is robust if agents include the lagged endogenous variable y_{t-1} in the PLMs.

5.3.1 Impact of HM Learning

We calibrate the parameter domains of (ϕ_π, ϕ_x) satisfying the learnability conditions under different degrees of heterogeneity and misspecification in the PLMs, separately. We focus on the HME with the observable steady state, which ensures the impact of HM learning, and exogenous variables with the same stochastic characteristics: $\rho_{ij} = \rho \in [0, 1)$, and $\varphi_i = \varphi \in [0, 1)$ for all

i, j. For robustness, we consider two cases of structural parameters: (a) $\alpha = 1$, $\kappa = 0.3$, and $\beta = 0.99$ (Clarida et al., 2000); (b) $\alpha = 1/0.157$, $\kappa = 0.024$, and $\beta = 0.99$ (Woodford, 1999). We set the autocorrelations φ of the shocks to be equal to 0.9 (Milani, 2008).

First, we calibrate the impact of heterogeneity in the PLMs. Figure 2 shows the parameter domains satisfying the learnability conditions under HM learning under different correlations of the exogenous variables. The domain under $\rho = 1.0$ represents the learnability condition under CS learning. We consider the existence of 10 idiosyncratic shocks ($n = 10$). In both examples of the structural parameters, the learnability domains under HM learning are significantly enlarged as the correlation ρ decreases. When $\rho = 0.7$, the domains cover almost all positive values of (ϕ_π, ϕ_x).[21]

Next, we examine the impact of misspecification in the PLMs in Figure 3. In each panel, the boundary of the domain ensuring the learnability under CS learning is shown for comparison. We focus on four cases of $n \in \{10, 100, 1000, 10000\}$ and set $\rho = 0.9$ to detect the impact of misspecification in the positive region of (ϕ_π, ϕ_x). In both examples, the learnability domains are enlarged under HM learning. Thus, the impact of misspecification in the PLMs is economically significant as well as the impact of heterogeneity. On the other hand, the impact of misspecification is largely unaffected even if the number of exogenous variables increases. This is because we have assumed the same correlations for exogenous variables $\{\rho_{ij} = \rho\}_{i \neq j}^{n}$, under which the evolutions of all unobservable variables continue to be correlated to the same degree, regardless of the number of exogenous variables. Thus, the updating of individual forecasts (that is, the learnability) is not significantly affected by the degree of misspecification.

The above results suggest that the impact of private information on the learnability is economically significant. In particular, if privately observable variables are weakly correlated (that is, information sets are highly heterogeneous), the learnability is likely to be ensured under a wide range of policy parameters. In such a situation, the central bank does not need to consider the learnability constraints. On the other hand, the impact on the learnability is not

[21] Fidrmuc and Korhonen (2003, Table 2) find the correlations of demand shocks between euro countries in the 1990s to be 0.65 at the highest. Those values might be seen as the proxies for the average of the correlations of idiosyncratic shocks at an individual level.

significantly changed by the number of privately observable variables (that is, the limitation of each information set).

5.3.2 Existence of Public Information y_{t-1}

Variations in y_{t-1} reflect past evolutions of not only observable but also unobservable exogenous variables. If agents have the information $\{y_s\}_{s=1}^{t-1}$, they may utilize y_{t-1} as public information for updating their forecasts.[22] We consider a type i agent's PLM that includes y_{t-1} into the PLM (15):

$$y_t = \bar{a} + b_i y_{t-1} + c_i w_{it} + \varepsilon_{it}, \tag{22}$$

where coefficient b_i is the $m \times m$ matrix of coefficients on y_{t-1}. Under HM learning with Eq. (22) for all i, the learnability of the HME might become similar to the learnability of the equilibrium under CS learning.

Figure 4 shows parameter domains satisfying the learnability condition of the HME with the observable steady state.[23] Compared with the results in Figure 2, the introduction of y_{t-1} into the PLMs reduces the learnability domains of the HME because y_{t-1} contributes to the updating of the parameters. However, the domain of the HME continues to be larger than the domain of the equilibrium under CS learning; that is, the updating of $E_t^* y_{t+1}$ under HM learning remains sluggish. This is because as long as $n-1 > m$ (that is, the number of unobservable exogenous variables for each agent is greater than that of observable endogenous variables),

[22]For example, Grossman and Stiglitz (1980) assume that agents infer from market prices the unobservable information of other agents. Marcet and Sargent (1989a) consider that firms in different industries, each of which has private information about its own capital stock, include the prices of outputs of all industries into their PLMs as public information in order to forecast the future prices of their outputs.

[23]Here, projection facilities are employed in the real-time learning process to keep agents' forecasts bounded. First, we assume that the eigenvalues of $\{b_i\}_{i=1}^n$ are strictly inside the unit circle so that agents' forecasts $\{E_{it} y_{t+s}\}_{s=1}^{\infty}$ for all i are asymptotically stationary (see Evans and Honkapohja, 2001, chapter 10). Second, we also assume that the diagonal elements of $\{b_i\}_{i=1}^n$ are strictly inside the range $(-1, 1)$ (see Mitra et al., 2013). When projection facilities are applied in this paper, parameter estimates are reset to their previous estimates (see Marcet and Sargent, 1989a). We find that there exists a unique equilibrium, that is, a unique set of parameters satisfying these constraints.

agents cannot perfectly identify the evolution of the unobservable variables by observing y_{t-1}. Thus, the strong learnability of the HME remains even if agents utilize the public information.

This result implies additionally that an equilibrium with a VAR(1) PLM, $y_t = \bar{a} + b y_{t-1} + \varepsilon_t$, which excludes all exogenous variables, is not less learnable than the equilibrium under CS learning. The VAR(1) PLM has been considered in the literature as a possible PLM when agents have only data for the endogenous variables (e.g., Marcet and Sargent, 1989a; Slobodyan and Wouters, 2012). In our model, the VAR(1) PLM can be reproduced by $n \simeq \infty$ and $\{\rho_{ij} = 0\}_{i,j=1}^{n}$, where the PLM (22) for each i is asymptotically equivalent to the VAR(1) PLM because the evolution of each exogenous variable makes a negligible contribution to forecasting the dynamics of the economy and the parameters $\{c_i\}_{i=1}^{n}$ are estimated to be zero asymptotically. Hence, the equilibrium under adaptive learning with the VAR(1) PLM is likely to have the same strong learnability as the HME.

6 Conclusions

This paper has investigated the learnability of an equilibrium with private information, which makes agents' information sets limited and heterogeneous. In the real economy, there might exist economic variables observable by some agents and unobservable by other agents. In such a situation, agents are constrained to form heterogeneously misspecified PLMs, and there exists a heterogeneous misspecification equilibrium (HME).

The paper finds that the learnability condition of the HME is not more stringent than the condition of the equilibrium under CS learning. The learnability of the HME is weakly increased as the degree of heterogeneity in PLMs caused by private information is increased, and if privately observable variables have similar stochastic characteristics, the learnability is also weakly increased as the degree of misspecification in the PLMs is increased. In a basic NK model, the paper finds that the central bank should follow the Taylor principle for learnability regardless of whether or not there exists private information. The numerical analysis indicates that the impact of private information on learnability is economically significant, and the strong learnability of the HME is found to be robust to including lagged endogenous variables into the PLMs as public information. These results are applicable to other frameworks of adaptive learning with misspecified and/or heterogeneous PLMs.

Potential issues remain for future research. Heterogeneous misspecification in adaptive learning may be investigated as a possible ingredient of the persistence of economic fluctuations. Adam (2005), Milani (2008), and Slobodyan and Wouters (2012) find that adaptive learning prolongs the response of an economy to fundamental shocks in NK DSGE models. This paper finds that heterogeneous misspecification in learning makes an aggregate forecast sluggish to exogenous variables. Thus, the responses of the HME to fundamental shocks are expected to be more persistent than the responses of an equilibrium under CS learning.

In addition, heterogeneous misspecification in learning may be applied to the analysis of the effect of financial frictions on the macroeconomy. Assenza and Berardi (2009) consider the impact of bankruptcy in a credit economy where borrowers and lenders form heterogeneous expectations. While their heterogeneity stems from different gain parameters in learning algorithms, this paper considers the other type of heterogeneity in terms of different forecasting models, which might be prevailing in financial markets where private information exists. Thus, HM learning might well describe the dynamics of a credit economy in the framework of adaptive learning.

Appendix

A Regularity Assumptions

Assumption 2

1. $\det(I_m - B) \neq 0$ *and* $\det(I_{mn} - \Phi \otimes B) \neq 0$.
2. Φ *is a diagonal and nonnegative matrix whose diagonal elements exist in the interval* $[0, 1)$.
3. Γ *is a nonnegative matrix, in which* $0 \leq \rho_{ij} \leq 1$ *for each* $i, j \in \{1, ..., n\}$.

Assumption 2.1 avoids the possibility that a nonexplosive fundamental REE could be indeterminate (see Honkapohja and Mitra, 2006, Proposition 1).

The diagonal representation of Φ in Assumption 2.2 simplifies the analysis by equating the eigenvalues of Φ with its diagonal elements existing in the interval $[0,1)$. Note that this assumption is not crucial for our analysis, because even if Φ were originally nondiagonal, Eq. (2) could be transformed to an equation that includes a diagonal autoregressive matrix by premultiplying Eq. (2) by the $n \times n$ matrix formed from the eigenvectors of Φ. The diagonal elements in the interval $[0,1)$ ensure the stationarity of w_t.

Neither is Assumption 2.3 crucial for our analysis because any linear model can be transformed to the system with $\Gamma \geq 0_{n\times n}$. For example, if any ρ_{ij} is negative in an original model, this negative correlation can be transformed to be positive by changing the sign of w_i (or w_j) and redefining the correlation between $-w_i$ and w_j as $\rho_{ij} \geq 0$. Applying this transformation to any negative correlation, the original model is transformed to the system with $\Gamma \geq 0_{n\times n}$.

B Derivations of ODE under HM Learning

Agent i for each $i \in \{1,\ldots,n\}$ forms $E^*_{it}y_{t+1}$ by using real-time learning with the PLM (7) and the information set $\{y_s, w_{is}\}_{s=1}^{t}$. We assume the *t-dating of expectations* considered by Evans and Honkapohja (2001, chapter 10): coefficient parameters ϕ_{it} at time t are estimated with past data until time t, $\{y_{is}, w_{is}\}_{s=1}^{t-1}$, and $E^*_{it}y_{t+1}$ is formed with ϕ_{it} and the current data $\{y_t, w_{it}\}$. The estimates of the coefficient parameters $\phi'_{it} = (a_{it}, c_{it})$ are given by the least-squares projection of y_{t-1} on $z'_{i,t-1} = (1, w_{i,t-1})$: $Ez_{i,t-1}\left(y_{t-1} - \phi'_{it}z_{i,t-1}\right)' = 0_{2\times m}$. Then, the updating rule of ϕ_{it} is shown by the RLS representation:

$$\phi_{it} = \phi_{i,t-1} + t^{-1}R^{-1}_{it}z_{i,t-1}\left(y_{t-1} - \phi'_{i,t-1}z_{i,t-1}\right)', \tag{B.1}$$

$$R_{it} = R_{i,t-1} + t^{-1}\left(z_{i,t-1}z'_{i,t-1} - R_{i,t-1}\right), \tag{B.2}$$

where $R_{it} = t^{-1}\sum_{s=1}^{t} z_{i,s-1}z'_{i,s-1}$, which is the updating of the matrix of the second moment of z_{it}. See Evans and Honkapohja (2001, section 10.3) for details regarding obtaining RLS equations to satisfy the orthogonality condition.

The stochastic recursive algorithm (SRA) for ϕ_{it} for each i is obtained by substituting the ALM (11) into Eq. (B.1):

$$\phi_{it} = \phi_{i,t-1} + t^{-1}R^{-1}_{it}z_{i,t-1}\begin{pmatrix} 1 & w_{1,t-1} & \cdots & w_{n,t-1}\end{pmatrix}\left[\begin{pmatrix} D_{0,t-1} & D_{1,t-1} & \cdots & D_{n,t-1}\end{pmatrix} - \begin{pmatrix} a_{i,t-1} & c^{+}_{i,t-1}\end{pmatrix}\right]',$$

where we denote $D_{0t} \equiv A + Ba_t$, $a_t \equiv \frac{1}{n}\sum_{i=1}^{n} a_{it}$ as the constant term of the aggregate PLM (9), $D_{it} \equiv \frac{1}{n}Bc_{it}\varphi_i + C_i$ for each $i \in \{1,...,n\}$, C_i as the i-th column of matrix C in Eq. (1), and $c_{it}^{+} \equiv \left(0_{m\times(i-1)}, c_{it}, 0_{m\times(n-i)}\right)$ as an $m \times n$ matrix in which the columns, except c_{it}, are zero vectors.

To obtain the ODEs for ϕ_i associated with the SRA, we have to calculate the unconditional expectations of the updating terms in the SRA. The convergence of the SRA is analyzed by Marcet and Sargent (1989b) in the stochastic approximation approach, which is also introduced by Evans and Honkapohja (2001, chapter 6). Denote the operator E as the expectation of variables for ϕ_i fixed, taken over the invariant distributions of w_t. Then, by letting $Ez_iz_j' = \lim_{t\to\infty} Ez_{it}z_{jt}'$ for any $i,j \in \{1,...,n\}$, the unconditional expectation of the updating term in Eq. (B.1) is transformed to

$$\begin{aligned}
& ER_i^{-1}z_{i,t-1}\begin{pmatrix} 1 & w_{1,t-1} & \cdots & w_{n,t-1}\end{pmatrix}\left[\begin{pmatrix} D_0 & D_1 & \cdots & D_n\end{pmatrix} - \begin{pmatrix} a_i & c_i^{+}\end{pmatrix}\right]' \\
= \; & R_i^{-1}\left(\left(Ez_{i,t-1}z_{i,t-1}'\right)\left[\begin{pmatrix} D_0 & D_i\end{pmatrix} - \begin{pmatrix} a_i & c_i\end{pmatrix}\right]' + E\begin{pmatrix} 1 \\ w_{i,t-1}\end{pmatrix}\left(\sum_{j=1}^{n} w_{j,t-1}D_j' - w_{i,t-1}D_i'\right)\right) \\
= \; & R_i^{-1}\begin{pmatrix} \left(Ez_{i,t-1}z_{i,t-1}'\right)\left[\begin{pmatrix} D_0 & D_i\end{pmatrix} - \begin{pmatrix} a_i & c_i\end{pmatrix}\right]' \\ +\begin{pmatrix} 0_{1\times m} \\ \sum_{j=1}^{n}\left(Ew_{i,t-1}w_{j,t-1}\right)D_j' - \left(Ew_{i,t-1}w_{i,t-1}\right)D_i'\end{pmatrix}\end{pmatrix} \\
= \; & R_i^{-1}\left(Ez_{i,t-1}z_{i,t-1}'\right)\begin{pmatrix} \left[\begin{pmatrix} D_0 & D_i\end{pmatrix} - \begin{pmatrix} a_i & c_i\end{pmatrix}\right]' \\ +\begin{pmatrix} 0_{1\times m} \\ \sum_{j=1}^{n}\left(Ew_{i,t-1}w_{i,t-1}\right)^{-1}\left(Ew_{i,t-1}w_{j,t-1}\right)D_j' - D_i'\end{pmatrix}\end{pmatrix}.
\end{aligned}$$

In addition, the expectation of the updating term in Eq. (B.2) is given by

$$Ez_iz_i' - R_i.$$

Hence, the ODEs for ϕ_i and R_i associated with the SRA are obtained as

$$\frac{d\phi_i}{d\tau} = R_i^{-1}\left(Ez_iz_i'\right)\left(T_i\left(a_i,c_i\right) - \phi_i'\right)', \tag{B.3}$$

$$\frac{dR_i}{d\tau} = Ez_iz_i' - R_i, \tag{B.4}$$

where

$$T_i\left(a_i,c_i\right) \equiv \begin{pmatrix} D_0 & \sum_{j=1}^{n} D_j\omega_{ij}\omega_{ii}^{-1}\end{pmatrix}.$$

A scalar ω_{ij} denotes the covariance of w_i and w_j; $\omega_{ij} \equiv \sigma_{ii}\rho_{ij}\sigma_{jj}$ for each $i,j \in \{1,...,n\}$. Furthermore, because R_i and Ez_iz_i' in Eq. (B.4) are asymptotically equal, $R_i^{-1}\left(Ez_iz_i'\right)$ in Eq. (B.3) globally converges to unity. Hence, the stability of the ODE for $\phi_i' = (a_i, c_i)$ in Eq. (B.3) is determined by smaller differential equations:

$$\frac{da_i}{d\tau} = D_0 - a_i, \tag{B.5}$$

$$\frac{dc_i}{d\tau} = \sum_{j=1}^{n} D_j\omega_{ij}\omega_{ii}^{-1} - c_i. \tag{B.6}$$

In the same manner, smaller ODEs for the parameters $\{\phi_j\}_{j \neq i}^n$ are obtained.

The ODEs (B.5)–(B.6) for all i are represented by the ODEs for the aggregate parameters (a, c) in Eq. (9). First, the ODEs for all a_is have the same form, and a is an arithmetic average of all a_is. Then, the convergence property of a is equivalent to that of a_i for each i; the ODEs for all a_is are represented by a single ODE for a that has the same form as that for a_i:

$$\frac{da}{d\tau} = T_a(a) - a,$$

where

$$T_a(a) \equiv D_0 = A + Ba.$$

Next, the ODEs for all c_is are represented by a single ODE for the aggregate parameter c. If the ODEs (B.6) for all i are multiplied by $\frac{1}{n}$ and combined in a single $m \times n$ matrix, the single ODE for c is obtained by:

$$\frac{dc}{d\tau} = T_c(c) - c,$$

where

$$\begin{aligned} T_c(c) &\equiv \begin{pmatrix} \frac{1}{n}\sum_{j=1}^n D_j \omega_{1j}\omega_{11}^{-1} & \cdots & \frac{1}{n}\sum_{j=1}^n D_j \omega_{nj}\omega_{nn}^{-1} \end{pmatrix} \\ &= (Bc\Phi + C)\left(\frac{1}{n}\Psi\right), \end{aligned}$$

and

$$\begin{aligned} \Psi &\equiv \begin{pmatrix} 1 & \omega_{12}\omega_{22}^{-1} & \cdots & \omega_{1n}\omega_{nn}^{-1} \\ \omega_{21}\omega_{11}^{-1} & 1 & \cdots & \omega_{2n}\omega_{nn}^{-1} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{n1}\omega_{11}^{-1} & \omega_{n2}\omega_{22}^{-1} & \cdots & 1 \end{pmatrix} \\ &= diag(\sigma_{ii})_{1 \leq i \leq n} \cdot \Gamma \cdot diag(\sigma_{ii})_{1 \leq i \leq n}^{-1}. \end{aligned}$$

The derivation is complete.

C Proof of Lemma 2

C.1 Proof of Lemma 2.1

To find that the eigenvalues of $\Phi\left(\frac{1}{n}\Psi\right)$ are real and exist in the interval $[0, 1)$, we use Trenkler (1995)'s lemma.

Lemma 3 *If X and Y are nonnegative definite matrices of the same dimension, then eigenvalues of XY are real and nonnegative, and they are zero if and only if $XY = 0$.*

Proof. If X is nonnegative definite, it may be written as $X = PP'$ for some matrix P. Then, the eigenvalues of $XY = PP'Y$ are those of $P'YP$ (plus possibly some zeros). $P'YP$ is obviously a nonnegative definite matrix. It follows that all eigenvalues of XY are nonnegative, and that they are zero iff $XY = 0$. The proof is complete. ■

As Φ and $\frac{1}{n}\Psi$ are nonnegative definite matrices by Assumptions 2.2 and 2.3, Lemma 3 yields eigenvalues of $\Phi\left(\frac{1}{n}\Psi\right)$ that are real and nonnegative. Notice that the diagonal elements of $\Phi\Psi$ are equal to the diagonal elements of Φ; then, $tr\left(\Phi\left(\frac{1}{n}\Psi\right)\right) < 1$. Therefore, all eigenvalues of $\Phi\left(\frac{1}{n}\Psi\right)$ are real and exist in the interval $[0, 1)$. The proof is complete.

C.2 Proof of Lemmas 2.2 & 2.3

According to the *Perron–Frobenius Theorem* (see Berman, A. and R. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, 1979, p.27), Kolotilina (Kolotilina, L. Y., "Bounds for the Perron root, Singularity/Nonsingularity Conditions, and Eigenvalue Inclusion Sets," *Numerical Algorithms*, Vol. 42, No. 3–4, 2006, pp. 247–280) shows the monotonicity property of the Perron root (Theorem 2.1): "Let A and B be nonnegative matrices of order $n \geq 1$ and let $A \geq B$. Then, $\lambda[A] \geq \lambda[B]$." In our paper, $\frac{d\left(\Phi\left(\frac{1}{n}\Psi\right)\right)}{d\rho_{ij}} \geq 0$ and $\frac{d\left(\Phi\left(\frac{1}{n}\Psi\right)\right)}{d\varphi_i} \geq 0$ for any i, j; thus $\frac{d\lambda\left[\Phi\left(\frac{1}{n}\Psi\right)\right]}{d\rho_{ij}} \geq 0$ and $\frac{d\lambda\left[\Phi\left(\frac{1}{n}\Psi\right)\right]}{d\varphi_i} \geq 0$ for any i, j. The proof is complete.

D Reproductions of Different Informational Structures

D.1 Asymmetric Information Sets

The asymmetric information sets of a full information set and a limited information set are reproduced in Assumption 1. Suppose that there exist two exogenous variables $\{w_{1t}, w_{2t}\}$ with $\rho_{12} = 0$ and $\frac{\sigma_{22}}{\sigma_{11}} \simeq 0$. In this case, the evolution of y_t is governed only by the evolution of w_{1t}, and because $E_t(w_{2t}y_t) \simeq 0$, agents of type 2 are unable to forecast the evolution of y_t by observing the evolution of w_{2t}. This structure is equivalent to the case where the agent of type 1 has full information on the economy and the agent of type 2 has a limited information set.

D.2 Overlapping Information Sets

The information sets that partly overlap with each other in terms of $\{w_{it}\}_{i=1}^{n}$ are equivalent to the case of the existence of highly correlated exogenous variables in Assumption 1.

D.3 Different Populations of Different Types of Agents

For example, the existence of two types of agents who have populations $\frac{2}{3}$ and $\frac{1}{3}$, respectively, are reproduced in Assumption 1. Suppose the existence of three exogenous variables $\{w_{it}\}_{i=1}^{3}$, each of which is privately observable for $\frac{1}{3}$ of agents, and the perfect correlation of two of three variables (e.g., w_1 and w_2). Then, the population of agents detecting the evolutions of w_1 and w_2 is $\frac{2}{3}$, and the population of agents observing the evolution of w_3 is $\frac{1}{3}$. This environment is equivalent to the existence of two types of agents with the populations of $\frac{2}{3}$ and $\frac{1}{3}$, respectively.

E Derivations of Learnability Conditions

Before proceeding, we provide a lemma that will be used to obtain the learnability conditions of equilibria under different learning rules. A simpler derivation under CS learning is shown in Bullard and Mitra (2002).

Lemma 4 *Define an $n \times n$ matrix X whose eigenvalues are all real and exist in the interval $[0,1)$. Given $\alpha \geq 0$, $\kappa \geq 0$, $0 \leq \beta < 1$, $\phi_\pi \geq 0$, $\phi_x \geq 0$, and $B = \begin{pmatrix} 1+\alpha\phi_x & \alpha\phi_\pi \\ -\kappa & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix}$, then the real parts of eigenvalues of $X \otimes B - I_{2n}$ are all negative if and only if:*

$$\kappa\left(\phi_\pi - \lambda[X]\right) + \phi_x\left(1-\beta\lambda[X]\right) > -\frac{\left(1-\lambda[X]\right)\left(1-\beta\lambda[X]\right)}{\alpha}. \tag{E.1}$$

Proof. Define the eigenvalues of X as $0 \leq \chi_i < 1$ for each $i \in \{1,...,n\}$ and the eigenvalues of B as δ_j for each $j \in \{1,2\}$. Then, $0 \leq \lambda[X] < 1$, and the eigenvalues of $X \otimes B$ are given by $\chi_i\delta_j$ for each i,j. First, we show $\lambda[B] \geq 0$ by calculating the characteristic equation of B: $q(x) = x^2 + p_1 x + p_0$, where $p_0 = \frac{\beta}{1+\kappa\alpha\phi_\pi+\alpha\phi_x} > 0$ and $p_1 = -\frac{1+\beta+\kappa\alpha+\alpha\beta\phi_x}{1+\kappa\alpha\phi_\pi+\alpha\phi_x} < 0$. According to the *Routh Theorem* (see Alpha C. Chiang, *Fundamental Methods of Mathematical Economics: Second Edition*, McGraw-Hill, 1974), the eigenvalues of B have all negative real parts; that is, $\lambda[B] < 0$, if and only if $|p_1|$ and $\begin{vmatrix} p_1 & 0 \\ 1 & p_0 \end{vmatrix}$ are all positive, that is, $p_1 > 0$ and $p_1 p_0 > 0$. The above $q(x)$ violates these conditions; therefore, $\lambda[B] \geq 0$. Here, let us prove Lemma 4. Because $\lambda[B] \geq 0$ and $\lambda[X] \geq 0$, $\lambda[X \otimes B] = \lambda[X]\lambda[B] = \lambda[\lambda[X]B]$, and hence

$\lambda[X \otimes B - I_{2n}] = \lambda[\lambda[X] B - I_2]$. Thus, the real parts of the eigenvalues of $X \otimes B - I_{2n}$ are all negative if and only if the eigenvalues of $\lambda[X] B - I_2$ have all negative real parts. The characteristic equation of $\lambda[X] B - I_2$ is $q(x) = x^2 + p_1 x + p_0$, where:

$$\begin{aligned} p_0 &= \frac{(1-\lambda[X])(1-\beta\lambda[X]) + \alpha(\kappa(\phi_\pi - \lambda[X]) + \phi_x(1-\beta\lambda[X]))}{1+\kappa\alpha\phi_\pi + \alpha\phi_x}, \\ p_1 &= \frac{(1-\lambda[X]) + (1-\beta\lambda[X]) + \kappa\alpha(2\phi_\pi - \lambda[X]) + \alpha\phi_x(2-\beta\lambda[X])}{1+\kappa\alpha\phi_\pi + \alpha\phi_x}. \end{aligned}$$

Note that $p_1 = p_0 + \frac{(1-\beta\lambda[X]) + \alpha(\kappa\phi_\pi + \phi_x)}{1+\kappa\alpha\phi_\pi+\alpha\phi_x}$; then $p_1 > p_0$. The eigenvalues of $\lambda[X] B - I_2$ have all negative real parts if and only if $p_1 > 0$ and $p_1 p_0 > 0$. As $p_1 > p_0$, the necessary and sufficient condition is given solely by $p_0 > 0$, that is, Eq. (E.1). The proof is complete. ■

Substituting Eq. (20) into Eqs. (18)–(19), the NK model is transformed into the form of the system (1)–(2) with $y_t = (x_t, \pi_t)'$, $w_t = (g_{1t}, ..., g_{nt})'$, and $B = \begin{pmatrix} 1+\alpha\phi_x & \alpha\phi_\pi \\ -\kappa & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix}$.

Under the unobservable steady state $\bar{a}$, the HME (13)–(14) is globally asymptotically stable if and only if the Jacobian (16) has the negative real parts of eigenvalues. In the NK model of the above form, $m = 2$, the eigenvalues of the Jacobian are equal to those of $I_n \otimes B - I_{2n}$. This case corresponds to $X = I_n$ in Lemma 4. As $\lambda[X] = 1$, the sufficient and necessary condition for the stability of the equilibrium is obtained by $\kappa(\phi_\pi - 1) + \phi_x(1-\beta) > 0$, which is also the learnability condition under CS learning.

In the same manner, under the observable steady state, the HME is stable if and only if the Jacobian (17) has the negative real parts of eigenvalues. This case corresponds to $X = \left(\Phi\left(\frac{1}{n}\Psi\right)\right)'$ in Lemma 4. If we define $\lambda^h \equiv \lambda\left[\Phi\left(\frac{1}{n}\Psi\right)\right]$, the stability condition is provided by $\kappa\left(\phi_\pi - \lambda^h\right) + \phi_x\left(1-\beta\lambda^h\right) > -\frac{(1-\lambda^h)(1-\beta\lambda^h)}{\alpha}$. Under CS learning, $X = \Phi'$. Then, by defining $\lambda^c \equiv \lambda[\Phi]$, the stability condition is obtained by $\kappa(\phi_\pi - \lambda^c) + \phi_x(1-\beta\lambda^c) > -\frac{(1-\lambda^c)(1-\beta\lambda^c)}{\alpha}$. Note that $0 \leq \lambda^h \leq \lambda^c < 1$.

References

Adam, K. (2005): "Experimental Evidence on the Persistence of Output and Inflation," *Economic Journal*, 117, 603–636.

Adam, K., G. W. Evans, and S. Honkapohja (2006): "Are Hyperinflation Paths Learnable?" *Journal of Economic Dynamics and Control*, 30, 2725–2748.

Airaudo, M. and L.-F. Zanna (2010): "Learning About Inflation Measures for Interest Rate Rules," *IMF Working Papers*, 1–45.

Allen, F. and D. Gale (2004): "Financial Intermediaries and Markets," *Econometrica*, 72, 1023–1061.

Anufriev, M., T. Assenza, C. Hommes, and D. Massaro (2013): "Interest Rate Rules and Macroeconomic Stability under Heterogeneous Expectations," *Macroeconomic Dynamics*, 17, 1574–1604.

Assenza, T. and M. Berardi (2009): "Learning in a Credit Economy," *Journal of Economic Dynamics and Control*, 33, 1159–1169.

Berardi, M. (2007): "Heterogeneity and Misspecifications in Learning," *Journal of Economic Dynamics and Control*, 31, 3203–3227.

——— (2009): "Monetary Policy with Heterogeneous and Misspecified Expectations," *Journal of Money, Credit, and Banking*, 41, 79–100.

——— (2011): "Fundamentalists vs. Chartists: Learning and Predictor Choice Dynamics," *Journal of Economic Dynamics and Control*, 35, 776–792.

Bovi, M. (2013): "Are the Representative Agent's Beliefs based on Efficient Econometric Models?" *Journal of Economic Dynamics and Control*, 37, 633–648.

Branch, W. A. (2004): "Restricted Perceptions Equilibria and Learning in Macroeconomics," in *Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model*, ed. by D. Colander, New York: Cambridge University Press, 135–160.

——— (2007): "Sticky Information and Model Uncertainty in Survey Data on Inflation Expectations," *Journal of Economic Dynamics and Control*, 31, 245–276.

Branch, W. A. and G. W. Evans (2006): "Intrinsic Heterogeneity in Expectation Formation," *Journal of Economic Theory*, 127, 264–295.

——— (2007): "Model Uncertainty and Endogenous Volatility," *Review of Economic Dynamics*, 10, 207–237.

Branch, W. A. and B. McGough (2009): "A New Keynesian Model with Heterogeneous Expectations," *Journal of Economic Dynamics and Control*, 33, 1036–1051.

——— (2010): "Dynamic Predictor Selection in a new Keynesian Model with Heterogeneous Expectations," *Journal of Economic Dynamics and Control*, 34, 1492–1508.

——— (2011): "Business Cycle Amplification with Heterogeneous Expectations," *Economic Theory*, 47, 395–421.

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

Brock, W. A. and C. H. Hommes (1997): "A Rational Route to Randomeness," *Econometrica*, 65, 1059–1095.

Bullard, J. and S. Eusepi (2014): "When Does Determinacy Imply Expectational Stability?" *International Economic Review*, 55, 1–22.

Bullard, J. and K. Mitra (2002): "Learning about Monetary Policy Rules," *Journal of Monetary Economics*, 49, 1105–1129.

Clarida, R., J. Gali, and M. Gertler (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 115, 147–180.

Ellison, M. and J. Pearlman (2011): "Saddlepath Learning," *Journal of Economic Theory*, 146, 1500–1519.

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

——— (2003a): "Adaptive Learning and Monetary Policy Design," *Journal of Money, Credit, and Banking*, 35, 1045–1072.

——— (2003b): "Expectations and the Stability Problem for Optimal Monetary Policies," *Review of Economic Studies*, 70, 807–824.

——— (2006): "Monetary Policy, Expectations and Commitment," *Scandinavian Journal of Economics*, 108, 15–38.

——— (2008): "Expectations, Learning and Monetary Policy: An Overview of Recent Research," *Centre for Dynamic Macroeconomic Analysis Working Paper Series*, 08/02.

Evans, G. W. and B. McGough (2005): "Stable Sunspot Solutions in Models with Predetermined Variables," *Journal of Economic Dynamics and Control*, 29, 601–625.

Fidrmuc, J. and I. Korhonen (2003): "Similarity of Supply and Demand Shocks between the Euro Area and the CEECs," *Economic Systems*, 27, 313–334.

Grossman, S. J. and J. E. Stiglitz (1980): "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, 70, pp. 393–408.

Guse, E. A. (2008): "Learning in a Misspecified Multivariate Self-referential Linear Stochastic Model," *Journal of Economic Dynamics and Control*, 32, 1517–1542.

——— (2010): "Heterogeneous Expectations, Adaptive Learning, and Evolutionary Dynamics," 74, 42–57.

Heinemann, M. (2009): "E-stability and Stability of Adaptive Learning in Models with Private Information," *Journal of Economic Dynamics and Control*, 33, 2001–2014.

Hommes, C. and M. Zhu (2014): "Behavioral Learning Equilibria," *Journal of Economic Theory*, 150, 778–814.

Honkapohja, S. and K. Mitra (2004a): "Are Non-fundamental Equilibria Learnable in Models of Monetary Policy," *Journal of Monetary Economics*, 51, 1743–1770.

——— (2004b): "Monetary Policy with Internal Central Bank Forecasting: A Case of Heterogenous Information," in *From National Gain to Global Markets, Essays in Honour of Paavo Okko on His 60th Birthday*, ed. by M. Widgren, Helsinki: Taloustieto, 99–108.

——— (2005): "Performance of Monetary Policy with Internal Central Bank Forecasting," *Journal of Economic Dynamics and Control*, 29, 627–658.

——— (2006): "Learning Stability in Economies with Heterogeneous Agents," *Review of Economic Dynamics*, 9, 284–309.

Marcet, A. and T. J. Sargent (1989a): "Convergence of Least-Squares Learning in Environments with Hidden State Variables and Private Information," *Journal of Political Economy*, 97, 1306–1322.

——— (1989b): "Convergence of Least-Squares Learning Mechanisms in Self-Referential Linear Stochastic Models," *Journal of Economic Theory*, 48, 337–368.

McCallum, B. T. (2004): "On the Relationship between Determinate and MSV Solutions in Linear RE Models," *Economics Letters*, 84, 55–60.

——— (2007): "E-stability vis-a-vis Determinacy Results for a Broad Class of Linear Rational Expectations Models," *Journal of Economic Dynamics and Control*, 31, 1376–1391.

Milani, F. (2008): "Learning, Monetary Policy Rules, and Macroeconomic Stability," *Journal of Economic Dynamics and Control*, 32, 3148–3165.

Mitra, K., G. W. Evans, and S. Honkapohja (2013): "Policy Change and Learning in the RBC Model," *Journal of Economic Dynamics and Control*, 37, 1947–1971.

Muto, I. (2011): "Monetary Policy and Learning from the Central Bank's Forecast," *Journal of Economic Dynamics and Control*, 35, 52–66.

Pfajfar, D. (2013): "Formation of Rationally Heterogeneous Expectations," *Journal of Economic Dynamics and Control*, 37, 1434–1452.

Slobodyan, S. and R. Wouters (2012): "Learning in an Estimated Medium-scale DSGE Model," *Journal of Economic Dynamics and Control*, 36, 26–46.

Stiglitz, J. E. and A. Weiss (1981): "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, 71, 393–410.

Taylor, J. B. (1993): "Discretion versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy*, 39, 195–214.

Trenkler, G. (1995): "Eigenvalues of the Product of Nonnegative Definite Matrices," *Econometric Theory*, 11, 808–808.

Woodford, M. (1999): "Optimal Monetary Policy Inertia," *Manchester School, Supplement*, 61, 1–35.

Table 1: Learnability Conditions under CS and HM Learning

	Unobservable SS	Observable SS
CS	$\kappa\left(\phi_\pi - 1\right) + \phi_x\left(1-\beta\right) > 0$	$\kappa\left(\phi_\pi - \lambda^c\right) + \phi_x\left(1-\beta\lambda^c\right) > -\frac{(1-\lambda^c)(1-\beta\lambda^c)}{\alpha}$
HM	$\kappa\left(\phi_\pi - 1\right) + \phi_x\left(1-\beta\right) > 0$	$\kappa\left(\phi_\pi - \lambda^h\right) + \phi_x\left(1-\beta\lambda^h\right) > -\frac{\left(1-\lambda^h\right)\left(1-\beta\lambda^h\right)}{\alpha}$

Note: The derivations of the learnability conditions are summarized in Appendix E. $\lambda^c \equiv \lambda\left[\Phi_n\right]$, $\lambda^h \equiv \lambda\left[\Phi_n\left(\frac{1}{n}\Psi_n\right)\right]$, and $0 \leq \lambda^h \leq \lambda^c < 1$.

Figure 1 Parameter Domains of Learnability Conditions

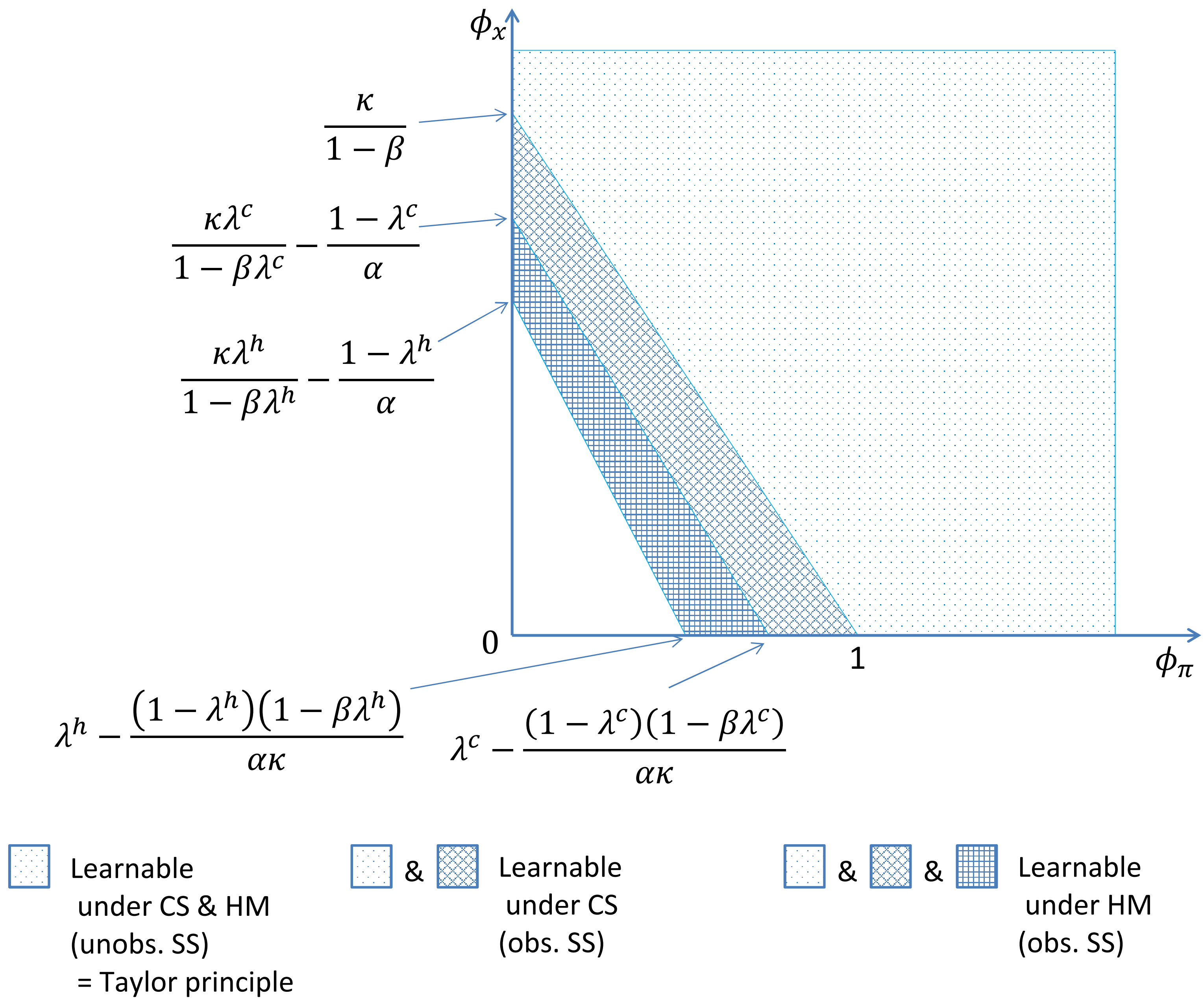
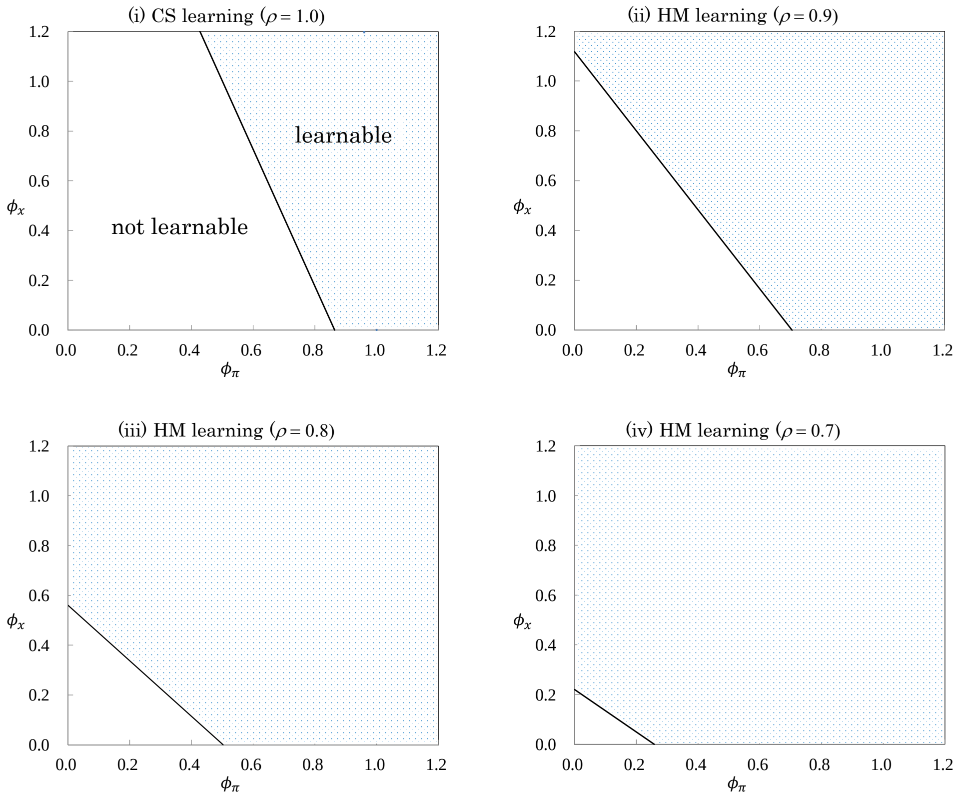
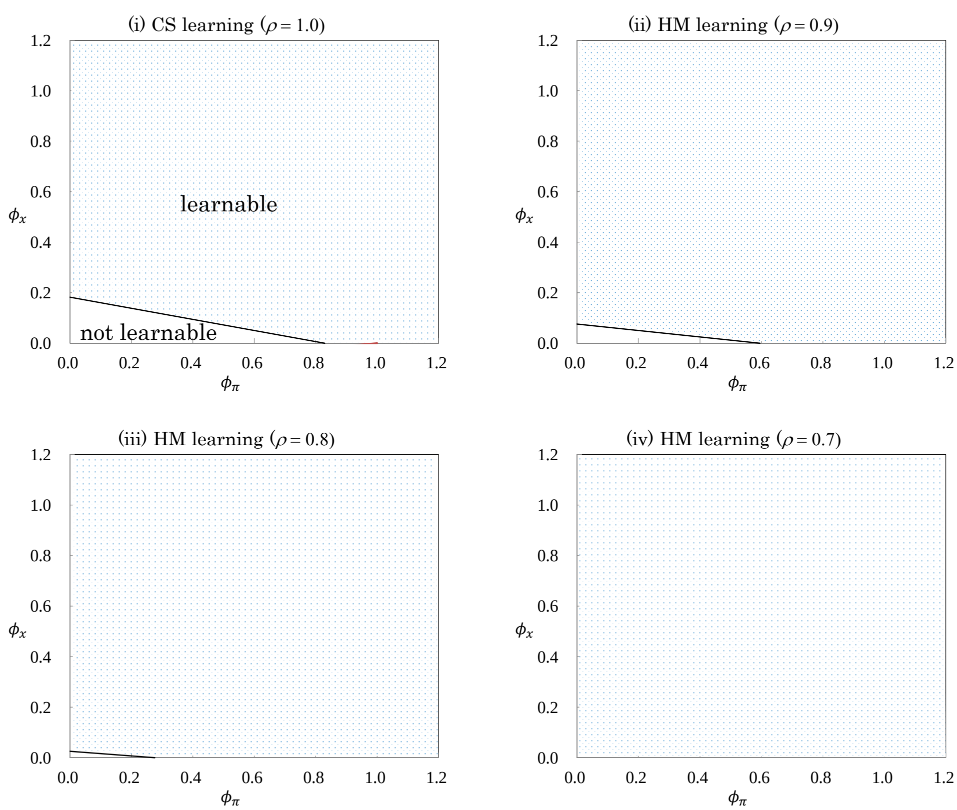


Figure 2 Heterogeneity and Learnability Conditions
(with the observable steady state)

(a) $\alpha = 1$, $\kappa = 0.3$, $\beta = 0.99$, $\varphi = 0.9$, $n = 10$



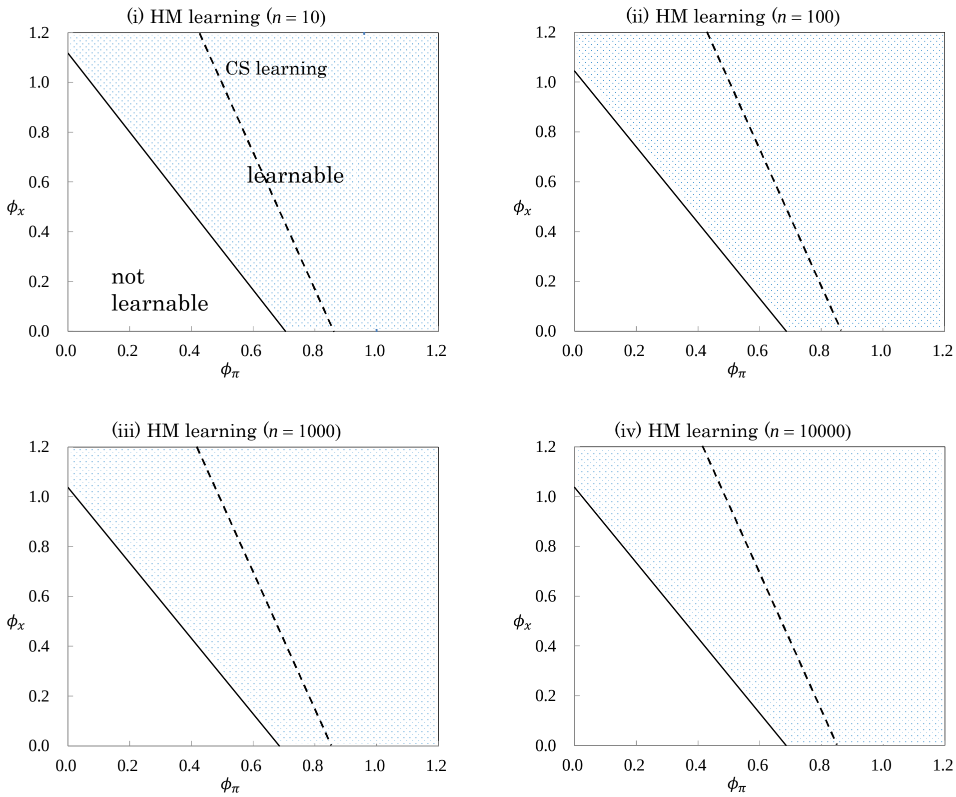
(b) $\alpha = 1/0.157$, $\kappa = 0.024$, $\beta = 0.99$, $\varphi = 0.9$, $n = 10$



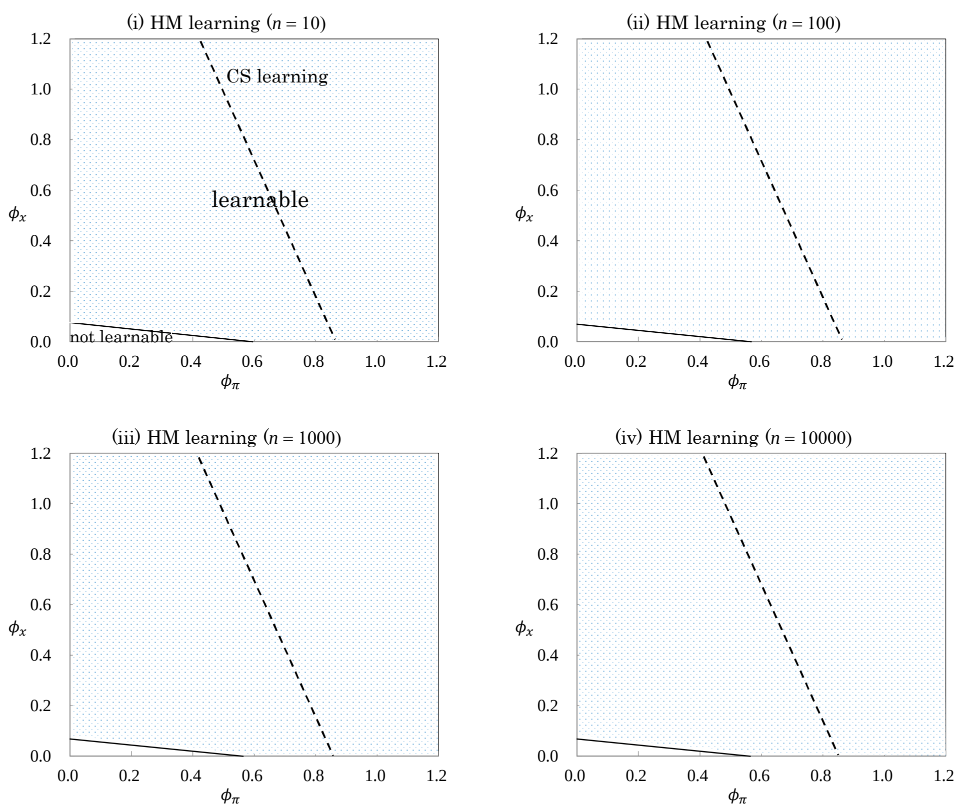
Note: Solid lines in each panel represent the boundaries of parameter domains satisfying the learnability condition under HM learning.

Figure 3 Misspecification and Learnability Conditions
(with the observable steady state)

(a) $\alpha = 1,\ \kappa = 0.3,\ \beta = 0.99,\ \varphi = 0.9,\ \rho = 0.9$



(b) $\alpha = 1/0.157,\ \kappa = 0.024,\ \beta = 0.99,\ \varphi = 0.9,\ \rho = 0.9$

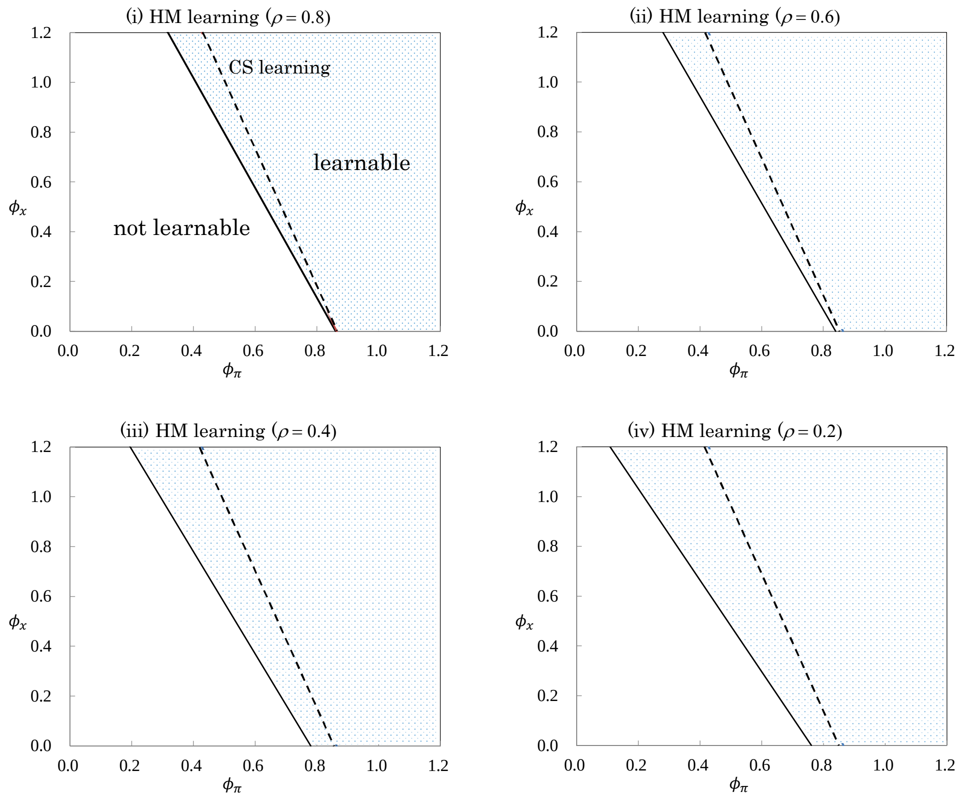


Note: Solid and dotted lines in each panel represent the boundaries of parameter domains satisfying the learnability conditions under HM and CS learning, respectively.

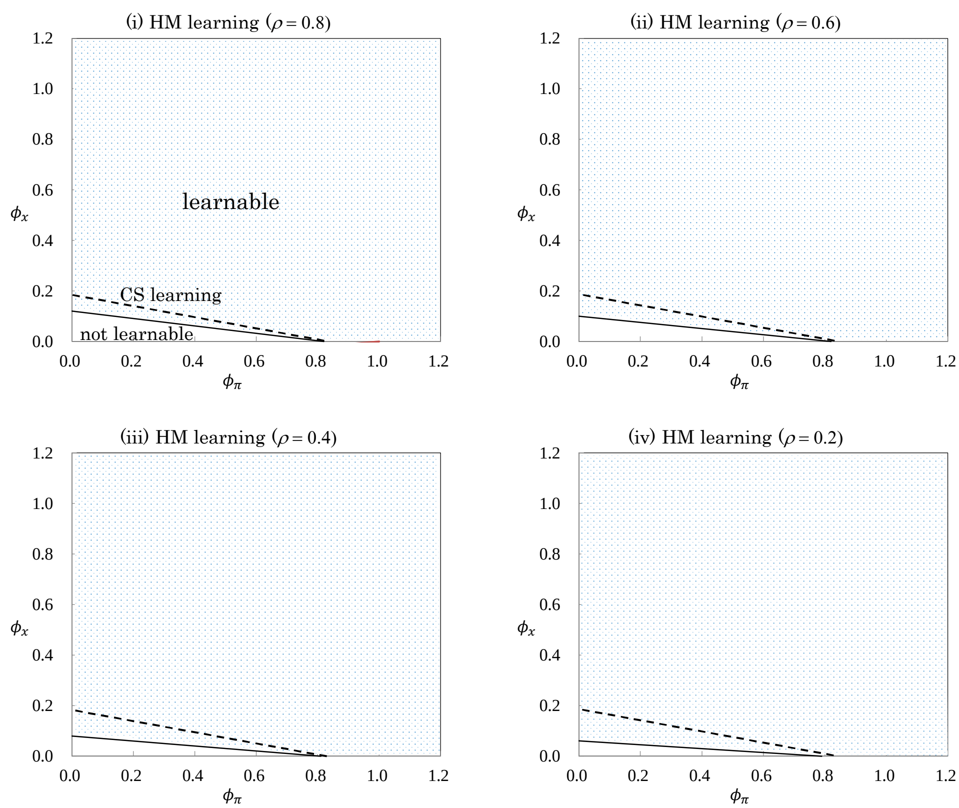
Figure 4 Heterogeneity and Learnability Conditions

(with the observable steady state and the PLMs with y_{t-1})

(a) $\alpha = 1$, $\kappa = 0.3$, $\beta = 0.99$, $\varphi = 0.9$, $n = 10$



(b) $\alpha = 1/0.157$, $\kappa = 0.024$, $\beta = 0.99$, $\varphi = 0.9$, $n = 10$



Note: Solid and dotted lines in each panel represent the boundaries of parameter domains satisfying the learnability conditions under HM and CS learning, respectively.