ESSAYS IN MACROECONOMICS AND DYNAMIC FACTOR MODELS

By

Ziyi Guo

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

August, 2013

Nashville, Tennessee

Approved:

Professor Mototsugu Shintani

Professor Mario J. Crucini

Professor Kevin X.D. Huang

Professor David C. Parsley

To my parents and my old sister of whose lives I am proud.

ACKNOWLEDGMENTS

I am especially grateful to my advisor Dr. Mototsugu Shintani, who gives me tremendous guidance and support on this and many other research. I cannot thank him enough for the countless discussions and insights he has passed on to me. I also especially thank my committe members - Dr. Mario J. Crucini, Dr. Kevin X.D. Huang and Dr. David C. Parsley for their valuable comments and useful discussions.

I also have been benefited from the insightful comments from my presentations at Vanderbilt University, the 2012 Midwest Macroeconomics Meetings at University of Notre Dame and the 2012 Midwest Econometrics Meetings at University of Kentucky. In particular, I thank Dr. Chris Benett, Eric Bond, William Collins, Yanqin Fan, Gregory Huffman, Tong Li, Jingfeng Lu, and Joel Rodrigue.

This work has been generously supported by College of Arts & Sciences Summer Research Award, Kirk Dornbush Summer Research Grant, and Graduate Research Assistantship at Vanderbilt University. Without these financial supports, this work would not have been possible. Also, without the professional assistance provided by Ms. Kathleen Finn, this work cannot be completed.

I am also grateful for all of the classmates and officemates who have made my years in graduate school a great experience. I am looking forward to interacting and working with many of these friends and colleagues in the future.

Finally, I would like to thank my parents, Turen Guo and Xinv Chen, as well as my old sister, Nina Guo, for their endless love and unconditional encouragement throughout my life. Without their support, I could not finish my degree in the end.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER I

NOISY INFORMATION IN AN INTERNATIONAL REAL BUSINESS CYCLE MODEL

Introduction

Standard international real business cycle (IRBC) models formulated by Backus, Kehoe, and Kydland (BKK, 1992, 1995) have been considered a natural starting point to assess the quantitative implications of dynamic stochastic general equilibrium (DSGE) models in an open economy environment. Since the standard IRBC model under assumptions of flexible prices and perfect competition cannot replicate all the observed characteristics of international business cycles, a number of extended models with more realistic features have been developed in the past two decades. Most importantly, incorporating monopolistic competition and sticky prices, along with the monetary sector in open economy DSGE models has been proven to be very successful in matching the data. In contrast to a large interest in the role of nominal rigidities, however, few studies have attempted to formally assess the quantitative implications of introducing informational frictions in the model. In this paper, we introduce a noisy information structure in an otherwise standard IRBC model and show that an extension in this direction is also useful in understanding some key features of international comovements of output, consumption, and labor.

We consider an imperfect information variant of a standard two-country bond-economy IRBC model similar to the one used in Baxter and Crucini (1995) and Heathcote and Perri (2002), except that we exclude capital accumulation from the model. While we believe that an open economy DSGE model with nominal rigidities is more realistic, we maintain the assumptions of perfect competition and flexible prices in this paper simply because they provide a reasonable benchmark in evaluating the pure effect of imperfect information on the international business cycle properties. In terms of explaining the international comovement, the original BKK model predicts negative (or near-zero) output correlation, near-perfect consumption correlation, and negative correlation of factors of production, all of which contradict the data. To improve the performance of the model, Baxter and Crucini (1995) and Kollman (1996) replaced the complete market assumption of the BKK model with the incomplete market assumption, so that consumers only have access to a real bond market. A convenient approach to ensure a unique stationary solution to an open economy model of incomplete market is to impose a (small) real cost of bond holding (see Heathcote and Perri, 2002, and Schmitt-Grohe and Uribe, 2003). According to the simulation results reported by Boileau and Normandin (2008, Table 1), under the stationary technology process with positive international spillovers, an incomplete market model with a tiny bond holding cost can yield positive international output correlation, but its magnitude is still less than the data.[1] As in the original BKK model, we focus on stationary technology shocks with international spillovers as a source of aggregate fluctuations. However, domestic firms are assumed to observe the current foreign technology with noise. We first show that when the information noise is sufficiently large, the model can match the positive output comovement in the data not only for the case of incomplete market but also for the case of perfect international risk sharing.

Even in the case of incomplete market where international consumption risk sharing is restricted, the standard IRBC models with stationary technology shocks are known to predict international consumption correlation higher than the international output correlation (see Boileau and Normandin, 2008, Table 1). The data, however, typically suggest that the former is lower than the latter (see Ambler, Cardia, and Zimmermann, 2004). To narrow the gap between output correlation and consumption correlation predicted by the model, several different channels have been emphasized in the literature. For example,

[1] Baxter and Crucini (1995) emphasized the better performance of the bond economy model when the technology is highly persistent and there is no international spillover.

the proposed channels include nontraded goods (Stockman and Tesar, 1995), endogenous incomplete market with limited enforcement (Kehoe and Perri, 2002), sticky prices (Chari, Kehoe, and McGrattan, 2002) and variable capital utilization (Baxter and Farr, 2005). In this paper, we highlight the information channel and show that the presence of a noisy information structure in the household sector helps to fill the gap between the cross country output correlation and consumption correlation.

In the recent global financial crisis of 2007-2009, employment and hours worked declined both in the US and Euro area. Such a positive comovement is not predicted by the standard IRBC models but can be generated in our imperfect information variant of the model. Furthermore, since the labor declined more in the US than in the Euro area, observed labor productivity increased in the US, which contrasts to the Euro area where near-zero or negative productivity growth was observed. The empirical observation of near-zero (or negative) correlation between productivity and hours worked has been viewed as a productivity-hours anomaly in the macroeconomic literature, since the standard real business cycle models predict positive response of hours worked to positive technology shocks, provided an upward sloping labor supply curve (see for example, Galí, 1999, and Christiano, Eichenbaum, and Vigfusson, 2003). To explain the negative productivity-hours correlation, Galí (1999) emphasizes the role of monetary policy shocks and sticky prices. In this paper, we show that negative productivity-hours correlation can also be predicted from the noisy information structure even if prices are flexible and that heterogenous observations in two regions can be obtained if the fraction of information-constrained consumers differs across regions.

We note that there are other studies that emphasize the role of imperfect information structures in open economy macroeconomic models. For example, Gourinchas and Tornell (2004) discuss distorted beliefs of investors, while Bacchetta and van Wincoop (2006) and Crucini, Shintani, and Tsuruga (2010), respectively, introduce the heterogenous information and sticky information structures in open economy monetary models. However, these studies mainly focus on explaining nominal and real exchange rate dynamics rather than the international comovement of real variables. Luo, Nie, and Young (2010) introduce the rational inattention to an intertemporal current account model. However, since the intertemporal current account model is a small open partial equilibrium model, it is not suitable for understanding cross-country correlations. In our paper, we introduce a noisy information structure in a two-country economy general equilibrium model with direct implications on cross-country comovements. Our approach is similar in spirit to Angeletos and La'O (2009), who introduce an imperfect common knowledge structure in a close-economy real business cycle model and show that the model can induce a negative short-run response of employment to productivity shocks. Unlike their model where heterogenous information across firms plays an important role, we assume homogeneous information across firms but only allow heterogenous information between countries. Even with such a simple information structure, the model still has rich implications on international business cycle features.

The remainder of the paper is organized as follows. Our two-country model with noisy information is introduced in Section 2. Section 3 discusses the implications of our model on output, consumption and labor in order. Section 4 extends the baseline model with capital accumulation. Section 5 concludes.

Model

Our baseline international real business model is a simplified version of the two-country bond-economy model of Baxter and Crucini (1995) from which we have eliminated capital accumulation. We introduce information noise to both firms and households in the baseline model and compare it with the case of perfect information. Foreign country variables are denoted by stars.

Firms

Firms in the domestic country produce the same final good as firms in the foreign country. Labor is internationally immobile but the labor market is competitive. Firms produce the output using a diminishing-returns-to-scale technology

$$Y_t = A_t N_t^{\theta} \tag{I.1}$$

$$Y_t^* = A_t^* N_t^{*\theta}$$

where $Y_t(Y_t^*)$ is the output in the home (foreign) country, $A_t(A_t^*)$ is the technology level in the home (foreign) country, $N_t(N_t^*)$ is labor employed in the home (foreign) country, and $\theta \in (0,1)$. Domestic and foreign firms maximize expected value of their profits, $\Pi_t = P_t Y_t - w_t N_t$ and $\Pi_t^* = P_t Y_t^* - w_t^* N_t^*$, respectively, where common price P_t of the final goods in two countries is normalized to one and $w_t(w_t^*)$ is the wage rate in the domestic (foreign) country. We assume firms in a country are owned by the residents of the same country so that the profits Π_t and Π_t^* are given to consumers in corresponding countries[2]. Technology follows the VAR(1) model given by

$$\begin{bmatrix} \log A_t \\ \log A_t^* \end{bmatrix} = \begin{bmatrix} \rho & \nu \\ \nu & \rho \end{bmatrix} \begin{bmatrix} \log A_{t-1} \\ \log A_{t-1}^* \end{bmatrix} + \begin{bmatrix} \epsilon_t \\ \epsilon_t^* \end{bmatrix} \tag{I.2}$$

where $\nu(>0)$ represents technology spillovers, and $\epsilon_t, \epsilon_t^* \sim N(0, 1/k_a)$ and $corr(\epsilon_t, \epsilon_t^*) \equiv \eta$. Domestic firms know their own level of technology, but receive a signal (with noise) for the technology level of firms in the foreign country. The signals received by home and foreign firms at the beginning of each period t are respectively given by

$$x_t = \log A_t^* + \upsilon_t$$

[2]Here we exclude the possiblity of cross-border ownerships of firms.

and

$$x_t^* = \log A_t + \upsilon_t^*$$

where $\upsilon_t, \upsilon_t^* \sim N(0, 1/k_x)$.

Households

Each country consists of two types of consumers. The first type (type 1) decides the consumption level based on the same information set as the firms located in the same country. The fraction of the type 1 consumers in the home (foreign) country is represented by $\kappa(\kappa^*)$. The remaining consumers choose their consumption level after the information on the foreign technology level is revealed. Households consume the final products and supply labor to firms located in the same country. Each type of consumer in the home country maximizes the expected value of the discounted sum of utility given by

$$\sum_{t=0}^{\infty} \beta^t [\frac{C_{it}^{1-\gamma}}{1-\gamma} - \frac{N_{it}^{1+\epsilon}}{1+\epsilon}]$$

conditional on the information available at the decision timing, where C_{it} and N_{it} are consumption and labor supply of type i ($i = 1, 2$) consumers, $\gamma(\geq 0)$ is the reciprocal of the intertemporal elasticity of substitution or relative risk aversion, $\epsilon(\geq 0)$ is the reciprocal of the Frisch elasticity of labor supply, and β is the discount factor. The international asset market is restricted to trade only non-contingent bonds. The household budget constraint is given by

$$C_{it} + Q_t B_{it+1} + \frac{\pi}{2} B_{it+1}^2 \leq \Pi_t + w_t N_{it} + B_{it}$$

where B_{it} is bonds held by the type i consumers, Q_t $(= (1 + r_t)^{-1})$ is the price of bonds in units of good, r_t is the world interest rate, and $(\pi/2)B_{it+1}^2$ is a quadratic holding cost of bonds with π being a small positive value. The household maximization problem is similarly defined for foreign consumers with preferences identical to domestic consumers.

For the timing of decisions made by firms and households, we follow the setting of Angeletos and La'O (2009) and consider each period in two stages. At the beginning of each time period (stage 1), firms and labor representatives of households meet and decide the production level based on the information set $\{\log A_t, x_t\} \cup \Omega_{t-1}$. All households make labor supply decisions at this stage. The type 1 consumers, κ fraction of households, also determine their consumption level (which cannot be adjusted in the next stage). Firms produce final goods. Then, at the end of each time period (stage 2), information on foreign productivity is revealed. The type 2 consumers, the remaining $1-\kappa$ fraction of households, make their consumption-saving decisions based on the updated information set $\Omega_t = \{\log A_t, \log A_t^*, x_t, x_t^*\} \cup \Omega_{t-1}$. The interest rate level and the real wage rate level are determined where the bond market and the labor market clear. Countries export or import goods in the world market.

Reis (2006) built a microfoundation of inattentive consumers, who update their information sporadically. Mankiw and Reis (2006) further considered the role of inattentive consumers in a general equilibrium framework. In our model, type 1 consumers play a role similar to that of inattentive consumers (planner) considered in Mankiw and Reis (2006), except that we allow our consumers to observe a signal. The presence of type 2 consumers, who make their consumption decision after all the information is revealed, is essential in closing our model so that $\kappa = 1$ case is excluded in the analysis. The timing of the decision made by type 2 consumers is important in avoiding strategic responses by firms and type 1 consumers. In the beginning of each period, neither firms nor type 1 consumers can observe prices to extract the information about the state of the economy. Since there is no strategic responses by firms, they make their production decisions based on their expected value of the price, conditional on their restricted information set. Likewise, type 1 consumers make their saving-borrowing decisions based on their conditional expectation of the interest rate.

Equilibrium

Labor is internationally immobile so that the labor market clearing condition for each country is respectively given by

$$N_t = \kappa N_{1t} + (1-\kappa) N_{2t}$$

and

$$N_t^* = \kappa^* N_{1t}^* + (1-\kappa^*) N_{2t}^*$$

Trade across countries is allowed so that the world goods-market clearing condition (resource constraint) is given by

$$Y_t - C_t + Y_t^* - C_t^* - \frac{\kappa\pi}{2} B_{1t+1}^2 - \frac{\kappa^*\pi}{2} B_{1t+1}^{*2} - \frac{(1-\kappa)\pi}{2} B_{2t+1}^2 - \frac{(1-\kappa^*)\pi}{2} B_{2t+1}^{*2} = 0$$

where

$$C_t = \kappa C_{1t} + (1-\kappa) C_{2t}$$

and

$$C_t^* = \kappa^* C_{1t}^* + (1-\kappa^*) C_{2t}^*.$$

Finally, the Walras' Law implies that the remaining bond market clears as

$$\kappa B_{1t} + \kappa^* B_{1t}^* + (1-\kappa) B_{2t} + (1-\kappa^*) B_{2t}^* = 0$$

so that bonds are in zero net supply at the world level.

Implications for International Business Cycles

International Output Correlation ($\kappa = \kappa^* = 0$)

We first solve the model and investigate its implication on the international output correlation when $\kappa = \kappa^* = 0$ so that all the consumers can decide their consumption levels after the information about foreign technology is revealed. This setting is convenient for comparing the implication of the model under incomplete market assumption and that of the model under complete market assumption. To solve the model, we log-linearize all the first-order conditions and then use the guess-verification approach. That is, we assume a policy function to take a linear form and plug it into the model to match the coefficients of the same state variables in the two sides of the equations.

Let $y_t = \log Y_t - \log Y$ ($y_t^* = \log Y_t^* - \log Y^*$) and $b_t = B_t/Y$ ($b_t^* = B_t^*/Y^*$) where variables with no subscript imply steady state values. We then have the following results on the level of output.

Proposition 1 *Suppose $\kappa = \kappa^* = 0$ under the incomplete market assumption. Then, (i) the equilibrium level of output in the home country and in the foreign country is given by*

$$\begin{aligned} y_t &= m_{-1} \log A_{t-1} + m_{-1}^* \log A_{t-1}^* + m \log A_t + m_x x_t + m_b b_t \\ y_t^* &= m_{-1} \log A_{t-1}^* + m_{-1}^* \log A_{t-1} + m \log A_t^* + m_x x_t^* + m_b b_t^* \end{aligned} \tag{I.3}$$

for some coefficients $(m_{-1}, m_{-1}^*, m, m_x, m_b)$; and

(ii) the equilibrium value of the coefficients $(m_{-1}, m_{-1}^*, m, m_x, m_b)$ satisfies the following properties: m_{-1} and m_{-1}^* approach zero as $k_a/k_x \to 0$; m approaches a positive value as $k_a/k_x \to 0$; m_x approaches a negative value as $k_a/k_x \to 0$ and approaches zero as $k_z/k_x \to \infty$; and m_b is invariant to k_a/k_x.

To illustrate the reason why the model with noisy information provides a quantitatively different result on international output correlation from the full information model, it is helpful to first consider the case of the complete market which has a closed form solution. For the complete market case, firms' problems are the same as before but the households maximize expected value of

$$\sum_{t=0}^{\infty} \beta^t \{ \frac{C_t^{1-\gamma}}{1-\gamma} - \frac{N_t^{1+\epsilon}}{1+\epsilon} + \frac{C_t^{*1-\gamma}}{1-\gamma} - \frac{N_t^{*1+\epsilon}}{1+\epsilon} \}$$

which is common across countries subject to the world resource constraint

$$Y_t - C_t + Y_t^* - C_t^* = 0.$$

For the full information case, the solution is given by

$$\begin{aligned} y_t &= m \log A_t + m^* \log A_t^* \\ y_t^* &= m \log A_t^* + m^* \log A_t \end{aligned} \tag{I.4}$$

with $m = \frac{1-\frac{\alpha}{2}}{1-\alpha}\zeta > 0$ and $m^* = \frac{\frac{\alpha}{2}}{1-\alpha}\zeta < 0$ where $\alpha = \frac{-\theta\gamma}{1+\epsilon-\theta} < 0, \zeta = \frac{1+\epsilon}{1+\epsilon-\theta} > 0$. Note that the combination of $m > 0$ and $m^* < 0$ explain the reason why the the domestic output responds negatively to foreign technology shocks. When the empirical performance of the model is evaluated, both data series and the simulated series are typically filtered either by using the Hodorick-Prescott filter or the first difference filter. In this paper, we employ the latter and focus on the international correlations in terms of the log growth rates $\Delta y_t = y_t - y_{t-1}$ and $\Delta y_t^* = y_t^* - y_{t-1}^*$. Our choice of filter here is convenient for computing the predicted correlation directly when a closed form solution is provided, as in the case of (I.4). Given the technology process (III.2) with a typical choice of parameters, it is straightforward to show that (I.4) yields negative correlation of Δy_t and Δy_t^*.

If the noisy information structure is introduced in this complete market model, we have the following result.

Proposition 2 $\kappa = \kappa^* = 0$ *under the complete market assumption. Then, (i) the equilibrium level of outputs in the home country and in the foreign country is given by*

$$\begin{aligned} y_t &= m_{-1} \log A_{t-1} + m_{-1}^* \log A_{t-1}^* + m \log A_t + m_x x_t \\ y_t^* &= m_{-1} \log A_{t-1}^* + m_{-1}^* \log A_{t-1} + m \log A_t^* + m_x x_t^* \end{aligned} \tag{I.5}$$

where

$$\begin{cases} m_{-1} = \frac{[\frac{\alpha}{2}\rho+(1-\frac{\alpha}{2})\nu](1-\frac{\alpha}{2})\zeta\alpha}{2(1-\alpha)[(1-\frac{\alpha}{2})^2+(1-\alpha)(k_a/k_x)^{-1}]} \\ m^*_{-1} = \frac{[\rho-\frac{\alpha}{2}(\rho-\nu)](1-\frac{\alpha}{2})\zeta\alpha}{2(1-\alpha)[(1-\frac{\alpha}{2})^2+(1-\alpha)(k_a/k_x)^{-1}]} \\ m = \frac{(1-\frac{\alpha}{2})\zeta(1+(k_a/k_x)^{-1})}{(1-\frac{\alpha}{2})^2+(1-\alpha)(k_a/k_x)^{-1}} \\ m_x = \frac{\frac{\alpha}{2}\zeta(k_a/k_x)^{-1}}{(1-\frac{\alpha}{2})^2+(1-\alpha)(k_a/k_x)^{-1}} \end{cases}$$

(ii) the equilibrium value of the coefficients $(m_{-1}, m^*_{-1}, m, m_x)$ *approaches* $(0, 0, \frac{1-\frac{\alpha}{2}}{1-\alpha}\zeta, \frac{\frac{\alpha}{2}}{1-\alpha}\zeta)$ *as* $k_a/k_x \to 0$*.*

If we compare the coefficients in (I.4) and (I.5), the output responds less to currently observed variables and more to old information. When the relative precision of information becomes worse, firms rely less on the signal x_t, and more on old information $\log A^*_{t-1}$ so that m^*_{-1} becomes more negative as k_a/k_x increases. Even if there is a positive technology shock in the home country, since foreign firms cannot directly observe it, they do not reduce their production level as much as the full information case. Consequently, the home firms do not increase their production as much as the full information case and m becomes smaller as k_a/k_x increases. Again, we can easily compute the correlation of Δy_t and Δy^*_t explicitly based on (I.5).

To better understand the difference of the impact of imperfect information on complete and incomplete markets, we conduct a simple calibration exercise using the results from Propositions 1 and 2. We set parameters at $\theta = 0.64$, $\epsilon = 0.5$, $\gamma = 2$ and $\beta = 0.99$, values that are commonly used in the literature. We set $\pi = .0001$ for the quadratic cost of bond holding to assure a unique steady state. For the parameters appear in technology process (III.2), we use our own estimated values based on the quarterly series of output and hours worked from the US and Euro area. For the hours worked series in the Euro area, we obtain quarterly average weekly or monthly hours of work in manufacturing from 1989Q1 to 2009Q4 for Austria, France, Germany and Spain from LABORSTA. We then convert these series to quarterly hours worked series in all sectors, by using the ratio of annual hours of worker in manufacturing sector to that in all sectors, for each country, obtained from OECD Main Economic Indicator. The hours worked series for the US is obtained from the BLS. Quarterly real GDP series, obtained from OECD Quarterly National Accounts is used to construct output series for the US and Euro area. We then transform the hours worked series and output series to $\log A_t$ using (I.1) combined with $\theta = 0.64$. Using the estimation procedure employed by BKK, we obtain $\rho = 0.931$, $\nu = 0.046$ and $\eta = 0.040$, values that are very close to the ones used by BKK. The output (growth) correlation of the US and Euro area from 1989Q1 to 2009Q4 is 0.54 when the Euro area is based on the four countries we used to construct $\log A_t$. When we expand the output series of the Euro area to those from 15 European countries (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Norway, Netherlands, Portugal, Spain, Sweden and the United Kingdom), the output correlation from the same period becomes 0.32.

Figure 1 shows how the predicted correlation of Δy_t and Δy_t^* changes in response to changes in the relative precision of information k_a/k_x under two different asset market assumptions. The left panel shows the complete market case based on (I.5) and the right panel shows the incomplete market case based on (I.3). When the information is perfect ($k_a/k_x = 0$), the output correlation is negative for the complete market. As k_a/k_x increases, the correlation monotonically increases and becomes positive. In case of the incomplete market, the output correlation is positive but is much smaller than what the data suggests. Again, the correlation increases as k_a/k_x increases. For both cases, the model with a sufficiently large noise matches the observed output correlation from the data (0.54 and 0.32).

An intuitive explanation on the role of restricted information in increasing output correlation is as follows. The main reason why standard IRBC models generate negative or near zero correlation of output is that the domestic and foreign firms respond to technology shocks in the opposite direction. For example, with a positive productivity shock in the home country, domestic firms increase their production, while foreign firms decrease their production. In contrast, if foreign firms do not directly observe a positive shock at home country, they do not reduce their production. Furthermore, as a result of excess supply caused by uninformed foreign firms, home firms do not increase production as much as the fully informed case. Combining the effect of weaker responses with positively correlated technologies across two countries can yield positive output correlation.

International Consumption Correlation ($\kappa = \kappa^* > 0$)

We now focus on the bond-economy IRBC model when there are two types of consumers. We show that introducing type 2 consumers in the economy will make the international correlation of consumption lower compared to the benchmark model with full information. To simplify the argument, we here maintain that the fraction of type 1 consumers is common across the countries. As in the previous subsection, we use the first difference filter to investigate the international consumption correlation. Typically, the data suggests that international consumption growth correlation is less than the international output growth correlation. For example, Obstfeld and Rogoff (2000) use the annual Penn World Table data over 1973 to 1992 and find that the average international correlation in real GDP growth rates is 0.53, while the average consumption growth correlation is 0.40. We also compute the consumption growth rate correlation based on the data from OECD Quarterly National Accounts. If we use four countries for the Euro area, consumption correlation is 0.46 compared to the outputs correlation of 0.54 during the period from 1989Q1 to 2009Q4. When we use 15 European countries to construct Euro aggregates, the consumption correlation is 0.26, but the outputs correlation is 0.32. In either case, consumption correlation is lower than the output correlation, which cannot be predicted by the standard full information model.[3]

To solve the model with $\kappa = \kappa^* > 0$, we need to combine an extended version of

[3] Pakko (2004) uses 10 country data from 1973:Q1 to 2002:Q4 and show that for all countries, the correlations of output growth rates is higher than that of consumption growth rates.

Sims' (2001) approach and the guess-verification approach. We decompose heterogeneous expectations into homogeneous expectation component and expectation error component. We then solve the model by treating as if the latter is an exogenous shock in the first step. In the second step, we use the method of undetermined coefficients to assure the endogenous expectation errors consistent with the solution from the first step (see technical appendix for details). All the parameter values are the same as before except that we set $\kappa = 0.6$. The solutions are obtained for both (y_t, y_t^*) and (c_t, c_t^*) where $c_t = \log C_t - \log C$ and $c_t^* = \log C_t^* - \log C^*$.

To understand the characteristics of the model, we compute the impulse response of consumption to one standard positive deviation of domestic technology shocks with three different choice of relative precision of information, $k_a/k_x = 0$, 1, and 25, which is shown in Figure 2[4]. Since the information is revealed at the end of each period, the effects of information precision become almost negligible after one period. In the perfect information case ($k_a/k_x = 0$), households in the home country increase their consumptions as their income increases. Households in the foreign country also increase consumption, since the spillover effects of the positive technology shocks make foreign households to borrow from the home country. When the information noise becomes large ($k_a/k_x = 1$, and 25), foreign households cannot predict the increase in income in the future and do not borrow as much as they should from the international asset market. Therefore, even if foreign firms produce relatively more than the perfect information case, foreign households still decrease their consumption. This asymmetric responses of c_t and c_t^* is even more amplified by taking the first difference Δc_t and Δc_t^*. This makes consumption growth correlation decreasing with respect to the magnitude of information noise.

Figure 3 demonstrates the dynamics of consumptions growth correlation and out-

[4] Since both information-constrained and unconstrained consumers have rational expectations, as long as κ is not extremely large the calibration of our exercise shows the response of interest rates to technology shocks is not unrealistic.

puts growth correlation in response to different degrees of information frictions. As in the case of $\kappa = \kappa^* = 0$, we can see information noise increases outputs growth correlation, and at the same time it reduces consumptions-growth correlation. When the relative precision of information (k_a/k_x) reaches 5, the consumption growth correlation becomes less than the output growth correlation which dramatically reduces the gap between the prediction of the model and the data.

International Productivity-Hours Dynamics ($\kappa \neq \kappa^*$)

In the recent global financial crisis of 2007-2009, employment and hours worked declined both in the US and Euro area. Such a positive comovement is not predicted by the standard IRBC models. Furthermore, since the labor declined more in the US than in Euro area, observed labor productivity increased in the US which contrast to the Euro area where near-zero or negative productivity growth was observed. This fact was first investigated by Ohanian (2010). The empirical observation of near-zero (or negative) correlation between productivity and hours worked has been viewed as a productivity-hours anomaly in the macroeconomic literature since the standard real business cycle model predicts a positive response of hours worked to positive technology shocks, provided an upward sloping labor supply curve (see Galí, 1999; Christiano, Eichenbaum and Vigfusson, 2003).

Let us first show that given a certain range of parameter values, our model can predict the positive comovement of labor input, which cannot be obtained in the full information model. In our data, the hours worked (growth) correlation between the US and Euro area based on four European countries is positive at 0.20. Using the same solution technique as before, we can obtain the solution for (n_t, n_t^*) where $n_t = \log N_t - \log N$ and $n_t^* = \log N_t^* - \log N^*$. Figure 4 shows the predicted international correlation of hours worked using the same set of parameter values as before. For the perfect information case with $k_a/k_x = 0$, the correlation is negative. The correlation is not monotonically increasing in k_a/k_x. However, it predicts the positive correlation when k_a/k_x lies between the values of 0.1 and 0.5.

We also solve the model when the fraction of information constrained consumers differs across the country. Figure 5 shows the predicted correlation of hour worked growth, $\Delta n_t(\Delta n_t^*)$, and measured productivity growth, $\Delta y_t - \Delta n_t(\Delta y_t^* - \Delta n_t^*)$, when $\kappa = 0.1$ and $\kappa^* = 0.7$. It shows that when k_a/k_x increases, the model can predict negative productivity-hours correlation in one region and positive productivity-hours correlation in the other region (Figure 6), where the former represents the Euro area and the latter represents the US.

The Model with Capital Accumulation

In this section, we extend our model, and show that our main results are consistent for models with and without capital as an input in the production functions. The model structure is the same as in section 2, and the only difference is firms' production functions and households' budget-constraint equations. The production functions is

$$Y_t = A_t K_t^{1-\theta} N_t^{\theta}$$

for the home country and

$$Y_t^* = A_t^* K_t^{*1-\theta} N_t^{*\theta}$$

for the foreign country, where $K_t(K_t^*)$ is the capital stock for the home (foreign) country. In addition to investing in the financial capital market, households also invest in the physical capital market. The households' budget constraint is

$$C_t + I_t + B_{t+1} + \frac{\pi}{2} B_{t+1}^2 \leq r_t K_t + W_t N_t + R_t B_t$$

for the home country and

$$C_t^* + I_t^* + B_{t+1}^* + \frac{\pi}{2} B_{t+1}^{*2} \leq r_t^* K_t + W_t^* N_t^* + R_t B_t^*$$

for the foreign country, if households can only trade real bonds across countries. If households can trade state-contingent bonds internationally, the households' budget constraint becomes

$$C_t + C_t^* + I_t + I_t^* = Y_t + Y_t^*,$$

where $I_t(I_t^*)$ is the investment in the physical capital in the home (foreign) country and $r_t(r_t^*)$ is the interest rate in the capital renting market in the home (foreign) country. The capital stock evolves according to

$$K_{t+1} = (1 - \delta) K_t + \phi(\frac{I_t}{K_t}) K_t$$

for the home country and

$$K_{t+1}^* = (1 - \delta) K_t^* + \phi(\frac{I_t^*}{K_t^*}) K_t^*$$

for the foreign country, where δ is the capital depreciation rate and the function $\phi(\cdot)$ implies an adjustment cost. The function $1/\phi^{'}$ is Tobin's q, which gives the number of units of output which must be foregone to increase the capital stock in a particular location by one unit.

The solution method is the same as the method used in section 3.2. We log-linearize the first-order conditions of the model first, and then use an approach combining an extended version of Sims' (2001) approach and the guess-verification approach. To calibrate the model, in addition to the parameters specified before, we need specify two additional parameters, δ and $\phi(\cdot)$. δ is set to 0.025 which indicates that capital depreciates at the rate of 2.5 percent per quarter. The solution method does not require us to specify the function form of ϕ but requires us to set the values of ϕ, ϕ' and ϕ'' in the steady state. We choose $\phi(\frac{I}{K}) = \delta$ and $\phi'(\frac{I}{K}) = 1$ so that the model with adjustment costs has the same steady state as the model without adjustment costs. $\phi''(\frac{I}{K})$ is set to equal -2.5 as in Christiano, Eichenbaum and Evans (2005).

The calibration results show that including capital accumulation in the model does not change the results of the model quantitatively significantly. In Figure 7, one can see when $\kappa = \kappa^* = 0$, information frictions increase the output correlation from around 0.14 to 0.52 for both the complete market case and incomplete market case. When we introduce information frictions into the consumption side by choosing $\kappa = \kappa^* = 0.7$, a slightly bigger value than the one we used for the model without capital accumulation, the consumption correlation declines from around 0.96 to around 0.40 when k_a/k_x increases from 0 to 10, and is exceeded by the output correlation (Figure 8). From Figure 9, one can see the implication of international comovement of labor inputs is much more significant when capital accumulation is allowed in the model. The hours worked correlation increases monotonically with the degree of information frictions. When $k_a/k_x > 1$, the hours worked correlation becomes positive. If we choose $\kappa = 0.1$ and $\kappa = 0.7$, the positive productivity-hours correlation in the home country and negative correlation in the foreign country can also be generated if the degree of information frictions is chosen to be large enough (Figure 10).

Conclusion

We introduced a noisy information structure into an otherwise standard international real business cycle model with two countries. When domestic firms observe current foreign technology with some noise, prediction of the model on international correlation turned out to be very different from that of a standard perfect information model. First, we found that the imperfect information model can explain positive output correlation both in complete and incomplete market models. Second, consumption correlation became smaller than output correlation when the precision of the information becomes worse in the presence of information constrained households. Third, the model can explain the observation of positive productivity-hours correlation in one country and negative correlation in the other country.

There are several directions in which our model can be extended. First, we can allow for information heterogeneity not only across the countries but also within a country. When firms in the same country face different signals about the foreign technology, the lagged foreign technology will have a role of public information, in addition to its role as the predictor of the current foreign technology. This may amplify the effect of noisy information and increase the predicted international output comovement. Second, we can introduce nominal shocks into the model and consider the possibility of confusion between nominal and real shocks. Third, we can investigate the role of the possible correlation of noise shocks across countries for the output correlations. Finally, it would be another contribution to the literature if one can estimate the parameters of variances of noise shocks and of proportions of information-constrained households within a country[5]. These extensions are left for future research.

[5] To the best of my knowledge, until now there is no estimation work on to what degree noise shocks can quantitatively explain business cycle fluctuations in a DSGE framework. Olivier J. Blanchard, Guido Lorenzoni, and Jean Paul L'Huillier (2012) use a structural vector autoregression (SVAR) and argue that noise shocks explain around half of business cycle fluctuations.

Appendix A

Proof of Proposition 1

The bond economy can be fully characterized by the following first-order conditions (normalize $P_t = 1$):

$$E_{ht}\omega_t - E_{ht}(A_t\theta N_t^{\theta-1}) = 0 \tag{I.6}$$

$$E_{ft}\omega_t^* - E_{ft}(A_t^*\theta N_t^{*\theta-1}) = 0 \tag{I.7}$$

$$Y_t - A_t N_t^{\theta} = 0 \tag{I.8}$$

$$Y_t^* - A_t^* N_t^{*\theta} = 0 \tag{I.9}$$

$$\lambda_t - C_t^{-\gamma} = 0 \tag{I.10}$$

$$\lambda_t\omega_t - N_t^{\epsilon} = 0 \tag{I.11}$$

$$\lambda_t(1 + \pi B_{t+1}) - E_t(\beta\lambda_{t+1}R_{t+1}) = 0 \tag{I.12}$$

$$B_{t+1} + \frac{\pi}{2}B_{t+1}^2 + C_t - Y_t - R_t B_t = 0 \tag{I.13}$$

$$\lambda_t^* - C_t^{*-\gamma} = 0 \tag{I.14}$$

$$\lambda_t^*\omega_t^* - N_t^{*\epsilon} = 0 \tag{I.15}$$

$$\lambda_t^*(1 + \pi B_{t+1}^*) - E_t(\beta\lambda_{t+1}^* R_{t+1}) = 0 \tag{I.16}$$

$$B_{t+1}^* + \frac{\pi}{2}B_{t+1}^{*2} + C_t^* - Y_t^* - R_t B_t^* = 0 \tag{I.17}$$

$$B_t + B_t^* = 0 \tag{I.18}$$

Boundary condition:

$$\lim_{t\to\infty} \beta^t \lambda_t B_t = 0 \tag{I.19}$$

and

$$\lim_{t\to\infty} \beta^t \lambda_t^* B_t^* = 0 \tag{I.20}$$

where $\lambda_t(\lambda_t^*)$ is the Lagrange multipliers. Combine equations (6), (8), (10) and (11) and log-linearize,

$$y_t = c_0 + \zeta \log A_t + \alpha E_{ht} c_t \tag{I.21}$$

and equations (7), (9), (14) and (15),

$$y_t^* = c_0^* + \zeta \log A_t^* + \alpha E_{ft} c_t^*. \tag{I.22}$$

where $\alpha = \frac{-\theta\gamma}{1+\epsilon-\theta}$ and $\zeta = \frac{1+\epsilon}{1+\epsilon-\theta}$, $y_t = \log Y_t - \log Y$ ($y_t^* = \log Y_t^* - \log Y^*$), $c_t = \log C_t - \log C$, ($c_t^* = \log C_t^* - \log C^*$). From equations (10), (12), (14), and (16)

$$\begin{aligned} c_t - E_t c_{t+1} &= \frac{1}{\gamma} E_t(-r_{t+1} + \pi b_{t+1}) \\ c_t^* - E_t c_{t+1}^* &= \frac{1}{\gamma} E_t(-r_{t+1} + \pi b_{t+1}^*). \end{aligned}$$

where $r_t = \log R_t - \log R$ and $b_t = B_t/Y$. With $c_t + c_t^* = y_t + y_t^*$,

$$c_t - E_t c_{t+1} = \frac{1}{2}[y_t + y_t^* - E_t(y_{t+1} + y_{t+1}^*) + \frac{2\pi}{\gamma} b_{t+1}].$$

Assume

$$\begin{aligned} c_t &= \frac{1}{2}(y_t + y_t^* + d_t) \\ c_t^* &= \frac{1}{2}(y_t + y_t^* - d_t) \end{aligned} \tag{I.23}$$

we have

$$d_t = E_t[d_{t+1} + \frac{2\pi}{\gamma} b_{t+1}] = E_t[d_{t+2} + \frac{2\pi}{\gamma} b_{t+1} + \frac{2\pi}{\gamma} b_{t+2}] = \cdots. \tag{I.24}$$

Since $B_{t+1} = Y_t - C_t - R_t B_t - \frac{\pi}{2} B_{t+1}^2$,

$$\begin{aligned} \beta^k B_{t+k} &= \beta^k[(Y_{t+k-1} - C_{t+k-1} - \frac{\pi}{2} B_{t+k}^2) + (Y_{t+k-2} - C_{t+k-2} - \frac{\pi}{2} B_{t+k-1}^2) R_{t+k-1} \\ &\quad + \cdots + R_{t+k-1} \cdots R_t B_t]. \end{aligned}$$

By using the boundary condition (19), $\lim_{k\to\infty} E_t\beta^k B_{t+k} = 0$, in the steady state $Y = C$ and $\beta R = 1$, $Y_t - C_t \approx Y(y_t - c_t)$, and (24),

$$d_t = (1-\beta)\{E_t[\sum_{k=0}^{\infty}\beta^k(y_{t+k} - y^*_{t+k}) + \sum_{k=0}^{\infty}\beta^{k+1}\frac{2\pi}{\gamma(1-\beta)}b_{t+1+k} + 2b_t]\}. \tag{I.25}$$

Use guess-verification approach and assume

$$y_t = m_{-1}\log A_{t-1} + m^*_{-1}\log A^*_{t-1} + m\log A_t + m_x x_t + m_b b_t \tag{I.26}$$

$$y^*_t = m_{-1}\log A^*_{t-1} + m^*_{-1}\log A_{t-1} + m\log A^*_t + m_x x^*_t + m_b b^*_t \tag{I.27}$$

and technology processes have the following vector-autoregressive form

$$\begin{bmatrix}\ln A_t \\ \ln A^*_t\end{bmatrix} = \begin{bmatrix}\rho & \nu \\ \nu & \rho\end{bmatrix}\begin{bmatrix}\ln A_{t-1} \\ \ln A^*_{t-1}\end{bmatrix} + \begin{bmatrix}\epsilon_t \\ \epsilon^*_t\end{bmatrix} \tag{I.28}$$

then

$$\begin{aligned} d_t &= (1-\beta)(m_{-1} - m^*_{-1})(\log A_{t-1} - \log A^*_{t-1}) + \frac{1-\beta}{1-\beta(\rho-\nu)}[\beta(m_{-1} - m^*_{-1}) \qquad \text{(I.29)} \\ &\quad + m - \beta m_x(\rho-\nu)](\log A_t - \log A^*_t) + (1-\beta)[m_x(x_t - x^*_t) + (2m_b + \frac{2\pi}{\gamma(1-\beta)}) \\ &\quad \cdot E_t\sum_{k=0}^{\infty}\beta^k b_{t+k} + (2 - \frac{2\pi}{\gamma(1-\beta)})b_t]. \qquad \text{(I.30)}\end{aligned}$$

Plug (25) into (24),

$$\begin{aligned} & E_t[\sum_{k=0}^{\infty}\beta^k(y_{t+k} - y^*_{t+k}) + (2m_b + \frac{2\pi}{\gamma(1-\beta)})\sum_{k=0}^{\infty}\beta^k b_{t+k} + (2 - \frac{2\pi}{\gamma(1-\beta)})b_t] \qquad \text{(I.31)} \\ = \; & E_t[\sum_{k=0}^{\infty}\beta^k(y_{t+1+k} - y^*_{t+1+k}) + (2m_b + \frac{2\pi}{\gamma(1-\beta)})\sum_{k=0}^{\infty}\beta^k b_{t+1+k} + 2b_{t+1}].\end{aligned}$$

Since

$$
\begin{aligned}
& E_t \sum_{k=0}^{\infty} \beta^k (y_{t+k} - y^*_{t+k}) \\
= \quad & E_t \sum_{k=0}^{\infty} \beta^k [(m_{-1} - m^*_{-1})(\log A_{t-1+k} - \log A^*_{t-1+k}) + m(\log A_{t+k} - \log A^*_{t+k}) + \\
& m_x(x_{t+k} - x^*_{t+k})] + 2m_b E_t \sum_{k=0}^{\infty} \beta^k b_{t+k} \\
\equiv \quad & U_t + 2m_b V_t,
\end{aligned}
$$

then equation (30) can be rewritten as

$$
(m_b + 1)V_t - (\beta + \frac{\pi}{\gamma} + m_b + 1)E_t V_{t+1} + \beta E_t V_{t+2} = -\frac{1}{2}(U_t - E_t U_{t+1}).
$$

Use Lag-operator, $b_t = V_t - \beta E_t V_{t+1}$ as initial condition and U_{t+k} as given,

$$
V_t = (1 - \frac{\beta}{\lambda_2})^{-1} \frac{\beta}{2\lambda_2(1 + m_b)} \{U_t - \frac{(1 - \lambda_1)E_t U_{t+1}}{1 - \lambda_1(\rho - \nu)}\} + \frac{1}{1 - \frac{\beta}{\lambda_2}} b_t \tag{I.32}
$$

Where $\lambda_1 < 1 < \lambda_2$ solve the equation $(m_b + 1)\lambda^2 - (\beta + \frac{\pi}{\gamma} + m_b + 1)\lambda + \beta = 0$. Plug (31) into (29), then (23), then (21) and compare the coefficients with (26),

$$
\begin{cases}
m_{-1} = \frac{\alpha(m_{-1}+m^*_{-1})}{2} + \frac{\alpha(1-\beta)(1+d_1)(m_{-1}-m^*_{-1})}{2} - \frac{\alpha(1-\beta)(1+d_1)(\beta(m_{-1}-m^*_{-1})+m-\beta m_x(\rho-\nu))k_a\nu}{2(1-\beta\rho+\beta\nu)(k_a+k_x)} \\
+\frac{\alpha m k_a \nu}{2(k_a+k_x)} + \frac{\alpha(1-\beta)d_2(m_{-1}-m^*_{-1}+\frac{(\rho-\nu)(\beta(m_{-1}-m^*_{-1})+m-m_x)}{(1-\beta\rho+\beta\nu)})k_a\nu}{2(k_a+k_x)}; \\
m^*_{-1} = \frac{\alpha(m_{-1}+m^*_{-1})}{2} - \frac{\alpha(1-\beta)(1+d_1)(m_{-1}-m^*_{-1})}{2} - \frac{\alpha(1-\beta)(1+d_1)(\beta(m_{-1}-m^*_{-1})+m-\beta m_x(\rho-\nu))k_a\rho}{2(1-\beta\rho+\beta\nu)(k_a+k_x)} \\
+\frac{\alpha m k_a \rho}{2(k_a+k_x)} + \frac{\alpha(1-\beta)d_2(m_{-1}-m^*_{-1}+\frac{(\rho-\nu)(\beta(m_{-1}-m^*_{-1})+m-m_x}{(1-\beta\rho+\beta\nu)})k_a\rho}{2(k_a+k_x)}; \\
m = \zeta + \frac{\alpha(m+m_x)}{2} - \frac{\alpha(1-\beta)(1+d_1)m_x}{2} + \frac{\alpha(1-\beta)(1+d_1)(\beta(m_{-1}-m^*_{-1})+m-\beta m_x(\rho-\nu))}{2(1-\beta\rho+\beta\nu)} \\
-\frac{1}{2}\alpha(1-\beta)d_2(m_{-1} - m^*_{-1} + \frac{(\rho-\nu)(\beta(m_{-1}-m^*_{-1})+m-m_x)}{1-\beta\rho+\beta\nu}); \\
m_x = \frac{\alpha m k_x}{2(k_x+k_a)} + \frac{\alpha m_x}{2} + \frac{\alpha(1-\beta)(1+d_1)m_x}{2} - \frac{\alpha(1-\beta)(1+d_1)(\beta(m_{-1}-m^*_{-1})+m-\beta m_x(\rho-\nu))k_x}{2(1-\beta\rho+\beta\nu)(k_a+k_x)} \\
+\frac{\alpha(1-\beta)d_2(m_{-1}-m^*_{-1}+\frac{(\rho-\nu)(\beta(m_{-1}-m^*_{-1})+m-m_x)}{(1-\beta\rho+\beta\nu)})k_x}{2(k_a+k_x)}; \\
m_b = \frac{\alpha(1-\beta)(1-\frac{\pi}{\gamma(1-\beta)})(1-\frac{\beta}{\lambda_2})+\alpha(1-\beta)\frac{\pi}{\gamma(1-\beta)}}{1-\frac{\beta}{\lambda_2}-\alpha(1-\beta)}.
\end{cases}
$$

where $\alpha = \frac{-\theta\gamma}{1+\epsilon-\theta}$, $\zeta = \frac{1+\epsilon}{1+\epsilon-\theta}$, $d_1 = (m_b + \frac{\pi}{\gamma(1-\beta)})(1-\frac{\beta}{\lambda_2})^{-1}\frac{\beta}{(\lambda_2(1+m_b))}$, $d_2 = (m_b + \frac{\pi}{\gamma(1-\beta)})(1-\frac{\beta}{\lambda_2})^{-1}\frac{\beta(1-\lambda_1)}{(\lambda_2(1+m_b)[1-\lambda_1(\rho-\nu)]}$. There are five undetermined variables and five equations, so we solve the coefficients $(m_{-1}, m^*_{-1}, m, m_x, m^*_x)$. We can easily verify that m_b is invariant to k_a/k_x.

For part (ii), when $k_a/k_x \to 0$, the above equations become

$$\begin{cases} m_{-1} = \frac{\alpha(m_{-1}+m^*_{-1})}{2} + \frac{\alpha(1-\beta)(1+d_1)(m_{-1}-m^*_{-1})}{2}; \\ m^*_{-1} = \frac{\alpha(m_{-1}+m^*_{-1})}{2} - \frac{\alpha(1-\beta)(1+d_1)(m_{-1}-m^*_{-1})}{2}; \\ m = \zeta + \frac{\alpha(m+m_x)}{2} - \frac{\alpha(1-\beta)(1+d_1)m_x}{2} + \frac{\alpha(1-\beta)(1+d_1)(\beta(m_{-1}-m^*_{-1})+m-\beta m_x(\rho-\nu))}{2(1-\beta\rho+\beta\nu)} \\ -\frac{1}{2}\alpha(1-\beta)d_2(m_{-1}-m^*_{-1}+\frac{(\rho-\nu)(\beta(m_{-1}-m^*_{-1})+m-m_x)}{1-\beta\rho+\beta\nu}); \\ m_x = \frac{\alpha(m+m_x)}{2} + \frac{\alpha(1-\beta)(1+d_1)m_x}{2} - \frac{\alpha(1-\beta)(1+d_1)(\beta(m_{-1}-m^*_{-1})+m-\beta m_x(\rho-\nu))}{2(1-\beta\rho+\beta\nu)} \\ +\frac{\alpha(1-\beta)d_2(m_{-1}-m^*_{-1}+\frac{(\rho-\nu)(\beta(m_{-1}-m^*_{-1})+m-m_x)}{(1-\beta\rho+\beta\nu)})}{2}; \\ m_b = \frac{\alpha(1-\beta)(1-\frac{\pi}{\gamma(1-\beta)})(1-\frac{\beta}{\lambda_2})+\alpha(1-\beta)\frac{\pi}{\gamma(1-\beta)}}{1-\frac{\beta}{\lambda_2}-\alpha(1-\beta)}. \end{cases}$$

so,

$$\begin{cases} m_{-1} = m^*_{-1} = 0; \\ m = \frac{\zeta}{1-\alpha} - \frac{\alpha[(1-\beta\rho+\beta\nu)-(1-\beta)(1+d_1)+(1-\beta)d_2(\rho-\nu)]\zeta}{2(1-\alpha)[(1-\beta\rho+\beta\nu)-\alpha(1-\beta)(1+d_1)+\alpha(1-\beta)d_2(\rho-\nu)]}; \\ m_x = \frac{\alpha[(1-\beta\rho+\beta\nu)-(1-\beta)(1+d_1)+(1-\beta)d_2(\rho-\nu)]\zeta}{2(1-\alpha)[(1-\beta\rho+\beta\nu)-\alpha(1-\beta)(1+d_1)+\alpha(1-\beta)d_2(\rho-\nu)]}; \\ m_b = \frac{\alpha(1-\beta)(1-\frac{\pi}{\gamma(1-\beta)})(1-\frac{\beta}{\lambda_2})+\alpha(1-\beta)\frac{\pi}{\gamma(1-\beta)}}{1-\frac{\beta}{\lambda_2}-\alpha(1-\beta)}. \end{cases}$$

Then, let us proof the case under the condition $\pi \to 0$, $m > 0$ and $m_x < 0$. Let us prove $-1 < m_b < 0$ as a preparation for later proofs. Since $0 < \lambda_1 < 1 < \lambda_2$, $0 < \beta < 1$,and $\alpha < 0$, $\frac{\beta}{\lambda_2} < 1$, if $\pi \to 0$, we have $m_b < 0$ and,

$$m_b + 1 = \frac{(1-\frac{\beta}{\lambda_2})(1-\alpha(1-\beta)) + \frac{\alpha\beta\pi}{\gamma\lambda_2}}{1-\frac{\beta}{\lambda_2}-\alpha(1-\beta)} > 0.$$

Therefore, we have $-1 < m_b < 0$. Next,

$$\begin{aligned} d_1 &= (m_b + \frac{\pi}{\gamma(1-\beta)})(1 - \frac{\beta}{\lambda_2})^{-1} \frac{\beta}{\lambda_2(1+m_b)} \\ &= \frac{\alpha(1-\beta) - \frac{\alpha\pi}{\gamma} + \frac{\pi}{\gamma(1-\beta)}}{1 - \frac{\beta}{\lambda_2} - \alpha(1-\beta)} \frac{\beta}{\lambda_2(1+m_b)} = \frac{\frac{\alpha\beta(1-\beta)}{\lambda_2} - \frac{\alpha\beta\pi}{\gamma\lambda_2} + \frac{\pi\beta}{\gamma\lambda_2(1-\beta)}}{-\frac{\alpha\beta(1-\beta)}{\lambda_2} + \frac{\alpha\beta\pi}{\gamma\lambda_2} + 1 - \frac{\beta}{\lambda_2}} < 0. \end{aligned}$$

Similarly, we can also have $d_2 < 0$.

$$-d_2(\rho - \nu) < -d_2 = -d_1 \frac{1-\lambda_1}{1-\lambda_1(\rho-\nu)} < -d_1,$$

so we have $-d_1 + d_2(\rho - \nu) > 0$. Therefore, m_x's numerator:

$$\begin{aligned} & \alpha[(1 - \beta\rho + \beta\nu) - (1-\beta)(1+d_1) + (1-\beta)d_2(\rho-\nu)]\zeta \\ = \; & \alpha\{\beta - \beta\rho + \beta\nu + (1-\beta)[-d_1 + d_2(\rho-\nu)]\} < 0. \end{aligned}$$

Furthermore,

$$d_1 + 1 = \frac{1 - \frac{\beta}{\lambda_2} + \frac{\pi\beta}{\gamma\lambda_2(1-\beta)}}{-\frac{\alpha\beta(1-\beta)}{\lambda_2} + \frac{\alpha\beta\pi}{\gamma\lambda_2} + 1 - \frac{\beta}{\lambda_2}} > 0.$$

m_x's denominator

$$2(1-\alpha)[(1 - \beta\rho + \beta\nu) - \alpha(1-\beta)(1+d_1) + \alpha(1-\beta)d_2(\rho-\nu)] > 0.$$

Overall,

$$m_x = \frac{\alpha[(1 - \beta\rho + \beta\nu) - (1-\beta)(1+d_1) + (1-\beta)d_2(\rho-\nu)]\zeta}{2(1-\alpha)[(1 - \beta\rho + \beta\nu) - \alpha(1-\beta)(1+d_1) + \alpha(1-\beta)d_2(\rho-\nu)]} < 0.$$

Note that $m + m_x = \frac{\zeta}{1-\alpha} > 0$, so $m > 0$. ■

Proof of Proposition 2

The complete-market economy can be fully characterized by the following first-order conditions:

$$E_{ht}\omega_t - E_{ht}(A_t\theta N_t^{\theta-1}) = 0 \tag{I.33}$$

$$E_{ft}\omega_t^* - E_{ft}(P_t A_t^*\theta N_t^{*\theta-1}) = 0 \tag{I.34}$$

$$Y_t - A_t N_t^{\theta} = 0 \tag{I.35}$$

$$Y_t^* - A_t^* N_t^{*\theta} = 0 \tag{I.36}$$

$$C_t^{-\gamma} - C_t^{*-\gamma} = 0 \tag{I.37}$$

$$N_t^{\epsilon} - C_t^{-\gamma}\omega_t = 0 \tag{I.38}$$

$$N_t^{*\epsilon} - C_t^{*-\gamma}\omega_t^* = 0 \tag{I.39}$$

$$C_t + C_t^* - A_t N_t^{\theta} - A_t^* N_t^{*\theta} = 0. \tag{I.40}$$

From equations (32), (34) and (37)

$$y_t = c_0 + \zeta \log A_t + \alpha E_{ht} c_t \tag{I.41}$$

and equations (33), (35) and (38)

$$y_t^* = c_0^* + \zeta \log A_t^* + \alpha E_{ft} c_t^* \tag{I.42}$$

where the constant terms $c_0 = c_0^* = \frac{\theta}{1+\varepsilon-\theta}$, $\alpha \equiv -\frac{\theta\gamma}{1+\epsilon-\theta}$ and $\zeta \equiv \frac{1+\epsilon}{1+\epsilon-\theta}$. Since $C_t = C_t^*$, we have:

$$c_t = c_t^* = \frac{1}{2}[y_t + y_t^*]. \tag{I.43}$$

Use guess-verification approach and assume

$$\begin{aligned} y_t &= m_{-1}\log A_{t-1} + m_{-1}^*\log A_{t-1}^* + m\log A_t + m_x x_t \\ y_t^* &= m_{-1}\log A_{t-1}^* + m_{-1}^*\log A_{t-1} + m\log A_t^* + m_x x_t^* \end{aligned}$$

and plug the two above equations into equation (42), and then (40), we have

$$
\begin{aligned}
& m_0 + m_{-1}\log A_{t-1} + m^*_{-1}\log A^*_{t-1} + m\log A_t + m_x x_t \\
= \; & c_0 + \zeta\log A_t + \alpha\{m_0 + \frac{m_{-1}+m^*_{-1}}{2}\log A_{t-1} + \frac{m_{-1}+m^*_{-1}}{2}\log A^*_{t-1} \\
& + \frac{m[k_a(\rho\log A^*_{t-1} + \nu\log A_{t-1}) + k_x x_t]}{2(k_a+k_x)} + \frac{m}{2}\log A_t + \frac{m_x}{2}x_t + \frac{m_x}{2}\log A_t\}.
\end{aligned}
$$

Compare the coefficients in the above equation,

$$m_{-1} = \alpha\frac{m_{-1}+m^*_{-1}}{2} + \frac{m\alpha\nu k_a}{2(k_a+k_x)} \tag{I.44}$$

$$m^*_{-1} = \alpha\frac{m_{-1}+m^*_{-1}}{2} + \frac{m\alpha\rho k_a}{2(k_a+k_x)} \tag{I.45}$$

$$m = \zeta + \alpha\frac{m+m_x}{2} \tag{I.46}$$

$$m_x = \alpha[\frac{m_x}{2} + \frac{mk_x}{2(k_a+k_x)}] \tag{I.47}$$

Solve equations (42) to (46), and ignore the constant term, we have, $\begin{cases} m = \frac{(1-\frac{\alpha}{2})\zeta(k_a+k_x)}{(1-\frac{\alpha}{2})^2k_a+(1-\alpha)k_x}; \\ m_x = \frac{\frac{\alpha}{2}\zeta k_x}{(1-\frac{\alpha}{2})^2k_a+(1-\alpha)k_x}; \\ m_{-1} = \frac{[\frac{\alpha}{2}\rho+(1-\frac{\alpha}{2})\nu](1-\frac{\alpha}{2})\zeta\alpha}{2(1-\alpha)[(1-\frac{\alpha}{2})^2k_a+(1-\alpha)k_x]}k_a; \\ m^*_{-1} = \frac{[(1-\frac{\alpha}{2})\rho+\frac{\alpha}{2}\nu](1-\frac{\alpha}{2})\zeta\alpha}{2(1-\alpha)[(1-\frac{\alpha}{2})^2k_a+(1-\alpha)k_x]}k_a. \end{cases}$.

For part (ii), when $k_a/k_x \to 0$, it is straightforward to have that the coefficients $(m_{-1}, m^*_{-1}, m, m_x)$ approaches $(0, 0, \frac{1-\frac{\alpha}{2}}{1-\alpha}\zeta, \frac{\frac{\alpha}{2}}{1-\alpha}\zeta)$. ■

Appendix B

This appendix briefly describes a method to solve a system of linear expectational difference equations with heterogeneous information. At first, we divide heterogeneous expectational operators into two components, the full information part and the expectation errors part. The expectation errors part is then treated as shocks to the model, and then we solve the system of linear expectational difference equations as the case with homogeneous information by Sims's (2001) method in the first step. In the second step, we use the method of undetermined coefficients to assure the endogenous expectation errors consistent with the solution from the first step. Let $y(t)$ be a vector $(k \times 1)$ which we are interested in. Then a typical system of linear rational expectational difference equations with heterogeneous expectational operators can be written as

$$\sum_{i=1}^{N} \Gamma_{i0} E_{it} y(t) = \sum_{i=1}^{N} \Gamma_{i1} E_{it-1} y(t-1) + C + \Psi z(t) + \Pi \eta(t) \tag{I.48}$$

t=1,...,T, where C is a vector $(k \times 1)$ of constants, $z(t)$ is a vector $(p \times 1)$ of exogenously evolving, possibly serially correlated, random disturbances, $\eta(t)$ is a vector $(q \times 1)$ of expectational errors, satisfying $E_{it}\eta(t+1) = 0$, E_{it} denotes expectational operator with information set Ω_{it}, Γ_{i0} and Γ_{i1} are $(k \times k)$ coefficient matrics, and Ψ and Π are $(k \times p)$ and $(k \times q)$ matrics.

Step 1. Divide heterogeneous expectational operators into two components, the full information part and the expectation errors part. After this treatment, the equations can be reorganized as

$$(\sum_{i=1}^{N} \Gamma_{i0}) E_t y(t) = (\sum_{i=1}^{N} \Gamma_{i0}) E_{t-1} y(t-1) + C + \Psi^* z(t)^* + \Pi \eta(t) \tag{I.49}$$

where

$$\Psi^* = [\Psi, -\Gamma_{10}, -\Gamma_{20}, ..., -\Gamma_{Ni}, \Gamma_{11}, \Gamma_{21}, ..., \Gamma_{N1}]$$

and

$$z(t)^* = [z(t), (E_{1t} - E_t) y(t), ..., (E_{Nt} - E_t) y(t), (E_{1t-1} - E_{t-1}) y(t-1), ..., (E_{Nt-1} - E_{t-1}) y(t-1)]' \tag{I.50}$$

where E_t denotes the expectation operator based on full information. (I.49) can also be solved by other standard methods (see Anderson (2008) for a survey).

Step 2. Undetermined Coefficients Method.

By using Sims's (2001) method, the solution of (I.49) can be characterized as

$$y(t) = \Theta_1 y(t-1) + \Theta_c + \Theta_0 z^*(t) + \Theta_y \sum_{s=1}^{\infty} \Theta_f^{s-1} Q_z E_t z^*(t+s) \tag{I.51}$$

where the coefficients are defined in equations (44) and (45) by Sims (2001). To use undetermined coefficients method, first, we list all the exogenous innovations by $\epsilon(t) \equiv (\epsilon_{1t}, \epsilon_{2t}, ..., \epsilon_{lt})$ which might affect the solution of $y(t)$. The state variables ϵ_{it}s could be technology innovations, information signals, or other type of shocks. We then assume

$$z^*(t) = \Upsilon \epsilon'(t) \tag{I.52}$$

where Υ is the $((p + kN) \times l)$ undetermined coefficients matrix. Remember that the top entry of z_t^* is z_t, so the first p row of Υ should also be known at this point and in total we have $kN * l$ unknown coefficients. Plug equation (I.52) into (I.51), and we can have $y(t)$. Then plug the $y(t)$ into the definition of $z^*(t)$ (I.50), and finally match the coefficients in equation (I.52). We will have exactly the same number of linear equations as of unknown variables, so we can exactly identify the unknown matrix Υ. Because it is a linear equations, the solution procedure will not take too much time by using regular matrix-based software.

Example: Noisy Information with International Business Cycles

When there exist both information constrained households and informed households, the model can be characterized by the following log-linear equations:

$$0 = \epsilon n_{nt} + \gamma c_{nt} - (\theta - 1)n_t - a_t \tag{I.53}$$

$$0 = c_{nt} - E_t c_{nt+1} - \frac{1}{\gamma}[p_t + \pi b_{nt+1} + (E_{ht}p_t - p_t) + (E_{ht}c_{nt+1} - E_t c_{nt+1})] \tag{I.54}$$

$$0 = \beta b_{nt+1} + c_{nt} - a_t - \theta n_{nt} - b_{nt} \tag{I.55}$$

$$0 = \epsilon \frac{n_t - \kappa n_{nt}}{1-\kappa} + \gamma c_{it} - a_t - (\theta - 1)n_t + (E_{ht}\omega_t - \omega_t) \tag{I.56}$$

$$0 = \frac{1}{\gamma}(p_t + \pi b_{it+1}) - c_{it} - E_t c_{it+1} \tag{I.57}$$

$$0 = \beta b_{it+1} + c_{it} - a_t - \theta \frac{n_t - \kappa n_{nt}}{1-\kappa} - b_{it} \tag{I.58}$$

$$0 = \epsilon n^*_{nt} + \gamma c^*_{nt} - (\theta - 1)n^*_t - a^*_t \tag{I.59}$$

$$0 = c^*_{nt} - E_t c^*_{nt+1} - \frac{1}{\gamma}[p_t + \pi b^*_{nt+1} + (E_{ft}p_t - p_t) + (E_{ft}c^*_{nt+1} - E_t c^*_{nt+1})] \tag{I.60}$$

$$0 = \beta b^*_{nt+1} + c^*_{nt} - a^*_t - \theta n^*_{nt} - b^*_{nt} \tag{I.61}$$

$$0 = \epsilon \frac{n^*_t - \kappa n^*_{nt}}{1-\kappa} + \gamma c^*_{it} - a^*_t - (\theta - 1)n^*_t + (E_{ft}\omega^*_t - \omega^*_t) \tag{I.62}$$

$$0 = c^*_{it} - E_t c^*_{it+1} - \frac{1}{\gamma}(p_t - \pi \frac{\kappa b_{nt+1} + (1-\kappa)b_{it+1} + \kappa b^*_{nt+1}}{1-\kappa}) \tag{I.63}$$

$$0 = \kappa b_{nt} + (1-\kappa)(c^*_{it} - a^*_t + b_{it} - \beta b_{it+1}) + \kappa b^*_{nt} - \theta(n^*_t - \kappa n^*_{nt}) - \beta\kappa(b_{nt+1} + b^*_{nt+1}) \tag{I.64}$$

$$0 = a_t - \rho a_{t-1} - \nu a^*_{t-1} - \epsilon_t \tag{I.65}$$

$$0 = a^*_t - \rho a^*_{t-1} - \nu a_{t-1} - \epsilon^*_t \tag{I.66}$$

So

$$
z_t^* = \begin{bmatrix} \epsilon_t \\ \epsilon_t^* \\ E_{ht}p_t - p_t \\ E_{ht}\omega_t - \omega_t \\ E_{ft}p_t - p_t \\ E_{ft}\omega_t^* - \omega_t^* \\ E_{ht}c_{nt+1} - E_t c_{nt+1} \\ E_{ft}c_{nt+1}^* - E_t c_{nt+1}^* \end{bmatrix}
$$

At first, we solve the models and get the equation (I.51), then assume

$$
z_t^* = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ p_1 & p_2 & p_3 & p_4 \\ \omega_1 & \omega_2 & \omega_3 & \omega_4 \\ p_1^* & p_2^* & p_3^* & p_4^* \\ \omega_1^* & \omega_2^* & \omega_3^* & \omega_4^* \\ c_1 & c_2 & c_3 & c_4 \\ c_1^* & c_2^* & c_3^* & c_4^* \end{vmatrix} * \begin{bmatrix} \epsilon_t \\ \epsilon_t^* \\ v_t \\ v_t^* \end{bmatrix}
$$

Use the algorithm discussed above, we solve the models.

Appendix C

We choose US versus the Euro as the two countries in our model. To construct the Euro aggregator we choose the following four countries: Austria, France, Germany, and Spain. The following files are used for average weekly hours worked per worker in manufacturing.

Austria: we collect Monthly hours of work per month in manufacturing data from LABORSTA. Then we use arithmetic mean of every three months to calculate the quarterly hours of work per month in manufacturing. The data cover wage earners from 1989M1 to 1995M12. After that the data cover employees[6].

France: we collect Quarterly hours of work per week in manufacturing data from LABORSTA. The data cover wage earners from 1989Q1 to 1992Q4. From 1993Q1 to 2009Q4, the main coverage is for employees.

Germany: before 2005, LABORSTA has two different sequences for Germany: Western Germany and Eastern Germany, but Western Germany covers both of two parts after 1990Q1. After 2005, there is only one Germany sequence. We choose the sequence of Western Germany to supplement the Germany sequence to have a complete data series of Germany. The data cover wage earners.

Spain: we collect quarterly hours of work per week data from LABORSTA. The data cover wage earners. There are two missing observations and we use the average of the two closest observations to replace them.

All the raw data are not seasonally adjusted. We use X-12-ARIMA to seasonally adjust them.

From above, we have the data of quarterly average hours of work per week or per month in manufacturing. We collect the data of average annual hours actually worked per worker in all sectors from OECD Main Economic Indicator. Then we use the quarterly average hours of worker per week or per month in manufacturing as proportions to divide the annual hours worked per worker in all sector to construct the quarterly hours of work in all sectors data.

[6] Because we use the series of quarterly hours of work as proportions to divide the series of annual hours of work, so we conjecture the change of coverage only has a minor effect.

Employment: we use quarterly average employment data from OECD Main Economic Indicator. The data is seasonally adjusted.

Output and Consumption: the data for quarterly GDP and quarterly consumption are from OECD Quarterly National Accounts.

The series for US quarterly average hours of work per week and employment are from BLS. The series for US consumption and GDP are from OECD Quarterly National Accounts.

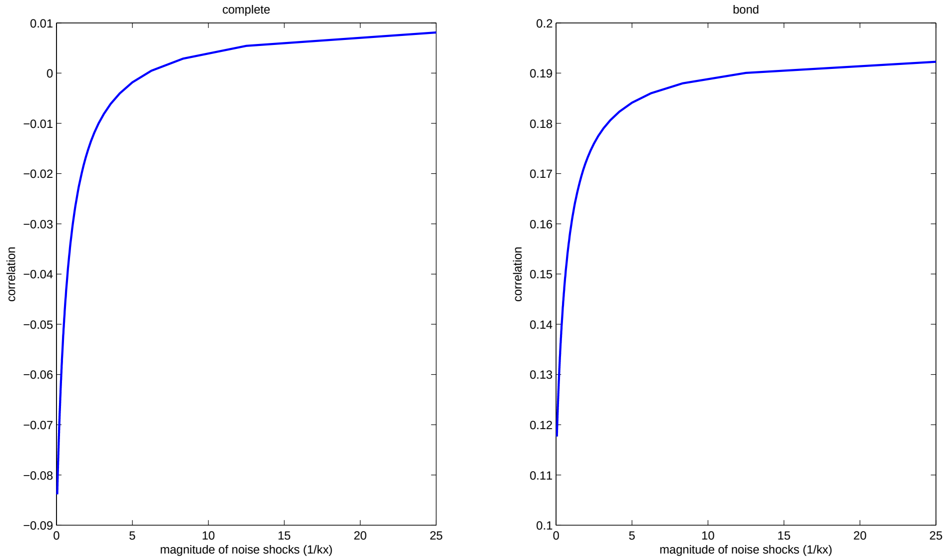

Figure 1: The correlations of outputs in different asset markets

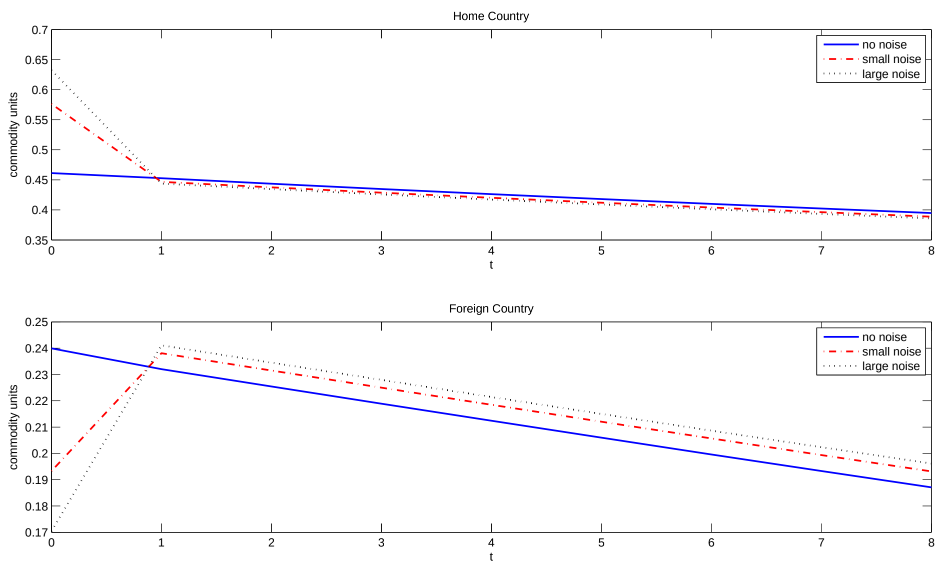

Figure 2: The impulse response of consumption

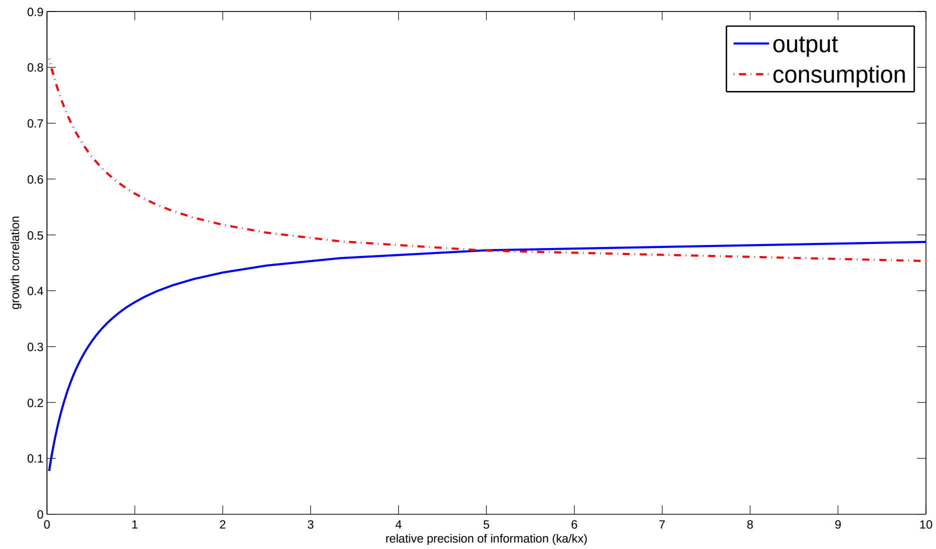

Figure 3: Outputs growth correlation and consumption growth correlation with different degrees of noise shocks($\kappa = \kappa^* = 0.6$)

Figure 4: Labor Productivity and Hours Worked during the Recent Financial Crisis

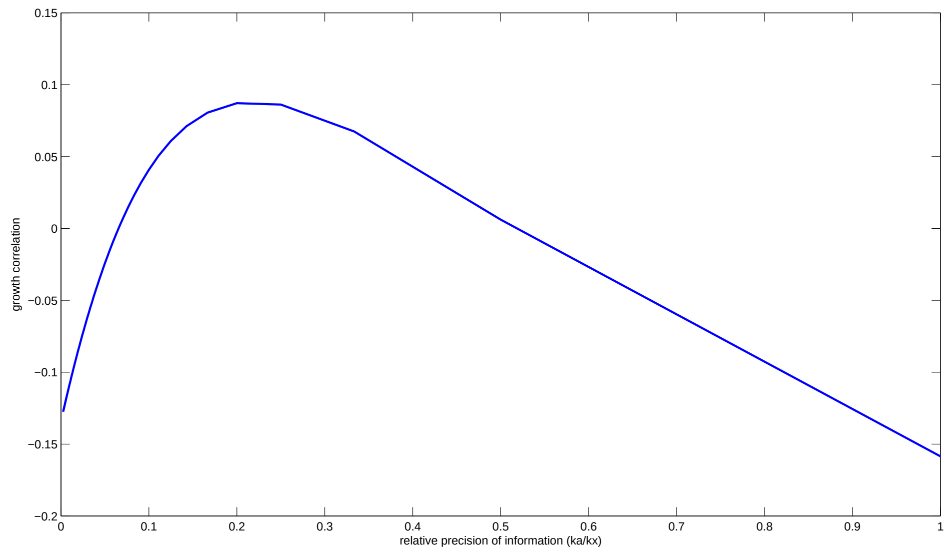

Figure 5: The correlations of hours worked growth

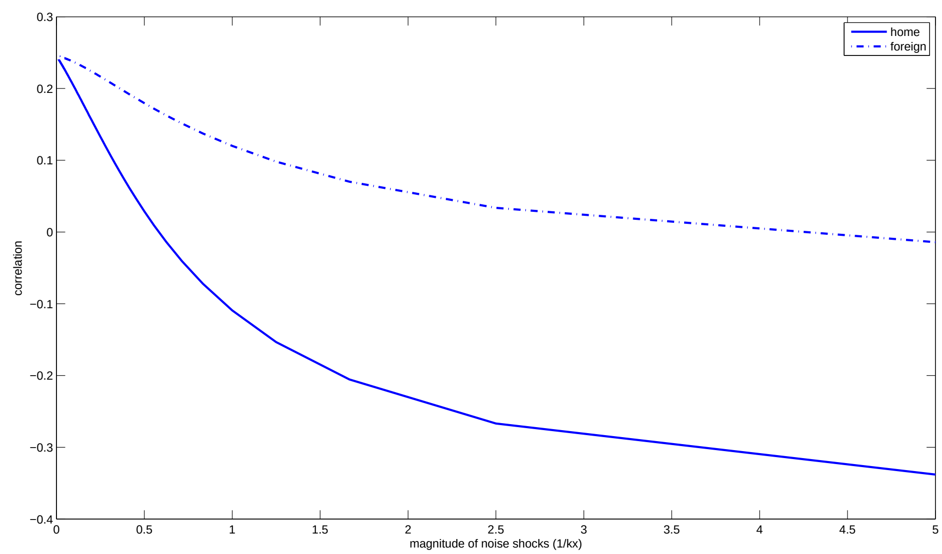

Figure 6: Productivity growth and hours growth correlation ($\kappa = 0.1$ and $\kappa^* = 0.7$)

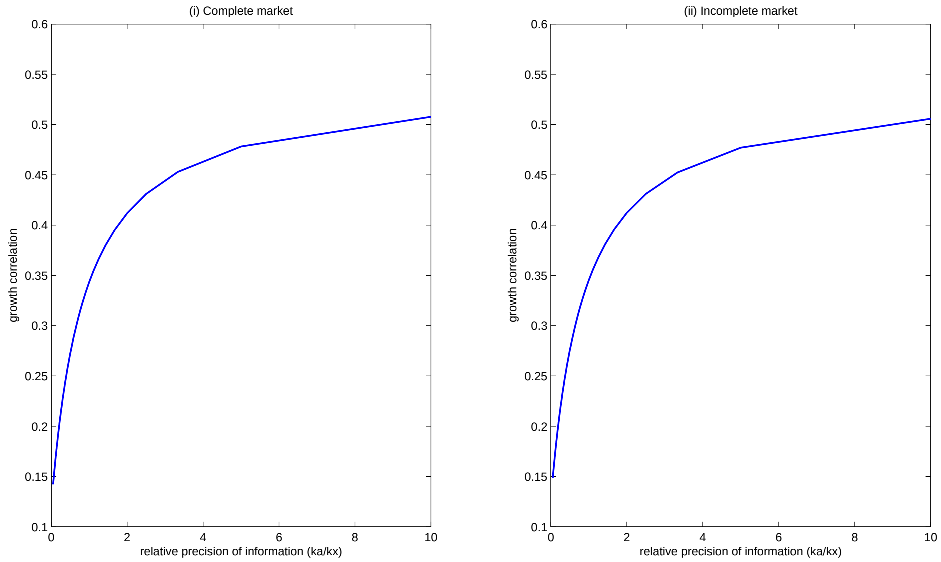

Figure 7: The correlations of outputs in different asset markets (models with capital)

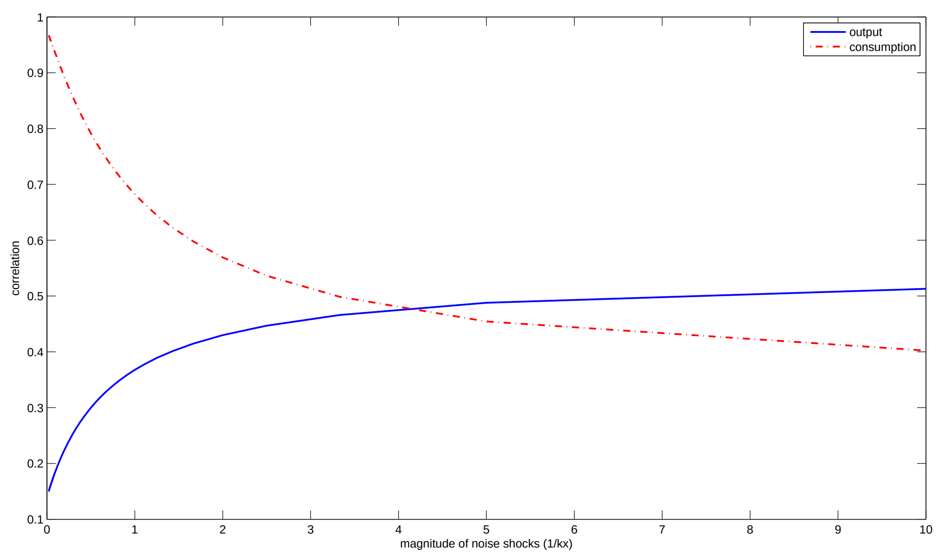

Figure 8: Outputs growth correlation and consumption growth correlation with different degrees of noise shocks (models with capital,$\kappa = \kappa^{*} = 0.7$)

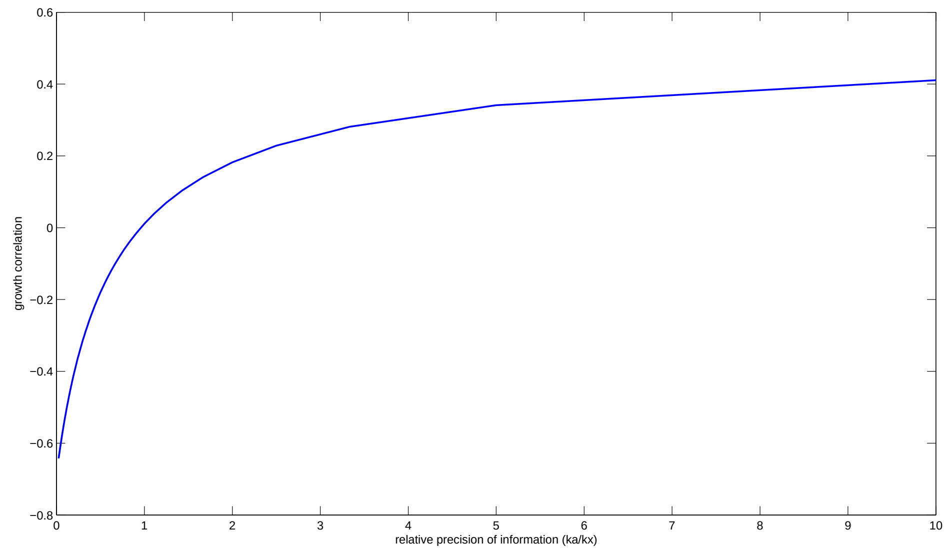

Figure 9: The correlations of hours worked growth (models with capital)

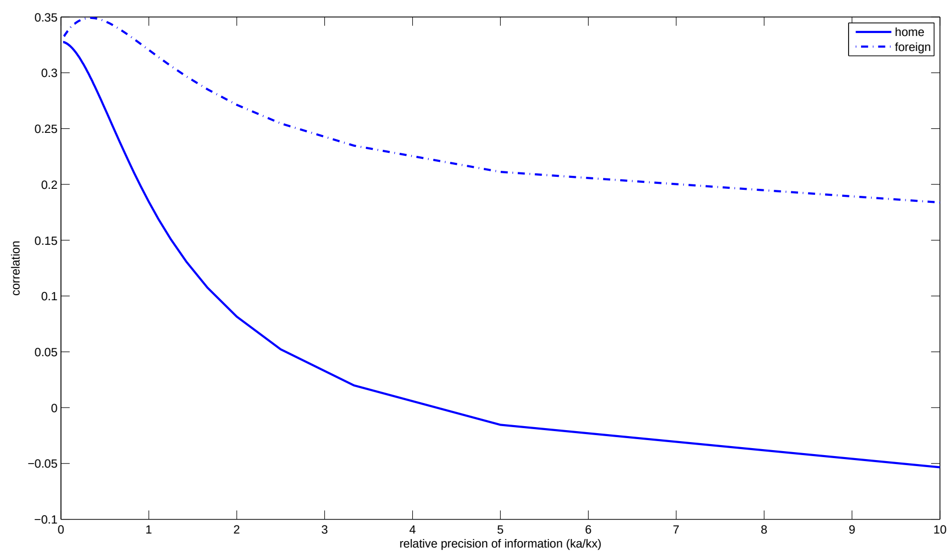

Figure 10: Productivity growth and hours growth correlation (models with capital, $\kappa = 0.1$ and $\kappa^* = 0.7$)

CHAPTER II

FINITE SAMPLE PERFORMANCE OF PRINCIPAL COMPONENTS ESTIMATORS FOR DYNAMIC FACTOR MODELS: ASYMPTOTIC AND BOOTSTRAP APPROXIMATIONS

Introduction

The estimation of dynamic factor models has become popular in macroeconomic analysis since influential works by Sargent and Sims (1977), Geweke (1977) and Stock and Watson (1989). Later studies by Stock and Watson (1998, 2002), Bai and Ng (2002) and Bai (2003) emphasize the consistency of the principal components estimator of unobservable common factors under the asymptotic framework with a large number of cross-sectional observations. This paper investigates the finite sample properties of two-step persistence estimators in dynamic factor models when unobservable common factors are estimated by the principal components method in the first step. The first-step estimation is followed by the estimation of autoregressive models of common factors in the second step. Using analytical results and simulation experiments, we evaluate the effect of the number of the series (N) relative to the time series observations (T) on the performance of the two-step estimator of a persistence parameter. Furthermore, we propose a simple bootstrap procedure that works well when N is relatively small.

In this paper, we focus on the persistence parameter of the common factor because of its empirical relevance in macroeconomic analysis. In modern macroeconomics literature, dynamic stochastic general equilibrium (DSGE) models predict that a small set of driving forces is responsible for covariation in macroeconomic variables. Theoretically, the persistence of the common factor often plays a key role on implications of these models. For example, in the real business cycle model, there is a well-known trade-off between the persistence of the technology shock and the performance of the model. When the shock becomes more persistent, the performance improves along some dimensions but deteriorates along other dimensions (King et al., 1988, Hansen, 1997, Ireland, 2001). In DSGE models with a monetary sector, the optimal monetary policy largely depends on the persistence of real shocks in the economy (Woodford, 1999). In open economy models, the welfare gain from the introduction of international risk-sharing becomes larger when the technology shock becomes more persistent (Baxter and Crucini, 1995). Since these common shocks are not directly observable, a dynamic factor model offers a simple robust statistical framework for measuring the persistence of the common components that cause macroeconomic fluctuations.[1]

Dynamic factor models have also been used to construct a business cycle index (e.g., Stock and Watson, 1989, Kim and Nelson, 1993) and to extract a measure of underlying, or core, inflation (e.g., Bryan and Cecchetti, 1993). In such applications, the persistence of a single factor can often be of main interest. For example, Clark (2006) examines the possibility of a structural shift in the persistence of a single common factor estimated using the first principal component of disaggregate inflation series. In this paper, we consider only the case in which a single common factor is generated from a univariate autoregressive (AR) model of order one. This specification keeps our problem simple since the persistence measure corresponds to the AR coefficient. However, in principle, the main idea of our approach can be applicable to AR models of higher order.[2]

The principal components estimation of the unobserved common factors is computationally simple and feasible with a large number of cross-sectional observations N. The method also allows for an approximate factor structure with possible cross-sectional correlations of idiosyncratic errors.[3] The large N asymptotic results obtained by Connor and Korajczyk (1986) and Bai (2003) imply $\sqrt{N}$-consistency of the principal components estimators of common factors up to a scaling constant. Therefore, if N is sufficiently large, we can treat the estimated common factor as if we directly observe the true common factor when conducting inference. However, since this argument is based on the asymptotic the-

[1] Recently, Boivin and Giannoni (2006) proposed estimating a dynamic factor model in which they impose the full structure of the DSGE model on the transition equation of the latent factors.

[2] In the case of AR models of higher order, however, there are several measures of persistence, including the sum of AR coefficients, largest characteristic root and first-order autocorrelation.

[3] The principle components estimator of the common factor with large N can also be used to estimate nonlinear models (Connor, Korajczyk and Linton, 2006, Diebold, 1998, Shintani, 2005, 2008) or to test the hypothesis of a unit root (Bai and Ng, 2004, and Moon and Perron, 2004).

ory, an approximation may not work when N is small relative to the time series observation T that is typically available in practice. Consistent with our theoretical prediction, the results from our Monte Carlo experiment using positively autocorrelated factors suggest the downward bias in the AR coefficient estimator and significant under-coverage of the naive confidence interval when N is small. The simulation results also suggest that a simple bootstrap procedure works well in correcting the bias and improves the performance of the confidence interval.

The bootstrap part of our analysis is closely related to recent studies by Gonçalves and Perron (2012) and Yamamoto (2012). Both papers also employ bootstrap procedures for the purpose of improving the finite sample performance of estimators of dynamic factor models. Gonçalves and Perron (2012) employ a bootstrap procedure in factor-augmented forecasting regression models with multiple factors. The factor-augmented forecasting regression models are very useful in utilizing information from many predictors without including too many regressors. This aspect is emphasized in Stock and Watson (1998, 2002), Marcellino, Stock and Watson (2003) and Bai and Ng (2006), among others. Gonçalves and Perron (2012) provide the first order asymptotic validity of their bootstrap procedure for factor-augmented forecasting regression models, but not in the context of estimation of the persistence parameter of the common factor. It should also be noted that, unlike their factor-augmented forecasting regression models with multiple factors, the bootstrap procedure for our univariate AR model of the common factor is not subject to scaling and rotation issues.[4] Yamamoto (2012) examines the performance of the bootstrap procedure applied to the factor-augmented vector autoregressive (FAVAR) models of Bernanke, Boivin and Eliasz (2005). While his multiple factor structure is more general than our single factor structure, his main focus is the identification of structural parameters in the FAVAR analysis using various identifying assumptions. In contrast, we are more interested in the role of parameters in the model in explaining the deviation from the large N asymptotics when N is small.

[4] To be more specific, under our normalizing assumption, the factor is estimated up to sign but the autoregressive coefficient can be identifed as the sign cancels out from both side of the autoregressive equation.

There are several simulation results available in the literature on the principal components estimator of dynamic factor models. Stock and Watson (1998) report the finite sample simulation results on the magnitude of the first-step estimation error of the common factor as well as the performance of an out-of-sample forecast based on the estimated factor relative to that of an infeasible forecast with a true factor. Boivin and Ng (2006) report similar performance measures in investigating the marginal effect of increasing N when there is a strong cross-sectional correlation of the errors. In addition, Stock and Watson (1998) and Bai and Ng (2002) find that information criteria designed to determine the number of the factors perform well in a finite sample. None of these studies, however, directly investigate the effect of N on the estimation of dynamic structure of the common factors.

The remainder of the paper is organized as follows: Section 2 reviews the asymptotic theory of the two-step estimator, and investigates the finite sample performance of the estimator in simulation. Section 3 considers a bootstrap approach to reduce the bias. Section 4 considers a bootstrap approach to improve the coverage performance of the confidence interval. Section 5 provides an empirical illustration of our procedures. Some concluding remarks are made in Section 6. All the proofs of theoretical results are provided in the Appendix.

Two-Step Estimation of the Autoregressive Model of Latent Factor

We begin our discussion by reviewing the literature of finite sample bias correction of an infeasible estimator of an AR(1) model, and then provide asymptotic properties of a two-step estimator of dynamic factor structure. Let x_{it} be an i-th component of N-dimensional multiple time series $X_t = (x_{1t}, \ldots, x_{Nt})'$ and $t = 1, ..., T$. A natural way to explain the comovement of x_{it}'s caused by a single factor, such as productivity shocks, is to use a simple one-factor model

$$x_{it} = \lambda_i f_t + e_{it} \tag{II.1}$$

for $i = 1, ..., N$, where λ_i's are factor loadings with respect to i-th series, f_t is a scalar common factor and e_{it}'s are possibly cross-sectionally correlated idiosyncratic shocks. If a dynamic structure is introduced by incorporating (i) a dynamic data generating process for f_t, (ii) lags of f_t in (II.1) or (iii) serial correlation in e_{it}'s, then the model becomes a dynamic factor model. In this paper, we limit our attention to a simple case with a single factor generated from a zero-mean linear stationary AR(1) model,

$$f_t = \rho f_{t-1} + \varepsilon_t \tag{II.2}$$

where $|\rho| < 1$, and ε_t is i.i.d. with $E(\varepsilon_t) = 0, E({\varepsilon_t}^2) = \sigma_\varepsilon^2$ and a finite fourth moment.

When f_t is directly observable, the AR parameter ρ can be estimated by ordinary least squares (OLS),

$$\widehat{\rho} = \left(\sum_{t=2}^{T+1} f_{t-1}^2 \right)^{-1} \sum_{t=2}^{T} f_{t-1} f_t. \tag{II.3}$$

Under the assumption described above, the limiting distribution of the OLS estimator (II.3) is given by

$$\sqrt{T}\,(\widehat{\rho} - \rho) \xrightarrow{d} N(0, 1-\rho^2), \tag{II.4}$$

as T tends to infinity, which justifies the use of the asymptotic confidence intervals for ρ. For example, the 90% confidence interval is typically constructed as

$$[\widehat{\rho} - 1.645 \times SE(\widehat{\rho}), \widehat{\rho} + 1.645 \times SE(\widehat{\rho})] \tag{II.5}$$

where $SE(\widehat{\rho})$ is the standard error of $\widehat{\rho}$ defined as $SE(\widehat{\rho}) = (\widehat{\sigma}_\varepsilon^2 / \sum_{t=2}^{T+1} f_{t-1}^2)^{1/2}$, $\widehat{\sigma}_\varepsilon^2 = (T-1)^{-1} \sum_{t=2}^{T} \widehat{\varepsilon}_t^2$ and $\widehat{\varepsilon}_t = f_t - \widehat{\rho} f_{t-1}$.

When T is small, the presence of bias of the OLS estimator (II.3) is well-known and several procedures have been proposed to reduce the bias in the literature. Using the approximation formula of the bias obtained in early studies by Hurwicz (1950), Marriott and Pope (1954) and Kendall (1954), one can construct a simple bias-corrected estimator. For example, in the current setting with a zero-mean restriction, the bias-corrected estimator is given by $\widehat{\rho}_{KBC} = T(T-2)^{-1}\widehat{\rho}$, which is a solution to the bias approximation formula $E(\widehat{\rho}) - \rho = -2T^{-1}\rho + O(T^{-2})$ for ρ with $E(\widehat{\rho})$ replaced by $\widehat{\rho}$. Alternatively, one can use the bootstrap method for the bias correction. A similar methodology was first employed by Quenouille (1949), who proposed a subsampling procedure to correct the bias. A bootstrap method for AR models based on resampling residuals was later formalized by Bose (1988) and was extended to the multivariate case by Kilian (1998), among others. In particular, the bias-corrected estimator is given by $\widehat{\rho}_{BC} = \widehat{\rho} - \widehat{bias}$ where the bootstrap bias estimate is $\widehat{bias} = B^{-1}\sum_{b=1}^{B}\widehat{\rho}_b^* - \widehat{\rho}$ and $\widehat{\rho}_b^*$ is the b-th AR estimate from the bootstrap sample and B is the number of bootstrap replications. By using either the Kendall-type bias correction or bootstrap bias correction procedures, the small T bias is reduced by the order of T^{-1}.

Table 1 reports the mean values of the OLS estimator $\widehat{\rho}$ along with the effective coverage rates of the nominal 90% conventional asymptotic confidence intervals (II.5) in 10,000 replications, using f_t generated from (II.2) with the AR parameter, $\rho = 0.5$ and 0.9 combined with $\varepsilon_t \sim iidN(0, 1-\rho^2)$.[5] The sample sizes are $T = 100$ and 200. The initial value f_t is drawn from the unconditional distribution of f_t, that is $N(0,1)$. In addition to the OLS estimator $\widehat{\rho}$, the mean values of the Kendall-type bias-corrected estimator $\widehat{\rho}_{KBC}$ and the bootstrap bias-corrected estimator $\widehat{\rho}_{BC}$ are also reported. For the bootstrap bias correction, we use $B = 499$. The results suggest that the coverage of conventional asymptotic confidence intervals seems very accurate for sample sizes $T = 100$ and 200. In addition, comparisons between two bias correction methods suggest that the small T bias of the OLS estimator ($\widehat{\rho}$) can be corrected reasonably well either by the Kendall-type correction ($\widehat{\rho}_{KBC}$) or the bootstrap-type correction ($\widehat{\rho}_{BC}$). In what follows, we use the results in Table 1 as a benchmark to evaluate the performance of the two-step estimator when the factor f_t is not known.

Let us now review the asymptotic property of the two-step estimator for the persistence parameter ρ when only x_{it} from (II.1) is observable. Under very general conditions, f_t can still be consistently estimated (up to scale) by using the first principal component of the $N \times N$ matrix $X'X$ where X is the $T \times N$ data matrix with t-th row X_t', or by using the first eigenvector of the $T \times T$ matrix XX'.[6] We denote this common factor estimator by $\widetilde{f}_t$

[5] Since our results are based on 10,000 replications, the standard error of 90% coverage rate in the simulation is about 0.003 ($\approx \sqrt{0.9 \times 0.1/10000}$).

[6] Since principal components are not scale-invariant, it is common practice to standardized all x_{it}'s to have zero sample mean and unit sample variance before applying the principal components method.

with a normalization $T^{-1}\sum_{t=1}^{T}\widetilde{f}_t^2=1$. Once $\widetilde{f}_t$ is obtained, we can replace f_t in (II.3) by $\widetilde{f}_t$ and the feasible estimator of ρ is

$$\widetilde{\rho}=\left(\sum_{t=2}^{T+1}\widetilde{f}_{t-1}^2\right)^{-1}\sum_{t=2}^{T}\widetilde{f}_{t-1}\widetilde{f}_t. \tag{II.6}$$

Below, we first show the asymptotic validity of this two-step estimator, followed by the examination of its finite sample performance using a simulation. To this end, we employ the following assumptions on the moment conditions for factors, factor loadings and idiosyncratic errors. Below, we let M be some finite positive constant.

Assumption F (factors): (i) $E|f_t|^4\leq M$ and (ii) $F'F/T\xrightarrow{p}\sigma_f^2=1$ where $F=[f_1,\cdots,f_T]'$ as $T\rightarrow\infty$.

Assumption FL (factor loadings): (i) $E|\lambda_i|^4\leq M$ and (ii) $\Lambda'\Lambda/N\xrightarrow{p}\sigma_\lambda^2>0$ where $\Lambda=[\lambda_1,\cdots,\lambda_N]'$ as $N\rightarrow\infty$.

Assumption E (errors): (i) For all (i,t), $E\left(e_{it}\right)=0$, $E\left|e_{it}\right|^8\leq M$, (ii) $E(e_{is}e_{it})=0$ for all $t\neq s$, and $N^{-1}\sum_{i,j=1}^{N}|\tau_{ij}|\leq M$ where $\tau_{ij}=E(e_{it}e_{jt})$, (iii) $E|N^{-1/2}\sum_{i=1}^{N}[e_{it}e_{is}-E(e_{it}e_{is})]|^4\leq M$ for all t and s and (iv) $(TN)^{-1}\sum_{t=1}^{T}\sum_{i,j=1}^{N}\lambda_i\lambda_je_{it}e_{jt}\xrightarrow{p}\Gamma>0$, as $N,T\rightarrow\infty$.

Since we focus on the AR(1) process of the factor, Assumption F is equivalent to the finite fourth moment condition of an i.i.d. error ε_t with variance $\sigma_\varepsilon^2=1-\rho^2$ given the stationarity condition $|\rho|<1$. Assumption FL can be replaced by the bounded deterministic sequence of factor loadings, but we only consider the case of random sequence in this paper. Assumption E allows cross-sectional correlation and heteroskedasticity but not serial correlation of idiosyncratic error terms. It should be noted that Assumption E can be replaced by a weaker assumption that allows serial correlations of idiosyncratic errors (see Bai, 2003, and Bai and Ng, 2002). Finally, we employ the following assumption on the relation among three random variables.

Assumption I (independence): The variables $\{f_t\}$, $\{\lambda_i\}$ and $\{e_{it}\}$ are three mutually independent groups. Dependence within each group is allowed.

The following proposition provides the asymptotic properties of the two-step estimator of the autoregressive coefficient.

Proposition 1. *Let x_{it} and f_t be generated from (II.1) and (II.2), respectively, and Assumptions F, FL, E and I hold. Then, as $T \to \infty$ and $N \to \infty$ such that $\sqrt{T}/N \to c$ where $0 \leq c < \infty$,*

$$\sqrt{T}\,(\widetilde{\rho} - \rho) \xrightarrow{d} N(-c\rho\sigma_{\lambda}^{-4}\Gamma, 1 - \rho^2). \tag{II.7}$$

The proposition is derived using the asymptotic framework employed by Bai (2003) and Gonçalves and Perron (2012) in their analysis of the factor-augmented forecasting regression model. In particular, it relies on the simultaneous limit theory where both N and T are allowed to grow simultaneously with a rate of N being at least as fast as $\sqrt{T}$. The bias term of order $T^{-1/2}$ is analogous to the bias term in the factor-augmented forecasting regression discussed by Ludvigson and Ng (2010) and Gonçalves and Perron (2012). Bai (2003) emphasizes that the factor estimation error has no effect on the estimation of the factor-augmented forecasting regression model if $\sqrt{T}/N$ is sufficiently small in the limit ($c = 0$). Similarly, in the context of estimating the autoregressive model of the common factor, the factor estimation error can be negligible for small $\sqrt{T}/N$. A special case of Proposition 1 with $c = 0$ implies

$$\sqrt{T}\,(\widetilde{\rho} - \rho) \xrightarrow{d} N(0, 1 - \rho^2) \tag{II.8}$$

as T tends to infinity, so that the limiting distribution of $\widetilde{\rho}$ in Theorem 1 is same as that of $\widehat{\rho}$ given by (II.4). In fact, we can also show the asymptotic equivalence of $\widetilde{\rho}$ and $\widehat{\rho}$ with their difference given by $\widetilde{\rho} - \widehat{\rho} = o_P(T^{-1/2})$.[7] Therefore, when the number of the series (N) is sufficiently large relative to the time series observations (T), the estimated factor $\widetilde{f}_t$ can be treated in exactly the same way as in the case of observable f_t. Combined with the consistency of the standard error, asymptotic confidence intervals analogues to (II.4) can be used for the two-step estimator $\widetilde{\rho}$. For example, the 90% confidence interval can be

[7] See the proof of Proposition 1.

constructed as

$$[\widetilde{\rho} - 1.645 \times SE(\widetilde{\rho}), \widetilde{\rho} + 1.645 \times SE(\widetilde{\rho})] \tag{II.9}$$

where $SE(\widetilde{\rho})$ is the standard error of $\widetilde{\rho}$ defined as $SE(\widetilde{\rho}) = (\widetilde{\sigma}_\varepsilon^2 / \sum_{t=2}^{T+1} \widetilde{f}_{t-1}^2)^{1/2}$, $\widetilde{\sigma}_\varepsilon^2 = (T-1)^{-1} \sum_{t=2}^{T} \widetilde{\varepsilon}_t^2$ and $\widetilde{\varepsilon}_t = \widetilde{f}_t - \widetilde{\rho}\widetilde{f}_{t-1}$.

When N is small (relative to T), however, the distribution of $\widetilde{\rho}$ may better be approximated by (II.7) in Proposition 1, rather than by (II.8). In such a case, the presence of bias term in (II.7) can result in bad coverage performance of a naive asymptotic confidence interval (II.9). Since the asymptotic bias term $-T^{-1/2} c\rho\sigma_\lambda^{-4}\Gamma$ can also be approximated by $-N^{-1}\rho\sigma_\lambda^{-4}\Gamma$, in what follows, we refer to this bias as the small N bias as opposed to the small T bias, $-2T^{-1}\rho$, discussed above. Within our asymptotic framework, the small N bias dominates the small T bias since the former is of order $T^{-1/2}$ and the latter is of order T^{-1}. However, it is interesting to note some similarity between the small N bias and the small T bias. For positive values of ρ, both types of bias are downward and increasing in ρ. However, the small N bias also depends on the dispersion of the factor loadings (σ_λ^2) and covariance structure of the factor loadings and idiosyncratic errors (Γ).

To examine the finite sample performance of the two-step estimator $\widetilde{\rho}$ in a simulation, we now generate x_{it} from (II.1) with the factor loading $\lambda_i \sim N(0,1)$, the serially and cross-sectionally uncorrelated idiosyncratic error $e_{it} \sim N(0, \sigma_e^2)$, and the factor f_t from the same data generating process as before. The relative size of the common component and idiosyncratic error in x_{it} is expressed in terms of the signal-to-noise ratio defined by $Var(\lambda_i f_t)/Var(e_{it}) = 1/\sigma_e^2$, which is controlled by changing σ_e^2. The set of values of the signal-to-noise ratio we consider is $\{0.5, 0.75, 1.0, 1.5, 2.0\}$. We also follow Bai and Ng (2006) and Gonçalves and Perron (2012) in considering the performance in the presence of cross-sectionally correlated errors where the correlation between e_{it} and e_{jt} is given by $0.5^{|i-j|}$ if $|i-j| \leq 5$. For a given value of T, the relative sample size N is set according to $N = [\sqrt{T}/c]$ for $c = \{0.5, 1.0, 1.5\}$ where $[x]$ is integer part of x. Therefore, sets of Ns under consideration are $\{7, 10, 20\}$ for $T = 100$ and $\{9, 14, 28\}$ for $T = 200$.

Table 2 reports the mean values of the two-step estimator $\widetilde{\rho}$, along with the effective coverage rates of the nominal 90% asymptotic confidence intervals (II.9). The theoretical result for $c = 0$ implies that the coverage probability of (II.9) should be close to 0.90 only if N is sufficiently large relative to T, but we are interested in examining its finite sample performance when N is small. The upper panel of the table shows the results with cross-sectionally uncorrelated errors, while the lower panel shows those with cross-sectionally correlated errors.

Overall, the point estimates of the two-step estimator $\widetilde{\rho}$ are clearly biased downward when N is small. Compared to the results for the infeasible estimator $\widehat{\rho}$ in Table 1, the magnitude of bias is much larger with $\widetilde{\rho}$ reflecting the fact that the theoretical order of the small N bias dominates that of the small T bias. In addition, consistent with the theoretical prediction in Proposition 1, the magnitude with bias increases when (i) ρ increases, (ii) c increases (or N decreases) and (iii) the signal-to-noise ratio decreases (or Γ increases). For the same parameter values for ρ, c and signal-to-noise ratio, the introduction of the cross-sectional correlation seems to increase the bias of $\widetilde{\rho}$. This effect does not show up in the first order asymptotics in Proposition 1 since it does not change the value of Γ. However, when the signal-to-noise ratio is highest, the difference in point estimates between cross-sectionally uncorrelated and cross-sectionally correlated cases is smallest.

The coverage performance of the standard asymptotic confidence intervals also becomes worse compared to the results in Table 1. For all the cases, the actual coverage frequency is much less than the nominal coverage rate of 90%. The closest coverage to the nominal rate is obtained when $\rho = 0.5$ is combined with a small c (a large N) and a large signal-to-noise ratio. In this case, there is about a 2 to 4% under-coverage. The deviation from the nominal rate becomes larger when ρ increases, c increases, the signal-to-noise ratio decreases and the cross-sectional correlation is introduced. The fact that the degree of under-coverage is in parallel relationship to the magnitude of the small N bias can also be explained by Proposition 1. When $-c\rho\sigma_{\lambda}^{-4}\Gamma$ in (II.7) is not negligible, the confidence interval (II.9), which is based on approximation (II.8), cannot be expected to perform well. In summary, the asymptotic confidence interval (II.9) may work well in terms of the coverage rate when N is as large as a half of T and when the AR parameter is not close to unity. Otherwise, the presence of the small N bias results in a poor coverage of the naive confidence interval. The effect of this downward bias becomes more severe as the AR parameter approaches to unity. In the next section, we consider the possibility of improving the performance of the two-step estimator when N is small, by approximating the true distribution using bootstrap procedures.

The Bootstrap Approach to Bias Correction

In the previous section, we conjectured that the presence of the small N bias is likely the main source of poor coverage of the asymptotic confidence interval when N is small. Recall that in the case of correcting the small T bias, an analytical bias formula is utilized to obtain $\widehat{\rho}_{KBC}$ while the bootstrap estimate of bias is used to construct $\widehat{\rho}_{BC}$. Similarly, we can either utilize the explicit bias function and correct the bias analytically using the formula in Proposition 1, or estimate the bias using the bootstrap method for the purpose of correction. For example, Ludvigson and Ng (2010) consider the former approach in reducing bias in the context of the factor-augmented forecasting regression model. Here we take the latter approach and employ the bootstrap procedure designed to work with cross-sectionally and serially uncorrelated errors. To be specific, we set $\tau_{ij} = 0$ for all $i \neq j$ in Assumption E(ii). However, in simulation, we also investigate its performance in the presence of cross-sectionally correlated errors ($\tau_{ij} \neq 0$). We first describe a simple bootstrap procedure for the bias correction.

Bootstrap Bias Correction I

1. Estimate factors and factor loadings using the principal components method and obtain residuals $\widetilde{e}_{it} = x_{it} - \widetilde{\lambda}_i \widetilde{f}_t$.

2. Recenter $\widetilde{e}_{it}$, $\widetilde{\lambda}_i$ and $\widetilde{f}_t$ around zero. Generate $x_{1t}^* = \lambda_1^* \widetilde{f}_t + e_{1t}^*$ for $t = 1, ..., T$ by first drawing λ_1^* from $\widetilde{\lambda}_i$ and then drawing e_{1t}^* for $t = 1, ..., T$ from $\widetilde{e}_{jt}$ given $\lambda_1^* = \widetilde{\lambda}_j$. Repeat the same procedure N times to generate all x_{it}^*'s for $i = 1, ..., N$.

3. Apply the principal components method to x_{it}^* and estimate $\widetilde{f}_t^*$.

4. Compute the bootstrap AR coefficient estimate $\widetilde{\rho}^*$ from $\widetilde{f}_t^*$.

5. Repeat steps 2 to 4 B times to obtain the bootstrap bias estimator $bias^* = B^{-1} \sum_{b=1}^{B} \widetilde{\rho}_b^* -$

$\widetilde{\rho}$ where $\widetilde{\rho}_b^*$ is the b-th bootstrap AR estimate and $\widetilde{\rho}$ is the AR estimate from $\widetilde{f}_t$. The bias-corrected estimator of ρ is given by $\widetilde{\rho}_{BC} = \widetilde{\rho} - bias^*$.

Beran and Srivastava (1985) have established the validity of applying the bootstrap procedure to the principal components analysis. Our procedure slightly differs from theirs in that we resample x_{it}^* using the estimated factor model in step 2.

In the implementation of the bootstrap, theoretically, it is possible that the first principal components cannot be computed for some bootstrap sample if an associated eigenvalue is extremely small. In such a case, we just set $\widetilde{\rho}^* = \widetilde{\rho}$ for the corresponding bootstrap sample. This modification, however, does not affect the asymptotic property of the bootstrap estimator of bias.

It should be noted that the procedure above is designed to evaluate the small N bias in the principal components method rather than the small T bias in the autoregression. In order to incorporate both the small T bias and the small N bias simultaneously, we may combine the procedure above with bootstrapping AR models. This possibility is considered in the second bootstrap bias correction method described below.

Bootstrap Bias Correction II

1. Estimate factors and factor loadings using the principal components method and obtain residuals $\widetilde{e}_{it} = x_{it} - \widetilde{\lambda}_i \widetilde{f}_t$.

2. Compute the AR coefficient estimate $\widetilde{\rho}$ from $\widetilde{f}_t$ and obtain residuals $\widetilde{\varepsilon}_t = \widetilde{f}_t - \widetilde{\rho}\widetilde{f}_{t-1}$.

3. Recenter $\widetilde{\varepsilon}_t$ around zero, if necessary, and generate ε_t^* by resampling from $\widetilde{\varepsilon}_t$. Then generate pseudo factors using $f_t^* = \widetilde{\rho} f_{t-1}^* + \varepsilon_t^*$.

4. Recenter $\widetilde{e}_{it}$ and $\widetilde{\lambda}_i$ around zero. Generate $x_{1t}^* = \lambda_1^* f_t^* + e_{1t}^*$ for $t = 1, ..., T$ by first drawing λ_1^* from $\widetilde{\lambda}_i$ and then drawing e_{1t}^* for $t = 1, ..., T$ from $\widetilde{e}_{jt}$ given $\lambda_1^* = \widetilde{\lambda}_j$. Repeat the same procedure N times to generate all x_{it}^*'s for $i = 1, ..., N$.

5. Apply the principal components method to x_{it}^* and estimate $\widetilde{f}_t^*$.

6. Compute the bootstrap AR coefficient estimate $\widetilde{\rho}^*$ from $\widetilde{f}_t^*$.

7. Repeat steps 2 to 6 B times to obtain the bootstrap bias estimator $bias^* = B^{-1}\sum_{b=1}^{B}\widetilde{\rho}_b^* - \widetilde{\rho}$ where $\widetilde{\rho}_b^*$ is the b-th bootstrap AR estimate and $\widetilde{\rho}$ is the AR estimate from $\widetilde{f}_t$. The bias-corrected estimator of ρ is given by $\widetilde{\rho}_{BC} = \widetilde{\rho} - bias^*$.

The second procedure for the bias correction involves a combination of bootstrapping principal components and bootstrapping the residuals in AR models (Freedman, 1984, and Bose, 1988). Note that our procedures employ the bootstrap bias correction based on a constant bias function. While this form of bias correction seems to be the one most frequently used in practice (e.g., Kilian, 1998), the performance of the bias-corrected estimator may be improved by introducing linear or nonlinear bias functions in the correction (see MacKinnon and Smith, 1998).

Let P^* denotes the probability measure induced by the bootstrap conditional on the original sample, and let E^* denotes expectation with respect to the distribution of the bootstrap sample conditional on the original sample. The asymptotic justification of using our bootstrap methods to correct the small N bias is established in the following proposition.

Proposition 2. *Let all the assumptions of Proposition 1 hold with $\tau_{ij} = 0$ for all $i \neq j$, $E|f_t|^8 \leq M$, $E|\lambda_i|^8 \leq M$, $E\left|e_{it}\right|^{16} \leq M$, and the bootstrap data be generated as described in Bootstrap Bias Correction I or in Bootstrap Bias Correction II. Then, as $T \to \infty$ and $N \to \infty$ such that $\sqrt{T}/N \to c$ where $0 \leq c < \infty$, $E^*(\widetilde{\rho}^* - \widetilde{\rho}) = -T^{-1/2}c\rho\sigma_\lambda^{-4}\Gamma + o_P(T^{-1/2})$.*

Proposition 2 implies the consistency of the bootstrap bias estimator $bias^*$ since $E^*(\widetilde{\rho}^* - \widetilde{\rho})$ can be accurately approximated by $bias^*$ with a suitably large value of B. The proposition also suggests that the bias-corrected estimator $\widetilde{\rho}_{BC} = \widetilde{\rho} - bias^*$ has the asymptotic bias of order smaller than $T^{-1/2}$. Since the consistency holds for both *Bootstrap Bias Correction I* and *Bootstrap Bias Correction II,* whether or not bootstrapping AR models is included in the procedure does not matter asymptotically.

Let us now conduct the simulation to evaluate the performance of the bootstrap bias correction method. The results of the simulation under the same specification as in Table 2 are shown in Table 3. For each specification, the true bias is first evaluated by using the mean value of $\tilde{\rho} - \rho$ in 10,000 replications. The asymptotic bias $-T^{-1/2} c\rho\sigma_{\lambda}^{-4}\Xi$ is also reported. The performance of bootstrap bias estimator based on *Bootstrap Bias Correction I* and *Bootstrap Bias Correction II* is evaluated by using the mean value of $bias^*$ in 10,000 replications. The number of the bootstrap replications is set at $B = 199$.

The results of the simulation can be summarized as follows. First, results turn out to be very similar between the cases of *Bootstrap Bias Correction I* and *Bootstrap Bias Correction II.* This finding suggests that the small T bias is almost negligible for the size of T we consider, which is consistent with the results in Table 1. Two bootstrap bias estimates match closely with the true bias for both the $\rho = 0.5$ and $\rho = 0.9$ cases unless the signal-to-noise ratio is too small. Second, while the direction of the changes of the theoretical bias is consistent with that of true bias, it only accounts for a fraction of the actual bias. In many cases, bootstrap bias estimates are much closer to the actual bias than the first-order term of the theoretical bias. Third, the bootstrap bias estimate does not seem to capture the effect of increased bias in the presence of the cross-sectional correlation. However, this is not surprising because our bootstrap procedure is designed for the case of cross-sectionally uncorrelated errors. Overall, the performance of the bootstrap correction method seems to be satisfactory.

The Bootstrap Approach to Confidence Intervals

Since the bootstrap bias correction method has been proven to be effective in simulation, we now turn to the issue of improving the performance of confidence intervals using a bootstrap approach. Recall that the deviation of the actual coverage rate of a naive asymptotic confidence interval (II.9) from the nominal rate is proportional to the size of bias in Table 2. Thus, it is natural to expect that recentered asymptotic confidence intervals using the bootstrap bias-corrected estimates improve the coverage accuracy. For example, the 90% confidence interval can be constructed as

$$[\widetilde{\rho}_{BC} - 1.645 \times SE(\widetilde{\rho}), \widetilde{\rho}_{BC} + 1.645 \times SE(\widetilde{\rho})]. \tag{II.10}$$

The asymptotic validity of the confidence interval (II.10) can easily be shown using the consistency result of the bootstrap bias estimator provided in Proposition 2.

Instead of using a bias-corrected estimator, we can directly utilize the bootstrap distribution of the estimator to construct bootstrap confidence intervals. Here we consider the percentile confidence interval based on the recentered bootstrap estimator $\widetilde{\rho}^* - \widetilde{\rho}$ as well as the percentile-t equal-tailed confidence interval based on the bootstrap t statistic defined as $t(\widetilde{\rho}^*) = (\widetilde{\rho}^* - \widetilde{\rho})/SE(\widetilde{\rho}^*)$ where $SE(\widetilde{\rho}^*)$ is the standard error of $\widetilde{\rho}^*$, which is asymptotically pivotal.[8] For example, the 90% percentile confidence interval and 90% percentile-t equal-tailed confidence interval can be constructed as

$$[\widetilde{\rho} - q_{0.95}(\widetilde{\rho}^* - \widetilde{\rho}), \widetilde{\rho} - q_{0.05}(\widetilde{\rho}^* - \widetilde{\rho})] \tag{II.11}$$

and

$$[\widetilde{\rho} - q_{0.95}(t(\widetilde{\rho}^*)) \times SE(\widetilde{\rho}), \widetilde{\rho} - q_{0.05}(t(\widetilde{\rho}^*)) \times SE(\widetilde{\rho})] \tag{II.12}$$

respectively, where $q_\alpha(x)$ denotes $100 \times \alpha$-th percentile of x. We now describe our procedure of constructing the bootstrap confidence intervals.

Bootstrap Confidence Interval

1. Follow either steps 1 to 3 in *Bootstrap Bias Correction I* or steps 1 to 5 in *Bootstrap Bias Correction II.*

2. Compute the bootstrap AR coefficient estimate $\widetilde{\rho}^*$ or $t(\widetilde{\rho}^*)$ from $\widetilde{f}_t^*$.

3. Repeat steps 1 to 2 B times to obtain the empirical distribution of $\widetilde{\rho}^* - \widetilde{\rho}$ to construct the percentile confidence interval and of $t(\widetilde{\rho}^*)$ to construct the percentile-t confidence interval.

Note that, as in Kilian's (1998) argument on vector autoregression, $\widetilde{\rho}$ in step 3 in *Bootstrap Bias Correction II* can be replaced by bias-corrected estimates $\widetilde{\rho}_{BC}$ without changing the limiting distribution of the bootstrap estimator. The following proposition provides the asymptotic validity of the bootstrap confidence intervals.

[8]See Hall (1992) on the importance of using asymptically pivotal statistics in achieving the higher order accuracy of the bootstrap confidence interval.

Proposition 3. *Let all the assumptions of Proposition 1 hold with $\tau_{ij} = 0$ for all $i \neq j$, and the bootstrap data be generated as described in Bootstrap Confidence Interval. Then, as* $T \to \infty$ and $N \to \infty$ *such that* $\sqrt{T}/N \to c$ *where* $0 \leq c < \infty$, $\sup_{x \in \Re} |P^*(\sqrt{T}(\tilde{\rho}^* - \tilde{\rho}) \leq x) - P(\sqrt{T}(\tilde{\rho} - \rho) \leq x)| \xrightarrow{P} 0$.

Proposition 3 implies the consistency of our bootstrap procedure in the sense that the limiting distribution of the bootstrap estimator $\tilde{\rho}^*$ is asymptotically equivalent to that of $\widetilde{\rho}$.[9] Since the limiting distribution of $\widetilde{\rho}$ is given by (II.7) in Proposition 1, the same distribution can be used to describe the limiting behavior of $\tilde{\rho}^*$. Since the coverage rate of the asymptotic confidence interval around the bias-corrected estimate, given by (II.10), approaches the nominal coverage rate in the limit, the same is true for the percentile bootstrap confidence interval. Similarly, we can modify Proposition 3 and replace $\tilde{\rho}^*$ and $\tilde{\rho}$ by their studentized statistics $t(\widetilde{\rho}^*)$ and $t(\widetilde{\rho}) = (\tilde{\rho} - \rho)/SE(\widetilde{\rho})$ and show the bootstrap consistency of $t(\widetilde{\rho}^*)$ and the asymptotic validity of the percentile-t confidence interval.[10]

Table 4 reports coverage of three confidence intervals based on the bootstrap applied to the two-step estimator $\widetilde{\rho}$ for $\rho = 0.5$ and $\rho = 0.9$ cases. Here, for the bootstrap bias correction method required in the confidence interval (II.10), we use *Bootstrap Bias Correction II.* Similarly, we report percentile and percentile-t confidence intervals based on *Bootstrap Confidence Interval* combined with *Bootstrap Bias Correction II.* The table shows that all three confidence intervals significantly improve over the naive asymptotic interval (II.9) in Table 2. Especially, when $T = 200$, $c = 0.5$ and $\rho = 0.5$, the coverage rates of all three bootstrap intervals are very close to each other and are nearly the nominal rate regardless of the signal-to-noise ratio. The percentile confidence interval (II.11) seems to work relatively well when $T = 100$. The percentile-t confidence interval (II.12) seems to dominate the bias-corrected confidence interval (II.10) for all the cases.

As in the case of the bias correction result, the performance of confidence intervals tends to improve when the signal-to-noise ratio increases. Likewise, the performance

[9] In general, signs of the coefficients in the factor forecasting regression cannot be identified, and Gonçalves and Perron (2012) argue the consistency of their bootstrap procedure in renormalized parameter space. In contrast, our result is not subject to the sign identification problem since slope coefficients in univariate AR models can still be identified.

[10] Note that we are not claiming here the higher order refinement of the percentile-t bootstrap confidence interval.

deteriorates when errors are cross-sectionally correlated. Yet, their coverage is much closer to the nominal rate when compared to the corresponding results for the naive asymptotic confidence interval. In summary, the percentile-t confidence interval works at least as well as the bias-corrected confidence interval but does not uniformly dominates the percentile confidence interval. Therefore, we suggest using three methods complementarily in practice.

Empirical Application to US Diffusion Index

In this section, we apply our bootstrap procedure to the analysis of a diffusion index based on a dynamic factor model. Stock and Watson (1998, 2002) extract common factors from 215 U.S. monthly macroeconomic time series and report that the forecasts based on such diffusion indexes outperform the conventional forecasts.[11] We use the same data source (and transformations) as Stock and Watson and sample period is from 1959:3 to 1998:12 giving a maximum number of time series observation $T = 478$. By excluding the series with missing observations, we first construct a balance panel with $N = 159$.[12] For the purpose of visualizing the effect of small N on the estimation of persistence parameter of the single common factor, we then generate multiple subsamples using the following procedure. Based on the full balanced panel, we select variables 1, 4, 7 and so on to construct a balanced panel subsample. Next, we construct another subsample by selecting variables 2, 5, 8 and so on. By repeating such a selection three times, we can construct three balanced panel data sets with $T = 478$ and $N = 53$. Similarly, we can select variables 1, 6, 11 and so on to construct five balanced panel with $T = 478$ and $N = 31$. Since the number of the series in the full balanced panel and the two subsamples are $N = 159$, 53 and 31, corresponding $\sqrt{T}/N$ are 0.14, 0.41 and 0.71. Since the values of $\sqrt{T}/N$ are not close to zero, the bootstrap method is likely more appropriate than the naive asymptotic approximation

[11] The list provided in Appendix B of Stock and Watson (2002) shows that the individual series are from 14 categories that consist of (1) real output and income; (2) employment and hours; (3) real retail, manufacturing and trade sales; (4) consumption; (5) housing starts and sales; (6) real inventories and inventory-sales ratios; (7) orders and unfilled orders; (8)stock prices; (9) exchange rates; (10) interest rates; (11) money and credit quantity aggregates; (12) price indexes; (13) average hourly earnings; and (14) miscellaneous.

[12] The number of the series in the full balanced panel differs from that of Stock and Watson (2002) due to the different treatment of outliers.

in the two-step estimation. Diffusion indexes, obtained as the cummurative sums of the first principal components of panel data sets, are shown in Figure 11. The bold line shows the estimated common factor using the full balanced panel with $N = 159$. The darker shaded area represents the range of common factor estimates among three subsamples with $N = 53$, while the lighter shaded area represents the range of common factor estimates among five subsamples with $N = 31$. As the asymptotic theory predicts, we observe that the variation among the indexes based on $N = 31$ is much larger than the variation among indexes based on $N = 53$.

In the next step, we estimate the dynamic structure of three diffusion indexes using the AR(1) specification. Table 5 reports the point estimates $\widetilde{\rho}$, naive 90% confidence intervals (II.9), bias-corrected estimates $\widetilde{\rho}_{BC}$ and 90% confidence intervals (II.10), which are based on the bootstrap bias-corrected estimates. The bias-corrected estimates are computed with the number of bootstrap replication $B = 799$. One notable observation from this empirical exercise is that the size of the bootstrap bias correction is substantial for all three cases with the size largest in the $N = 31$ case and smallest in the $N = 159$ case. In addition, the non-overlapping range between the naive and bootstrap intervals seems to be wider when N is smaller. These observations are consistent with our finding in the Monte Carlo section.

Conclusion

In this paper, we examined the finite sample properties of the two-step estimator of the persistence parameter in dynamic factor models when unobservable common factors are estimated by the principal components methods in the first step. As a result of the simulation experiment with small N, we found that the AR coefficient estimator of a positively autocorrelated factor is biased downward, and the bias is larger for a more persistent factor. This finding is consistent with the theoretical prediction. The property of the small N bias somewhat resembles that of the small T bias of the AR estimator. However, the bias caused by the small N is also present in the large T case. When there is a possibility of such a downward bias, a bootstrap procedure proposed in the paper is effective in correcting bias and controlling the coverage rate of the confidence interval.

Using a large number of series in the dynamic factor analysis has become a very popular approach in applications. The finding of this paper suggests that practitioners need to pay attention to the relative size of N and T before relying too much on a naive asymptotic approximation. Finally, it would be interesting to extend the experiments to allow for higher order AR models and nonlinear factor dynamics.

Appendix: Proofs

Proof of Proposition 1

The principal components estimator $\widetilde{F} = \left[\widetilde{f}_1, \cdots, \widetilde{f}_T\right]'$ is the first eigenvector of the $T \times T$ matrix XX' with normalization $T^{-1}\sum_{t=1}^{T}\widetilde{f}_t^2 = 1$, where

$$X = \begin{bmatrix} X_1' \\ \vdots \\ X_T' \end{bmatrix} = \begin{bmatrix} x_{11} & \cdots & x_{N1} \\ \vdots & \ddots & \vdots \\ x_{1T} & \cdots & x_{NT} \end{bmatrix}.$$

By definition, $(1/TN)XX'\widetilde{F} = \widetilde{F}v_{NT}$ where v_{NT} is the largest eigenvalue of $(1/TN)XX'$. Let $\gamma_{st} = N^{-1}\sum_{i=1}^{N}E(e_{is}e_{it})$, $\zeta_{st} = N^{-1}\sum_{i=1}^{N}(e_{is}e_{it} - E(e_{is}e_{it}))$, $\eta_{st} = N^{-1}f_s\sum_{i=1}^{N}\lambda_i e_{it}$, and $\xi_{st} = N^{-1}f_t\sum_{i=1}^{N}\lambda_i e_{is}$. Following the proof of Theorem 5 in Bai (2003), the estimation error of the factor can be decomposed as

$$\begin{aligned}
\widetilde{f}_t - H_{NT}f_t &= v_{NT}^{-1}[T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\gamma_{st} + T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\zeta_{st} + T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\eta_{st} + T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\xi_{st}] \\
&= O_P\left(N^{-1/2}\delta_{NT}^{-1}\right) + O_P\left(N^{-1/2}\delta_{NT}^{-1}\right) + O_P\left(N^{-1/2}\right) + O_P\left(N^{-1/2}\delta_{NT}^{-1}\right) \\
&= O_P\left(N^{-1/2}\right)
\end{aligned}$$

where $H_{NT} = (\widetilde{F}'F/T)(\Lambda'\Lambda/N)v_{NT}^{-1}$ and $\delta_{NT} = \min\{\sqrt{N}, \sqrt{T}\}$. From Bai's (2003) Lemma A.3, we have $\operatorname*{p\lim}_{T,N\to\infty} v_{NT} = \sigma_\lambda^2\sigma_f^2 = v$ and $\operatorname*{p\lim}_{T,N\to\infty} H_{NT}^2 = \operatorname*{p\lim}_{T,N\to\infty} (\widetilde{F}'F/T)(\Lambda'\Lambda/N)^2(F'\widetilde{F}/T)v_{NT}^{-2} = v\sigma_\lambda^2 v^{-2} = \sigma_\lambda^2(\sigma_\lambda^2\sigma_f^2)^{-1} = \sigma_f^{-2} = 1$.

If f_t is observable,

$$\begin{aligned}
\sqrt{T}\left(\widehat{\rho} - \rho\right) &= \sqrt{T}\left(\sum_{t=2}^{T+1} f_{t-1}^2\right)^{-1}\left(\sum_{t=2}^{T} f_{t-1}f_t - \rho\sum_{t=2}^{T+1} f_{t-1}^2\right) \\
&= \sqrt{T}\left(\sum_{t=2}^{T+1} f_{t-1}^2\right)^{-1}\sum_{t=2}^{T} f_{t-1}\varepsilon_t - \rho\sqrt{T}\left(\sum_{t=2}^{T+1} f_{t-1}^2\right)^{-1} f_T^2 \\
&= T^{-1/2}\sum_{t=2}^{T} f_{t-1}\varepsilon_t + o_P(1)
\end{aligned}$$

since $T^{-1}\sum_{t=1}^{T} f_t^2 = 1 + o_P(1)$. If f_t is replaced with $\widetilde{f}_t$, we have

$$
\begin{aligned}
\sqrt{T}\,(\widetilde{\rho}-\rho) &= \sqrt{T}\left(\sum_{t=2}^{T+1}\widetilde{f}_{t-1}^2\right)^{-1}(\sum_{t=2}^{T}\widetilde{f}_{t-1}\widetilde{f}_t - \rho\sum_{t=2}^{T+1}\widetilde{f}_{t-1}^2) \\
&= T^{-1/2}\sum_{t=2}^{T}\widetilde{f}_{t-1}\left(\widetilde{f}_t-\rho\widetilde{f}_{t-1}\right) - T^{-1/2}\rho\widetilde{f}_T^2 = T^{-1/2}\sum_{t=2}^{T}\widetilde{f}_{t-1}\left(\widetilde{f}_t-\rho\widetilde{f}_{t-1}\right) + o_P(1) \\
&= T^{-1/2}H_{NT}\sum_{t=2}^{T}\widetilde{f}_{t-1}\varepsilon_t + T^{-1/2}\sum_{t=2}^{T}\widetilde{f}_{t-1}\left\{\widetilde{f}_t - H_{NT}f_t - \rho\left(\widetilde{f}_{t-1}-H_{NT}f_{t-1}\right)\right\} + o_P(1) \\
&= T^{-1/2}H_{NT}^2\sum_{t=2}^{T}f_{t-1}\varepsilon_t - T^{-1/2}\rho\sum_{t=2}^{T}\widetilde{f}_{t-1}\left(\widetilde{f}_{t-1}-H_{NT}f_{t-1}\right) \\
&\quad + T^{-1/2}\sum_{t=2}^{T}\widetilde{f}_{t-1}\left(\widetilde{f}_t-H_{NT}f_t\right) + T^{-1/2}H_{NT}\sum_{t=2}^{T}\left(\widetilde{f}_{t-1}-H_{NT}f_{t-1}\right)\varepsilon_t + o_P(1).
\end{aligned}
$$

We next show (i) $T^{-1}\rho\sum_{t=2}^{T}\widetilde{f}_{t-1}(\widetilde{f}_{t-1}-H_{NT}f_{t-1}) = 2\rho v^{-2}N^{-1}\Gamma + o_P(\delta_{NT}^{-2})$; (ii) $T^{-1}\sum_{t=2}^{T}\widetilde{f}_{t-1}(\widetilde{f}_t - H_{NT}f_t) = \rho v^{-2}N^{-1}\Gamma + o_P(\delta_{NT}^{-2})$; and (iii) $T^{-1}H_{NT}\sum_{t=2}^{T}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})\varepsilon_t = o_P(\delta_{NT}^{-2})$.

We decompose the left-hand-side of (i) as,

$$
\begin{aligned}
& T^{-1}\rho\sum_{t=2}^{T}\widetilde{f}_{t-1}(\widetilde{f}_{t-1}-H_{NT}f_{t-1}) \\
= \;& T^{-1}\rho\sum_{t=2}^{T}(\widetilde{f}_{t-1}-H_{NT}f_{t-1})^2 + T^{-1}\rho\sum_{t=2}^{T}H_{NT}f_{t-1}(\widetilde{f}_{t-1}-H_{NT}f_{t-1}) \\
= \;& \rho(A+B).
\end{aligned}
$$

For A, we have,

$$
\begin{aligned}
A &= T^{-1}\sum_{t=2}^{T}v_{NT}^{-2}[T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\gamma_{st-1} + T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\zeta_{st-1} + T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\eta_{st-1} + T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\xi_{st-1}]^2 \\
&= v_{NT}^{-2}T^{-1}\sum_{t=2}^{T}(A_{0t}+A_{1t}+A_{2t}+A_{3t})^2
\end{aligned}
$$

where $A_{0t} = T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\gamma_{st-1}$, $A_{1t} = T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\zeta_{st-1}$, $A_{2t} = T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\eta_{st-1}$ and $A_{3t} = T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\xi_{st-1}$.

First,

$$\begin{aligned} T^{-1}\sum_{t=2}^{T} A_{0t}^2 &= T^{-1}\sum_{t=2}^{T}[T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\gamma_{st-1}]^2 \leq \{[T^{-1}\sum_{s=1}^{T}\tilde{f}_s^2][T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T}\gamma_{st-1}^2]\} \\ &= [T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T}\gamma_{st-1}^2] = O_P(T^{-1}), \end{aligned}$$

since $T^{-2}E\sum_{t=2}^{T}\sum_{s=1}^{T}\gamma_{st-1}^2 = T^{-2}E\sum_{s=2}^{T}\gamma_{ss}^2 = O(T^{-1})$. Second,

$$\begin{aligned} T^{-1}\sum_{t=2}^{T} A_{1t}^2 &= T^{-1}\sum_{t=2}^{T}[T^{-1}\sum_{s=1}^{T}\widetilde{f}_s\zeta_{st-1}]^2 = T^{-1}\sum_{t=2}^{T}[T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s + H_{NT}f_s)\zeta_{st-1}]^2 \\ &\leq T^{-1}\sum_{t=2}^{T}[T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)\zeta_{st-1}]^2 + T^{-1}\sum_{t=2}^{T}[T^{-1}\sum_{s=1}^{T}H_{NT}f_s\zeta_{st-1}]^2 \\ &= O_P(\delta_{NT}^{-2}N^{-1}) = o_P(\delta_{NT}^{-2}), \end{aligned}$$

since

$$\begin{aligned} T^{-1}\sum_{t=2}^{T}[T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)\zeta_{st-1}]^2 &\leq [T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)^2][T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T}\zeta_{st-1}^2] \\ &= O_P(\delta_{NT}^{-2}N^{-1}), \end{aligned}$$

where the last equality follows from Assumption E(iii), and

$$\begin{aligned} T^{-1}E\sum_{t=2}^{T}[T^{-1}\sum_{s=1}^{T}f_s\zeta_{st-1}]^2 &= T^{-1}E\sum_{t=2}^{T}[T^{-2}\sum_{s=1}^{T}\sum_{l=1}^{T}f_sf_l\zeta_{st-1}\zeta_{lt-1}] \\ &= T^{-1}\sum_{t=2}^{T}T^{-2}\sum_{s=1}^{T}\sum_{l=1}^{T}E[f_sf_l]E[\zeta_{st-1}\zeta_{lt-1}] \\ &\leq MT^{-2}\sum_{s=2}^{T}\sum_{l=1}^{T}E[f_sf_l] = O(T^{-1}), \end{aligned}$$

provided $\sigma_f^2 = 1$ and $T^{-1}\sum_{s=1}^{T}\sum_{l=1}^{T}E[f_sf_l] = O(1)$. Third,

$$\begin{aligned} T^{-1}\sum_{t=2}^{T} A_{2t}^2 &= T^{-3}\sum_{t=2}^{T}[\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)\eta_{st-1} + \sum_{s=1}^{T}H_{NT}f_s\eta_{st-1}]^2, \\ &= T^{-1}\sum_{t=2}^{T}(A_{21t} + A_{22t})^2 = T^{-1}\sum_{t=2}^{T}(A_{21t}^2 + A_{22t}^2 + 2A_{21t}A_{22t}) \end{aligned}$$

where $A_{21t} = T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)\eta_{st-1}$ and $A_{22t} = T^{-1}\sum_{s=1}^{T}H_{NT}f_s\eta_{st-1}$.We have

$$T^{-1}\sum_{t=2}^{T}A_{21t}^2 \leq [T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)^2][T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T}\eta_{st-1}^2] = O_P(\delta_{NT}^{-2}N^{-1}),$$

where the last equality follows from

$$\begin{aligned} T^{-2}E\sum_{t=2}^{T}\sum_{s=1}^{T}\eta_{st-1}^2 &= T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T}E(N^{-1}f_s\sum_{i=1}^{N}\lambda_i e_{it-1})^2 \\ &= T^{-1}\sigma_f^2\sum_{t=2}^{T}N^{-2}E\sum_{i=1}^{N}\lambda_i^2 e_{it-1}^2 = O(N^{-1}); \end{aligned}$$

and

$$\begin{aligned} T^{-1}\sum_{t=2}^{T}A_{22t}^2 &= H_{NT}^2 T^{-3}\sum_{t=2}^{T}(\sum_{s=1}^{T}f_s N^{-1}f_s\sum_{i=1}^{N}\lambda_i e_{it-1})(\sum_{s=1}^{T}f_s N^{-1}f_s\sum_{i=1}^{N}\lambda_i e_{it-1}) \\ &= H_{NT}^2(T^{-1}\sum_{s=1}^{T}f_s^2)^2 T^{-1}\sum_{t=2}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i e_{it-1})^2 \\ &= N^{-1}\Gamma + o_P(\delta_{NT}^{-2}) = O_P(\delta_{NT}^{-2}); \end{aligned}$$

and

$$T^{-1}\sum_{t=2}^{T}A_{21t}A_{22t} \leq [T^{-1}\sum_{t=2}^{T}A_{21t}]^{1/2}[T^{-1}\sum_{t=2}^{T}A_{22t}]^{1/2} = O_P(\delta_{NT}^{-2}N^{-1/2}).$$

Therefore, $T^{-1}\sum_{t=2}^{T}A_{2t}^2 = N^{-1}\Gamma + o_P(\delta_{NT}^{-2}) = O_P(\delta_{NT}^{-2})$. Fourth,

$$\begin{aligned} T^{-1}\sum_{t=2}^{T}A_{3t}^2 &= T^{-1}\sum_{t=2}^{T}T^{-2}[\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)\xi_{st-1} + \sum_{s=1}^{T}H_{NT}f_s\xi_{st-1}]^2, \\ &= T^{-1}\sum_{t=2}^{T}(A_{31t} + A_{32t})^2 = T^{-1}\sum_{t=2}^{T}(A_{31t}^2 + A_{32t}^2 + 2A_{31t}A_{32t}) = o_P(\delta_{NT}^{-2}) \end{aligned}$$

where $A_{31t} = T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)\xi_{st-1}$ and $A_{32t} = T^{-1}\sum_{s=1}^{T}H_{NT}f_s\xi_{st-1}$. The proof of $T^{-1}\sum_{t=1}^{T}A_{31t}^2 = o_P(\delta_{NT}^{-2})$ and $T^{-1}\sum_{t=1}^{T}A_{31t}A_{32t} = o_P(\delta_{NT}^{-2})$ is similar to the proof of $T^{-1}\sum_{t=1}^{T}A_{21t}^2 = o_P(\delta_{NT}^{-2})$ and $T^{-1}\sum_{t=1}^{T}A_{21t}A_{22t} = o_P(\delta_{NT}^{-2})$. For the remaining term,

$$T^{-1}\sum_{t=2}^{T}A_{32t}^2 = H_{NT}^2(T^{-1}\sum_{t=2}^{T}f_{t-1}^2)(T^{-1}\sum_{s=1}^{T}N^{-1}\sum_{i=1}^{N}f_s\lambda_i e_{is})^2 = O_P((NT)^{-1}),$$

since

$$\begin{aligned} E(T^{-1}N^{-1}\sum_{s=1}^{T}\sum_{i=1}^{N} f_s\lambda_i e_{is})^2 &= E[T^{-2}N^{-2}\sum_{s=1}^{T}\sum_{t=1}^{T} f_s f_t \sum_{i=1}^{N}\lambda_i^2 e_{is}^2] \\ &= M\sigma_\lambda^2 E[T^{-2}N^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T} f_s f_t] = O((NT)^{-1}). \end{aligned}$$

Using the Cauchy-Schwartz inequality, we can show that $T^{-1}\sum_{t=2}^{T} A_{1t}A_{2t} = o_P(\delta_{NT}^{-2})$, $T^{-1}\sum_{t=2}^{T} A_{1t}A_{3t} = o_P(\delta_{NT}^{-2})$, and $T^{-1}\sum_{t=2}^{T} A_{2t}A_{3t} = o_P(\delta_{NT}^{-2})$. By combining all the results, we have $A = T^{-1}\sum_{t=2}^{T}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})^2 = v_{NT}^{-2}N^{-1}\Gamma + o_P(\delta_{NT}^{-2}) = v^{-2}N^{-1}\Gamma + o_P(\delta_{NT}^{-2})$. For B,

$$\begin{aligned} B &= H_{NT}v_{NT}^{-1}T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T}[f_{t-1}\widetilde{f}_s\gamma_{st-1} + f_{t-1}\widetilde{f}_s\zeta_{st-1} + f_{t-1}\widetilde{f}_s\eta_{st-1} + f_{t-1}\widetilde{f}_s\xi_{st-1}] \\ &= H_{NT}v_{NT}^{-1}(B_0 + B_1 + B_2 + B_3). \end{aligned}$$

First,

$$\begin{aligned} B_0 &= T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T} f_{t-1}\widetilde{f}_s\gamma_{st-1} \le [T^{-1}\sum_{s=1}^{T}\widetilde{f}_s^2]^{1/2}[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T} f_{t-1}\gamma_{st-1})^2]^{1/2} \\ &= O_P(T^{-1}), \end{aligned}$$

where

$$\begin{aligned} T^{-1}E\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T} f_{t-1}\gamma_{st-1})^2 &= T^{-3}E\sum_{s=1}^{T}T^{-2}\sum_{t=2}^{T}\sum_{l=2}^{T} f_{t-1}f_{l-1}\gamma_{st-1}\gamma_{sl-1} \\ &= O(T^{-2}). \end{aligned}$$

Second,

$$\begin{aligned} B_1 &= T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T} f_{t-1}\widetilde{f}_s\zeta_{st-1} \le [T^{-1}\sum_{s=1}^{T}\widetilde{f}_s^2]^{1/2}[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T} f_{t-1}\zeta_{st-1})^2]^{1/2} \\ &= O_P((NT)^{-1/2}). \end{aligned}$$

Third,

$$\begin{aligned}
B_2 &= T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T} f_{t-1}\widetilde{f}_s\eta_{st-1} \leq [T^{-1}\sum_{s=1}^{T}\widetilde{f}_s^2]^{1/2}[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T} f_{t-1}\eta_{st-1})^2]^{1/2} \\
&= [T^{-1}\sum_{s=1}^{T}\widetilde{f}_s^2]^{1/2}[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T} f_{t-1}N^{-1}\sum_{i=1}^{N} f_s\lambda_i e_{it-1})^2]^{1/2} \\
&= [T^{-1}\sum_{s=1}^{T}\widetilde{f}_s^2]^{1/2}[(T^{-1}\sum_{s=1}^{T} f_s^2)(T^{-1}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N} f_{t-1}\lambda_i e_{it-1})^2]^{1/2} \\
&= O_P((NT)^{-1/2}).
\end{aligned}$$

Fourth,

$$\begin{aligned}
B_3 &= T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T} f_{t-1}\widetilde{f}_s\xi_{st-1} = T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T} f_{t-1}(\widetilde{f}_s - H_{NT}f_s + H_{NT}f_s)\xi_{st-1} \\
&= T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T} f_{t-1}(\widetilde{f}_s - H_{NT}f_s)\xi_{st-1} + T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T} f_{t-1}H_{NT}f_s\xi_{st-1} \\
&= B_{31} + B_{32}.
\end{aligned}$$

For B_{31},

$$\begin{aligned}
B_{31} &= [T^{-1}\sum_{t=2}^{T} f_{t-1}^2][T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)N^{-1}\sum_{i=1}^{N}\lambda_i e_{is}] \\
&= T^{-2}\sum_{s=1}^{T} v_{NT}^{-1}\sum_{t=1}^{T}[\widetilde{f}_t\gamma_{ts} + \widetilde{f}_t\zeta_{ts} + \widetilde{f}_t\eta_{ts} + \widetilde{f}_t\xi_{ts}]N^{-1}\sum_{i=1}^{N}\lambda_i e_{is} + o_P(\delta_{NT}^{-2}) \\
&= v_{NT}^{-1}(B_{310} + B_{311} + B_{312} + B_{313}) + o_P(\delta_{NT}^{-2}),
\end{aligned}$$

where

$$\begin{aligned}
B_{310} &= T^{-2}\sum_{s=1}^{T}\sum_{t=1}^{T}\widetilde{f}_t\gamma_{ts}N^{-1}\sum_{i=1}^{N}\lambda_i e_{is} \\
&\leq [T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=1}^{T}\widetilde{f}_t\gamma_{ts})^2]^{1/2}[T^{-1}\sum_{s=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i e_{is})^2]^{1/2} = O_P(\delta_{NT}^{-1}N^{-1}),
\end{aligned}$$

and

$$\begin{aligned}
B_{311} &= T^{-2}\sum_{s=1}^{T}\sum_{t=1}^{T}\widetilde{f}_t\zeta_{ts}N^{-1}\sum_{i=1}^{N}\lambda_i e_{is} \\
&\leq [T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=1}^{T}\widetilde{f}_t\zeta_{ts})^2]^{1/2}[T^{-1}\sum_{s=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i e_{is})^2]^{1/2} = O_P(\delta_{NT}^{-1}N^{-1}),
\end{aligned}$$

and

$$\begin{aligned} B_{312} &= T^{-2}\sum_{s=1}^{T}\sum_{t=1}^{T}\widetilde{f}_t\eta_{ts}N^{-1}\sum_{i=1}^{N}\lambda_i e_{is} = (T^{-1}\sum_{t=1}^{T}\widetilde{f}_t f_t)[T^{-1}\sum_{s=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i e_{is})^2] \\ &= H_{NT}N^{-1}\Gamma + o_P(\delta_{NT}^{-2}). \end{aligned}$$

and

$$B_{313} = T^{-2}\sum_{s=1}^{T}\sum_{t=1}^{T}(\widetilde{f}_t - H_{NT}f_t + H_{NT}f_t)\xi_{ts}N^{-1}\sum_{i=1}^{N}\lambda_i e_{is} = O_P(\delta_{NT}^{-1}N^{-1}),$$

since

$$\begin{aligned} & T^{-2}\sum_{s=1}^{T}\sum_{t=1}^{T}(\widetilde{f}_t - H_{NT}f_t)\xi_{ts}N^{-1}\sum_{i=1}^{N}\lambda_i e_{is} \\ \leq\ & [T^{-1}\sum_{t=1}^{T}(\widetilde{f}_t - H_{NT}f_t)^2]^{1/2}[T^{-1}\sum_{t=1}^{T}(T^{-1}\sum_{s=1}^{T}\xi_{ts}N^{-1}\sum_{i=1}^{N}\lambda_i e_{is})^2]^{1/2} \\ \leq\ & [T^{-1}\sum_{t=1}^{T}(\widetilde{f}_t - H_{NT}f_t)^2]^{1/2}[(T^{-2}\sum_{t=1}^{T}\sum_{s=1}^{T}\xi_{ts}^2)T^{-1}\sum_{s=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i e_{is})^2]^{1/2} \\ =\ & O_P(\delta_{NT}^{-1}N^{-1}T^{-1/2}), \end{aligned}$$

and

$$\begin{aligned} T^{-2}\sum_{s=1}^{T}\sum_{t=1}^{T}H_{NT}f_t\xi_{ts}N^{-1}\sum_{i=1}^{N}\lambda_i e_{is} &= H_{NT}T^{-2}[\sum_{s=1}^{T}N^{-1}\sum_{i=1}^{N}f_s\lambda_i e_{is}]^2. \\ &= O_P((NT)^{-1}). \end{aligned}$$

Thus, $B_{31} = H_{NT}v_{NT}^{-1}N^{-1}\Gamma + o_P(\delta_{NT}^{-2})$. For B_{32},

$$\begin{aligned} B_{32} &= T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T}f_{t-1}H_{NT}f_s\xi_{st-1} = T^{-2}H_{NT}\sum_{t=2}^{T}\sum_{s=1}^{T}f_{t-1}f_sN^{-1}\sum_{i=1}^{N}f_{t-1}\lambda_i e_{is} \\ &= (T^{-1}H_{NT}\sum_{t=2}^{T}f_{t-1}^2)(T^{-1}N^{-1}\sum_{s=1}^{T}\sum_{i=1}^{N}f_s\lambda_i e_{is}) = O_P((NT)^{-1/2}). \end{aligned}$$

Therefore, $B_3 = H_{NT}v_{NT}^{-1}N^{-1}\Gamma + o_P(\delta_{NT}^{-2})$. By combining all the results for B_1, B_2 and B_3, we have $B = T^{-1}\sum_{t=2}^{T}H_{NT}f_{t-1}(\tilde{f}_{t-1} - H_{NT}f_{t-1}) = H_{NT}^2v_{NT}^{-2}N^{-1}\Gamma + o_P(\delta_{NT}^{-2}) = v^{-2}N^{-1}\Gamma + o_P(\delta_{NT}^{-2})$. Thus, $T^{-1}\rho\sum_{t=2}^{T}\widetilde{f}_{t-1}(\widetilde{f}_{t-1} - H_{NT}f_{t-1}) = \rho(A+B) = \rho(v^{-2}N^{-1}\Gamma + v^{-2}N^{-1}\Gamma) + o_P(\delta_{NT}^{-2}) = 2\rho v^{-2}N^{-1}\Gamma + o_P(\delta_{NT}^{-2})$.

To show (ii), we can decompose the left-hand-side of (ii) as

$$T^{-1}\sum_{t=2}^{T}\widetilde{f}_{t-1}(\widetilde{f}_t - H_{NT}f_t) = T^{-1}\sum_{t=2}^{T}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})(\widetilde{f}_t - H_{NT}f_t) + H_{NT}T^{-1}\sum_{t=2}^{T}f_{t-1}(\widetilde{f}_t - H_{NT}f_t).$$

The proof is almost the same as the proof of (i). We only mention the difference. To show $T^{-1}\sum_{t=1}^{T}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})(\widetilde{f}_t - H_{NT}f_t) = o_P(\delta_{NT}^{-2})$, we need use

$$\begin{aligned}
& T^{-1}\sum_{t=2}^{T}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})(\widetilde{f}_t - H_{NT}f_t) \\
= \;& v_{NT}^{-2}H_{NT}^2T^{-3}\sum_{t=2}^{T}(\sum_{s=1}^{T}f_sN^{-1}f_s\sum_{i=1}^{N}\lambda_ie_{it-1})(\sum_{s=1}^{T}f_sN^{-1}f_s\sum_{i=1}^{N}\lambda_ie_{it}) + o_P(\delta_{NT}^{-2}) \\
= \;& v_{NT}^{-2}H_{NT}^2T^{-3}(\sum_{s=1}^{T}f_s^2)^2\sum_{t=2}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_ie_{it-1})(N^{-1}\sum_{i=1}^{N}\lambda_ie_{it}) + o_P(\delta_{NT}^{-2}) \\
= \;& o_P(\delta_{NT}^{-2}).
\end{aligned}$$

To show $H_{NT}T^{-1}\sum_{t=2}^{T}f_{t-1}(\widetilde{f}_t - H_{NT}f_t) = \rho v^{-2}N^{-1}\Gamma + o_P(\delta_{NT}^{-2})$, we need use

$$\begin{aligned}
H_{NT}T^{-1}\sum_{t=2}^{T}f_{t-1}(\widetilde{f}_t - H_{NT}f_t) &= [T^{-1}\sum_{t=2}^{T}f_{t-1}f_t][T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)N^{-1}\sum_{i=1}^{N}\lambda_ie_{is}] + o_P(\delta_{NT}^{-2}) \\
&= \rho T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s - H_{NT}f_s)N^{-1}\sum_{i=1}^{N}\lambda_ie_{is} + o_P(\delta_{NT}^{-2}) \\
&= \rho H_{NT}N^{-1}\Gamma + o_P(\delta_{NT}^{-2}).
\end{aligned}$$

To obtain the result (iii), we first decompose the left-hand-side of (iii) as

$$\begin{aligned}
& T^{-1}H_{NT}\sum_{t=2}^{T}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})\varepsilon_t \\
= \;& H_{NT}v_{NT}^{-1}T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T}[\widetilde{f}_s\gamma_{st-1}\varepsilon_t + \widetilde{f}_s\zeta_{st-1}\varepsilon_t + \widetilde{f}_s\eta_{st-1}\varepsilon_t + \widetilde{f}_s\xi_{st-1}\varepsilon_t] \\
= \;& H_{NT}v_{NT}^{-1}(C_0 + C_1 + C_2 + C_3).
\end{aligned}$$

For C_0,

$$\begin{aligned}
C_0 &= T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T}\widetilde{f}_s\gamma_{st-1}\varepsilon_t \le (T^{-1}\sum_{s=1}^{T}\widetilde{f}_s^2)^{1/2}[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T}\gamma_{st-1}\varepsilon_t)^2]^{1/2} \\
&= [T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T}\gamma_{st-1}\varepsilon_t)^2]^{1/2} = O_P((NT)^{-1/2}),
\end{aligned}$$

where the last equality follows from

$$\begin{aligned}
T^{-1}\sum_{s=1}^{T}E(T^{-1}\sum_{t=2}^{T}\gamma_{st-1}\varepsilon_t)^2 &= T^{-1}\sum_{s=1}^{T}E(T^{-1}\sum_{t=2}^{T}[N^{-1}\sum_{i=1}^{N}E(e_{it-1}e_{is})]\varepsilon_t)^2 \\
&= \sigma_\varepsilon^2 T^{-3}\sum_{s=1}^{T}\sum_{t=2}^{T}[N^{-1}\sum_{i=1}^{N}E(e_{it-1}e_{is})]^2 \\
&= O((NT)^{-1}).
\end{aligned}$$

For C_1,

$$\begin{aligned}
C_1 &= T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T}\widetilde{f}_s\zeta_{st-1}\varepsilon_t \leq (T^{-1}\sum_{s=1}^{T}\widetilde{f}_s^2)^{1/2}[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T}\zeta_{st-1}\varepsilon_t)^2]^{1/2} \\
&= O_P((NT)^{-1/2}),
\end{aligned}$$

where the last equality follows from

$$\begin{aligned}
T^{-1}\sum_{s=1}^{T}E(T^{-1}\sum_{t=2}^{T}\zeta_{st-1}\varepsilon_t)^2 &= T^{-1}\sum_{s=1}^{T}E(T^{-1}\sum_{t=2}^{T}[N^{-1}\sum_{i=1}^{N}(e_{it-1}e_{is}-E(e_{it-1}e_{is}))]\varepsilon_t)^2 \\
&= \sigma_\varepsilon^2 T^{-1}\sum_{s=1}^{T}T^{-2}E\sum_{t=2}^{T}[N^{-1}\sum_{i=1}^{N}(e_{it-1}e_{is}-E(e_{it-1}e_{is}))]^2 \\
&= O((NT)^{-1}).
\end{aligned}$$

For C_2,

$$C_2 = T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T}\widetilde{f}_s\eta_{st-1}\varepsilon_t \leq (T^{-1}\sum_{s=1}^{T}\widetilde{f}_s^2)^{1/2}[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T}\eta_{st-1}\varepsilon_t)^2]^{1/2} = O_P((NT)^{-1/2}),$$

where the last equality follows from

$$\begin{aligned}
T^{-1}\sum_{s=1}^{T}E(T^{-1}\sum_{t=2}^{T}\eta_{st-1}\varepsilon_t)^2 &= \sigma_\varepsilon^2 T^{-1}\sum_{s=1}^{T}E(T^{-2}\sum_{t=2}^{T}\eta_{st-1}^2) \\
&= \sigma_\varepsilon^2 T^{-1}\sum_{s=1}^{T}E[T^{-2}\sum_{t=2}^{T}(N^{-1}\sum_{i=1}^{N}f_s\lambda_i e_{it-1})^2] \\
&= \sigma_\varepsilon^2\sigma_f^2 T^{-2}E[\sum_{t=2}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i e_{it-1})^2] = O((NT)^{-1}).
\end{aligned}$$

Similarly, we can show $C_3 = O_P((NT)^{-1/2})$. By combining all the results for C_0, C_1, C_2 and C_3, we have $T^{-1}H_{NT}\sum_{t=2}^{T}(\widetilde{f}_{t-1} - H_{NT}f_{t-1})\varepsilon_t = O_P((NT)^{-1/2})$. Finally, we use

$H_{NT}^2 - 1 = o_P(1)$ and $T^{1/2}N^{-1} - c = o(1)$ to obtain

$$\sqrt{T}\left(\widetilde{\rho} - \rho\right) = T^{-1/2}\sum_{t=2}^{T} f_{t-1}\varepsilon_t - c\rho v^{-2}\Gamma + o_P(1).$$

The desired result follows from the central limit theorem applied to the martingale difference sequence $f_{t-1}\varepsilon_t$ with $E(f_{t-1}^2\varepsilon_t^2) = 1 - \rho^2$ combined with Slutsky's theorem.

■

Proof of Proposition 2.

In this proof, we only derive the limiting behavior of $E^*\left(\widetilde{\rho}^* - \tilde{\rho}\right) = E^*[\left(\sum_{t=2}^{T+1}\widetilde{f}_{t-1}^{*2}\right)^{-1} \sum_{t=2}^{T}\widetilde{f}_{t-1}^*\widetilde{f}_t^*] - \tilde{\rho}$ based on *Bootstrap Bias Correction II* because the proof for *Bootstrap Bias Correction I* is similar but simpler. The bootstrap principal components estimator $\widetilde{F}^* = \left[\widetilde{f}_1^*, \cdots, \widetilde{f}_T^*\right]'$ is the first eigenvector of the $T \times T$ matrix $X^*X^{*\prime}$ with normalization $T^{-1}\sum_{t=1}^{T}\widetilde{f}_t^{*2} = 1$, where the bootstrap sample is given by

$$X^* = \begin{bmatrix} X_1^{*\prime} \\ \vdots \\ X_T^{*\prime} \end{bmatrix} = \begin{bmatrix} x_{11}^* & \cdots & x_{N1}^* \\ \vdots & \ddots & \vdots \\ x_{1T}^* & \cdots & x_{NT}^* \end{bmatrix}.$$

Analogous to the original version, we have $(1/TN)X^*X^{*\prime}\tilde{F}^* = v_{NT}^*\tilde{F}^*$ where v_{NT}^* is the largest eigenvalue of $(1/TN)X^*X^{*\prime}$. Let $\zeta_{st}^* = N^{-1}\sum_{i=1}^{N} e_{is}^* e_{it}^*$, $\eta_{st}^* = N^{-1}f_s^*\sum_{i=1}^{N}\lambda_i^* e_{it}^*$, and $\xi_{st}^* = N^{-1}f_t^*\sum_{i=1}^{N}\lambda_i^* e_{is}^* = \eta_{ts}^*$. The estimation error of the factor can be decomposed as

$$\tilde{f}_t^* - H_{NT}^* f_t^* = v_{NT}^{*-1}T^{-1}\sum_{s=1}^{T}\tilde{f}_s^*\zeta_{st}^* + v_{NT}^{*-1}T^{-1}\sum_{s=1}^{T}\tilde{f}_s^*\eta_{st}^* + v_{NT}^{*-1}T^{-1}\sum_{s=1}^{T}\tilde{f}_s^*\xi_{st}^*$$

where $H_{NT}^* = (\tilde{F}^{*\prime}F^*/T)(\Lambda^{*\prime}\Lambda^*/N)v_{NT}^{*-1}$. From Lemma C.1 of Gonçalves and Perron (2012), with mutual independence of f_t, λ_i, and e_{it} and $E|f_t|^8 \leq M$, $E|\lambda_i|^8 \leq M$, $E|e_{it}|^{16} \leq M$, we have (a) $T^{-1}\sum_{t=1}^{T}|\widetilde{f}_t - H_{NT}f_t|^8 = O_P(1)$; (b) $N^{-1}\sum_{i=1}^{N}|\tilde{\lambda}_i - H_{NT}^{-1}\lambda_i|^8 = O_P(1)$; and (c) $(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{e}_{it}^8 = O_P(1)$. (a), (b) and (c) imply that $E^*(e_{it}^{*8}) = (NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{e}_{it}^8 =$

$O_P(1)$,

$$E^*\lambda_i^{*8} = N^{-1}\sum_{i=1}^N \tilde{\lambda}_i^8 \leq 8N^{-1}(\sum_{i=1}^N |\tilde{\lambda}_i - H_{NT}^{-1}\lambda_i|^8 + \sum_{i=1}^N |H_{NT}^{-1}\lambda_i|^8) = O_P(1),$$

and

$$\begin{aligned}
E^*\varepsilon_t^{*8} &= T^{-1}\sum_{t=1}^T (\widetilde{f}_t - \tilde{\rho}\widetilde{f}_{t-1})^8 \\
&= T^{-1}\sum_{t=1}^T [\widetilde{f}_t - H_{NT}f_t + H_{NT}f_t - \tilde{\rho}(\widetilde{f}_{t-1} - H_{NT}f_{t-1}) - \tilde{\rho}H_{NT}f_{t-1}]^8 \\
&\leq 4^7T^{-1}\sum_{t=1}^T [(\widetilde{f}_t - H_{NT}f_t)^8 + (H_{NT}f_t)^8 + \tilde{\rho}^4(\widetilde{f}_{t-1} - H_{NT}f_{t-1})^8 + (\tilde{\rho}H_{NT}f_{t-1})^8] \\
&= O_P(1).
\end{aligned}$$

We denote $S_T^* = o_{P^*}(\alpha_T^{-1})$ if the bootstrap statistic S_T^* satisfies $P^*(\alpha_T|S_T^*| > \delta) = o_P(1)$ for any $\delta > 0$ as $\alpha_T \to \infty$. We have $v_{NT}^* = v^* + o_{P^*}(1)$, where $v^* = \Sigma_\Lambda^*\Sigma_F^*$, $\Sigma_\Lambda^* = \widetilde{\Lambda}'\widetilde{\Lambda}/N \to^P v$ and $\Sigma_F^* = \widetilde{F}'\widetilde{F}/T = 1$, and $H_{NT}^{*2} - 1 = o_{P^*}(1)$ because

$$H_{NT}^{*2} = (\tilde{F}^{*'}F^*/T)(\Lambda^{*\prime}\Lambda^*/N)^2(F^{*\prime}\tilde{F}^*/T)v_{NT}^{*-2} = \Sigma_F^{*-1} + o_{p^*}(1).$$

Note that v_{NT}^* is the largest eigenvalue of a positive semi-definite matrix $(1/TN)X^*X^{*\prime}$ and $v_{NT}^* \to^{P^*} v > 0$. By the construction of our bootstrap procedure, v_{NT}^* has a lower bound and $(TN)^{-1}X^*X'^*$ is non-zero for all bootstrap samples. Because v_{NT}^* is greater than some small positive number ϵ in our bootstrap procedure, we have $E^*v_{NT}^{*-4} = v^{-4} + o_P(1)$. Even if there is no lower bound, $P^*(v_{NT}^* \geq \epsilon) \to^P 1$ holds and thus the effect of such a modification in our procedure on the distributions of random variables λ_i^*, f_t^*, and e_{it}^* is asymptotically negligible. We can also show $E^*H_{NT}^{*2} = 1 + o_P(1)$ and $E^*H_{NT}^{*4} = 1 + o_P(1)$ by using the following argument. We have

$$\begin{aligned}
& E^*[T^{-1}\sum_{t=1}^T(\tilde{f}^*_{t-1} - H^*_{NT} f^*_{t-1})^2] \\
= \;& T^{-1}E^*\{\sum_{t=1}^T v_{NT}^{*-2}[T^{-1}\sum_{s=1}^T \widetilde{f}^*_s \zeta^*_{st-1} + T^{-1}\sum_{s=1}^T \widetilde{f}^*_s \eta^*_{st-1} + T^{-1}\sum_{s=1}^T \widetilde{f}^*_s \xi^*_{st-1}]^2\} \\
= \;& v_{NT}^{*-2} E^*\{T^{-1}\sum_{t=1}^T (A^*_{1t} + A^*_{2t} + A^*_{3t})^2\} \\
\leq \;& \epsilon^{-2} E^*\{T^{-1}\sum_{t=1}^T (A^*_{1t} + A^*_{2t} + A^*_{3t})^2\},
\end{aligned}$$

where $A^*_{1t} = T^{-1}\sum_{s=1}^T \widetilde{f}^*_s \zeta^*_{st-1}$, $A^*_{2t} = T^{-1}\sum_{s=1}^T \widetilde{f}^*_s \eta^*_{st-1}$ and $A^*_{3t} = T^{-1}\sum_{s=1}^T \widetilde{f}^*_s \xi^*_{st-1}$.

First,

$$\begin{aligned}
E^*\{T^{-1}\sum_{t=1}^T A^{*2}_{1t}\} &= E^*\{T^{-1}\sum_{t=1}^T [T^{-1}\sum_{s=1}^T \widetilde{f}^*_s \zeta^*_{st-1}]^2\} \\
&\leq E^*\{[T^{-1}\sum_{s=1}^T \widetilde{f}^{*2}_s][T^{-2}\sum_{t=1}^T\sum_{s=1}^T \zeta^{*2}_{st-1}]\} \\
&= \{E^*[T^{-2}\sum_{t=1}^T\sum_{s=1}^T \zeta^{*2}_{st-1}]\} = T^{-2}\sum_{t=1}^T\sum_{s=1}^T N^{-2}E^*[\sum_{i=1}^N e^{*2}_{is} e^{*2}_{it}] \\
&= O_P(\delta_{NT}^{-2}),
\end{aligned}$$

provided $E^* e^{*4}_{it} = O_P(1)$. Second,

$$\begin{aligned}
E^*\{T^{-1}\sum_{t=1}^T A^{*2}_{2t}\} &= E^*\{T^{-1}\sum_{t=1}^T (T^{-1}\sum_{s=1}^T \tilde{f}^*_s N^{-1} f^*_s \sum_{i=1}^N \lambda^*_i e^*_{it-1})^2\} \\
&= E^*\{T^{-1}\sum_{t=1}^T (T^{-1}\sum_{s=1}^T \tilde{f}^*_s f^*_s)^2 (N^{-1}\sum_{i=1}^N \lambda^*_i e^*_{it})^2\} \\
&\leq E^*\{(T^{-1}\sum_{s=1}^T \tilde{f}^{*2}_s)(T^{-1}\sum_{s=1}^T f^{*2}_s) T^{-1}\sum_{t=1}^T (N^{-1}\sum_{i=1}^N \lambda^*_i e^*_{it})^2\} \\
&= E^*[(T^{-1}\sum_{s=1}^T f^{*2}_s)][T^{-1}\sum_{t=1}^T N^{-2} E^* \sum_{i=1}^N \lambda^{*2}_i e^{*2}_{it}] \\
&= E^*[(T^{-1}\sum_{s=1}^T f^{*2}_s)][T^{-1}\sum_{t=1}^T N^{-2} \sum_{i=1}^N E^*\lambda^{*2}_i E^* e^{*2}_{it}] \\
&= O_P(\delta_{NT}^{-2}),
\end{aligned}$$

which follows from $E^*[T^{-1}\sum_{s=1}^T f_s^{*2}] = T^{-1}\sum_{s=1}^T \tilde{f}_s^2 = 1$ and $E^*\lambda_i^{*4} = O_P(1)$. Third,

$$\begin{aligned} E^*\{T^{-1}\sum_{t=1}^T A_{3t}^{*2}\} &= E^*\{(T^{-1}\sum_{t=1}^T f_t^{*2})(T^{-1}\sum_{s=1}^T \tilde{f}_s^* N^{-1}\sum_{i=1}^N \lambda_i^* e_{is}^*)^2\} \\ &\leq E^*\{(T^{-1}\sum_{s=1}^T \tilde{f}_s^{*2})(T^{-1}\sum_{t=1}^T f_t^{*2})[T^{-1}\sum_{s=1}^T (N^{-1}\sum_{i=1}^N \lambda_i^* e_{is}^*)^2]\} \\ &= \{E^*(T^{-1}\sum_{t=1}^T f_t^{*2})[T^{-1}N^{-2}E^*\sum_{s=1}^T\sum_{i=1}^N \lambda_i^{*2} e_{is}^{*2}]\} \\ &= O_P(\delta_{NT}^{-2}). \end{aligned}$$

Using the Cauchy-Schwartz inequality, we can show that $E^*[T^{-1}\sum_{t=1}^T A_{1t}^* A_{2t}^*] = O_P(\delta_{NT}^{-2})$, $E^*[T^{-1}\sum_{t=1}^T A_{1t}^* A_{3t}^*] = O_P(\delta_{NT}^{-2})$, and $E^*[T^{-1}\sum_{t=1}^T A_{2t}^* A_{3t}^*] = O_P(\delta_{NT}^{-2})$. Therefore, we have $E^*[T^{-1}\sum_{t=1}^T(\tilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*)^2] = O_P(\delta_{NT}^{-2})$. Since $H_{NT}^{*2} = 1 + o_{P^*}(1)$, with Markov's inequality we have $E^* H_{NT}^{*2} \geq 1 + o_P(1)$. Also, with $E^*[T^{-1}\sum_{t=1}^T(\tilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*)^2] = O_P(\delta_{NT}^{-2})$,

$$\begin{aligned} E^* H_{NT}^{*2} &= E^*[T^{-1}\tilde{F}^{*\prime}\tilde{F} - T^{-1}(\tilde{F}^* - H_{NT}^*\tilde{F})'\tilde{F}]^2 \\ &= E^*(T^{-1}\tilde{F}^{*\prime}\tilde{F})^2 + E^*[T^{-1}(\tilde{F}^* - H_{NT}^*\tilde{F})'\tilde{F}]^2 - 2E^*[T^{-2}\tilde{F}^{*\prime}\tilde{F}(\tilde{F}^* - H_{NT}^*\tilde{F})'\tilde{F}] \\ &\leq E^*(T^{-1}||\tilde{F}^{*\prime}||||\tilde{F}||) + o_P(1) = 1 + o_P(1). \end{aligned}$$

Therefore, $E^* H_{NT}^{*2} = 1 + o_P(1)$.Similarly, with $E^* e_{it}^{*8}$, $E^*\lambda_i^{*8}$, and $E^*\varepsilon_i^{*8}$ bounded in probability, we can obtain $E^*[T^{-1}\sum_{t=1}^T(\tilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*)^4] = O_P(\delta_{NT}^{-4})$ and $E^* H_{NT}^{*4} = 1 + o_P(1)$.

The bootstrap bias estimator can be decomposed as

$$\begin{aligned}
& E^*\left(\widetilde{\rho}^* - \tilde{\rho}\right) \\
= \;& E^*[\left(\sum_{t=2}^{T+1} \widetilde{f}_{t-1}^{*2}\right)^{-1} (\sum_{t=2}^{T} \widetilde{f}_{t-1}^{*} \widetilde{f}_{t}^{*} - \tilde{\rho} \sum_{t=2}^{T+1} \widetilde{f}_{t-1}^{*2})] \\
= \;& T^{-1} E^*[\sum_{t=2}^{T} \widetilde{f}_{t-1}^{*} \left(\widetilde{f}_{t}^{*} - \tilde{\rho} \widetilde{f}_{t-1}^{*}\right)] - T^{-1} \tilde{\rho} E^* \widetilde{f}_{T}^{*2} \\
= \;& T^{-1} E^*[\sum_{t=2}^{T} \widetilde{f}_{t-1}^{*} \left(\widetilde{f}_{t}^{*} - \tilde{\rho} \widetilde{f}_{t-1}^{*}\right)] + o_P(T^{-1/2}) \\
= \;& T^{-1} E^*[\sum_{t=2}^{T} \widetilde{f}_{t-1}^{*} \left\{\widetilde{f}_{t}^{*} - H_{NT}^{*} f_{t}^{*} - \tilde{\rho} \left(\widetilde{f}_{t-1}^{*} - H_{NT}^{*} f_{t-1}^{*}\right)\right\} + H_{NT}^{*} \sum_{t=2}^{T} \widetilde{f}_{t-1}^{*} \varepsilon_{t}^{*}] + o_P(T^{-1/2}) \\
= \;& T^{-1} E^*[H_{NT}^{*2} \sum_{t=2}^{T} f_{t-1}^{*} \varepsilon_{t}^{*}] - T^{-1} E^*[\tilde{\rho} \sum_{t=2}^{T} \widetilde{f}_{t-1}^{*} \left(\widetilde{f}_{t-1}^{*} - H_{NT}^{*} f_{t-1}^{*}\right)] \\
& + T^{-1} E^*[\sum_{t=2}^{T} \widetilde{f}_{t-1}^{*} \left(\widetilde{f}_{t}^{*} - H_{NT}^{*} f_{t}^{*}\right)] + T^{-1} E^*[H_{NT}^{*} \sum_{t=2}^{T} \left(\widetilde{f}_{t-1}^{*} - H_{NT}^{*} f_{t-1}^{*}\right) \varepsilon_{t}^{*}] + o_P(T^{-1/2}).
\end{aligned}$$

Note that third equality follows from $T^{-1} E^* \widetilde{f}_{T}^{*2} = o_P(T^{-1/2})$ which can be shown by using the decomposition $T^{-1} E^* H_{NT}^{*2} f_{T}^{*2} + T^{-1} E^* (\widetilde{f}_{T}^{*} - H_{NT}^{*} f_{T}^{*})^2$. The leading term can be written as

$$\begin{aligned}
& T^{-1} E^*[\left(H_{NT}^{*2} - 1\right) \sum_{t=2}^{T} f_{t-1}^{*} \varepsilon_{t}^{*} + \sum_{t=2}^{T} f_{t-1}^{*} \varepsilon_{t}^{*}] = T^{-1} E^*[\left(H_{NT}^{*2} - 1\right) \sum_{t=2}^{T} f_{t-1}^{*} \varepsilon_{t}^{*}] \\
\leq \;& \{E^*[H_{NT}^{*4} - 2H_{NT}^{*2} + 1] T^{-2} E^*[\sum_{t=2}^{T} f_{t-1}^{*} \varepsilon_{t}^{*}]^2\}^{1/2} \\
= \;& \{E^*[H_{NT}^{*2} - 1]^2 T^{-2} E^*[\sum_{t=2}^{T} f_{t-1}^{*2} \varepsilon_{t}^{*2}]\}^{1/2} = o_P(T^{-1/2}),
\end{aligned}$$

where we use the fact $E^*[H_{NT}^{*2} - 1]^2 = o_P(1)$. In what follows, we show that (i) $T^{-1} E^*[\tilde{\rho} \sum_{t=2}^{T} \widetilde{f}_{t-1}^{*} \left(\widetilde{f}_{t-1}^{*} - H_{NT}^{*} f_{t-1}^{*}\right)] = 2N^{-1} \rho v^{-2} \Gamma + o_P(T^{-1/2})$; (ii) $T^{-1} E^*[\tilde{\rho} \sum_{t=2}^{T} \widetilde{f}_{t-1}^{*} \left(\widetilde{f}_{t}^{*} - H_{NT}^{*} f_{t}^{*}\right)] = N^{-1} \rho v^{-2} \Gamma + o_P(T^{-1/2})$; and (iii) $T^{-1} E^*[H_{NT}^{*} \sum_{t=2}^{T} (\widetilde{f}_{t-1}^{*} - H_{NT}^{*} f_{t-1}^{*}) \varepsilon_{t}^{*} = o_P(T^{-1/2})$. The proof of (i) to (iii) is similar to the proof of Proposition 1. For (i),

$$T^{-1} E^* \{\tilde{\rho} \sum_{t=2}^{T} \tilde{f}_{t-1}^{*} (\tilde{f}_{t-1}^{*} - H_{NT}^{*} f_{t-1}^{*})\} = \tilde{\rho}(A^* + B^*)$$

where $A^* = E^*\{T^{-1}\sum_{t=2}^{T}(\tilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*)^2\} = v_{NT}^{*-2}E^*\{T^{-1}\sum_{t=1}^{T}(A_{1t}^* + A_{2t}^* + A_{3t}^*)^2\}$ and $B^* = E^*\{T^{-1}\sum_{t=2}^{T} H_{NT}^* f_{t-1}^*(\tilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*)\}$. To show $A^* = v^{-2}N^{-1}\Gamma + o_P(\delta_{NT}^{-2})$, we can focus on dominant term A_{2t}^* which is analogous to the dominant term A_{2t} in the proof of Proposition 1.

$$
\begin{aligned}
E^*\{v_{NT}^{*-2}T^{-1}\sum_{t=2}^{T} A_{2t}^{*2}\} &= T^{-3}E^* v_{NT}^{*-2}\sum_{t=2}^{T}[\sum_{s=1}^{T}(\widetilde{f}_s^* - H_{NT}^* f_s^*)\eta_{st-1}^* + \sum_{s=1}^{T} H_{NT}^* f_s^* \eta_{st-1}^*]^2, \\
&= T^{-1}E^* v_{NT}^{*-2}\sum_{t=2}^{T}(A_{21t}^* + A_{22t}^*)^2 \\
&= T^{-1}E^* v_{NT}^{*-2}\sum_{t=2}^{T}(A_{21t}^{*2} + A_{22t}^{*2} + 2A_{21t}^* A_{22t}^*),
\end{aligned}
$$

where $A_{21t}^* = T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s^* - H_{NT}^* f_s^*)\eta_{st-1}^*$ and $A_{22t}^* = T^{-1}\sum_{s=1}^{T} H_{NT}^* f_s^* \eta_{st-1}^*$. Furthermore,

$$
\begin{aligned}
T^{-1}E^* v_{NT}^{*-2}\sum_{t=2}^{T} A_{21t}^{*2} &\leq \epsilon^{-2}E^*\{[T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s^* - H_{NT}^* f_s^*)^2][T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T}\eta_{st-1}^{*2}]\} \\
&= \epsilon^{-2}\{E^*[T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s^* - H_{NT}^* f_s^*)^2]^2 E^*[T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T}\eta_{st-1}^{*2}]^2\}^{1/2} \\
&= O_P((TN)^{-1/2}),
\end{aligned}
$$

and

$$
T^{-1}E^* v_{NT}^{*-2}\sum_{t=2}^{T} A_{22t}^{*2} = T^{-1}E^*(v_{NT}^{*-2}H_{NT}^{*2} - v^{-2})H_{NT}^{*-2}\sum_{t=2}^{T} A_{22t}^{*2} + T^{-1}E^* v^{-2}H_{NT}^{*-2}\sum_{t=2}^{T} A_{22t}^{*2},
$$

where

$$
\begin{aligned}
& T^{-1}E^* v^{-2}H_{NT}^{*-2}\sum_{t=2}^{T} A_{22t}^{*2} = v^{-2}E^*[T^{-3}\sum_{t=2}^{T}(\sum_{s=1}^{T} f_s^* N^{-1} f_s^* \sum_{i=1}^{N}\lambda_i^* e_{it-1}^*)^2] \\
= \; & E^*\{v^{-2}(T^{-1}\sum_{s=1}^{T} f_s^{*2})^2 T^{-1}\sum_{t=2}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i^* e_{it-1}^*)^2\} \\
= \; & v^{-2}E^*(T^{-1}\sum_{s=1}^{T} f_s^{*2})^2 E^*[T^{-1}\sum_{t=2}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i^* e_{it-1}^*)^2] \\
= \; & v^{-2}N^{-1}\Gamma + o_P(T^{-1/2}) = O_P(\delta_{NT}^{-2}),
\end{aligned}
$$

where the last equality follows from

$$\begin{aligned} T^{-1}E^*\sum_{t=2}^{T}(N^{-1/2}\sum_{i=1}^{N}\lambda_i^* e_{it}^*)^2 &= T^{-1}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\widetilde{\lambda}_i^2\widetilde{e}_{it}^2 \\ &= \Gamma + o_P(1), \end{aligned}$$

and

$$\begin{aligned} & T^{-1}E^*(v_{NT}^{*-2}H_{NT}^{*2}-v^{-2})H_{NT}^{*-2}\sum_{t=2}^{T}A_{22t}^{*2} \\ = \quad & E^*(v_{NT}^{*-2}H_{NT}^{*2}-v^{-2})[T^{-3}\sum_{s=1}^{T}f_s^{*2}\sum_{t=2}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i^* e_{it-1}^*)^2] \\ \leq \quad & E^*(v_{NT}^{*-2}H_{NT}^{*2}-v^{-2})^2 E^*[T^{-3}\sum_{s=1}^{T}f_s^{*2}\sum_{t=2}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i^* e_{it-1}^*)^2]^2 \\ = \quad & o_P(\delta_{NT}^{-2}), \end{aligned}$$

and

$$\begin{aligned} T^{-1}E^*v_{NT}^{*-2}\sum_{t=2}^{T}A_{21t}^*A_{22t}^* &\leq E^*\{v_{NT}^{*-2}[T^{-1}\sum_{t=2}^{T}A_{21t}^{*2}]^{1/2}[T^{-1}\sum_{t=2}^{T}A_{22t}^{*2}]^{1/2}\} \\ &\leq \epsilon^{-2}\{E^*[T^{-1}\sum_{t=2}^{T}A_{21t}^{*2}]E^*[T^{-1}\sum_{t=2}^{T}A_{22t}^{*2}]\}^{1/2} = o_P(\delta_{NT}^{-2}). \end{aligned}$$

Therefore, $T^{-1}E^*v_{NT}^{*-2}\sum_{t=2}^{T}A_{2t}^{*2} = v^{-2}N^{-1}\Gamma + o_P(\delta_{NT}^{-2}) = O_P(\delta_{NT}^{-2})$. By combining all the results, $A^* = v^{-2}N^{-1}\Gamma + o_P(\delta_{NT}^{-2})$. For B^*, we have

$$\begin{aligned} B^* &= E^*\{H_{NT}^*v_{NT}^{*-1}T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T}[f_{t-1}^*\widetilde{f}_s^*\zeta_{st-1}^* + f_{t-1}^*\widetilde{f}_s^*\eta_{st-1}^* + f_{t-1}^*\widetilde{f}_s^*\xi_{st-1}^*]\} \\ &= (B_1^* + B_2^* + B_3^*). \end{aligned}$$

First,

$$
\begin{aligned}
B_1^* &= T^{-2}E^*[H_{NT}^* v_{NT}^{*-1}\sum_{t=2}^{T}\sum_{s=1}^{T} f_{t-1}^*\widetilde{f}_s^*\zeta_{st-1}^*] \\
&\le E^*\{[T^{-1}H_{NT}^* v_{NT}^{*-2}\sum_{s=1}^{T}\widetilde{f}_s^{*2}]^{1/2}[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T} f_{t-1}^*\zeta_{st-1}^*)^2]^{1/2}\} \\
&= E^*\{[H_{NT}^* v_{NT}^{*-2}][T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T} f_{t-1}^* N^{-1}\sum_{i=1}^{N} e_{is}^* e_{it-1}^*)^2]\}^{1/2} \\
&\le \{\epsilon^{-2}E^*[H_{NT}^{*2}]E^*[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T} f_{t-1}^* N^{-1}\sum_{i=1}^{N} e_{is}^* e_{it-1}^*)^2]\}^{1/2} \\
&= O_P((NT)^{-1/2}).
\end{aligned}
$$

Second,

$$
\begin{aligned}
B_2^* &= T^{-2}E^*[H_{NT}^* v_{NT}^{*-1}\sum_{t=2}^{T}\sum_{s=1}^{T} f_{t-1}^*\widetilde{f}_s^*\eta_{st-1}^*] \\
&\le E^*\{[H_{NT}^{*2} v_{NT}^{*-2}T^{-1}\sum_{s=1}^{T}\widetilde{f}_s^{*2}][T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T}\sum_{s=1}^{T} f_{t-1}^*\eta_{st-1}^*)^2]\}^{1/2} \\
&= E^*\{H_{NT}^{*2} v_{NT}^{*-2}[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T} f_{t-1}^*\eta_{st-1}^*)^2]\}^{1/2} \\
&\le \{\epsilon^{-2}E^*[H_{NT}^{*2}]E^*[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T} f_{t-1}^*\eta_{st-1}^*)^2]\}^{1/2} \\
&= O_P((NT)^{-1/2}).
\end{aligned}
$$

Third,

$$
\begin{aligned}
B_3^* &= T^{-2}E^*[H_{NT}^* v_{NT}^{*-1}\sum_{t=2}^{T}\sum_{s=1}^{T} f_{t-1}^*\widetilde{f}_s^*\xi_{st-1}^*] \\
&= T^{-2}E[H_{NT}^* v_{NT}^{*-1}\sum_{t=2}^{T}\sum_{s=1}^{T} f_{t-1}^*(\widetilde{f}_s^* - H_{NT}^* f_s^* + H_{NT}^* f_s^*)\xi_{st-1}^*] \\
&= T^{-2}E^*\{H_{NT}^* v_{NT}^{*-1}[\sum_{t=2}^{T}\sum_{s=1}^{T} f_{t-1}^*(\widetilde{f}_s^* - H_{NT}^* f_s^*)\xi_{st-1}^* + \sum_{t=2}^{T}\sum_{s=1}^{T} f_{t-1}^* H_{NT}^* f_s^*\xi_{st-1}^*]\} \\
&= B_{31}^* + B_{32}^*.
\end{aligned}
$$

For B^*_{31},

$$\begin{aligned}
B^*_{31} &= E^*\{H^*_{NT}v^{*-1}_{NT}[T^{-1}\sum_{t=2}^{T} f^{*2}_{t-1}][T^{-1}\sum_{s=1}^{T}(\widetilde{f}^*_s - H^*_{NT}f^*_s)N^{-1}\sum_{i=1}^{N}\lambda^*_i e^*_{is}]\} \\
&= T^{-2}E^*\{\sum_{s=1}^{T} H^*_{NT}v^{*-2}_{NT}\sum_{t=1}^{T}[\widetilde{f}^*_t\zeta^*_{ts} + \widetilde{f}^*_t\eta^*_{ts} + \widetilde{f}^*_t\xi^*_{ts}]N^{-1}\sum_{i=1}^{N}\lambda^*_i e^*_{is}\} \\
&= (B^*_{311} + B^*_{312} + B^*_{313}),
\end{aligned}$$

where

$$\begin{aligned}
B^*_{311} &= T^{-2}E^*[H^*_{NT}v^{*-2}_{NT}\sum_{s=1}^{T}\sum_{t=1}^{T}\widetilde{f}^*_t\zeta^*_{ts}N^{-1}\sum_{i=1}^{N}\lambda^*_i e^*_{is}] \\
&\leq E^*\{[(T^{-1}H^*_{NT}v^{*-2}_{NT}\sum_{t=1}^{T}\widetilde{f}^*_t)^2]^{1/2}[T^{-3}\sum_{t=1}^{T}(\sum_{s=1}^{T}\zeta^*_{ts}N^{-1}\sum_{i=1}^{N}\lambda^*_i e^*_{is})^2]^{1/2}\} \\
&\leq E^*\{[H^{*2}_{NT}v^{*-4}_{NT}]^{1/2}[T^{-3}\sum_{t=1}^{T}(\sum_{s=1}^{T}\zeta^*_{ts}N^{-1}\sum_{i=1}^{N}\lambda^*_i e^*_{is})^2]^{1/2}\} \\
&\leq \epsilon^{-2}\{E^*\{[H^{*2}_{NT}]E^*[T^{-1}\sum_{t=1}^{T}(T^{-1}\sum_{s=1}^{T}\zeta^*_{ts}N^{-1}\sum_{i=1}^{N}\lambda^*_i e^*_{is})^2]\}^{1/2} \\
&= O_P((NT)^{-1/2}),
\end{aligned}$$

and

$$\begin{aligned}
B^*_{312} &= E^*[H^*_{NT}v^{*-2}_{NT}T^{-2}\sum_{s=1}^{T}\sum_{t=1}^{T}\widetilde{f}^*_t\eta^*_{ts}N^{-1}\sum_{i=1}^{N}\lambda^*_i e^*_{is}] \\
&= E^*\{H^*_{NT}v^{*-2}_{NT}(T^{-1}\sum_{t=1}^{T}\widetilde{f}^*_t f^*_t)[T^{-1}\sum_{s=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda^*_i e^*_{is})^2]\} \\
&= E^*\{v^{-2}(T^{-1}\sum_{t=1}^{T} f^{*2}_t)[T^{-1}\sum_{s=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda^*_i e^*_{is})^2]\} \\
&\quad + E^*\{(H^{*2}_{NT}v^{*-2}_{NT} - v^{-2})(T^{-1}\sum_{t=1}^{T} f^{*2}_t)[T^{-1}\sum_{s=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda^*_i e^*_{is})^2]\} \\
&\quad + E\{H^*_{NT}v^{*2}_{NT}[T^{-1}\sum_{t=1}^{T}(\widetilde{f}^*_t - H^*_{NT}f^*_t)f^*_t][T^{-1}\sum_{s=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda^*_i e^*_{is})^2]\} \\
&= v^{-2}N^{-1}\Gamma + o_P(\delta^{-2}_{NT}) + o_P(\delta^{-2}_{NT}),
\end{aligned}$$

and

$$\begin{aligned} B^*_{313} &= T^{-2}E[H^*_{NT}v^{*-2}_{NT}\sum_{s=1}^{T}\sum_{t=1}^{T}(\widetilde{f}^*_t - H^*_{NT}f^*_t + H^*_{NT}f^*_t)\xi^*_{ts}N^{-1}\sum_{i=1}^{N}\lambda^*_i e^*_{is}] \\ &= O_P(\delta^{-1}_{NT}N^{-1}), \end{aligned}$$

since

$$\begin{aligned} & T^{-2}E^*[H^*_{NT}v^{*-2}_{NT}\sum_{s=1}^{T}\sum_{t=1}^{T}(\widetilde{f}^*_t - H^*_{NT}f^*_t)\xi^*_{ts}N^{-1}\sum_{i=1}^{N}\lambda^*_i e^*_{is}] \\ \leq\ & E^*\{[H^{*2}_{NT}v^{*-4}_{NT}T^{-1}\sum_{t=1}^{T}(\widetilde{f}^*_t - H^*_{NT}f^*_t)^2]^{1/2}[T^{-1}\sum_{t=1}^{T}(T^{-1}\sum_{s=1}^{T}\xi^*_{ts}N^{-1}\sum_{i=1}^{N}\lambda^*_i e^*_{is})^2]^{1/2}\} \\ \leq\ & \{E^*[T^{-1}H^{*2}_{NT}v^{*-4}_{NT}\sum_{t=1}^{T}(\widetilde{f}^*_t - H^*_{NT}f^*_t)^2]E^*[T^{-1}\sum_{t=1}^{T}(T^{-1}\sum_{s=1}^{T}\xi^*_{ts}N^{-1}\sum_{i=1}^{N}\lambda^*_i e^*_{is})^2]\}^{1/2} \\ \leq\ & \epsilon^{-2}\{E^*[T^{-1}H^{*2}_{NT}\sum_{t=1}^{T}(\widetilde{f}^*_t - H^*_{NT}f^*_t)^2]E^*[T^{-1}\sum_{t=1}^{T}(T^{-1}\sum_{s=1}^{T}\xi^*_{ts}N^{-1}\sum_{i=1}^{N}\lambda^*_i e^*_{is})^2]\}^{1/2} \\ \leq\ & \epsilon^{-2}\{E^*H^{*4}_{NT}E^*[T^{-1}\sum_{t=1}^{T}(\widetilde{f}^*_t - H^*_{NT}f^*_t)^2]^2E^*[T^{-1}\sum_{t=1}^{T}(T^{-1}\sum_{s=1}^{T}\xi^*_{ts}N^{-1}\sum_{i=1}^{N}\lambda^*_i e^*_{is})^2]\}^{1/2} \\ =\ & O_P((NT)^{-1/2}), \end{aligned}$$

and

$$\begin{aligned} & T^{-2}E^*[H^{*2}_{NT}v^{*-2}_{NT}\sum_{s=1}^{T}\sum_{t=1}^{T}f^*_t\xi^*_{ts}N^{-1}\sum_{i=1}^{N}\lambda^*_i e^*_{is}] \\ =\ & E^*\{H^{*2}_{NT}v^{*-2}_{NT}[T^{-1}\sum_{s=1}^{T}N^{-1}\sum_{i=1}^{N}f^*_s\lambda^*_i e^*_{is}]^2\}. = O_P((NT)^{-1/2}). \end{aligned}$$

Thus, $B^*_{31} = v^{-2}N^{-1}\Gamma + o_P(\delta^{-2}_{NT})$. For B^*_{32},

$$\begin{aligned} B^*_{32} &= E^*[H^{*2}_{NT}v^{*-1}_{NT}T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T}f^*_{t-1}f^*_s\xi^*_{st-1}] \\ &= E^*[H^{*2}_{NT}v^{*-1}_{NT}T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T}f^*_{t-1}f^*_sN^{-1}\sum_{i=1}^{N}f^*_{t-1}\lambda^*_i e^*_{is}] \\ &= E^*[H^{*2}_{NT}v^{*-1}_{NT}(T^{-1}\sum_{t=2}^{T}f^{*2}_{t-1})(T^{-1}N^{-1}\sum_{s=1}^{T}\sum_{i=1}^{N}f^*_s\lambda^*_i e^*_{is})] = O_P((NT)^{-1/2}). \end{aligned}$$

Therefore, $B^*_3 = v^{-2}N^{-1}\Gamma + o_P(\delta^{-2}_{NT})$. By combining all the results for B^*_1, B^*_2 and B^*_3, we have $B^* = v^{-2}N^{-1}\Gamma + o_P(\delta^{-2}_{NT})$. Thus, $E^*[T^{-1}\tilde{\rho}\sum_{t=2}^{T}\widetilde{f}^*_{t-1}(\widetilde{f}^*_{t-1} - H^*_{NT}f^*_{t-1})] = \tilde{\rho}(A^* + B^*) =$

$2\rho v^{-2}N^{-1}\Gamma + o_P(\delta_{NT}^{-2})$ which completes the proof of (i).

To show (ii), we use a similar decomposition as (i),

$$\begin{aligned}
& T^{-1}E^*\tilde{\rho}\sum_{t=2}^{T}\widetilde{f}_{t-1}^*(\widetilde{f}_t^* - H_{NT}^* f_t^*) \\
= \quad & E^*[T^{-1}\tilde{\rho}\sum_{t=2}^{T}(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*)(\widetilde{f}_t^* - H_{NT}^* f_t^*) + T^{-1}H_{NT}^*\tilde{\rho}\sum_{t=2}^{T} f_{t-1}^*(\widetilde{f}_t^* - H_{NT}^* f_t^*)].
\end{aligned}$$

Since the proof is almost the same as the proof of (i), we only mention the difference. To show $T^{-1}E^*\sum_{t=2}^{T}(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*)(\widetilde{f}_t^* - H_{NT}^* f_t^*) = o_P(\delta_{NT}^{-2})$, we need to use

$$\begin{aligned}
& T^{-1}E^*\sum_{t=2}^{T}(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*)(\widetilde{f}_t^* - H_{NT}^* f_t^*) \\
= \quad & E^*[H_{NT}^{*2}v_{NT}^{*-2}T^{-3}\sum_{t=2}^{T}(\sum_{s=1}^{T} f_s^* N^{-1} f_s^* \sum_{i=1}^{N}\lambda_i^* e_{it-1}^*)(\sum_{s=1}^{T} f_s^* N^{-1} f_s^* \sum_{i=1}^{N}\lambda_i^* e_{it}^*)] + o_P(\delta_{NT}^{-2}) \\
= \quad & E^*[H_{NT}^{*2}v_{NT}^{*-2}T^{-3}(\sum_{s=1}^{T} f_s^{*2})^2\sum_{t=1}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i^* e_{it-1}^*)(N^{-1}\sum_{i=1}^{N}\lambda_i^* e_{it}^*)] + o_P(\delta_{NT}^{-2}) \\
= \quad & o_P(\delta_{NT}^{-2}).
\end{aligned}$$

To show $E^*[H_{NT}^* T^{-1}\sum_{t=2}^{T} f_{t-1}^*(\widetilde{f}_t^* - H_{NT}^* f_t^*)] = \rho v^{-2}N^{-1}\Gamma + o_P(\delta_{NT}^{-2})$, we need to use

$$\begin{aligned}
& E^*[H_{NT}^* T^{-1}\sum_{t=2}^{T} f_{t-1}^*(\widetilde{f}_t^* - H_{NT}^* f_t^*)] \\
= \quad & E^*\{[T^{-1}\sum_{t=2}^{T} f_{t-1}^* f_t^*][T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s^* - H_{NT}^* f_s^*)N^{-1}\sum_{i=1}^{N}\lambda_i^* e_{is}^*]\} + o_P(\delta_{NT}^{-2}) \\
= \quad & E^*[\tilde{\rho}[T^{-1}\sum_{t=2}^{T} f_{t-1}^{*2}]T^{-1}\sum_{s=1}^{T}(\widetilde{f}_s^* - H_{NT}^* f_s^*)N^{-1}\sum_{i=1}^{N}\lambda_i^* e_{is}^*] + o_P(\delta_{NT}^{-2}) \\
= \quad & \rho v^{-2}N^{-1}\Gamma + o_P(\delta_{NT}^{-2}).
\end{aligned}$$

To obtain the result (iii), we have

$$\begin{aligned}
& T^{-1}E^*[H_{NT}^*\sum_{t=2}^{T}(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*)\varepsilon_t^*] \\
= \quad & E^*\{H_{NT}^* v_{NT}^{*-1}T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T}[\widetilde{f}_s^*\zeta_{st-1}^*\varepsilon_t^* + \widetilde{f}_s^*\eta_{st-1}^*\varepsilon_t^* + \widetilde{f}_s^*\xi_{st-1}^*\varepsilon_t^*]\} \\
= \quad & C_1^* + C_2^* + C_3^*.
\end{aligned}$$

For C_1^*,

$$
\begin{aligned}
C_1^* &= T^{-2}E^*[H_{NT}^* v_{NT}^{*-1} \sum_{t=2}^{T}\sum_{s=1}^{T} \widetilde{f}_s^* \zeta_{st-1}^* \varepsilon_t^*] \\
&\leq E^*\{[T^{-1}\sum_{s=1}^{T}(H_{NT}^* v_{NT}^{*-1}\widetilde{f}_s^*)^2]^{1/2}[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T}\zeta_{st-1}^*\varepsilon_t^*)^2]^{1/2}\} \\
&\leq \{E^*[H_{NT}^{*2} v_{NT}^{*-2}]E^*[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T}\zeta_{st-1}^*\varepsilon_t^*)^2]\}^{1/2} \\
&= O_P((NT)^{-1/2}),
\end{aligned}
$$

where the last equality follows from

$$
\begin{aligned}
T^{-1}\sum_{s=1}^{T}E^*(T^{-1}\sum_{t=2}^{T}\zeta_{st-1}^*\varepsilon_t^*)^2 &= T^{-1}\sum_{s=1}^{T}E^*(T^{-1}\sum_{t=2}^{T}[N^{-1}\sum_{i=1}^{N}e_{it-1}^* e_{is}^*]\varepsilon_t^*)^2 \\
&= \sigma^{*2}T^{-1}\sum_{s=1}^{T}T^{-2}E^*\sum_{t=2}^{T}[N^{-1}\sum_{i=1}^{N}e_{it-1}^* e_{is}^*]^2 \\
&= O_P((NT)^{-1}).
\end{aligned}
$$

For C_2^*,

$$
\begin{aligned}
C_2^* &= E^*\{H_{NT}^* v_{NT}^{*-1} T^{-2}\sum_{t=2}^{T}\sum_{s=1}^{T}\widetilde{f}_s^*\eta_{st-1}^*\varepsilon_t^*\} \\
&\leq E^*\{(T^{-1}\sum_{s=1}^{T}H_{NT}^{*2}v_{NT}^{*-2}\widetilde{f}_s^{*2})^{1/2}[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T}\eta_{st-1}^*\varepsilon_t^*)^2]^{1/2}\} \\
&\leq \{E^*(H_{NT}^{*2}v_{NT}^{*-2})E^*[T^{-1}\sum_{s=1}^{T}(T^{-1}\sum_{t=2}^{T}\eta_{st-1}^*\varepsilon_t^*)^2]\}^{1/2} \\
&= O_P((NT)^{-1/2}),
\end{aligned}
$$

where the last equality follows from

$$
\begin{aligned}
T^{-1}\sum_{s=1}^{T}E^*(T^{-1}\sum_{t=2}^{T}\eta_{st-1}^*\varepsilon_t^*)^2 &= \sigma^{*2}T^{-1}\sum_{s=1}^{T}E^*(T^{-2}\sum_{t=2}^{T}\eta_{st-1}^{*2}) \\
&= \sigma^{*2}T^{-1}\sum_{s=1}^{T}E^*[T^{-2}\sum_{t=2}^{T}(N^{-1}\sum_{i=1}^{N}f_s^*\lambda_i^* e_{it-1}^*)^2] \\
&= \sigma^{*2}\sigma_f^{*2}T^{-2}E^*[\sum_{t=2}^{T}(N^{-1}\sum_{i=1}^{N}\lambda_i^* e_{it-1}^*)^2] = O_P((NT)^{-1}).
\end{aligned}
$$

Similarly, we can show $C_3^* = O_P((NT)^{-1/2})$. By combining all the results for C_1^*, C_2^* and C_3^*, we have $E^*[T^{-1}H_{NT}^* \sum_{t=2}^T (\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*)\varepsilon_t^*] = O_P((NT)^{-1/2})$ which completes the proof of *(iii)*. Finally, the desired result follows from

$$E^*\left(\widetilde{\rho}^* - \tilde{\rho}\right) = -2N^{-1}\rho v^{-2}\Gamma + N^{-1}\rho v^{-2}\Gamma + o_P(T^{-1/2})$$

combined with $v = \sigma_\lambda^2$ and $T^{1/2}N^{-1} - c = o(1)$. ■

Proof of Proposition 3.

The dominant term of the bootstrap estimation error can be decomposed as

$$\begin{aligned}
&\sqrt{T}\left(\widetilde{\rho}^* - \tilde{\rho}\right) = T^{-1/2}\sum_{t=2}^T \widetilde{f}_{t-1}^*\left(\widetilde{f}_t^* - \tilde{\rho}\widetilde{f}_{t-1}^*\right) + o_{P^*}(1) \\
= \;& T^{-1/2}\sum_{t=2}^T \widetilde{f}_{t-1}^*\left\{\widetilde{f}_t^* - H_{NT}^* f_t^* - \tilde{\rho}\left(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*\right)\right\} + T^{-1/2}H_{NT}^*\sum_{t=2}^T \widetilde{f}_{t-1}^*\varepsilon_t^* + o_{P^*}(1) \\
= \;& T^{-1/2}H_{NT}^{*2}\sum_{t=2}^T f_{t-1}^*\varepsilon_t^* - T^{-1/2}\tilde{\rho}\sum_{t=2}^T \widetilde{f}_{t-1}^*\left(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*\right) + o_{P^*}(1) \\
& + T^{-1/2}\sum_{t=2}^T \widetilde{f}_{t-1}^*\left(\widetilde{f}_t^* - H_{NT}^* f_t^*\right) + T^{-1/2}H_{NT}^*\sum_{t=2}^T\left(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*\right)\varepsilon_t^* + o_{P^*}(1).
\end{aligned}$$

The leading term can be written as

$$T^{-1/2}\left(H_{NT}^{*2} - 1\right)\sum_{t=2}^T f_{t-1}^*\varepsilon_t^* + T^{-1/2}\sum_{t=2}^T f_{t-1}^*\varepsilon_t^* = T^{-1/2}\sum_{t=2}^T f_{t-1}^*\varepsilon_t^* + o_{P^*}(1).$$

The last equality follows from the fact that $H_{NT}^{*2} - 1 = o_{P^*}(1)$. $E^*(e_{it}^{*4})$, $E^*(\lambda_i^{*4})$, and $E^*(\varepsilon_t^{*4})$ are bounded in probability because of mutual independence of f_t, λ_i, and e_{it} and $E|f_t|^4 \leq M$, $E|\lambda_i|^4 \leq M$, and $E|e_{it}|^8 \leq M$. Analogous to the proofs of Propositions 1 and 2, we have (i) $T^{-1}\tilde{\rho}\sum_{t=2}^T \widetilde{f}_{t-1}^*\left(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*\right) = -2\rho v^{-2}N^{-1}\Gamma + o_P^*(\delta_{NT}^{-2})$; (ii) $T^{-1}\tilde{\rho}\sum_{t=2}^T \widetilde{f}_{t-1}^* \cdot \left(\widetilde{f}_t^* - H_{NT}^* f_t^*\right) = \rho v^{-2}N^{-1}\Gamma + o_P^*(\delta_{NT}^{-2})$; and (iii) $T^{-1}H_{NT}^*\sum_{t=2}^T(\widetilde{f}_{t-1}^* - H_{NT}^* f_{t-1}^*)\varepsilon_t^* = o_P^*(\delta_{NT}^{-2})$. Therefore,

$$\sqrt{T}\left(\widetilde{\rho}^* - \tilde{\rho}\right) = T^{-1/2}\sum_{t=2}^T f_{t-1}^*\varepsilon_t^* - c\rho\sigma_\lambda^{-4}\Gamma + o_P^*(1).$$

We apply the bootstrap central limit theorem to the term $T^{-1/2}\sum_{t=2}^T f_{t-1}^*\varepsilon_t^*$. Since $E^*[f_{t-1}^*\varepsilon_t^* | f_{t-2}^*\varepsilon_{t-1}^*, ...] = 0$, we can use the central limit theorem for the martingale difference sequence under the bootstrap probability measure and thus $P^*(\sqrt{T}(\tilde{\rho}^* - \tilde{\rho}) \leq x)$ approaches normal distribution function with mean $-c\rho\sigma_\lambda^{-4}\Gamma$ and variance $E^*(f_{t-1}^{*2}\varepsilon_t^{*2}) = T^{-1}\sum_{t=2}^T \tilde{f}_{t-1}^2\tilde{\varepsilon}_t^2$ under the bootstrap probability measure. Combining it with $T^{-1}\sum_{t=2}^T \tilde{f}_{t-1}^2\tilde{\varepsilon}_t^2 \to^P E(f_{t-1}^2\varepsilon_t^2) = 1-\rho^2$, we have $P^*(\sqrt{T}(\tilde{\rho}^* - \tilde{\rho}) \leq x) - P(\sqrt{T}(\tilde{\rho} - \rho) \leq x) \to^P 0$ for any x. By using Polya's theorem, we have the uniform convergence result. ■

Table 1: AR Estimation

ρ	T	Estimator $\hat{\rho}$	$\hat{\rho}_{KBC}$	$\hat{\rho}_{BC}$	Coverage Rate
0.5	100	0.49	0.50	0.50	0.90
	200	0.50	0.50	0.50	0.90
0.9	100	0.88	0.90	0.90	0.90
	200	0.89	0.90	0.90	0.90

Note: Mean values of the OLS estimator ($\hat{\rho}$), the Kendall-type bias-corrected estimator ($\hat{\rho}_{KBC}$) and the bootstrap bias-corrected estimator ($\hat{\rho}_{BC}$) and coverage rates of the asymptotic confidence interval (5) in 10,000 replications.

Table 2: Two-Step AR Estimation

			$\tilde{\rho}$					Coverage Rate				
ρ	T	c	S/N =0.5	0.75	1	1.5	2	S/N=0.5	0.75	1	1.5	2
			(A) No cross-sectional correlation									
0.5	100	0.5	0.42	0.43	0.44	0.45	0.46	0.77	0.82	0.84	0.86	0.86
		1	0.36	0.39	0.41	0.42	0.44	0.58	0.69	0.73	0.79	0.82
		1.5	0.32	0.36	0.38	0.40	0.41	0.44	0.56	0.64	0.73	0.76
	200	0.5	0.45	0.46	0.46	0.47	0.47	0.79	0.83	0.85	0.87	0.88
		1	0.41	0.43	0.44	0.45	0.46	0.60	0.70	0.75	0.81	0.83
		1.5	0.36	0.40	0.41	0.43	0.44	0.39	0.53	0.61	0.71	0.77
0.9	100	0.5	0.73	0.77	0.79	0.81	0.82	0.25	0.40	0.47	0.58	0.63
		1	0.62	0.68	0.71	0.75	0.77	0.07	0.13	0.22	0.32	0.41
		1.5	0.54	0.61	0.65	0.70	0.73	0.03	0.07	0.11	0.19	0.27
	200	0.5	0.80	0.82	0.83	0.85	0.85	0.27	0.42	0.51	0.62	0.70
		1	0.72	0.76	0.78	0.81	0.82	0.05	0.12	0.21	0.35	0.43
		1.5	0.65	0.70	0.73	0.77	0.79	0.01	0.04	0.08	0.16	0.25
			(B) Cross-sectional correlation									
0.5	100	0.5	0.40	0.42	0.44	0.45	0.45	0.71	0.79	0.81	0.84	0.86
		1	0.29	0.35	0.38	0.41	0.42	0.39	0.55	0.65	0.73	0.78
		1.5	0.21	0.28	0.32	0.37	0.39	0.24	0.39	0.49	0.61	0.69
	200	0.5	0.44	0.45	0.46	0.47	0.48	0.75	0.82	0.84	0.86	0.87
		1	0.37	0.41	0.43	0.45	0.46	0.44	0.61	0.69	0.78	0.82
		1.5	0.28	0.34	0.38	0.41	0.43	0.22	0.38	0.49	0.64	0.70
0.9	100	0.5	0.67	0.74	0.77	0.80	0.81	0.19	0.33	0.43	0.54	0.61
		1	0.44	0.56	0.63	0.70	0.74	0.05	0.11	0.17	0.29	0.38
		1.5	0.32	0.44	0.52	0.62	0.67	0.02	0.06	0.10	0.18	0.24
	200	0.5	0.78	0.81	0.83	0.84	0.85	0.22	0.37	0.48	0.61	0.66
		1	0.63	0.71	0.76	0.80	0.81	0.05	0.11	0.18	0.31	0.40
		1.5	0.46	0.59	0.65	0.73	0.76	0.01	0.05	0.08	0.16	0.23

Note: Mean values of the two-step estimator ($\tilde{\rho}$) and coverage rates of the asymptotic confidence interval (10) in 10,000 replications. S/N denotes the signal-to-noise ratio.

Table 3: Bootstrap Bias Corrections

			$T = 100$					$T = 200$				
ρ	c		S/N =0.5	0.75	1	1.5	2	S/N =0.5	0.75	1	1.5	2
			(A) No cross-sectional correlation									
0.5	0.5	bias	-0.08	-0.07	-0.06	-0.05	-0.05	-0.05	-0.04	-0.04	-0.03	-0.02
		asy bias	-0.05	-0.03	-0.03	-0.02	-0.01	-0.04	-0.02	-0.02	-0.01	-0.01
		bias I*	-0.05	-0.04	-0.03	-0.02	-0.02	-0.04	-0.03	-0.02	-0.01	-0.01
		bias II*	-0.07	-0.06	-0.05	-0.05	-0.04	-0.05	-0.04	-0.03	-0.03	-0.03
	1	bias	-0.13	-0.11	-0.10	-0.07	-0.07	-0.09	-0.08	-0.06	-0.04	-0.04
		asy bias	-0.10	-0.07	-0.05	-0.03	-0.03	-0.07	-0.05	-0.04	-0.02	-0.02
		bias I*	-0.08	-0.07	-0.06	-0.05	-0.04	-0.06	-0.05	-0.04	-0.04	-0.03
		bias II*	-0.09	-0.09	-0.08	-0.07	-0.07	-0.07	-0.06	-0.05	-0.05	-0.04
	1.5	bias	-0.18	-0.14	-0.12	-0.10	-0.08	-0.13	-0.10	-0.09	-0.07	-0.06
		asy bias	-0.15	-0.10	-0.07	-0.05	-0.04	-0.11	-0.07	-0.05	-0.04	-0.03
		bias I*	-0.09	-0.08	-0.08	-0.07	-0.06	-0.08	-0.07	-0.07	-0.06	-0.05
		bias II*	-0.10	-0.10	-0.10	-0.09	-0.08	-0.09	-0.08	-0.08	-0.07	-0.06
0.9	0.5	bias	-0.17	-0.13	-0.11	-0.10	-0.09	-0.10	-0.08	-0.07	-0.06	-0.05
		asy bias	-0.09	-0.06	-0.04	-0.03	-0.02	-0.06	-0.04	-0.03	-0.02	-0.02
		bias I*	-0.09	-0.08	-0.07	-0.05	-0.04	-0.07	-0.05	-0.05	-0.03	-0.03
		bias II*	-0.13	-0.12	-0.10	-0.09	-0.08	-0.09	-0.07	-0.07	-0.05	-0.05
	1	bias	-0.28	-0.22	-0.19	-0.15	-0.13	-0.18	-0.14	-0.12	-0.09	-0.08
		asy bias	-0.18	-0.12	-0.09	-0.06	-0.05	-0.13	-0.08	-0.06	-0.04	-0.03
		bias I*	-0.14	-0.13	-0.11	-0.10	-0.09	-0.12	-0.10	-0.09	-0.07	-0.06
		bias II*	-0.17	-0.16	-0.15	-0.14	-0.13	-0.13	-0.12	-0.10	-0.09	-0.08
	1.5	bias	-0.36	-0.29	-0.24	-0.20	-0.17	-0.26	-0.20	-0.17	-0.13	-0.11
		asy bias	-0.27	-0.18	-0.14	-0.09	-0.07	-0.19	-0.13	-0.10	-0.06	-0.05
		bias I*	-0.15	-0.15	-0.15	-0.14	-0.12	-0.15	-0.14	-0.12	-0.11	-0.09
		bias II*	-0.18	-0.18	-0.18	-0.17	-0.16	-0.16	-0.15	-0.14	-0.13	-0.11

Table 3 (continued)

ρ	c		$T = 100$ S/N =0.5	0.75	1	1.5	2	$T = 200$ S/N =0.5	0.75	1	1.5	2
		(B) Cross-sectional correlation										
0.5	0.5	bias	-0.10	-0.08	-0.07	-0.05	-0.05	-0.06	-0.04	-0.04	-0.03	-0.02
		asy bias	-0.05	-0.03	-0.03	-0.02	-0.01	-0.04	-0.02	-0.02	-0.01	-0.01
		bias I*	-0.05	-0.04	-0.03	-0.02	-0.02	-0.04	-0.03	-0.02	-0.02	-0.01
		bias II*	-0.07	-0.06	-0.05	-0.05	-0.04	-0.05	-0.04	-0.03	-0.03	-0.03
	1	bias	-0.21	-0.15	-0.12	-0.09	-0.08	-0.14	-0.09	-0.07	-0.05	-0.04
		asy bias	-0.10	-0.07	-0.05	-0.03	-0.03	-0.07	-0.05	-0.04	-0.02	-0.02
		bias I*	-0.06	-0.07	-0.06	-0.05	-0.05	-0.06	-0.05	-0.04	-0.04	-0.03
		bias II*	-0.08	-0.08	-0.08	-0.07	-0.07	-0.07	-0.06	-0.05	-0.05	-0.04
	1.5	bias	-0.29	-0.22	-0.18	-0.14	-0.10	-0.23	-0.16	-0.13	-0.09	-0.07
		asy bias	-0.15	-0.10	-0.07	-0.05	-0.04	-0.11	-0.07	-0.05	-0.04	-0.03
		bias I*	-0.07	-0.08	-0.08	-0.07	-0.07	-0.07	-0.07	-0.07	-0.06	-0.05
		bias II*	-0.08	-0.09	-0.09	-0.09	-0.09	-0.08	-0.08	-0.08	-0.07	-0.06
0.9	0.5	bias	-0.22	-0.17	-0.13	-0.11	-0.09	-0.12	-0.08	-0.07	-0.06	-0.05
		asy bias	-0.09	-0.06	-0.04	-0.03	-0.02	-0.06	-0.04	-0.03	-0.02	-0.02
		bias I*	-0.09	-0.08	-0.07	-0.05	-0.04	-0.07	-0.05	-0.05	-0.03	-0.03
		bias II*	-0.12	-0.11	-0.10	-0.09	-0.08	-0.09	-0.07	-0.07	-0.05	-0.05
	1	bias	-0.45	-0.35	-0.26	-0.18	-0.16	-0.27	0.18	-0.14	-0.10	-0.08
		asy bias	-0.18	-0.12	-0.09	-0.06	-0.04	-0.13	-0.08	-0.06	-0.04	-0.03
		bias I*	-0.11	-0.11	-0.11	-0.10	-0.09	-0.11	-0.10	-0.09	-0.07	-0.06
		bias II*	-0.13	-0.14	-0.14	-0.13	-0.13	-0.12	-0.11	-0.10	-0.09	-0.08
	1.5	bias	-0.57	-0.45	-0.37	-0.28	-0.23	-0.45	-0.31	-0.25	-0.18	-0.14
		asy bias	-0.27	-0.18	-0.14	-0.09	-0.07	-0.19	-0.13	-0.10	-0.06	-0.05
		bias I*	-0.12	-0.13	-0.14	-0.13	-0.12	-0.12	-0.13	-0.12	-0.11	-0.09
		bias II*	-0.13	-0.15	-0.16	-0.16	-0.16	-0.13	-0.14	-0.14	-0.13	-0.11

Note: The actual bias (bias), bootstrap bias estimator based on Method I (bias I*) and bootstrap bias estimator based on Method II (bias II*) are mean values in 10,000 replications. The asymptotic bias (asy bias) is $-T^{-1/2}c\rho\sigma_{\lambda}^{-4}\Gamma$. S/N denotes the signal-to-noise ratio.

Table 4: Coverage Rate of Bootstrap Confidence Intervals

			$T=100$					$T=200$				
ρ	c		S/N=0.5	0.75	1	1.5	2	S/N=0.5	0.75	1	1.5	2
			(A) No cross-sectional correlation									
0.5	0.5	Bc	0.85	0.86	0.86	0.86	0.87	0.86	0.87	0.87	0.87	0.88
		Per	0.87	0.87	0.88	0.87	0.88	0.88	0.88	0.88	0.89	0.89
		Per-t	0.86	0.87	0.87	0.87	0.88	0.87	0.88	0.87	0.88	0.89
	1	Bc	0.77	0.80	0.83	0.83	0.85	0.81	0.83	0.84	0.87	0.86
		Per	0.80	0.84	0.86	0.86	0.87	0.86	0.87	0.87	0.88	0.87
		Per-t	0.79	0.83	0.85	0.85	0.87	0.84	0.85	0.86	0.87	0.87
	1.5	Bc	0.68	0.75	0.77	0.80	0.82	0.72	0.79	0.81	0.82	0.85
		Per	0.73	0.80	0.81	0.84	0.85	0.78	0.83	0.85	0.85	0.87
		Per-t	0.72	0.79	0.80	0.84	0.84	0.75	0.82	0.84	0.84	0.87
0.9	0.5	Bc	0.78	0.82	0.83	0.84	0.84	0.84	0.87	0.88	0.89	0.89
		Per	0.90	0.93	0.93	0.93	0.93	0.95	0.95	0.95	0.94	0.93
		Per-t	0.80	0.86	0.87	0.88	0.88	0.86	0.90	0.90	0.90	0.89
	1	Bc	0.60	0.70	0.75	0.79	0.80	0.70	0.80	0.83	0.84	0.86
		Per	0.74	0.84	0.88	0.91	0.93	0.87	0.93	0.94	0.95	0.95
		Per-t	0.62	0.73	0.79	0.84	0.87	0.72	0.82	0.86	0.89	0.89
	1.5	Bc	0.45	0.60	0.66	0.73	0.76	0.50	0.64	0.71	0.76	0.80
		Per	0.60	0.74	0.79	0.87	0.88	0.70	0.84	0.90	0.92	0.93
		Per-t	0.48	0.63	0.70	0.79	0.82	0.53	0.70	0.79	0.84	0.88
			(B) Cross-sectional correlation									
0.5	0.5	Bc	0.81	0.84	0.86	0.86	0.87	0.85	0.86	0.87	0.88	0.88
		Per	0.83	0.86	0.87	0.87	0.88	0.87	0.88	0.88	0.89	0.88
		Per-t	0.82	0.85	0.87	0.87	0.88	0.86	0.86	0.87	0.89	0.89
	1	Bc	0.58	0.71	0.78	0.82	0.84	0.68	0.79	0.83	0.86	0.87
		Per	0.62	0.75	0.81	0.84	0.86	0.71	0.83	0.86	0.88	0.87
		Per-t	0.61	0.73	0.80	0.84	0.86	0.69	0.81	0.84	0.87	0.87
	1.5	Bc	0.45	0.60	0.67	0.75	0.78	0.48	0.64	0.73	0.78	0.82
		Per	0.48	0.64	0.70	0.78	0.81	0.52	0.69	0.78	0.83	0.85
		Per-t	0.47	0.63	0.69	0.77	0.81	0.51	0.66	0.76	0.81	0.84
0.9	0.5	Bc	0.62	0.73	0.78	0.81	0.83	0.75	0.83	0.85	0.87	0.88
		Per	0.73	0.85	0.89	0.91	0.92	0.87	0.93	0.93	0.93	0.93
		Per-t	0.62	0.76	0.81	0.85	0.86	0.73	0.82	0.87	0.88	0.89
	1	Bc	0.32	0.48	0.59	0.70	0.74	0.43	0.63	0.71	0.80	0.82
		Per	0.42	0.60	0.73	0.83	0.87	0.57	0.77	0.84	0.91	0.93
		Per-t	0.33	0.50	0.63	0.75	0.79	0.44	0.63	0.72	0.82	0.85
	1.5	Bc	0.21	0.36	0.46	0.60	0.65	0.24	0.42	0.56	0.65	0.71
		Per	0.29	0.46	0.58	0.72	0.78	0.36	0.57	0.71	0.81	0.87
		Per-t	0.23	0.39	0.50	0.65	0.71	0.26	0.44	0.60	0.71	0.78

Note: Coverage rates of three nominal 90% confidence intervals in 10,000 replications. Bc denotes the bootstrap bias corrected asymptotic confidence interval (11), Per denotes the percentile bootstrap confidence interval (12) and Per-t denotes the percentile-t equal-tailed bootstrap confidence interval (13). S/N denotes the signal-to-noise ratio.

Table 5: AR(1) Estimates of the US diffusion index

Series	$\tilde{\rho}$	Asymptotic Confidence interval	$\tilde{\rho}_{BC}$	Bootstrap confidence intervals Bc	Per	Per-t
			(A) Full sample ($N = 159$)			
1	0.66	(0.60, 0.71)	0.69	(0.64, 0.75)	(0.64, 0.76)	(0.64, 0.75)
			(B) Long subsample ($N = 53$)			
1	0.65	(0.60, 0.71)	0.74	(0.69, 0.80)	(0.68, 0.82)	(0.68, 0.80)
2	0.58	(0.52, 0.64)	0.66	(0.60, 0.72)	(0.59, 0.74)	(0.59, 0.72)
3	0.68	(0.63, 0.73)	0.78	(0.72, 0.83)	(0.71, 0.86)	(0.71, 0.83)
average	0.64	(0.58, 0.69)	0.73	(0.67, 0.79)	(0.66, 0.80)	(066., 0.78)
			(C) Short subsample ($N = 31$)			
1	0.57	(0.51, 0.63)	0.75	(0.69, 0.81)	(0.66, 0.84)	(0.65, 0.80)
2	0.83	(0.79, 0.87)	0.95	(0.91, 1.00)	(0.88, 1.06)	(0.88, 0.99)
3	0.63	(0.58, 0.69)	0.75	(0.69, 0.80)	(0.67, 0.83)	(0.67, 0.80)
4	0.55	(0.49, 0.61)	0.65	(0.58, 0.71)	(0.57, 0.73)	(0.57, 0.71)
5	0.54	(0.48, 0.60)	0.67	(0.61, 0.74)	(0.59, 0.77)	(0.59, 0.75)
average	0.62	(0.57, 0.68)	0.75	(0.70, 0.81)	(0.67, 0.84)	(0.67, 0.81)

Note: The sample period is from 1959:3 to 1998:12 ($T = 478$). $c = \sqrt{T}/N$ is 0.14, 0.41 and 0.71, respectively, for series A, B and C. The first confidence interval next to $\tilde{\rho}$ is the 90% asymptotic confidence interval (10). For the boostrap confidence intervals, Bc denotes the 90% bootstrap bias corrected asymptotic confidence interval (11), Per denotes the 90% percentile interval (12) and Per-t denotes the 90% percentile-t equal-tailed interval (13).

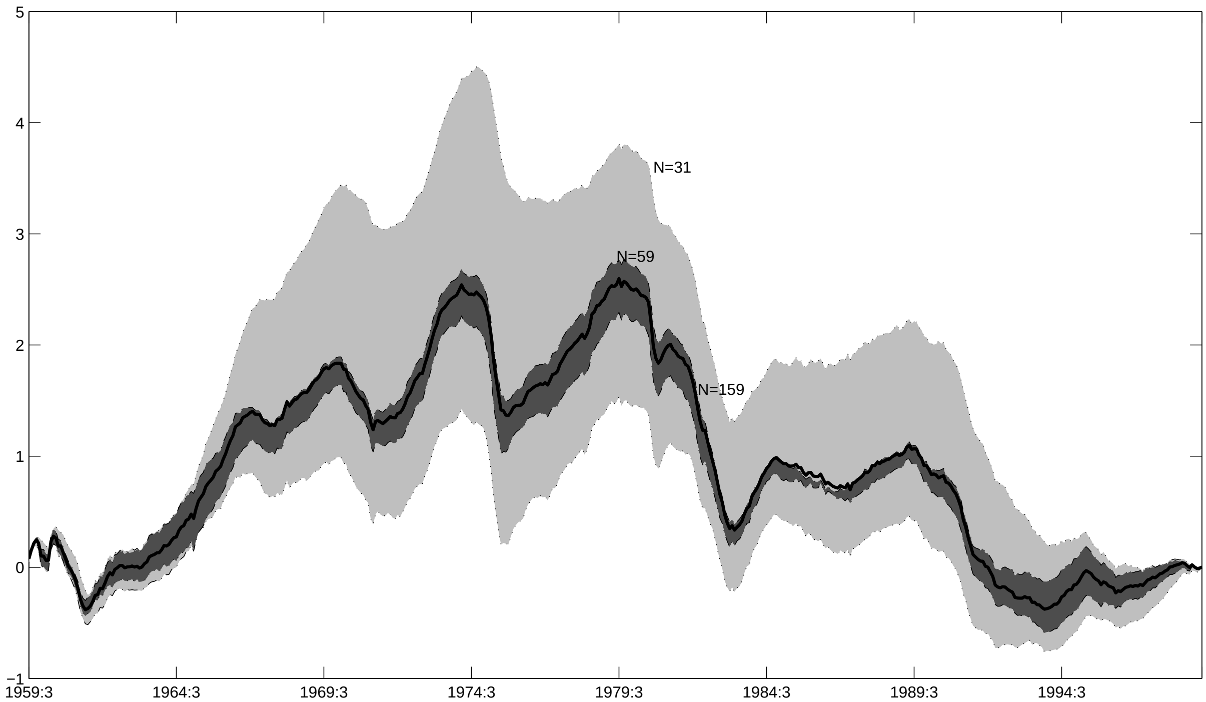

Figure 11: US Diffusion Index

CHAPTER III

INFORMATION HETEROGENEITY, HOUSING DYNAMICS AND THE BUSINESS CYCLE

Introduction

The recent financial crisis that started in the U.S. in December 2007 has demonstrated the importance of the housing sector in macroeconomic modeling. In response to the recession, a growing literature has tried to incorporate the housing sector into standard macroeconomic models to explain stylized facts in the housing market and the business cycle.[1] However, there are two facts that existing quantitative macroeconomic models have difficulty explaining: house prices are highly volatile and closely correlated with the business cycle, which is at odds with the evidence that rental prices are relatively stable and almost uncorrelated with the business cycle; and residential investment leads the business cycle while nonresidential investment moves contemporaneously with the business cycle.

The main goal of this paper is to present an alternative model to quantitatively explain these two facts. To incorporate the housing sector into the standard dynamic stochastic general equilibrium (DSGE) model, one usually assumes that firms need a collateral asset to secure their external financing as in Kiyotaki and Moore (1997), and specifies the collateral asset as houses, such as Iacoviello (2005), and Liu, Wang, and Zha (2011) et al. These types of models succeed in explaining either the close correlation between house prices and nonresidential investment or the close correlation between house prices and consumption, but fails in explaining the contrast between the high volatility of house prices and the low volatility of rental prices. Figure 12 illustrates the cyclical components of house prices and rental prices with the business cycle for the United States from 1975Q1 to 2010Q3[2]

[1] Iacoviello (2010) is a recent survey. A inexhaustive reading list should include Iacoviello (2005), Davis and Heathcote (2005, 2007), Piazzesi and Schneider (2009), Iacoviello and Neri (2010), Liu, Wang, and Zha (2011), Sterk (2010), Burnside, Eichenbaum, and Rebelo (2011), Caplin and Leahy (2011), Chaney, Sraer, and Thesmar (2012), Chatterjee and Eyigungor (2011), Favilukis, Ludvigson, and Nieuwerburgh (2011), Mian and Sufi (2011), Kiyotaki, Michaelides, and Nikolov (2011), and Rupert and Wasmer (2012).

[2] In this paper, we collect the data of output, consumption, residential investment, and nonresidential investment from the St. Louis Fed.

(all data are log-linearized and filtered using the Hodrick-Prescott filter). House prices are closely correlated with the business cycle and their correlation with U.S. GDP is around 0.52. In contrast, rental prices are almost uncorrelated with the business cycle and their correlation with U.S. GDP is less than 0.06. Furthermore, house prices are much more volatile than output and their standard deviation is around 1.55 times of the standard deviation of output. However, rental prices are much less volatile and their standard deviation is only 0.46 times of the standard deviation of output. To explain the difference between the volatility of house prices and the volatility of rental prices, in addition to incorporating financial frictions as in Liu, Wang, and Zha (2011), we further incorporate information frictions into the standard DSGE model, and demonstrate that information heterogeneity plays a key role in quantitative macroeconomic analysis of housing dynamics.

In the standard DSGE model with financial frictions, houses can be viewed as assets (see equation (20) in Liu, Wang, and Zha (2011)). If we define the rental prices as the marginal rate of substitution (MRS) between housing consumption and goods consumption, the asset pricing theory implies that house prices are determined by the discounted sum of future rents. With consumption smoothing, the model predicts that the volatility of house prices is much lower than the volatility of output (see Liu, Wang, and Zha (2011) for a detailed discussion). However, if households have heterogeneous information about the future average MRS between housing consumption and goods consumption, house prices will also be determined by households' expectations of other households' expectations of the future average MRS, households' expectations of other households' expectations of other households' expectations of the future average MRS, and so on. Therefore, higher-order expectations of the future average MRS play a potential role in determining the fluctuations of house prices. Our calibration exercise shows that information heterogeneity increases the relative volatility of house prices to output by more than 50% and explain the disconnect between house prices and the discounted sum of future rents compared with the full information case. However, our model still has a difficulty in predicting house prices having a higher volatility than output.

We assume households' information sets differ in two respects. First, households have dispersed information of the total factor productivity (TFP). Second, households have idiosyncratic information of the aggregate preferences on houses. When house prices rise, households are confused by whether this rise is driven by an improvement in TFP or an increase in the aggregate demand. Because of rational confusion, an improvement in TFP has an amplified effect on house prices[3]. Thus, information heterogeneity generates a higher volatility of house prices, and breaks down the close correlation between house prices and rental prices.

The other fact which standard macroeconomic models have difficulty in explaining is the lead-lag relationship between residential investment and nonresidential investment over the business cycle. Figure 13 displays the normalized cyclical components of residential and nonresidential investment over the business cycle for the United States from 1975Q1 to 2010Q3, and shows that residential investment leads the business cycle while nonresidential investment moves contemporaneously with the business cycle. The reason why standard real macroeconomic models have difficulty in explaining the lead-lag relationship is because nonresidential capital produces market consumption and investment goods, whereas residential capital produces only home consumption goods (e.g. Fisher, 2007). The asymmetry in how many goods to substitute away from residential capital provides a strong incentive to substitute away from residential capital toward nonresidential capital after a productivity shock. In our model, with incomplete information firms cannot fully observe the true TFP shocks, so the model generates a dampened response of nonresidential investment to TFP shocks. On the other side, since the amplified response of house prices mainly comes from the rising demand of real estate from households, the response of residential investment to TFP shocks is dampened, but to a smaller degree. In total, the correlation between lead residential investment and nonresidential investment increases, as does the correlation between lead residential investment and output. Our calibration shows that the correlation between lead residential investment and nonresidential investment increases from a negative value to a large positive value.

[3]The idea of rational confusion has long existed in the noisy rational expectation literature. For example, Bulow and Klemperer (1994) use this idea to explain the worldwide stock market crash of 1987. Bacchetta and van Wincoop (2006) claim that such rational confusion plays a key role in explaining the exchange rate disconnect puzzle and matching the evidence on micro trading activities.

The paper is related to several strands of the literature. First, it is related to the literature incorporating financial frictions into models of business cycles (see Gertler and Kiyotaki, 2010, for a survey). Within this strand, there is a large body of work that specifies houses as a collateral asset, and investigates frictions in the house market affecting the business cycle[1]. For example, Iacoviello (2005) introduces collateral constraints tied to home values into a standard monetary business cycle model and shows that houses contribute to the amplification and propagation of demand shocks. In terms of the labor market, Rupert and Wesmer (2012) incorporate frictions in housing mobility into a standard searching and matching model to investigate the difference of unemployment rates between the U.S. and Europe. Sterk (2011) studies the effect of the housing bust in 2007 on the unemployment rate of the recent financial crisis. However, these models either do not consider the disconnect between house prices and rental prices or have difficulty in explaining it. Liu, Wang, and Zha (2011) estimate a real business cycle model with land as a collateral asset in firms' credit constraints, and claim that a shock originated from households' preferences on houses is important in determining land prices and the business cycle. In their model, the housing demand shock explains more than 90% of the observed fluctuations of land prices, and other shocks make almost no contributions, which seems counterintuitive[4].

In this paper, we investigate information frictions in explaining the high volatility of house prices. Trading with information frictions in the housing market has been considered in the literature for a long time (see Himmelberg, Mayer, and Sinar, 2005, for a survey). For recent evidence, Piazzesi and Schneider (2009) propose a search model with transaction costs and show that a small portion of momentum trades generates a high volatility of house prices. Burnside, Eichenbaum, and Rebelo (2011) develop a model with heterogeneous expectations and show that changes in expectations can generate the boom-bust cycles in the housing market. However, these models are not in a micro-founded general equilibrium framework, and therefore are not suitable for a quantitative analysis of the interaction of information frictions and the housing dynamics over the business cycle. To the best of my knowledge, this paper is the first to introduce imperfect information into

[4]The other shocks include a patience shock, permanent and transitory shocks to neutral technology, permanent and transitory shocks to biased technology, a labor supply shock, and a collateral shock.

the standard DSGE model with a housing sector, and shows information heterogeneity has the potential to explain the aforementioned puzzles in both the housing market and the macroeconomy. Our paper is also the first one to introduce information frictions to explain the lead-lag relationship between residential investment and business investment. Previous literature investigating the lead-lag relationship includes Benhabib, Rogerson, and Wright (1991), Greenwood and Hercowitz (1991), Chang (2000), Gomme, Kydland, and Rupert (2001) Davis and Heathcote (2005), Fisher (2007), et al.

Finally, this paper also contributes to the growing interests in investigating imperfect information in macroeconomics[5]. In their seminal work, Phelps (1970) and Lucas (1972) demonstrate that the dispersion of information can help nominal shocks generate fluctuations in real variables. Recently, Morris and Shin (2002) investigate strategic interactions in a global game framework; Mankiw and Reis (2002) consider the case that agents update their information sets sporadically; and Sims (2003) formalizes the idea of information frictions by assuming limited capacity for processing information. Our work is more closely related with La'O (2010), which also studies the interaction of information frictions with financial frictions. However, our work differs from La'O's work in three aspects. First, our work directly investigates the information frictions in the housing market and the spillover effects from the housing market to the business cycle. Second, we build our model in a dynamic stochastic general equilibrium framework and thus can quantitatively evaluate the contribution of information heterogeneity to both the housing market and the business cycle. Finally, La'O's work focuses on how the interactions of financial frictions and information frictions affect noise shocks as an independent source of the business cycle fluctuations.

The remainder of the paper is organized as follows. Section 2 provides empirical evidence about the two facts in the housing market and the business cycle. Section 3 introduces the model with both financial frictions and information frictions. Section 4 discusses the implications of our model regarding house prices, residential investment, and

[5]Mankiw and Reis (2010) provide a recent survey. An inexhaustive list includes Phelps (1970), Lucas (1972), Townsend (1983), Mankiw and Reis (2002), Morris and Shin (2002), Sims (2003), Woodford (2003), Bacchetta and van Wincoop (2006), Nimark (2008), Lorenzoni (2009), Machowiak and Wiederholt (2009), Angeletos and La'O (2010), Graham and Wright (2010), and Guo and Shintani (2011), Crucini, Shintani, and Tsuruga (2010, 2012).

nonresidential investment over the business cycle. Section 5 presents additional evidence from the survey data. Finally, section 6 concludes.

Empirical Motivation

In this section, we empirically present the two facts that existing macroeconomic models have difficulty in explaining: the disconnect between house prices and the discounted sum of future rents; and the lead-lag relationship between residential investment and nonresidential investment. To investigate the disconnect between house prices and the discounted sum of future rents, we consider the user-cost approach, an approach commonly used in the literature (see Mayer, 2011, for a survey)[6]. This approach takes the simple non-arbitrage condition that the rent-price ratio should be equal to the user cost of housing, which is the sum of the after-tax equivalent-risk opportunity cost of capital and the expectation of future house prices appreciation excluding maintenance cost. This implies that the following relationship holds at each point in time:

$$\frac{R_t}{P_t} = \alpha_0 + \alpha_1 i_t + \alpha_2 \frac{(1-\delta_h)P_{t+1} - P_t}{P_t} + \varepsilon_t, \tag{III.1}$$

where R_t is the rental price for a representative home for one year at time t, P_t is the corresponding purchase price of the same home, i_t is the opportunity cost of capital, δ_h is the home depreciation rate, and ε_t is white noise.

We collect house prices and rent data from 1960Q1 to 2010Q3 from the Federal Housing Finance Agency (FHFA) home price index, and use the data with the same period from the Case-Shiller-Weiss (CSW) home price index as a robustness check. The FHFA series is well-known for its broad geographic coverage, but it covers only conventional mortgages. On the other hand, the CSW series covers both conventional and unconventional mortgages (see Davis and Heathcote (2007) for a detailed description of the data set). By assuming that the risk premium of house price fluctuations is constant, we take the federal funds rate to approximate the opportunity cost of capital. To introduce maintenance

[6] There are three alternative approaches commonly used in the literature: the user-cost methodology which compares the present discounted value of future rents with house prices; the construction-cost approach that compares the cost of constructiong a new home with house prices; and the affodability approach which compares the ability of potential buyers of the house with house prices.

costs, we assume that houses depreciate at a constant rate $\delta_h = 0.01$ as in Iacoviello and Neri (2010). Table 6 presents the regression results of equation (III.1). The results show that appreciation in house prices has almost no explanatory power in the fluctuations of the rent-price ratio. One percent increases in house prices predict around 0.09 increases in rent-price ratio for the FHFA series, and around 0.02 increases for the CSW series. The null hypothesis $\alpha_2 = 1$ is rejected at any significance level for both of the two data sets. Thus, the regression results confirm the disconnect between house prices and the discounted sum of future rents.

The second fact that we want to investigate is the lead-lag relationship between residential investment and nonresidential investment over the business cycle. The literature in home production has demonstrated that residential investment leads the business cycle and nonresidential investment lags the business cycle for the U.S. economy (see Davis, 2010, for a survey). In sharp contrast, Kydland, Rupert, and Šustek (2012) empirically show that the lead-lag relationship in the developed countries only holds for the two Western-Hemishpere countries: USA and Canada, and in other developed economies there is no such a clear feature of the lead-lag relationship between either residential investment or nonresidential investment and the business cycle. We reconsider the fact and calculate the correlations among the lead (lag) residential investment, the lead (lag) business investment, and the lead (lag) output for the following countries and periods: Austria (1988Q1-2012Q2), Finland (1975Q1-2012Q2), France (1978Q1-2012Q2), Netherlands (1988Q1-2012Q2), the U.K. (1970Q1-2012Q2), the EU (1988Q1-2012Q2), Australia (1959Q3-2012Q2), Canada (1981Q1-2012Q2), and the U.S. (1960Q1-2012Q2)[7]. All the data are logged and Hodrick-Prescott filtered. In Table 7, our main results confirm the leading (lagged) role of residential (nonresidential) investment over the business cycle in the U.S. and Canada. In other developed countries, there is no clear order among the second moments except Finland, which also shares this feature to some extent. One interesting thing in our calculation is that if we aggregate the five countries in the Europe together, the aggregate will also somewhat perform like the U.S. and Canada.

[7]The EU is aggregated by the five following countries: Austria, Finland, France, Netherlands, and the U.K.. We collect the data for the European countries from the Eurostat, for Canada from the OECD, for Australia from Australian Bureau of Statistics, and for the U.S. from the St. Louis Fed.

To further investigate the causality effect between residential and nonresidential investment, we conduct a bivariate vector autoregression (VAR) with a Granger-causality test for these two types of investment. To apply the Granger-causality test, we first test whether the two series have a unit-root process by the Dickey-Fuller test. If the two series are of $I(1)$, we further test whether the two are cointegrated. If we cannot detect a cointegration relationship between the two series, the following formulation is used in testing the null hypotheses:

$$\begin{aligned} \Delta I_t^s &= \alpha_0 + \sum_{i=1}^{k} \alpha_{1i} \Delta I_{t-i}^s + \sum_{\iota=1}^{k} \alpha_{2i} \Delta I_{t-i} + \varepsilon_{1t} \\ \Delta I_t &= \beta_0 + \sum_{i=1}^{k} \beta_{1i} \Delta I_{t-i}^s + \sum_{\iota=1}^{k} \beta_{2i} \Delta I_{t-i} + \varepsilon_{2t}. \end{aligned} \tag{III.2}$$

Failing to reject the H_0: $\alpha_{21} = \alpha_{22} = ... = \alpha_{2k} = 0$ implies that nonresidential investment does not Granger cause residential investment. Likewise, failing to reject H_0: $\beta_{12} = \beta_{12} = ... = \beta_{1k} = 0$ implies that residential investment does not Granger cause nonresidential investment. If the series are cointegrated, we need to incorporate an error correction term in testing the null hypotheses:

$$\begin{aligned} \Delta I_t^s &= \alpha_0 + \delta_1 (I_t^s - \lambda I_t) + \sum_{i=1}^{k} \alpha_{1i} \Delta I_{t-i}^s + \sum_{\iota=1}^{k} \alpha_{2i} \Delta I_{t-i} + \varepsilon_{1t} \\ \Delta I_t &= \beta_0 + \delta_2 (I_t^s - \lambda I_t) + \sum_{i=1}^{k} \beta_{1i} \Delta I_{t-i}^s + \sum_{\iota=1}^{k} \beta_{2i} \Delta I_{t-i} + \varepsilon_{2t}, \end{aligned} \tag{III.3}$$

in which δ_1 and δ_2 denote speeds of adjustment. Failing to reject the H_0: $\alpha_{21} = \alpha_{22} = ... = \alpha_{2k} = 0$ and $\delta_1 = 0$ implies that nonresidential investment does not Granger cause residential investment. Likewise, failing to reject H_0: $\beta_{12} = \beta_{12} = ... = \beta_{1k} = 0$ and $\delta_2 = 0$ implies that residential investment does not Granger cause nonresidential investment.

The data we use in testing equation (III.2) or (III.3) are the same as in Table 7. However, we conduct the Granger-causality test for the period from 1984Q1 to 2005Q4 in the U.S. as a robustness check to avoid the potential problem of structural changes, since this period is well-known for its low volatility of the business cycle in contrast to other periods. The lag parameter k is selected by the Akaike information criterion (AIC). Table 8 shows the fact that in the U.S. and Canada residential investment Granger causes nonresidential investment and nonresidential investment does not Granger cause residential investment. This fact is very clear in Canada, but in the U.S., we can reject the null hypothesis that residential investment does not Granger cause nonresidential investment at any significance level, whereas we cannot reject the null hypothesis that nonresidential investment does not Granger cause residential investment for the period from 1984Q1 to 2005Q4 at 5% significance level, and for the period from 1960Q1 to 2010Q3 at 1% significance level. In other developed countries, there is no such feature similar as in the U.S. and Canada, except in Australia and the U.K. In contrast to the lead-lag relationship that the European aggregate shares with the U.S. and Canada, we cannot see such a similarity for the Granger causality of the two types of investment between the two regions.

The Basic Model

To quantitatively explain the two facts in a dynamic general equilibrium framework, we build our model in the style of Liu, Wang, and Zha (2011) with real estate production and information heterogeneity. The model in Liu, Wang, and Zha (2011) is a variant of standard real business cycle models that include a feature of credit frictions (Kiyotaki and Moore, 1997). We add a real estate production sector into the model, and assume agents are endowed with heterogeneous information instead of perfect information. Following Iacoviello (2005), Iacoviello and Neri (2010), Kiyotaki, Michaelides, and Nikolov (2011), and Liu, Wang and Zha (2011), we assume two types of agents in the economy: a representative impatient entrepreneur and a continuum of patient households. The representative entrepreneur owns two types of firms: a continuum of residential firms and a continuum of nonresidential firms. The whole economy is segmented geographically and endowed with a continuum of islands. Each island $i \in [0, 1]$ contains one residential firm, one nonresidential firm, and one household. The residential firm hires labors from the household, and accumulates residential structures to build houses. The nonresidential firm also hires labor from the household, accumulates nonresidential capital, and combines with real estate input to produce final goods. The household provides labor services, saves for next period, and consumes final goods and housing services. The final goods can be used to finance residential investment and nonresidential investment, whereas real estate can only be used for residence.

Entrepreneurs

The representative entrepreneur owns a continuum of residential firms and a continuum of nonresidential firms. On each island resides one residential firm and one nonresidential firm. The residential firm and the nonresidential firm maximize their expected profits and return the profits to the entrepreneur. The nonresidential firm i takes a Cobb-Douglas constant-to-scale technology that uses labor, capital, and housing as input, according to

$$Y_{it} = K_{it}^{\mu_k} (A_t A_{it} N_{it}^{'k})^{v_k} H_{it}^{'1-\mu_k - v_k},$$

where Y_{it} is the output, A_t is the aggregate technology level, A_{it} is the firm-specific technology level, K_{it} is capital produced at the end of last period, $H_{it}^{'}$ is the real estate input, and $N_{it}^{'k}$ is the labor input in the nonresidential market. μ_k and $1 - \mu_k - v_k$ measure output share of capital and real estate respectively. The residential firm i also takes a Cobb-Douglas constant-to-scale technology that uses labor, residential structures, and land as input, according to

$$H_{it}^0 = S_{it}^{\mu_h} (A_t A_{it} N_{it}^{'h})^{v_h} L_{it}^{1-\mu_h - v_h},$$

where H_{it}^0 is newly built housing, S_{it} are residential structures, L_{it} is the land endowment, and $N_{it}^{'h}$ is the labor input in the residential market. μ_h and $1-\mu_h-v_h$ measure output share of residential structures and land respectively. The representative entrepreneur borrows B_{it} from household i in the asset market, invests I_{it} in the nonresidential capital market and I_{it}^s in the residential structure market, produces consumption final goods by purchasing real estate input $\Delta(H_{it})$ and hiring workers N_{it}^k, constructs houses by using the land endowment L_{it}, the labor input N_{it}^h, and the residential structure S_{it}, and consumes $C_t^{'}$ to maximize its expected utility according to

$$\max E \sum_{t=0}^{\infty} \beta'^t \frac{C_t'^{1-\gamma}}{1-\gamma}$$

$$s.t. \quad C_t' + \int_I [(N_{it}'^k + N_{it}'^h)W_{it} - \frac{\pi_{it}B_{it+1}}{R_{it}} + P_t(H_{it}' - (1-\delta_h)H_{it-1}') + I_{it} + I_{it}^s - K_{it}^{\mu_k}(A_{it}A_tN_{it}'^k)^{v_k}H_{it}'^{1-\mu_k-v_k} - P_tS_{it}^{\mu_h}(A_{it}A_tN_{it}'^h)^{v_h}L_{it}^{1-\mu_h-v_h} + B_{it}]di = 0$$

where β' is the discount factor of the entrepreneur, γ measures the relative risk aversion, W_{it} is the wage that the entrepreneur pays for workers from the household i, π_{it} is the island-specific bond-holding shock, R_{it} is the island-specific interest rate, δ_h is the discount factor of houses, and P_t is house prices. The island-specific bond-holding shock π_{it} serves one and only one role, to slow down the learning of agents in island i from the bond market. To replace the assumption of the island-specific bond-holding shock, one can introduce another aggregate shock, such as a patience shock to the entrepreneur, to serve a similar role. For simplicity, we do not consider adding another aggregate shock. As in Kiyotaki and Moore (1997), we assume the entrepreneur needs collateral to secure its borrowings

$$B_{it+1} \leq mE_{it}(P_{t+1}H_{it}'), \tag{III.4}$$

where m indicates that if borrowers repudiate their debt obligations, lenders can liquidate the borrowers' real estate assets but have to pay a proportional transaction cost $(1-m)P_{t+1}H_{it}$. Allowing capital as an additional collateral asset will amplify the effect of credit constraints since the entrepreneur will be more leveraged. We will discuss this later as a robustness check. Nonresidential capital accumulation follows the law of motion

$$K_{it+1} = (1-\delta_k)K_{it} + \Phi_1(\frac{I_{it}}{K_{it}})K_{it},$$

and similarly, residential structure accumulation follows the law of motion

$$S_{it+1} = (1-\delta_s)S_{it} + \Phi_2(\frac{I_{it}^s}{S_{it}})S_{it},$$

where δ_k and δ_s are the discount factors of nonresidential capital and of residential structures respectively, and $\Phi_1(\cdot)$ and $\Phi_2(\cdot)$ denote the adjustment cost functions of nonresidential capital and of residential structures respectively.

Households

We assume one household resides on each island i. The household i consumes the final goods, utilitzes the housing services, and provides labor services to the residential firm and the nonresidential firm. The household maximizes its expected discounted sum of utility conditional on its own information set Ω_{it} by

$$\max E_i \sum_{t=0}^{\infty} \beta^t [\ln C_{it} + \chi_0 \chi_t \chi_{it} \ln H_{it} - \psi N_{it}],$$

where C_{it} is goods consumption, H_{it} is the housing consumption, N_{it} is the labor services provided by the household, β is the discount factor, χ_t and χ_{it} denote the aggregate and the idiosyncratic housing preference shocks respectively, and χ_0 and ψ are constant parameters. We assume households' discount factor $\beta > \beta^{'}$, which indicates that households are more patient than the entrepreneur and inclined to save. The household i's budget constraint is given by

$$C_{it} + P_t(H_{it} - (1-\delta_h)H_{it-1}) + \frac{B_{it+1}}{R_{it}} - W_{it}N_{it} - B_{it} = 0.$$

Market Clearing

The economy has four markets in total: goods, labor, bond and housing. To clear the goods market, we have

$$C_t^{'} + \int_I [C_{it} + I_{it} + I_{it}^s] di = \int_I Y_{it} di.$$

We assume labor is immobile across islands, so in island i we have

$$N_{it} = N_{it}^{\prime k} + N_{it}^{\prime h}.$$

To clear the bond market, we have

$$B_{it} + B_{it}^{\prime} = 0.$$

Finally, to clear the housing market, we have

$$\int_I [H_{it} + H'_{it} - (1-\delta_h)(H_{it-1} + H'_{it-1})]di = \int_I H^0_{it} di.$$

Shocks

Our model includes two aggregate shocks and three idiosyncratic shocks. The aggregate shocks follow $AR(1)$ processes in logs,

$$\begin{aligned} \log A_t &= \rho_a \log A_{t-1} + u^a_t \\ \log \chi_t &= \rho_\chi \log \chi_{t-1} + u^\chi_t, \end{aligned}$$

where $u^a_t \sim N(1, \sigma^2_a)$, and $u^\chi_t \sim N(1, \sigma^2_\chi)$. The idiosyncratic shocks also follow the $AR(1)$ processes in logs

$$\begin{aligned} \log A_{it} &= \rho_{ai} \log A_{it-1} + u^a_{it} \\ \log \chi_{it} &= \rho_{\chi i} \log \chi_{it-1} + u^\chi_{it} \\ \log \pi_{it} &= \rho_{\pi i} \log \pi_{it-1} + u^\pi_{it}, \end{aligned}$$

where $u^a_{it} \sim N(1, \sigma^2_{ai})$, $u^\pi_{it} \sim N(1, \sigma^2_{\pi i})$, and $u^\chi_{it} \sim N(1, \sigma^2_{\chi i})$. We also assume the law of large numbers applies for the distribution of all the three types of idiosyncratic shocks, as is common in the literature.

The Information Structure and the Equilibrium

At each period t, the representative entrepreneur has full information. However, the final goods firm i, the real estate firm i, and the household i can only obtain information from their market activities: idiosyncratic preferences series on houses $\{\chi_{t-s}\chi_{it-s}\}^\infty_{s=0}$, wage series $\{W_{it-s}\}^\infty_{s=0}$, interest rate series $\{R_{it-s}\}^\infty_{s=0}$, and house prices series $\{P_{t-s}\}^\infty_{s=0}$. The information set for agents in island i is denoted as

$$\Omega_{it} = \{\{\chi_{t-s}\chi_{it-s}\}^\infty_{s=0}, \{W_{t-s}\}^\infty_{s=0}, \{R_{it-s}\}^\infty_{s=0}, \{P_{t-s}\}^\infty_{s=0}\}.$$

We assume the parameters and the model structure are common knowledge, which indicates our model is in line with the framework of noisy rational expectation models.

The equilibrium is defined as follows:

1. Given prices and information restrictions, the allocations solve the utility maximization problem of the entrepreneur and of the household i and the profit maximization problem of the final goods firm i and the real estate firm i.

2. All markets clear, and $\{P_{t-s}, R_{it-s}, W_{it-s}\}_{s=0}^{\infty}$ are the market clearing house prices, interest rates of bonds, and wages, respectively.

Economic Implications

In our model, we assume residential firms, nonresidential firms, and households do not have full information about the economic fundamentals and differ in their information sets for different islands. Instead of an ad hoc assumption of perfect information, we assume agents can only extract information about the true economic fundamentals from their idiosyncratic market activities. With information heterogeneity, agents make their decisions based on their forecasts of not just true economic fundamentals but also forecasts of other agents' actions, forecasts of other agents' forecasts of other agents' actions, etc. In this section, we show that higher-order beliefs play a potential explanatory role in the two facts: the disconnect between house prices and rental prices, and the lead-lag relationship between residential investment and nonresidential investment over the business cycle.

Solving a dynamic general model with dispersed information requires dealing with the well-known "infinite regress" problem (Townsend, 1983), since higher-order beliefs are crucial for the decisions of agents and depend on the entire history of shocks. The literature has solved this type of model by either truncating the dependence of equilibrium actions on higher order beliefs (Nimark, 2008) or by assuming private information is revealed after an ad hoc period T (Lorenzoni, 2009). We take the second approach, and assume that after $T =$ 30 periods all of the shocks are observed by agents across islands. The choice of T is based on two considerations: saving computational time and not affecting the results significantly if the value of T is increased. We assume all the shocks are relatively small in magnitude, so the inequality in (III.4) is always binding. Without the problem of occasional binding, one can solve the model by log-linearizing around the steady state. After log-linearization, we solve the linear equations by combining Sims's (2001) method and the guess-verification approach. In the model economy, agents on island i are integrated into two aggregate markets: the final goods market and the housing market. Therefore, decisions of agents are affected by two aggregate variables: consumption of the representative entrepreneur $C_t^{'}$ and house prices P_t. In the first step, we guess the aggregate variables, $C_t^{'}$ and P_t, to be linear functions of aggregate shocks; in the second step, we plug these two variables into the equations and solve the equations using Sims's (2001) method; in the third step, we update expectation operators of agents on island i by their information set Ω_{it}; finally, we verify the guess of linear functions of $C_t^{'}$ and P_t by minimizing their distance with the updated variables $C_t^{'}$ and P_t. The appendix provides a detailed description of the method.

To calibrate the model, we choose the parameters commonly used in the literature (e.g. Iacoviello and Neri, 2010). β and $\beta^{'}$ are set to 0.9925 and 0.97 respectively. Relative risk aversion, γ, is set to 2. The housing preference parameter χ_0 is set to 0.1 and the disutility on labor ψ is set to 1. The entrepreneurial "loan-to-value ratio" m is set to 0.89 to match the empirical debt to GDP ratio in the U.S. data. The nonresidential capital share in the output production function is set to $\mu_k = 0.63$, and the house share is set to $1-\mu_k-v_k = 0.05$. For the real estate production function, the share of residential structures is set to $\mu_h = 0.1$, and the share of land is set to $1-\mu_h-v_h = 0.1$. The discount factors for houses, residential structures, and nonresidential capital are set to $\delta_h = 0.01$, $\delta_s = 0.25$, and $\delta_k = 0.03$ respectively. These three discount factors, combined with the capital share in the goods production function and real estate production function, imply that nonresidential investment accounts for around 30% of the total output, residential investment accounts for about 6% of the total output, and the value of house stocks is about 1.80 time the total output. The solution method does not require us to specify the functional form of Φ_1 and Φ_2, but needs us to set the values of Φ_1, Φ_1', Φ_1'', Φ_2, Φ_2', and Φ_2'' in the steady state. We choose $\Phi_1(\frac{I}{K}) = \delta_k$, $\Phi_1'(\frac{I}{K}) = 1$, $\Phi_2(\frac{I^s}{S}) = \delta_s$, and $\Phi_1'(\frac{I^s}{S}) = 1$, so that the model with adjustment costs has the same steady state as the model without adjustment costs. We set the second-order derivative of the adjustment cost function of residential investment $\Phi_2''(\frac{I^s}{S}) = -2.5$, the same as that of nonresidential investment $\Phi_1''(\frac{I}{K}) = -2.5$. The later is chosen as in Christiano, Eichenbaum, and Evans (2005)[8].

There are two aggregate and three idiosyncratic $AR(1)$ shock processes in total. Two parameters are crucial for the $AR(1)$ processes: the persistence and the variance of the shocks. The persistence and the variance of the shocks affect the response of business cycle variables in two different ways: first, the shocks to the model are directly affected; second, the precision of agents' information and agents' information updating process are altered. For the aggregate technology shock process, we assume a persistent shock process and set $\rho_a = 0.95$ as in Fisher (2005). Similarly, the autocorrelation in the aggregate housing preference shock is assumed to be $\rho_\chi = 0.95$. We choose $\sigma_a^2 = 0.00984^2$ to match the volatility of output, and $\sigma_\chi^2 = \frac{1}{10^2}\sigma_a^2$ to weaken the effect of housing preference shocks and focus on technology shocks as a main driving force of the business cycle fluctuations[9]. Since our interest is in the role of information heterogeneity in matching aggregate business cycle variables, we choose the persistence and the variance of idiosyncratic shocks to maximize the effect of information heterogeneity on house prices, and ignore the empirical micro-level cross-sectional facts. For the idiosyncratic bond-specific shock processes, we set $\rho_{\pi i} = 0$ and $\sigma_{\pi i}^2 = \infty$ for one and only one reason: to screen the information contained by the real interest rate. For the idiosyncratic technology shock and the idiosyncratic housing preference shock, we set $\rho_{ai} = 0.001$, $\rho_{\pi i} = 0.001$, $\sigma_{ai}^2 = 100^2\sigma_a^2$ and $\sigma_{\chi i}^2 = 100^2\sigma_a^2$. The high magnitude of idiosyncratic shocks implies that agents extract information mainly from house prices instead of idiosyncratic variables, such as island-specific wages and island-specific technology shocks. This assumption of a large magnitude of idiosyncratic shocks relative to aggregate shocks has been used in the literature (Maćkowiak and Wiederholt, 2009).

[8]In the literature, one usually pins down the parameters $\Phi_1''(\frac{I}{S})$ and $\Phi_1''(\frac{I^s}{S})$ by matching the volatility of nonresidential investment and residential investment in the data, Unfortunately, our solving procedure can find a convergence point only for certain ranges of parameters values. Of course, this is left for future work.

[9]In our model, a low magnitude of the housing preference shocks is enough to confuse the rational agents. Nimark (2008) makes a similar assumption that the variance of the transitory labor supply shock is $\frac{1}{100}$ of other aggregate shocks, such as the technology shock.

To evaluate the model's performance, we turn on all the shocks and simulate the model $1,000$ times with 142 periods in each simulation. The simulated data are then filtered with the Hodrick-Prescott filter. The average second moments of all the simulations and their empirical counterparts are reported in Table 9. Our model confirms the main arguments in Liu, Wang, and Zha (2011) that collateral constraints in nonresidential investment play a key role in explaining the close correlation between house prices and other business cycle variables. All of the correlations between house prices and other business cycle variables for the simulated data are well above their empirical counterparts. In comparison with the model with full information, two facts stand out for the model with heterogeneous information: first, information heterogeneity amplifies the response of business cycle variables to technology shocks[10]; second, the correlation between lead residential investment and nonresidential investment increases significantly from a negative value to a large positive value, and exceeds the correlation between lag residential investment and nonresidential investment. Similarly, the correlation between lead residential investment and output increases significantly from a small positive value to a large positive value, and exceeds the correlation between lag residential investment and output.

What drives house prices fluctuations?

Table 9 shows that information heterogeneity amplifies the response of business cycle variables to technology shocks, especially for house prices, whose standard deviation in the model with heterogeneous information is about twice the standard deviation in the model with full information. In contrast, the standard deviation of goods consumption increases slightly. These two together indicate that our model might be able to explain the puzzle of the disconnect between house prices and rental prices, since the later is closely correlated with final goods consumption. As discussed in Liu, Wang, and Zha (2011), the main reason why standard DSGE models with a housing sector cannot predict a high

[10]Since the standard deviation of housing preference shocks is one-tenth of the standard deviation of technology shocks, the role of housing preference shocks in our calibration is limited.

volatility of house prices can be illustrated by the Euler equation of households,

$$P_t = \beta(1-\delta_h)E_{it}\frac{C_{it}}{C_{it+1}}P_{t+1} + \frac{\chi_0\chi_t\chi_{it}C_{it}}{H_{it}}.$$

If we define rental prices as the marginal rate of substitutions between goods consumption and housing service consumption as

$$R_{it}^h = \frac{\chi_0\chi_t\chi_{it}C_{it}}{H_{it}},$$

house prices can be expressed as

$$P_t = \beta(1-\delta_h)\int_I E_{it}\frac{C_{it}}{C_{it+1}}P_{t+1}di + R_t^h, \tag{III.5}$$

where $R_t^h = \int_I R_{it}^h di$ denotes the aggregate rental prices. We further write house prices recursively,

$$P_t = \beta(1-\delta_h)\bar{E}_t\frac{C_{it}}{C_{it+1}}P_{t+1} + R_t^h = \sum_{k=0}^{\infty}\beta^k(1-\delta_h)^k\bar{E}_t^{(k)}R_{t+k}^h\frac{C_t}{C_{t+1+k}}.$$

where $\bar{E}_t^0(P_t) = P_t$, $\bar{E}_t^1(P_{t+1}) = \bar{E}_t(P_{t+1})$, and higher-order expectations are defined as,

$$\bar{E}_t^k(P_{t+k}) = \bar{E}_t\bar{E}_{t+1}\cdots\bar{E}_{t+k}(P_{t+k}).$$

Therefore, house prices at time t depend on rental prices at time t, the average expectation at time t of rental prices at time $t+1$, the average expectation at time t of the average expectation at time $t+1$ of rental prices at time $t+2$, etc. In the case of complete information, the average expectation at t of the average expectation at $t+1$ of rental prices at $t+2$ coincides with the average expectation at t of the average expectation of rental prices at $t+2$, i.e. $\bar{E}_t\bar{E}_{t+1}\cdots\bar{E}_{t+k}(P_{t+k}) = \bar{E}_t(P_{t+k})$, and therefore equation (III.5) collapses to

$$P_t = \sum_{k=0}^{\infty}\beta^k(1-\delta_h)^kE_tR_{t+k}^h\frac{C_t}{C_{t+1+k}}.$$

Since households smoothly allocate their consumption period by period, the model with full information fails to predict a high volatility of house prices. However, in the case of imperfect information, equation $\bar{E}_t\bar{E}_{t+1}\cdots\bar{E}_{t+k}(P_{t+k}) = \bar{E}_t(P_{t+k})$ does not hold. In other words, even though rental prices are relatively stable, house prices might still be volatile since house prices are also determined by higher-order expectations of future rental prices. Figure 14 displays the response of house prices to one positive standard deviation of technology shocks in the models with full information and in the model with heterogeneous information. Information heterogeneity initially dampens the technology shocks, but amplifies and propagates the technology shocks after three quarters. Unfortunately, our model still fails to generate a higher volatility of house prices than output. To illustrate how information heterogeneity affects house prices, we plot the average expectation of next-period house prices for both the full information case and the heterogeneous information case in Figure 15, since equation (III.5) shows that it is crucial in determining house prices in this period. The figure displays that the model with heterogeneous information is accompanied by higher average expectations of house prices.

To rigorously prove that information heterogeneity can explain the disconnect between house prices and rental prices, we test the user-cost equation as in (III.1) using the simulated data. The results in Table 10 show that the null hypothesis $\alpha_2 = 1$ cannot be rejected by the model with full information, but is rejected by the model with heterogeneous information at 5% significance level. In sum, even though the model with heterogeneous information cannot predict house prices having a higher volatility than output, it explains the disconnect puzzle between house prices and rental prices to some level.

Implications for Investment Dynamics

The other prediction of our model is the lead-lag relationship among nonresidential investment, residential investment, and output. Empirical studies have documented that residential investment leads the business cycle, but nonresidential investment lags the business cycle, and the two types of investment are positively correlated with each other (see Gangopadhyay and Hatchondo, 2009, for a survey). However, standard real business cycle models with home production predict the opposite and even a large negative value for the correlation between the contemporaneous residential investment and nonresidential investment. To match the data, several different channels have been emphasized in the literature, including adjustment costs in capital accumulation (Chang, 2000), time-to-build in nonresidential investment (Gomme, Kydland, and Rupert, 2001), multiple-market sectors (Davis and Heathcote, 2005), and a direct role for household capital as an input in market production (Fisher, 2007). In this paper, we highlight the information channel and show that the presence of information heterogeneity has a potential to explain the lead-lag relationship between residential investment and nonresidential investment.

As emphasized by Fisher (2007), real business cycle models with home production can predict the lead-lag relationship between residential investment and nonresidential investment, if home product enters the production function of market goods with a reasonable share. In our model, real estate enters the production function of final goods in two different ways: first, it directly enters the production function with a share of output equal to $1 - \mu_k - v_k = 0.05$; second, it serves as collateral for nonresidential investment. Since the share in our model is lower than the share of 0.14 in Fisher (2007), our model with full information cannot explain the lead-lag relationship, but it does predict a positive correlation of 0.69 between the contemporaneous residential investment and nonresidential investment as shown in the panel B of Table 9. The panel also shows information heterogeneity plays a key role in generating the positive correlation between lead residential investment and nonresidential investment. When there is no information frictions, the model predicts a negative correlation of -0.04, which is much less than the correlation between lead nonresidential investment and residential investment of 0.58. In contrast, when there is information heterogeneity, the correlation between lead residential investment and nonresidential investment increases to a significantly positive value 0.51, larger than the correlation of 0.38 between the lead nonresidential investment and residential investment. However, our model still produces a larger correlation between the contemporaneous residential investment and nonresidential investment, which is at odds with the data.

In the standard real business cycle model with home production, firms increase their production and nonresidential investment immediately in response to TFP shocks. Whereas real estate firms increase residential investment gradually. Therefore, the model predicts a negative correlation between lead residential investment and nonresidential investment. In the model with information heterogeneity, both residential firms and nonresidential firms are partially informed about the size of TFP shocks, and therefore both firms postpone their investment in response to TFP shocks. However, if the amplified house prices are mainly caused by rising demand from households, real estate firms will have a stronger incentive to increase residential investment in response to TFP shocks since the marginal revenue of real estate production increases. As shown in last subsection, the main reason the response of house prices is amplified is the breakdown in households' Euler equation (III.5). In our calibration, we find aggregate housing demand from households $H_t = \int_I H_{it} di$ decreases by much less in the model with information heterogeneity compared with the model with full information. Accordingly, residential investment will decrease by much less, and the correlation between lead residential investment and nonresidential investment increases. With the delayed response of nonresidential investment, our model predicts a hump-shaped response of output to one standard deviation of TFP shocks. In the case of imperfect information, the response of output initially increases at a slow speed and peaks in several periods. The hump-shaped response of output confirms the finding in Nimark (2008) that imperfect information provides a potential explanation for the contrast between a positive autocorrelation of output in the data and a negative autocorrelation of output in the real business cycle theory (Cogley and Nason, 1995). The one-period-lag autocorrelation increases from -0.10 to 0.04, although not significantly.

Empirical Evidence from Survey Data

A difficulty in the literature of imperfect information is that it is hard to provide empirical evidence to test the model. A prediction of our model is that if we define expectation errors of real variables as the difference between the average expectation of real variables and the corresponding realized variables, the expectation errors should be correlated with the business cycle. For instance, the model predicts that the forecast errors of output are positively correlated with the business cycle in response to TFP shocks with a correlation of 0.052, since firms are partially informed about the shocks and agents' expectations of output tend to underreact. As other variables, such as house prices, are also positively responded to TFP shocks, if one identifies an independent shock in the expectation errors of output, a vector autoregression (VAR) should perform as this shock positively causes other real variables, such house prices, output and investment.

To confirm this prediction, we run a three-variable VAR with expectation errors of output, output, and house prices to consider the partial derivatives of output and house prices at various horizons with respect to shocks in the expectation errors of output. We compare the results from an empirical VAR to those arising from application of the same VAR specification to data generated from our model with information heterogeneity. To measure the average expectation of output, we collect data from the Survey of Professional Forecasters (SPF). The data cover the period from 1975Q1 to 2010Q3. We take the median forecasts of real GDP in the coming quarter as the forecast of output. We define the expectation errors as the percentage deviation of the realized real GDP from the forecast of real GDP. To see how innovations in the expectation errors affect other variables, we run the VAR with four lags and the expectation errors ordered first. Figure 15 shows the empirical impulse responses to shocks in expectation errors of output from the trivariate VAR. The shaded areas represent one-standard-error bias-corrected bootstrap confidence bands of Kilian (1998). The figure shows that one percent increases in agents' expectation errors are followed by around 0.05 increases in house prices and 0.4 increases in real GDP.

To run a similar trivariate VAR for the model, we collect simulated data with a length of 142 observations. The average expectations of real variables are directly calculated, as agents' information sets are clearly defined. Similarly, we define the expectation errors of output as the percent deviation between the average expectation of output and the true output. The correlation between the expectation errors and output is also a positive value of 0.042. Figure 16 plots the impulse response to one positive standard deviation of shocks in expectation errors of output from the trivariate VAR for the simulated data. The responses in the simulated data are as similar as the responses in the empirical data, although they differ in magnitude. A one percent increase in agents' expectation errors is followed by around a 0.05 percent increase in house prices and a 0.05 percent increase in output. The main difference between the data sets is that in the simulated data both house prices and output respond with a hump shape, but in the empiricial data, we do not observe such a hump.

To check the robustness of the results, we have repeated the VAR exercise using different variables or different numbers of variables. For instance, we have replaced the expectation errors of output by the expectation errors of nonresidential investment, and replaced output by nonresidential investment. We have also extended the three-variable VAR to a five-variable VAR by adding consumption and nonresidential investment. All of the regressions report similar qualitative results.

Conclude

The recent standard real business cycle models with financial frictions succeed in explaining the close correlations among house prices, consumption, and investment. However, the models cannot explain two facts: the disconnect between house prices and rental prices, and the lead-lag relationship between residential investment and nonresidential investment. We introduce information heterogeneity into a standard real business cycle model with real estate production and financial frictions. By assuming that agents are rationally confused about the sources of shocks, the model generates an amplified response of house prices to technology shocks, which explain the disconnect puzzle. Since the amplified response mainly comes from the rising demand of real estate from households, the model potentially explains the lead-lag relationship between residential investment and nonresidential investment.

There are several directions in which our paper can be improved[11]. In our model, although we show information heterogeneity amplifies the response of house prices to technology shocks, the volatility of house prices is still much lower compared to the data. One can introduce monetary shocks into the model and investigate the confusion between real shocks and nominal shocks, since nominal shocks can also be viewed as pure demand shocks and therefore may serve a similar role to housing demand shocks in our model. Second, we could apply the method of minimization of distance between the simulated second moments and the empirical second moments to pin down parameters for our calibration instead of choosing ad hoc values. Third, our model extends the standard real business cycle model in three directions: residential production, financial frictions, and information frictions. It

[11] Our solution method can only solve the model using certain ranges of parameters values. Of course, this is the most central issue to address.

is more intuitive to extend the model step by step, so one can clearly discuss how each extension affects the model. All of these are left for future work.

Appendix: Solving a DSGE model with heterogeneous information

The solving procedure consists of four steps in total.

- **Step one:** shut down all the shocks, solve the model in the steady state, and log-linearize the model around the steady state. In our model, there are two aggregate variables which affect agents' decisions: housing prices P_t and the aggregate consumption of the entrepreneur C_t. The later one also determines the stochastic discount factor. We assume the two aggregate variables are a linear function of aggregate shocks $\Xi_t = \{\{u_{t-i}^a\}_{i=1}^T, \{u_{t-i}^\chi\}_{i=1}^T\}$, $C_t = CC * [u_t^a, u_{t-1}^a, ..., u_{t-T}^a, u_t^\chi, u_{t-1}^\chi, ..., u_{t-T}^\chi]'$, and $P_t = PP * [u_t^a, u_{t-1}^a, ..., u_{t-T}^a, u_t^\chi, u_{t-1}^\chi, ..., u_{t-T}^\chi]'$.

- **Step two**: replace the goods market clearing condition and the housing market clearing condition by the two above equations of the definitions C_t and P_t, and solve the linear difference equations as a typical rational expectation model.

- **Step three**: from Step two, we have

  $$Y_{it} = G_1 Y_{it-1} + \Theta_c + \Theta_0 z_{it}^*,$$

  and then apply an expectation operator to both sides of the above equation conditional on the information set Ω_{it}

  $$Y_{it} = G_1 Y_{it-1} + \Theta_c + \Theta_0 E_{it} z_{it}^*.$$

  To derive $E_{it} z_{it}^*$, we should first keep in mind that the signals s_{it} island i receives are linear functions of z_{it}, given by,

  $$s_{it} = \Gamma z_{it}.$$

  By Kalman filter updating, we have

  $$E_{it} z_{it} \equiv E(z_{it}|s_{it}) = \Sigma\Gamma'(\Gamma\Sigma\Gamma')^{-1} s_{it} = \Sigma\Gamma'(\Gamma\Sigma\Gamma')^{-1}\Gamma z_{it}.$$

- **Step four**: plug the solved individual variables into the goods market clearing condition and the housing market clearing condition, derive the updated C_t^* and P_t^*, and match the distance between (C_t, P_t) and (C_t^*, P_t^*). If the distance is zero or close enough to zero, we solve the model. In our calibration, the square root of the distance is less than 10^{-3}, although we cannot find the exact solution.

Table 6: House price appreciation and rental prices

	α_0	α_1	α_2
The FHFA series	0.0449**	0.0022**	0.0899**
The CSW series	0.0439**	0.0024**	0.0191**

** indicates rejection at 1% significance level.

Table 7: Second Moments - Empirical lead-lag correlations

	Austria	FIN	FRA	NET	UK	EU	AUS	CAN	US
$\rho(I^s_{t-1}, I_t)$	-0.359	0.453	0.576	0.227	0.210	0.301	0.355	0.398	0.503
$\rho(I^s_t, I_t)$	-0.268	0.378	0.618	0.567	0.094	0.288	0.267	0.228	0.289
$\rho(I^s_{t+1}, I_t)$	-0.161	0.202	0.448	0.138	-0.029	0.182	0.137	0.018	0.021
$\rho(I^s_{t-1}, Y_t)$	0.047	0.669	0.540	0.378	0.467	0.722	0.519	0.640	0.689
$\rho(I^s_t, Y_t)$	0.029	0.668	0.595	0.489	0.513	0.715	0.578	0.580	0.571
$\rho(I^s_{t+1}, Y_t)$	0.019	0.560	0.604	0.463	0.454	0.618	0.503	0.378	0.345
$\rho(I_{t-1}, Y_t)$	0.381	0.452	0.082	0.416	-0.063	0.495	0.335	0.491	0.498
$\rho(I_t, Y_t)$	0.473	0.653	0.186	0.584	0.007	0.596	0.479	0.662	0.724
$\rho(I_{t+1}, Y_t)$	0.484	0.737	0.261	0.610	0.089	0.621	0.510	0.745	0.797

$I^s_t, I_t,$ and $Y_t$denote residential investment, nonresidential investment and output respectively.

Table 8: The causality test between residential and business investments

Country		$I^s_t \nrightarrow I_t$		$I_t \nrightarrow I^s_t$	
	Lag	χ^2 Value	p Value	χ^2 Value	p Value
Austria	4	4.120	0.390	8.199	0.085
Finland	6	13.63	0.034	12.318	0.055
France	6	116.52	0.000	99.495	0.000
Netherlands	4	5.311	0.257	7.454	0.114
UK	2	8.121	0.017	5.052	0.080
Euro	2	2.331	0.312	5.874	0.061
Australia	4	22.649	0.000	5.303	0.258
Canada	2	10.190	0.006	5.611	0.060
USA (1960Q1~2012Q2)	4	181.9	0.000	13.8	0.014
USA (1984Q1~2005Q4)	2	158.8	0.000	5.1	0.076

Table 9: Business cycle statistics for the models

	U.S. Data	Full info.	Hetero info.
	A. Basic statistics		
σ_y	1.42	1.08	1.31
σ_c/σ_y	0.62	0.48	0.36
σ_i/σ_y	2.54	2.35	2.73
σ_{i^s}/σ_y	5.05	2.64	2.54
σ_p/σ_y	1.55	0.49	0.64
$\rho(y_t, p_t)$	0.52	0.77	0.87
$\rho(c_t, p_t)$	0.47	0.90	0.94
$\rho(i_t, p_t)$	0.59	0.97	0.98
	B. Investment dynamics		
$\rho(I^s_{t-1}, Y_t)$	0.77	0.17	0.58
$\rho(I^s_t, Y_t)$	0.73	0.83	0.66
$\rho(I^s_{t+1}, Y_t)$	0.32	0.65	0.44
$\rho(I^s_{t-1}, I_t)$	0.84	-0.04	0.51
$\rho(I^s_t, I_t)$	0.71	0.69	0.59
$\rho(I^s_{t+1}, I_t)$	0.29	0.58	0.38
$\rho(I_{t-1}, Y_t)$	0.75	0.42	0.43
$\rho(I_t, Y_t)$	0.89	0.97	0.98
$\rho(I_{t+1}, Y_t)$	0.60	0.20	0.47

Table 10: House price appreciation and rental prices in simulated data

	α_0	α_1	α_2
U.S. Data	0.0449**	0.0022**	0.0899**
Full info.	0.0148	1.3487	1.4195
Hetero info.	0.0141	1.8102*	0.4895*

** and * indicate rejection at 1% and 10% significance level respectively.

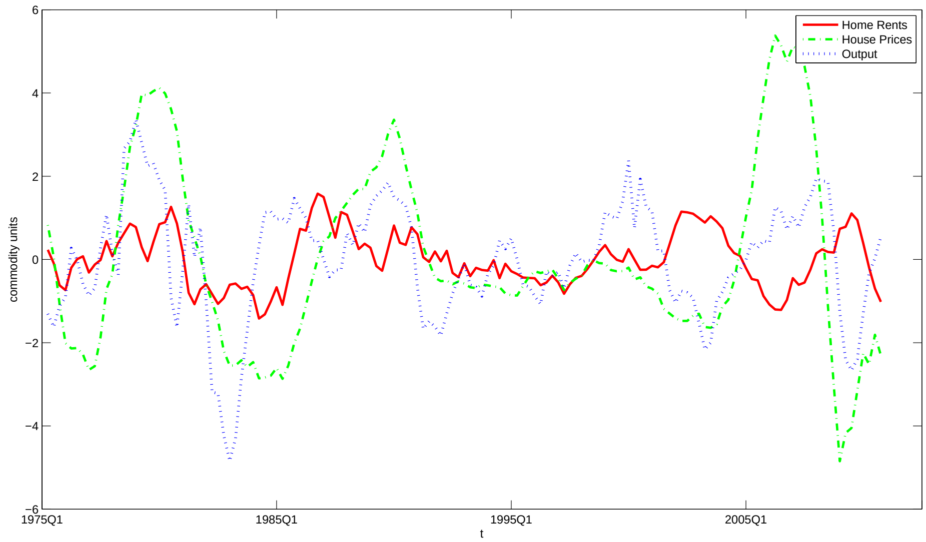

Figure 12: Home rents and house prices with the business cycle.

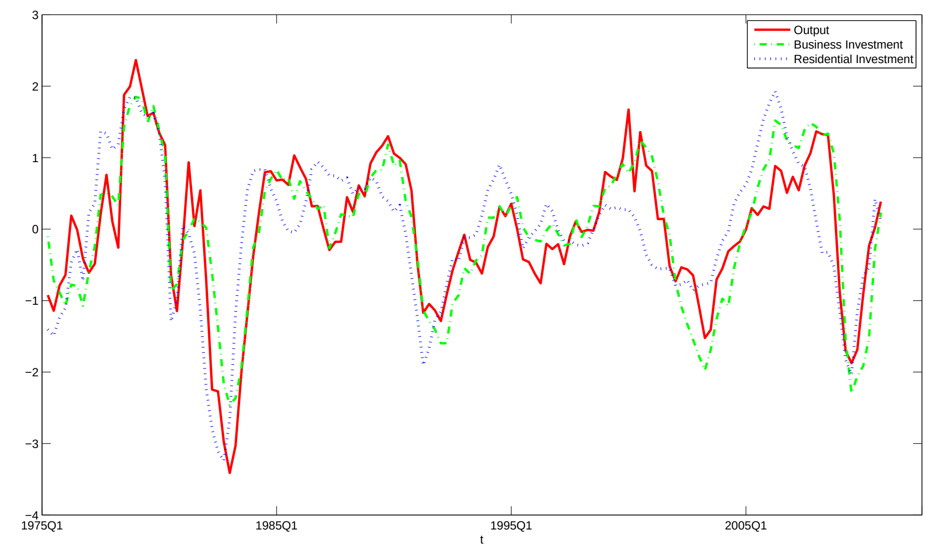

Figure 13: Residential investment and nonresidential investment with the business cycle

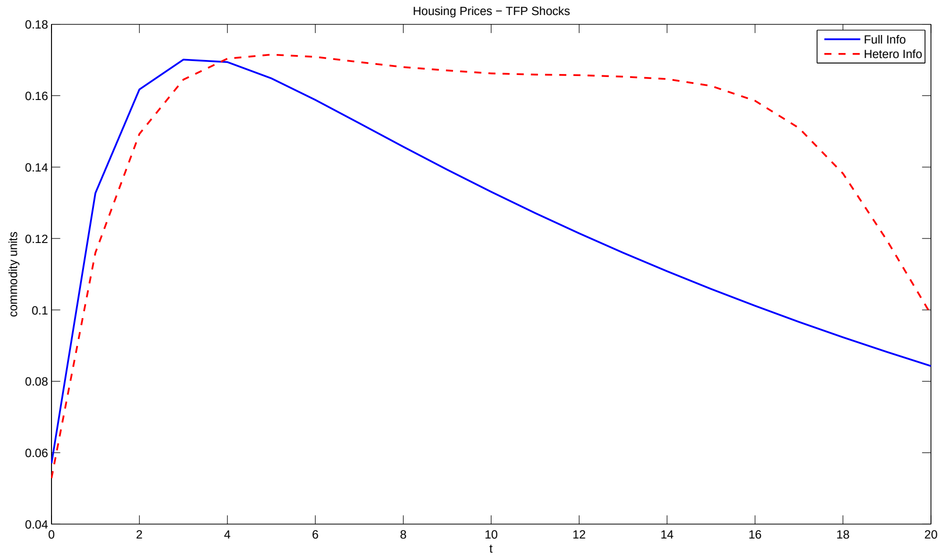

Figure 14: House prices in response to TFP shocks

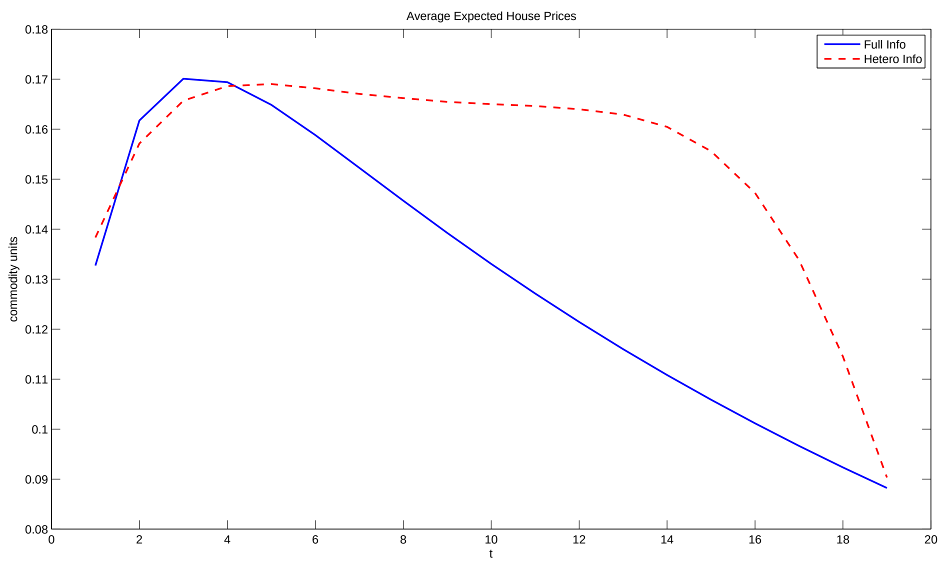

Figure 15: Average expectation of next-period house prices

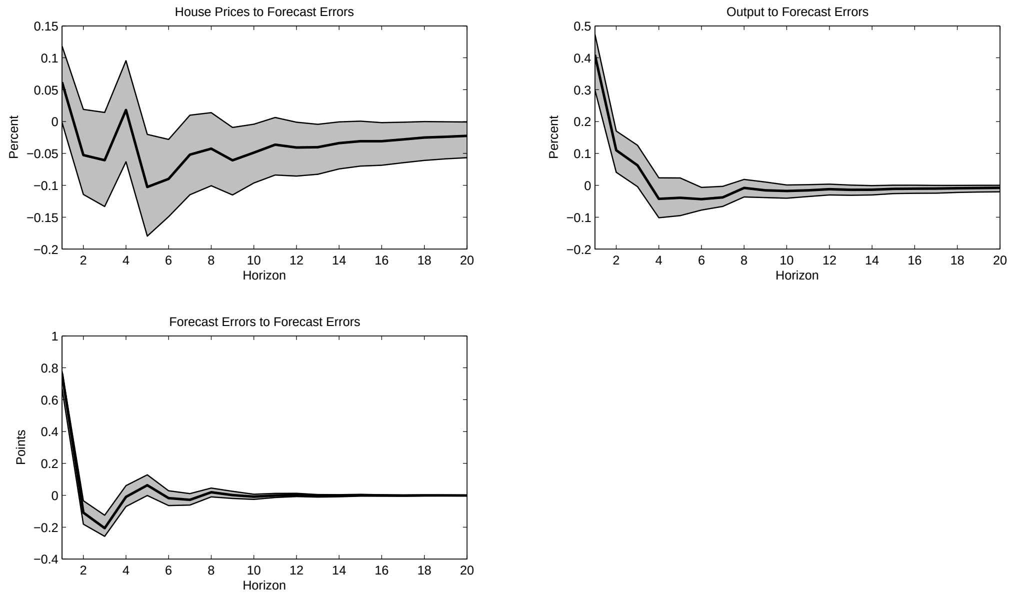

Figure 16: Empirical evidences from SVAR

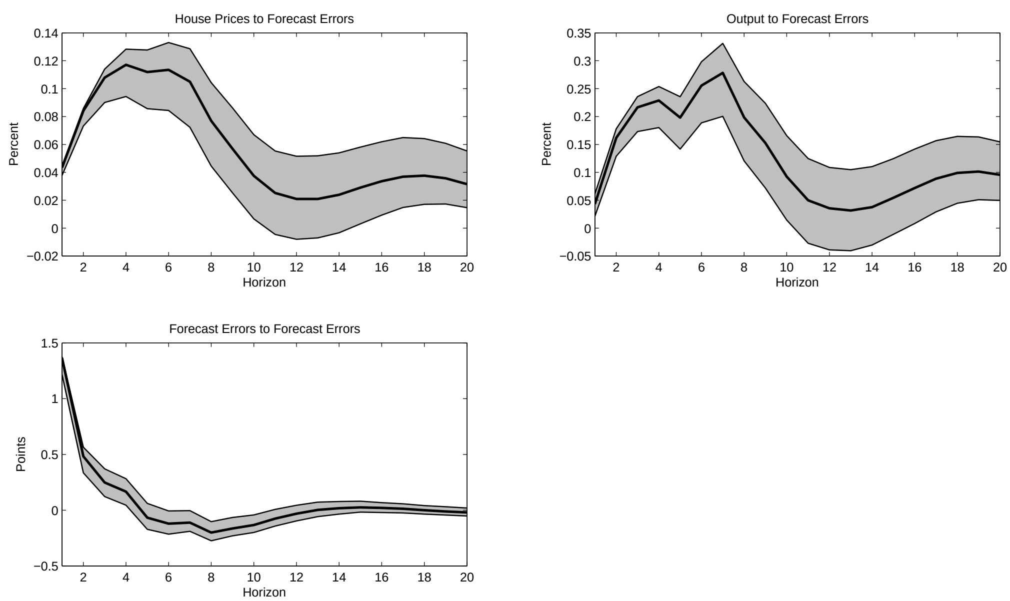

Figure 17: Simulation evidences from SVAR

BIBLIOGRAPHY

[1] Ambler, Steven, Emanuela Cardia and Christian Zimmermann, 2004. International business cycles: what are the facts? Journal of Monetary Economics, 51(2), pp. 257-276.

[2] Anderson, Gary, 2008. Solving linear rational expectations models: a horse race. Computational Economics, 31, pp. 95-113.

[3] Angeletos, George-Marios, and Jennifer La'O, 2009. Noisy business cycles? In Daron Acemoglu, Kenneth Rogoff, and Michael Woodford (Eds.), *NBER Macroeconomics Annual 2009*, 24. pp. 653-701.

[4] Bacchetta, Philippe, and Eric van Wincoop, 2006. Can information heterogeneity explain the exchange rate determination puzzle?. American Economic Review, 96(3), pp. 552-576.

[5] Backus, David K., Patrick J. Kehoe, and Finn E. Kydland, 1992. International real business cycles. Journal of Political Economy, 101, pp. 745-775.

[6] Backus, David K., Patrick J. Kehoe, and Finn E. Kydland,1995. International business cycles: theory and evidence. in Cooley, T.F. (Ed.), Frontiers of Business Cycle Research. Princeton University Press, Princeton, NJ, pp. 331-356.

[7] Bai, Jusan, 2003. Inferential theory for factor models of large dimensions. Econometrica, 71(1), pp. 135-171.

[8] Bai, Jusan and Serena Ng, 2002. Determining the number of factors in approximate factor models. Econometrica, 70(1), pp. 191-221.

[9] Bai, Jusan and Serena Ng, 2004. A PANIC attack on unit roots and cointegration. Econometrica, 72(4), pp. 1127-1177.

[10] Bai, Jusan and Serena Ng, 2006. Confidence intervals for diffusion index forecasts and inference with factor-augmented regressions. Econometrica, 74(4), pp. 1133-1150.

[11] Baxter, Marianne, and Mario J. Crucini, 1993. Explaining saving-investment correlations. American Economic Review, 83, pp. 416-436.

[12] Baxter, Marianne, and Mario J. Crucini, 1995. Business cycles and the asset structure of foreign trade. International Economic Review, 36, pp. 821-854.

[13] Benhabib, Jess, Richard Rogerson, and Randall Wright, 1991. Homework in macroeconomics: household production and aggregate fluctuations, Journal of Political Economy, 99, pp. 1166-1187.

[14] Beran, Rudolf and Muni S. Srivastava, 1985. Bootstrap tests and confidence regions for functions of a covariance matrix. Annals of Statistics, 13, pp. 95-115.

[15] Bernanke, Ben, Jean Boivin and Piotr Eliasz, 2005. Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. Quarterly Journal of Economics, 120(1), pp. 387-422.

[16] Bryan, Michael F. and Stephen G. Cecchetti, 1993. The consumer price index as a measure of inflation, Federal Reserve Bank of Cleveland Economic Review, pp. 15–24.

[17] Boileau, Martin and Michel Normandin, 2008. Closing international real business cycle models with restricted financial markets. Journal of International Money and Finance, 27, pp. 733-756.

[18] Boivin, Jean and Marc Giannoni, 2006. DSGE models in a data-rich environment. NBER Working Paper no. 12772.

[19] Boivin, Jean and Serena Ng, 2006. Are more data always better for factor analysis? Journal of Econometrics, 132, pp. 169-194.

[20] Bose, Arup, 1988. Edgeworth correction by bootstrap in autoregressions. Annals of Statistics, 16, pp. 1709-1722.

[21] Bulow, Jeremy, and Paul Klemperer, 1994. Rational frenzies and crashes. Journal of Political Economy, 102, pp. 1-23.

[22] Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo, 2011. Understanding booms and busts in housing markets, mimeo.

[23] Caplin, Andrew, and John Leahy, 2011. Trading frictions and house price dynamics. Journal of Money, Credit, and Banking, 43, pp. 283-303.

[24] Chaney, Thomas, David Sraer, and David Thesmar, 2012. The collateral channel: how real estate shocks affect corporate investment, American Economic Review, 102, pp. 2381-2409.

[25] Chang, Yongsung, 2000. Comovement, excess volatility, and home production. Journal of Monetary Economics, 46, pp. 385-396.

[26] Chari, V.V., Patrick J. Kehoe and Ellen R. McGrattan, 2002. Can sticky price models generate volatile and persistent real exchange rates?. Review of Economic Studies, 69, pp. 533-563.

[27] Chatterjee, Satyajit, and Burcu Eyigungor, 2011. A quantitative analysis of the US housing and mortgage markets and the foreclosure crisis, mimeo.

[28] Christiano, Lawrence J., Martin Eichenbaum, and Evans, 2005, Nominal rigidities and the dynamic effects of a shock to monetary policy. Journal of Political Economy, 113, pp. 1-45.

[29] Christiano, Lawrence J., Martin Eichenbaum, and Robert Vigfusson, 2003. What happens after a technology shock? NBER Working Paper 9819.

[30] Clark, Todd E., 2006. Disaggregate evidence on the persistence of consumer price inflation. Journal of Applied Econometrics, 21, pp. 563–587.

[31] Cogley, Timothy, and James M. Nason, 1995. Output dynamics in real-business-cycle models. American Economic Review, 85, pp. 492-511.

[32] Connor, Gonnor and Robert A. Korajczyk, 1986. Performance measurement with the Arbitrage Pricing Theory: A new framework for analysis. Journal of Financial Economics, 15, 373-394.

[33] Connor, Gonnor, Robert A. Korajczyk and Oliver B. Linton, 2006. The common and specific components of dynamic volatility. Journal of Econometrics 132(1), pp. 231-255.

[34] Crucini, Mario J., Mototsugu Shintani, and Takayuki Tsuruga, 2010. Accounting for persistence and volatility of good-level real exchange rates: the role of sticky information. Journal of International Economics, 81(1), pp. 48-60.

[35] Crucini, Mario J., Mototsugu Shintani, and Takayuki Tsuruga, 2012. Noisy information, distance and law of one price dynamics across US cities, mimeo.

[36] Davis, Morris A., and Jonathan Heathcote, 2005. Housing and the business cycle. International Economic Review, 46, pp. 751–84.

[37] Davis, Morris A., and Jonathan Heathcote, 2007. The price and quantity of residential land in the United States. Journal of Monetary Economics, 54, pp. 2595-2620.

[38] Davis, J. Scott, and Kevin X.D. Huang, 2011. International real business cycles with endogenous markup variability. Journal of International Economics, 85, pp. 302-315.

[39] Diebold, Francis X., 2000. "Big Data" dynamic factor models for macroeconomic measurement and forecasting. manuscript presented at World Congress of the Econometric Society 2000.

[40] Favilukis, Jack, Sydney C. Ludvigson, and Stijn Van Nieuwerburgh, 2011. The macroeconomic effects of housing wealth, housing finance, and the limited risk-sharing in general equilibrium, mimeo.

[41] Fisher, Jonas D., 2007. Why does household investment lead business investment over the business cycle?. Journal of Political Economy, 115, pp. 141-168.

[42] Freedman, David, 1984. On bootstrapping two-stage least-squares estimates in stationary linear models. Annals of Statistics, 12, pp. 827-842.

[43] Galí, Jordi, 1999. Technology, employment, and the business cycle: do technology shocks explain aggregate fluctuations? American Economic Review, 89(1), pp. 249-271.

[44] Gangopadhyay, Kausik, and Juan Carlos Hatchondo, 2009. The behavior of household and business investment over the business cycle, Federal Reserve Bank of Richmond Economic Quarterly, 95, pp. 269-288.

[45] Gertler, Mark, and Nobuhiro Kiyotaki, 2010. Financial intermediation and credit policy in business cycle analysis. *The Handbook of Monetary Economics*, ed. by Benjamin M. Friedman, and Michael Woodford, vol. 3B, pp. 1341-1393. Elsevier Science B.V., Amsterdam, North Holland.

[46] Geweke, John, 1977. The dynamic factor analysis of economic time-series models. in D. J. Aigner and A.S. Goldberger (eds.), Latent Variable in Socioeconomic Models, Amsterdam, North-Holland, pp. 365-387.

[47] Gomme, Paul, Finn E. Kydland, and Peter Rupert, 2001. Home production meets time to build, Journal of Political Economy, 109, pp. 1115-1131.

[48] Gonçalves, Silvia and Benoit Perron, 2012. Bootstrapping factor-augmented regression models, working paper, University of Montreal.

[49] Gourichas, Pierre-Olivier and Aaron Tornell, 2004. Exchange rate puzzles and distorted beliefs. Journal of International Economics, 64(2), pp. 303-333.

[50] Graham, Liam, and Stephen Wright, 2010. Information, heterogeneity and market incompleteness. Journal of Monetary Economics, 57, pp. 164-174.

[51] Greenwood, Jeremy, and Zvi Hercowitz, 1991. The allocation of capital and time over the business cycle. Journal of Political Economy, 99, pp. 1188-1214.

[52] Guo, Zi-Yi, and Mototsugu Shintani, 2011. Noisy information, risk sharing and international business cycles, mimeo.

[53] Hall, Peter, 1992. The Bootstrap and Edgeworth Expansion. New York: Springer-Verlag.

[54] Hansen, Gary D., 1997. Technical progress and aggregate fluctuations. Journal of Economic Dynamics and Control, 21, pp. 1005-1023.

[55] Heathcote, Jonathan and Fabrizio Perri, 2002. Financial autarky and international business cycles. Journal of Monetary Economics, 49, pp. 601-627.

[56] Himmelberg, Charles, Christopher Mayer, and Todd Sinai, 2005. Assessing high house prices: bubbles, fundamentals and misperceptions, Journal of Economic Perspectives, 19, pp. 67-92.

[57] Hurwicz, Leonid, 1950. Least squares bias in time series, in T. Koopmans (ed.) Statistical Inference in Dynamic Economic Models. Wiley, New York, pp. 365-383.

[58] Iacoviello, Matteo, 2005. House prices, borrowing constraints, and monetary policy in the business cycle. American Economic Review, 95, pp. 739-764.

[59] Iacoviello, Matteo, 2010. Housing in DSGE models: findings and new directions, in Bandt, O. de; Knetsch, Th.; Peňalosa, J.; Zollino, F. (Eds.), *Housing markets in Europe: a macroeconomic perspective*, Berlin, Hidelberg: Springer-Verlag, pp. 3-16.

[60] Iacoviello, Matteo, and Stefano Neri, 2010. Housing market spillovers: evidence from an estimated DSGE model. American Economic Journal: Macroeconomics, 2, pp. 125-164.

[61] Ireland, Peter N., 2001. Technology shocks and the business cycle: On empirical investigation. Journal of Economic Dynamics and Control, 25, pp. 703–719.

[62] Kehoe, Patrick J. and Fabrizio Perri, 2002. International business cycles with endogenous incomplete markets. Econometrica, 70, pp. 907-928.

[63] Kendall, Manfred G., 1954. Note on bias in the estimation of autocorrelation. Biometrika, 41, pp. 403-403.

[64] Kilian, Lutz, 1998. Small-sample confidence intervals for impulse response functions. Review of Economics and Statistics, 80, pp. 218-230.

[65] Kim, Chang-Jin and Charles R. Nelson, 1998. Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime-switching. Review of Economics and Statistics, 80, pp. 188-201.

[66] King, Robert G., Charles I. Plosser and Sergio T. Rebelo, 1988. Production, growth and business cycles: I. the basic neoclassical model. Journal of Monetary Economics, 21, pp. 195-232.

[67] Kiyotaki, Nobuhiro, and John Moore, 1997. Credit cycles. Journal of Political Economy, 105, pp. 211-248.

[68] Kiyotaki, Nobuhiro, Alexander Michaelides, and Kalin Nikolov, 2011. Winners and losers in housing markets. Journal of Money, Credit, and Banking, 43, pp. 255-296.

[69] Kollmann, Robert, 1996. Incomplete asset markets and the cross-country consumption correlation puzzle. Journal of Economic Dynamics and Control, 20, 945-961.

[70] La'O, Jennifer, 2010. Collateral constraints and noisy fluctuations, mimeo.

[71] Liu, Zheng, Pengfei Wang, and Tao Zha, 2013. Land-price dynamics and macroeconomic fluctuations. Econometrica, forthcoming.

[72] Lorenzoni, Guido, 2009. A theory of demand shocks. American Economic Review, 99, pp. 2050-2084.

[73] Lucas, Robert Jr., 1972. Expectations and the neutrality of money. Journal of Economic Theory, 4, pp. 103-124.

[74] Ludvigson, Sydney C. and Serena Ng, 2010. A factor analysis of bond risk premia. In A. Ullah and D. E. A. Giles (eds.), Handbook of Empirical Economics and Finance, pp. 313-372.

[75] Luo, Yulei, Jun Nie and Eric R. Young, 2012. Robustness, information-processing constraints, and the current account in small open economies. Journal of International Economics, 88, pp. 104-120. .

[76] MacKinnon, James G. and Anthony A. Smith, 1998. Approximate bias correction in econometrics. Journal of Econometrics, 85, pp. 205-230.

[77] Mackowiak, Bartosz A., and Mirko Wiederholt, 2009. Optimal sticky prices under rational inattention. American Economic Review, 99. pp. 769-803.

[78] Mankiw, N. Gregory, and Ricardo Reis, 2002. Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips curve. Quarterly Journal of Economics, 117, pp. 1295-1328.

[79] Mankiw, N. Gregory, and Ricardo Reis, 2011. Imperfect Information and Aggregate Supply. The *Handbook of Monetary Economics*, edited by B. Friedman and M. Woodford, vol. 3A, pp. 182-230., Elsevier-North Holland.

[80] Marcellino, Marcellino, James H. Stock and Mark W. Watson, 2003. Macroeconomic forecasting in the Euro area: Country specific versus area-wide information. European Economic Review, 47, pp. 1-18.

[81] Marriott, Frank. H. C., and James A. Pope, 1954. Bias in the estimation of autocorrelations. Biometrika, 41, pp. 390–402.

[82] Mayer, Christopher, 2011. Housing bubbles: a survey. Annual Review of Economics, 3, pp. 559-577.

[83] Mian, Atif, and Amir Sufi, 2011. House prices, home equity-based borrowing, and the U.S. household leverage crisis. American Economic Review, 101, pp. 2132-2156.

[84] Moon, Hyungsik R. and Benoit Perron, 2004. Testing for a unit root in panels with dynamic factors. Journal of Econometrics 122, pp. 81-126.

[85] Morris, Stephen and Hyun Song Shin, 2002. Social value of public information. American Economic Review, 92, pp. 1521-1534.

[86] Nimark, Kristoffer, 2008. Dynamic pricing and imperfect common knowledge. Journal of Monetary Economics, 55, pp. 365-382.

[87] Obstfeld, Maurice and Kenneth Rogoff, 2000. The six major puzzles in international macroeconomics: is there a common cause? in Bernanke, Ben; Rogoff, Kenneth

(Ed.), NBER Macroeconomics Annual 2000, The MIT Press, Cambridge MA, pp. 339-390.

[88] Pakko, Michael, 2004. A spectral analysis of the cross-country consumption correlation puzzle. Economics Letters, 84, pp. 341-347.

[89] Phelps, Edmund S., 1970. Introduction: the new macroeconomics in employment and inflation theory. *In Microeconomic Foundations of Employment and Inflation Theory*, edited by Edmund et al. Phelps. New York: Norton.

[90] Piazzesi, Monika, and Martin Schneider, 2009. Momentum traders in the housing market: survey evidence and a search model. American Economic Review: Papers & Proceedings, 99, pp. 406-411.

[91] Quenouille, Michael H., 1949. Approximate tests of correlation in time series. Journal of the Royal Statistical Society, Series B, 11, pp. 68–83.

[92] Rupert, Peter, and Etienne Wasmer, 2012. "Housing and the labor market: time to move and aggregate unemployment," Journal of Monetary Economics, 59, pp. 24-36.

[93] Sargent, Thomas J. and Christopher A. Sims, 1977. Business cycle modeling without pretending to have too much a priori theory. In C. Sims (ed.), New Methods of Business Cycle Research. Minneapolis: Federal Reserve Bank of Minneapolis.

[94] Schmitt-Grohé, Stephanie and Martín Uribe, 2003. Closing small open economy models. Journal of International Economics, 61, pp. 163-185.

[95] Shintani, Mototsugu, 2005. Nonlinear forecasting analysis using diffusion indexes: An application to Japan. Journal of Money, Credit, and Banking, 37(3), pp. 517-538.

[96] Shintani, Mototsugu, 2008. A dynamic factor approach to nonlinear stability analysis. Journal of Economic Dynamics and Control, 32(9), pp. 2788-2808.

[97] Sims, Christopher. A., 2001. Solving linear rational expectations models. Computational Economics, 20, pp. 1-20.

[98] Sims, Christopher A., 2003. Implications of rational inattention. Journal of Monetary Economics, 50, pp. 665-690.

[99] Sterk, Vincent, 2010. Home equity, mobility, and macroeconomic fluctuations. De Nederlandsche Bank and University of Amsterdam, mimeo.

[100] Stock, James H. and Mark W. Watson, 1989. New indexes of coincident and leading economic indicators. In O. Blanchard and S. Fischer (eds.), NBER Macroeconomics Annual (Cambridge, Mass. MIT Press).

[101] Stock, James H. and Mark W. Watson, 1998. Diffusion indexes. NBER Working Paper no. 6702.

[102] Stock, James H. and Mark W. Watson, 2002. Macroeconomic forecasting using diffusion indexes. Journal of Business and Economic Statistics 20, pp. 147-162.

[103] Stockman, Alan and Linda Tesar, 1995. Tastes and technology in a two-country model of the business cycle: explaining international comovements. American Economic Review, 85, pp. 168-185.

[104] Townsend, Robert M., 1983. Forecasting the forecasts of others. Journal of Political Economy, 91, pp. 546-588.

[105] Woodford, Michael D. 1999. Optimal monetary policy inertia, NBER Working Paper no. 7261.

[106] Woodford, Michael D., 2003. Imperfect common knowledge and the effects of monetary policy, In *Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, edited by Phillipe Aghion et al.

[107] Yamamoto, Yohei, 2012. Bootstrap inference for impulse response functions in factor-augmented vector autoregressions. working paper, University of Alberta.