# Dynamic Information Acquisition in an infinite-horizon framework

Zhifeng CAI *

University of Minnesota, Federal Reserve Bank of Minneapolis

January 8, 2016

**Abstract**

This paper studies a repeated Grossman and Stiglitz (1980)'s model with arbitrary horizon. Each generation of investors are born observing the contemporaneous stock price and choose whether to acquire information about the underlying fundamental. There are three main results. First, there is always a unique equilibrium if the model horizon is finite, and there may exist multiple equilibrium in information acquisition if the model horizon is infinite. Second, if horizon is infinite, there exists multiple steady states as well as a continuum nonstationary equilibria. Third, in the long run, almost all the nonstationary equilibria converges to the intermediate steady state which has interesting comparative statics. In particular, at that steady state, fewer investors choose to become informed if information becomes cheaper to acquire. Two applications are included to illustrate the potential of the infinite-horizon model in accounting for financial market time-series data.

*Address: 4-101 Hanson Hall, 1925 Fourth Street South, Minneapolis, MN 55414, USA, e-mail: *caixx162@umn.edu*. The author is grateful to Jonathan Heathcote, VV Chari, Manuel Amador for invaluable guidance. The author also thanks Hengjie Ai, Anmol Bhandari, Zhen Huo, Larry Jones, Erzon Luttmer, Chris Phelan, David Rahman, Aldo Rustichini, Junyan Shen, Laura Veldkamp, and seminar participants at the Public Workshop of University of Minnesota, Math Econ Workshop of University of Minnesota, and Seminar at the Federal Reserve Bank of Minneapolis for helpful comments.

# 1 Introduction

Investors in the financial market acquire private information and trade in order to "beat the market". How do investors' information acquisition decisions interact with each other? Grossman and Stiglitz (1980), in a static model, provides a classical view to this question: the fact that privately-acquired information is partially revealed through prices means that the larger is the share of informed investors today, the smaller is the return to information acquisition. Thus information is a *static substitute* in that its value decreases with the share of informed investors *today*[1].

Since then, people have proposed various ways to generate multiplicity in information acquisition. In particular, there is a literature that studies static information acquisition in finite-horizon models. In this literature, information market only opens at the very beginning. Recently, **?** proposes an infinite-horizon model with dynamic information acquisition where information market opens every period, and identifies multiplicity. Thus a question remains: if information market is allowed to open every period, is the assumption of inifinite-horizon essential for generating multiplicity?

To answer this question, this paper constructs a repeated Grossman and Stiglitz (1980) economy with arbitrary horizon. Every period a new generation of investors are born, which lives for only two periods. They observe contemporaneous stock prices and are allowed acquire information about the underlying fundamental at some cost. The economy ends at some arbitrary period $T \geq 2$. If $T = 2$, the economy collapses to the one considered in Grossman and Stiglitz (1980).

My first main result is that when $T < \infty$ there is always a unique equilibrium, whereas when $T = \infty$, there may exist multiple equilibria. Thus the assumption of infinite-horizon is essential to generate multiplicity in information acqusition, when the information market is allowed to open every period. The source of multiplicity here is similar to **?**: information

[1] Manzano and Vives (2011) shows that Grossman and Stiglitz (1980)'s substitutability result is robust

acquisition emerges naturally not only as a *static substitute*, but also as a *dynamic complement*, *i.e.* agents' incentive to acquire information not only decreases with the share of informed agents *today* (which corresponds to the classical insight provided by Grossman and Stiglitz (1980)), but also increases with the share of informed agents *in the future*. This is because as more agents get informed in the future, future resale stock price becomes more sensitive to the fundamental. This creates more uncertainty for today's uninformed agents (as they do not know the fundamental) and urge them to become informed today. Thus multiplicity arises due to a self-fulfilling prophecy: larger share of agents become informed today because they expect larger share of informed agents tomorrow.

This logic, however, breaks down when $T < \infty$. The reason is that in such a case one can pin down the share of informed investors through backward-induction. For example, the $T-1$ generation of investors' incentive to acquire information is no longer endogenous because there is no future generation of investors. With the $T-1$ generation's information choice pinned down, the $T-2$ generation's information choice is pinned down as well. The logic goes on and thus there is a unique equlibrium.

Next, I focus on characterizing equilibria of the infinite-horizon model. I proved that there exists multiple steady states as well as a continuum of nonstationary equilibria. The existence of nonstationary equilibria hinges on the assumption that uninformed investors observe only a finite history of information (in this case, only the current stock price). At each nonstationary equilibria, price coefficients are time-varying. But agents have perfect foresight and thus no uncertainty with respect to these price coefficients.

In the long run, almost all the identified nonstationary equilibria converge to the intermediate steady state where the value of information is upward-sloping instead of downward sloping as in Grossman and Stiglitz (1980). This produces interesting comparative statics. For example, fewer investors choose to become informed if the cost of acquiring information is lower.

Lastly, I include two examples to illustrate how the model can be used to think about time-serios data. In one application, I illustrate how a belief shock in the model could induce a sudden and persistent rise of uncertainty in the financial market. In the other application I show that the model can account for the fact that passive investing has become more popular in recent decades, even though new information and communications technologies have reduced the cost of information acquisition.

**Literature Review** The theory is related to the literature of Noisy Rational Expectation models with endogenous information acquisition (Grossman and Stiglitz, 1980; Verrecchia, 1982; Veldkamp, 2006a,b; Chamley, 2007; Barlevy and Veronesi, 2007; Ganguli and Yang, 2009; Cespa and Vives, 2014),etc. Grossman and Stiglitz (1980) and Verrecchia (1982) obtains the classical result of strategic substitution in information acquisition. Later works identifies various sources of strategic complementarity in information acquisition. Barlevy and Veronesi (2007) argues that with correlated fundamentals and noise trading complementarity may arise. Ganguli and Yang (2009) illustrates that complementarity may result when agents own private information about their endowment. Veldkamp (2006a,b) generates complementary by embedding an increasing-return-to-scale information production sector into an otherwise standard noisy rational expectation model. All of the above-mentioned models are static in nature.

Second, there is also a literature that studies multiplicity in finite-horizon economy (Froot et al., 1992; Chamley, 2007; Cespa and Vives, 2014; Avdis, 2014; Zhang, 2012) and argues that short-term trading leads to multiplicity. The literature typically relies on special assumptions to obtain strategic complementarity in information acquisition. In particular, Chamley (2007) generates multiple equilibria by departing from the traditional CARA-Gaussian framework. This is not necessary for my theory. Froot et al. (1992), Zhang (2012) and Avdis (2014), although set up in different ways, all share a common assumption that information acquisition is only allowed in the initial period. In my setup agents are allowed to acquire information every period, allowing for dynamic interaction of information acquisition. Lastly, the infinite-horizon setup distinguishes my theory from this literature in general as it allows me to study long-run implications and time-series properties of the model[2].

Third, the theory is also related to the literature that studies asymmetric information in infinite-horizon models with long-lived assets, pioneered by Wang (1993, 1994) and Campbell and Kyle (1993). It is particularly related to models that study overlapping generations of investors (Spiegel, 1998; Bacchetta and Van Wincoop, 2006; Watanabe, 2008; Biais et al., 2010; Albagli, 2015)[3]. My model differs from previous models in two aspects. First, in my model the information acquisition choice is endogenous as opposed to the rest of the literature where information is given exogenously. Due to this respect there exists multiple equilibria associated with different fractions of informed investors. Second, I depart from the steady state analysis usually employed in this literature and also study properties of nonstationary equilibria. The model can be viewed as a first step towards studying the impact of information choice in this class of models.

The paper is structured as follows. Section 2 sets up the model with arbitrary horizon and defines an equilibrium. Section 3 proves the uniqueness in the finite-horizon economy. Section 4 proves the multiplicity in the infinite-horizon economy and provides intuition of why complementarity dominates substitutability. Section 5 studies the dynamics of the model. Section 6 discusses two applications of the theory, one being the huge and persistent spike of uncertainty during the Great Recession, the other being the recent passive funds growth. Section 7 checks robustness issues. Section 8 concludes. Proofs can be found in appendix.

[2]A more subtle difference is that, all the previous works assume away interim dividend payout to focus on the implications of short-run price fluctuations. In my model, however, interim dividend payout is very important in terms of generating multiplicity. Thus agents in my model can be interpreted as finitely-lived long-term investors who care not only about future resale stock price, but also about future dividend. In fact, Spiegel (1998); Watanabe (2008); Biais et al. (2010) all calibrate their models to annual data.

[3]This literature identifies high volatility equilibria and low volatility equilibria with different stock price sensitivity with respect to noise trader risks.

# 2 An Economy with Horizon $T$

This section describes the physical environment for an economy with arbitrary horizon $T$, where $T$ can take any integer value from 1 to $\infty$. When $T = \infty$ I refer to this economy as infinite-horizon. Otherwise $T < \infty$ it is finite-horizon. Roughly, it can be understood as a repeated version of Grossman and Stiglitz (1980). Indeed, when $T = 1$, it collapses to the classic Grossman and Stiglitz (1980) economy.

The economy is populated by a continuum of overlapping-generation risk-averse agents who consume a single consumption goods. The goods is treated as numeraire. There are two assets in the economy: a bond in perfect elastic supply, paying return $R$[4]; and a stock in fixed supply (normalize to 1) which pays dividend $D_t = \theta + \varepsilon_t$ each period (except the initial period). $\theta \sim \mathcal{N}(\mu, \sigma_\theta^2)$ is drawn by nature before the world begins and is constant over time. In later sections of the article I call $\theta$ the stock's 'fundamental'. $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ is *i.i.d.* noise in stock return.

Investors live for two periods. The initial generation is endowed with 1 unit of stock and $w$ units of risk-free bonds. Later generations are endowed with only $w$ units of risk-free bonds. They have CARA utility so initial endowment does not matter for their information acquisition choice or portfolio choice, as the utility function admits no wealth effect. Initially they observe nothing but the prior of $\theta$. They have an option to observe $\theta$ at some cost $\chi$. I refer to $\chi$ the information cost. After the information acquisition decision is made the market opens and trade occurs. Investors form their portfolio. In the next period, they receive dividend and interest payments, liquidate their position in the asset market, exit and consume their wealth.

As is standard in the noisy rational expectation literature, I introduce noise trader risk to prevent asset price from full revealing. I model it as overlapping generations of noise traders mechanically trade stocks in the first period and automatically reverse their trade in the next.

[4] Alternatively one can interpret the bond as a storage technology without nonnegative constraint.

This guarantees independence of noise trader risk across periods. The assumption simplifies the analysis and is not essential to my result. The modeling device is widely adopted in the literature (Allen et al., 2006; Gao, 2008; Brown and Jennings, 1989).

The time line is as follows:

- Stage 1: Nature draws $\theta \sim \mathcal{N}(\mu, \sigma^2)$
- Stage 2: OLG investors trade
  - In period 1:
    1. A continuum of period-1 investors are born, endowed with 1 share of stock and $w$ unit bonds. Investors have CARA utility.
    2. Period-1 investors decide whether to observe $\theta$ at cost $\chi$
    3. Market opens. Two groups of agents trade:
       (a) Period-1 investors
       (b) Period-1 noise traders. Their demand is a random variable $x_1 \sim \mathcal{N}(0, \sigma_x^2)$
  - In period $1 < t \leq T$:
    1. A continuum of period-$t$ investors are born, endowed with 0 share of stocks and $w$ unit bonds. Investors have CARA utility.
    2. $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ is realized, $D_t = \theta + \varepsilon_t$ is paid out to shareholders. Interest on bonds is paid out.
    3. Period-$t$ investors decide whether to observe $\theta$ at cost $\chi$
    4. Market opens. Four groups of agents trade:
       (a) Period-$t$ investors
       (b) Period-$t$ noise traders. Their demand is a random variable $x_t \sim \mathcal{N}(0, \sigma_x^2)$

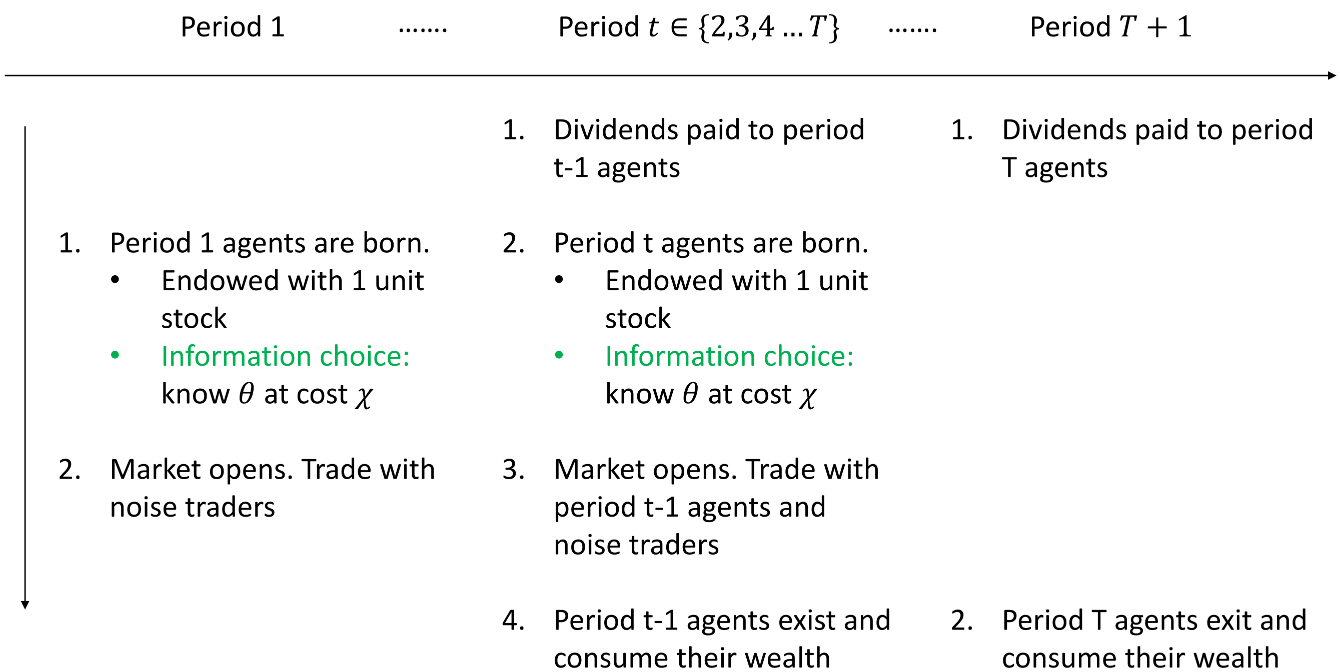



Figure 1: Timeline of Economy with Horizon $T$

(c) Period-$t-1$ investors. They liquidate their position

(d) Period-$t-1$ noise traders. They reverse their trade last period. Their demand is given by $-x_{t-1}$.

5. period-$t-1$ investors exit and consume their wealth

– In period $T+1$

1. $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ is realized, $D_t = \theta + \varepsilon_t$ is paid out to shareholders. Interest on bonds is paid out.

2. period-$T$ investors exit and consume their wealth

Several simplifying assumptions are made to make the analysis transparent. First, noise trader shock is *i.i.d.* over time. Second, the value of fundamental $\theta$ is time invariant. Third, and perhaps most crucially, uninformed agents cannot observe past prices and dividends to infer current value of fundamental. I will relax these assumptions in later sections and show that none of these assumptions is essential to the result.

## 2.1 Individual agents' problem

Agents make two choices sequentially. First, they make their information acquisition choice:

$$V_t = \max\{V_t^I, V_t^U\}$$

Where $V_t^I$ denotes expected utility for agents acquiring information (later I call them informed agents); $V_t^U$ denotes expected utility for agents not acquiring information (later I call them uninformed agents). $V_t^I$ and $V_t^U$ are determined by agents' portfolio choice:

$$\begin{aligned}
V_t^i &= \int_P V_t^i(P) dF_t(P) \\
V_t^i(P) &= \max_{e_t, b_t, c(P_{t+1}, \theta, \varepsilon_{t+1})} \int_{P_{t+1}, \theta, \varepsilon_{t+1}} U(c(P_{t+1}, \theta, \varepsilon_{t+1})) dH(P_{t+1}, \theta, \varepsilon_{t+1} | I_t^i) \\
e_t P_t + b_t &\leq w + P_t \mathbb{1}_{\{t=1\}} - \mathbb{1}\{i = I\}\chi \\
c(P', \theta, \varepsilon') &\leq (D(\theta, \varepsilon_{t+1}) + P_{t+1} \mathbb{1}_{\{t \neq T\}}) e + R b_t
\end{aligned}$$

Where $I_t^U = \{P_t\}, I_t^I = \{P_t, \theta\}$, $U(c) = -\exp(-\alpha c)$, $\alpha$ is the risk averse parameter. $D(\theta, \varepsilon_{t+1}) = \theta + \varepsilon_{t+1}$, $F, H$ are equilibrium objects as price is determined in general equilibrium.

The equation says the following. $V^i(P)$ is the expected utility of each type of agents after market opens and current price is realized. Observing current stock price $P$ agents of each type choose equity holding $e$, risk free bond holding $b$ and a contingent consumption plan to maximize their expected utility $V^i(P)$. The tradeoff here is that if agents choose to acquire information, then they need to incur information cost $\chi$. Meanwhile he observes the true value of $\theta$.

Note that for generation 1 agents, besides $w$ units of bonds they are also endowed with 1 unit of stocks, thus their endowment is given by $w + P_t$. The endowment structure is not essential as agents' have CARA utility. For the last generation $T$ agents, their stocks have no resale value.

Thus, $P_{T+1} = 0$.

## 2.2 Equilibrium Definition

Let $\lambda_t$ be the equilibrium fraction of informed agents in period $t$.

**Definition 2.1** *Denote* $s_t = \{\theta, x_t\}$, *An equilibrium of an economy with horizon* $T$ *is* $\{P_t(s_t), \lambda_t, \{s_{ti}(s_t), b_{ti}(s_t)\}_{i=U,I}\}_{t=1}^{T}$ *s.t:*

1. $e_{ti}(s_t), b_{ti}(s_t)$ *solves uninformed and informed agents' problem given* $P_t(s_t)$.

2. *Market clears:* $\lambda_t e_{tI}(s_t) + (1 - \lambda_t)e_{tU}(s_t) + x(s_t) = 1, \forall s_t, t$

3. $V_{tU} = V_{tI}$ *if* $\lambda_t \in (0, 1)$*; if* $\lambda_t = 0$*,* $V_{tU} \geq V_{tI}$*; if* $\lambda_t = 1$*,* $V_{tU} \leq V_{tI}$

The last condition guarantees that agents' information choice is optimal. That is, in equilibrium it must be the case that no one is willing to deviate in their information choices. For instance, if there is positive fraction of both informed and uninformed investors ($\lambda_t \in (0, 1)$), it has to be the case that expected utility of the informed and uninformed are equalized.

It is challenging to solve Noisy Rational Expectation Model with general, potentially nonlinear, price functions. Hence in later analysis, I accord with the literature and restrict myself in the class of linear equilibrium.

**Definition 2.2** *A linear equilibrium is an equilibrium where price functions are linear with respect to their arguments. i.e. there exists* $\{A_t, B_t, C_t\}_{t=1}^{T}$ *such that*

$$P_t(\theta, x) = A_t + B_t\theta + C_t x.$$

Except otherwise noted, in later sections I will restrict the attention to linear equilibrium.

# 3 Uniqueness in Finite Horizon Economy

In this section I briefly analyze equilibrium in finite-horizon economy. The main result here is that finite-horizon economy always yields a unique equilibrium. This result is intuitive. Note that when $T = 1$, the economy collapses to the static model used in Grossman and Stiglitz (1980). And it is well-known that there exists a unique equilibrium in that environment.

For $1 < T < \infty$, uniqueness follows directly from backward induction. To see this, note that the generation $T$ agents essentially live in a one-period static world. Thus their decision rule is uniquely pinned down, so as the period-$T$ equilibrium price function. With a uniquely pinned down period-$T$ price function, generation $T-1$ agents also lives in a static world and thus period $T-1$ equilibrium price function is also uniquely pinned down. Thus the backward induction process goes on and as a result equilibrium is unique. I summarize the above reasoning in the following proposition:

**Proposition 3.1** *When $T < \infty$, there is a unique linear equilibrium in the economy with horizon $T$.*

**Proof.** When $T = 1$, the model is identical to that of Grossman and Stiglitz (1980) (with some slight differences on distribution of endowments, these differences are not essential). Thus there is a unique equilibrium when $T = 1$(Grossman and Stiglitz, 1980, Theorem 3 on page 398). When $T > 1$, consider the problem faced by generation $T$ agents. They essentially live in a one-period world and thus the number of informed agents in period $T$, $\lambda_T$, and the price function $P_T(\theta, x_t)$ is uniquely pinned down. Consider the problem faced by generation $T-1$ agents. They also live in a static world with future stock payoff $D(\theta, \varepsilon_t) + P_T(\theta, x_t)$. Thus there is a uniquely pinned down $\lambda_{T-1}$ and $P_{T-1}(\theta, x_t)$. The iteration goes on and therefore there exists a unique equilibrium. ■

# 4 Multiplicity in Infinite Horizon Economy

In this section I analyse the case $T = \infty$. As in the previous section, I allow agents to make information choice at each date. The uniqueness result is overturned. There is a new source of multiplicity: there exists multiple equilibria associated with different equilibrium paths of agents' information choices.

Multiplicity in infinite-horizon overlapping-generation economy is not new (Spiegel, 1998; Watanabe, 2008; Biais et al., 2010; Albagli, 2015). Note that all the works take information choices as exogenous and the multiplicity arises because of agents' forward-looking *portfolio decision*. Here I endogenize the information choice and identify a new source of multiplicity due to the forward-looking nature of *information acquisition decision*.

The backward induction approach is no longer valid as there is no meaningful last period. Thus I take a different approach to characterize equilibria. Before getting into the details, let me define steady states in this economy, which is a subset of equilibria where the equilibrium price functions (and thus agents' information choices) are time-invariant:

**Definition 4.1** *Denote* $s = \{\theta, x\}$*, A steady state of an economy with horizon* $T = \infty$ *is* $\{P(s), \lambda, \{s_i(s), b_i(s)\}_{i=U,I}\}$ *s.t:*

1. $e_i(s), b_i(s)$ *solves uninformed and informed agents' problem given* $P(s)$*.*
2. *Market clears:* $\lambda e_I(s) + (1-\lambda)e_U(s) + x(s) = 1, \forall s_t, t$ *.*
3. $V_U = V_I$ *if* $\lambda \in (0,1)$*; if* $\lambda = 0$*,* $V_U \geq V_I$*; if* $\lambda = 1$*,* $V_U \leq V_I$

**Remark 1** *If* $\{P(s), \lambda, \{s_i(s), b_i(s)\}_{i=U,I}\}$ *is a steady state, then* $\{P(s_t), \lambda_t, \{s_i(s_t), b_i(s_t)\}_{i=U,I}\}_{t=1}^{\infty}$ *is an equilibrium.*

## 4.1 Proving the multiplicity result

The purpose of this section is to establish the existence of multiple steady states associated with different fraction of informed investors in infinite horizon economy(theorem 1). To do so, I take the following steps:

1. First define steady states in an economy where information(i.e. fraction of informed investors $\lambda$) is exogenous. Denote it exogenous-information steady state $\Phi(\lambda)$ (definition 4.2 and definition 7.1).

2. At each $\Phi(\lambda)$ compute the difference of the utility of the informed and uninformed. Denote it value of information $\pi(\lambda)$ (definition 4.4).

3. If $\pi(\lambda)$ is equal to some measure of the utility cost of acquiring information (unless at boundary), $\Phi(\lambda)$ is a steady state (lemma 4.2).

4. If there are multiple such $\lambda$, we find multiple steady states (lemma 4.3 and theorem 1).

**Step 1: Define exogenous-information steady state $\Phi(\lambda)$**

Intuitively, exogenous-information steady state is just steady state in an economy where agents' information choice is *exogenous*, as studied in Spiegel (1998); Watanabe (2008); Biais et al. (2010):

**Definition 4.2** *An exogenous-information steady state given $\lambda$ is $\{P(s), \lambda, \{s_i(s), b_i(s\}_{i=U,I}\}$ such that it satisfies condition 1 and 2 stated in definition A.1.*

As is well known in this literature, there exists multiple exogenous-information linear steady states given any $\lambda$. To make exposition transparent, I will focus on the *low volatility* one[5] in the main part of the analysis. Note that I do not take any stand on which equilibria one *should* select, as both low-volatility and high-volatility equilibria have desirable properties. The purpose of focusing on low-volatility equilibrium is to, loosely speaking, 'fix' the multiplicity arising due to agents' portfolio choice and show that there is a new source of multiplicity associated with agents' information choice.

**Definition 4.3** *A low-volatility exogenous-information steady state $\Phi(\lambda)$ given $\lambda$ is $\{P(s), \lambda, \{s_i(s), b_i(s\}_{i=U,I}\}$ such that*

1. *It is an exogenous-information steady state. In particular, there exists $\{A, B, C\}$ such that*

   $$P(s) = A + B\theta(s) + Cx(s).$$

[5] the low volatility steady state may also be the reasonable one to focus on since it has properties analogous to those found in the infinitely lived agents models of Campbell and Kyle (1993) and Wang (1993, 1994).

*2.*

$$
\begin{array}{rclll}
C & = & \frac{R-\sqrt{R^2-4\alpha^2\sigma_x^2(\sigma_\varepsilon^2+\sigma_\theta^2)}}{2\alpha\sigma_x^2} & \textit{if} & \lambda=0 \\
C & = & \frac{\lambda(B+1)/B-\sqrt{[\lambda(B+1)/B]^2-4\alpha^2\sigma_x^2\sigma_\varepsilon^2}}{2\alpha\sigma_x^2} & \textit{if} & \lambda\in(0,1]
\end{array}
$$

Note that corresponds to each $\lambda$ there is at most a unique low-volatility exogenous-information steady state $\Phi(\lambda)$.

One may wonder about the existence of exogenous-information steady state. As in Spiegel (1998), I provide condition such that exogenous-information steady states exist at least locally near $\lambda = 0$.

**Assumption 1**

$$R^2 - 4\alpha^2\sigma_x^2(\sigma_\varepsilon^2 + \sigma_\theta^2) > 0$$

This is a standard assumption in the literature. The following lemma illustrates the usefulness of assumption 1.

**Lemma 4.1** *Under assumption 1 and for $\lambda$ sufficiently small, a low-volatility exogenous-information steady state $\Phi(\lambda)$ exists.*

**Proof.** See appendix. ■

In later analysis, I assume that assumption 1 holds.

**Step 2: defining value of information $\pi(\lambda)$**

The value of information is the ratio of expected utilities of the informed and uninformed, at each low-volatility exogenous-information steady state indexed by $\lambda$.

**Definition 4.4** *Given $\Phi(\lambda)$, denote the expected utility of the informed $W^I$ and uninformed $W^U$. Define:*

$$\pi(\lambda) = W^U/W^I$$

*Where $W^i, i = I, U$ are given by:*

$$\begin{aligned} W^i &= \int_{\theta,x} W^i(P(\theta,x))df(\theta,x) \\ W^i(P) &= \max_{e,b,c(\theta,\varepsilon',x')} \int_{\theta,\varepsilon',x'} U(c(\theta,\varepsilon',x'))dh(\theta,\varepsilon',x'|I^i) \\ eP+b &\leq w \\ c(\theta,\varepsilon',x') &\leq (D(\theta,\varepsilon')+P(\theta,x')e+Rb) \end{aligned}$$

*Where $I^U = \{P\}, I^I = \{P,\theta\}$.*

**Step 3: Comparing value of information $\pi(\lambda)$ with some measure of information cost**

The next lemma shows that the value of information function $\pi(\lambda)$ allows us to directly compare expected gain from acquiring information to the cost of acquiring information and determine whether $\phi(\lambda)$ is an steady state.

**Lemma 4.2** $\forall\lambda \in (0,1)$, *$\Phi(\lambda)$ is a steady state if and only if:*

$$\pi(\lambda) = \exp(\alpha R\chi)$$

*For $\lambda = 0$ (1), $\Phi(\lambda)$ is a steady state if and only if:*

$$\pi(\lambda) \leq (\geq) \exp(\alpha R\chi)$$

**Proof.** Pick the case $\lambda \in (0,1)$. The other cases are similar to verify. It can be shown that under CARA utility: $V^U = W^U$; $V^I = W^I e^{\alpha R\chi}$. Thus if $\pi(\lambda) = e^{\alpha R\chi}$ holds, then:

$$\frac{V^U}{V^I} = \frac{W^U}{W^U e^{\alpha R\chi}} = \frac{\pi(\lambda)}{e^{\alpha R\chi}} = 1$$

Thus all the conditions for a steady state holds for $\phi(\lambda)$. $\phi(\lambda)$ is a steady state. ■

**Step 4: Proving multiplicity**

The goal of the last step is to show that (under some conditions) there exists multiple values of $\lambda$ satisfying conditions stated in lemma 4.2. Therefore by lemma 4.2, there exists multiple steady states.

To do so, let me state the following lemma:

**Lemma 4.3** *Under some regularity condition B.16,*

$$\frac{d\pi(\lambda)}{d\lambda} > 0 \text{ for } \lambda \text{ sufficiently small}$$

**Proof.** See appendix. ■

The lemma says that incentive of people to become informed, $\pi(\lambda)$, increases as there are more informed investors. This is in sharp contrast with the classical substitution effect in Grossman and Stiglitz (1980).

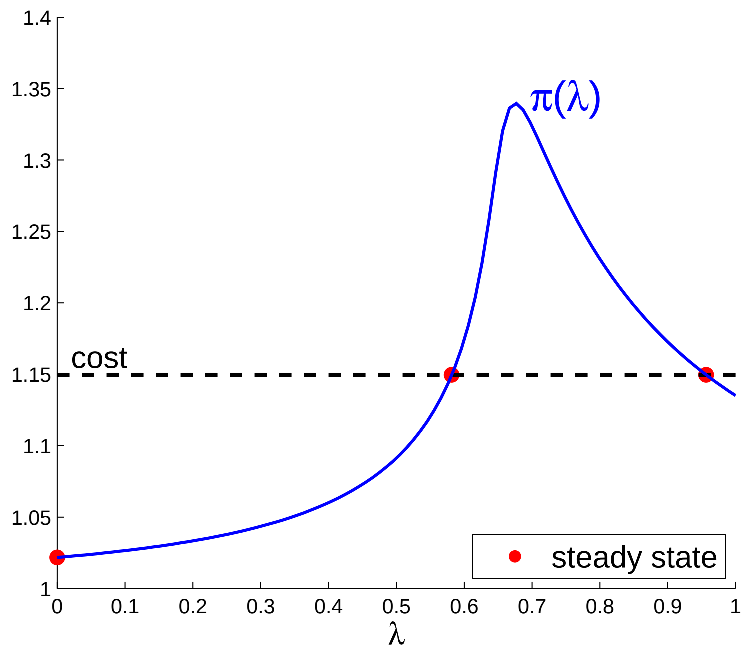
The next theorem establishes the multiplicity result:

**Theorem 1** *Under some regularity condition B.16, there exits $\chi$ such that multiple steady states exist.*

**Proof.** See appendix. ■

**Remark 2** *B.16 is satisfied when either $\sigma_\theta$ or $\sigma_x$ is sufficiently small. It is also satisfied under fairly general parameterizations, including those values used in the literature, say by Veldkamp (2006a).*

In figure 11 I numerically solve and plot function $\pi(\lambda)$. The crucial feature of $\pi(\lambda)$ is that it is monotonically increasing in $\lambda$ for $\lambda$ sufficiently small (lemma 4.3). Pick $\chi$ such that it crosses the upward sloping proportion of $\pi(\lambda)$. This gives us the middle steady state. Now, $\lambda = 0$ is another steady state because the cost of acquiring information is strictly greater than the value of information (lemma 4.2). Thus we obtain another steady state. Lastly, depending on whether $\pi(\lambda)$ may cross the information cost line the second time, we obtain another interior (boundary) steady state. Overall speaking, as long as there is an upward sloping proportion of $\pi(\lambda)$, there exists a




*Parameter values:* $\sigma_\theta^2 = 0.01, \sigma_\varepsilon^2 = 0.3, \sigma_x^2 = 0.4, \chi^U = 1.01$.
*Note that three steady states exist:* $\lambda = 0, 0.74, 1$

Figure 2: Steady-state gain from becoming informed: $\pi(\lambda)$

level of information cost $\chi$ such that multiple steady states exist.

One can see that the crucial step leading up to the multiplicity result is the existence of an upward sloping proportion of $\pi(\lambda)$. Intuitively, this means that information is more valuable when there are more informed investors. This is not straightforward, given that there is static substitutability arising due to market learning. Thus the next section is devoted to explaining agents' incentive to acquire information and in particular why there is an upward sloping proportion of $\pi(\lambda)$.

## 4.2 Inspecting the Mechanism

How do value of information change as the number of informed agents changes? Why is the value of information $\pi(\lambda)$ locally increasing for $\lambda$ sufficiently small? This section is devoted to explaining the intuition.

Suppose you live in an exogenous-information steady state with $\lambda$ number of informed agents. Now consider a perturbation from $\lambda$ to $\lambda + \Delta$. How does this perturbation impact your incentive to acquire information, that is, the value of information? First, there are more informed agents today, so today's price becomes more informative of fundamental $\theta$, reducing the value of information. Second, there are more informed agents tomorrow, so future price also becomes more sensitive to the fundamental, increasing the value of information today. This intuition motivates us to isolate substitutability and complementarity by considering an economy where there are $\lambda_1$ number of informed agents today and $\lambda_2$ number of informed agents from tomorrow on.

**Definition 4.5** $\forall \lambda_1, \lambda_2$, *an equilibrium in an infinite-horizon economy with* $\lambda_1$ *informed agents in period 1 and* $\lambda_2$ *informed agents in later periods is* $\{P_t(s_t), \lambda_t, \{s_{ti}(s_t), b_{ti}(s_t)\}_{i=U,I}\}_{t=1}^{\infty}$ *such that condition 1 and condition 2 in definition 2.1 are satisfied.*

In this environment, we are interested in the value of information for period-1 agents:

**Definition 4.6** *Denote expected utility for period-1 uninformed and informed agents* $W_1^U(\lambda_1, \lambda_2), W_1^I(\lambda_1, \lambda_2)$. *Then the period-1 value of information is defined as:*

$$\pi_d(\lambda_1, \lambda_2) = \frac{W_1^U(\lambda_1, \lambda_2)}{W_1^I(\lambda_1, \lambda_2)}$$

**Remark 3** $\pi(\lambda) = \pi_d(\lambda, \lambda)$. *i.e. Value of information in an exogenous informaiton steady state is just the value of information when there are constant fraction of informed agents both today and tomorrow on.*

The fact that $\pi(\lambda) = \pi_d(\lambda, \lambda)$ allows one to decompose the effect of perturbing $\lambda$ into perturbing today's $\lambda$ and future $\lambda$:

$$\frac{d\pi(\lambda)}{d\lambda} = \underbrace{\frac{\partial \pi_d}{\partial \lambda_1}(\lambda, \lambda)}_{\text{static substitution effect}} + \underbrace{\frac{\partial \pi_d}{\partial \lambda_2}(\lambda, \lambda)}_{\text{dynamic complementary effect}} \tag{4.1}$$

The next proposition makes it clear that locally near $\lambda = 0$, substitutability is dominated by complementarity:

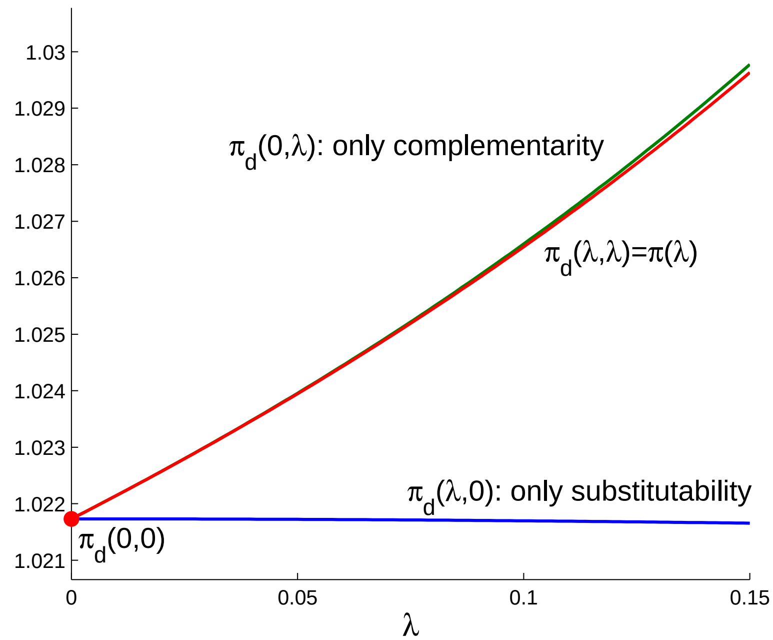



Figure 3: The value of information at the exogenous information steady state with no informed agents is given by: $\pi(0) = \pi_d(0, 0)$. Increasing today's $\lambda$ (first argument of $\pi_d$) reduces value of information (substitutability). Increasing future $\lambda$ (second argument $\pi_d$) increases value of information (complementarity). When varying both today and future $\lambda$, complementarity dominates substitutability. Thus (locally) value of information increases.

**Proposition 4.1** *At $\lambda = 0$:*

$$\frac{\partial \pi_d}{\partial \lambda_1}(\lambda, \lambda) = 0$$

*Under some regularity condition B.16:*

$$\frac{\partial \pi_d}{\partial \lambda_2}(\lambda, \lambda) > 0$$

**Proof.** See appendix. ■

Given equation 4.1 and proposition 4.1, lemma 4.3 can be trivially verified. Proposition 4.1 is illustrated in figure 3.

Why would perturbing future number of informed agents have larger impact on the value of information than perturbing current number of informed agents? To understand this, we need to further characterize the value of information function $\pi_d(\lambda_1, \lambda_2)$. The next lemma states that the value of information is just the relative stock *payoff* (sum of future stock price and future dividend) uncertainty faced by uninformed and informed agents, conditional on their information sets.

**Lemma 4.4**

$$\pi_d(\lambda_1, \lambda_2) = \sqrt{\frac{Var(D_2 + P_2|P_1)}{Var(D_2 + P_2|P_1, \theta)}}$$

**Proof.** See appendix ■

$P_2 + D_2$ is the period-2 stock payoff, including future stock price $P_2$ and $D_2$. Period-1 uninformed agents observe $P_1$, thus $Var(D_2 + P_2|P_1)$ is the stock payoff uncertainty faced by the uninformed agents. Informed agents also observe $\theta$, thus their stock return uncertainty is given by $Var(D_2 + P_2|\theta, P_1)$. Both conditional variances are functions of $\lambda_1, \lambda_2$:

**Lemma 4.5**

$$\begin{aligned} Var(P_2 + D_2|P_1, \theta) &= C_2(\lambda_2)^2\sigma_x^2 + \sigma_\varepsilon^2 \\ Var(P_2 + D_2|P_1) &= \underbrace{(B_2(\lambda_2) + 1)^2\sigma_\theta^2 + C_2(\lambda_2)^2\sigma_x^2 + \sigma_\varepsilon^2}_{Unconditional\ uncertainty} - \underbrace{\Omega(\lambda_1, \lambda_2)[(B_2(\lambda_2) + 1)^2\sigma_\theta^2]}_{Uncertainty\ reduction\ term} \end{aligned}$$

*Where:*

1. $\Omega(\lambda_1, \lambda_2) = \dfrac{B_1(\lambda_1, \lambda_2)^2\sigma_\theta^2}{B_1(\lambda_1, \lambda_2)^2\sigma_\theta^2 + C_1(\lambda_1, \lambda_2)^2\sigma_x^2}$.

2. $B_1(\lambda_1, \lambda_2), C_1(\lambda_1, \lambda_2), B_2(\lambda_2), C_2(\lambda_2)$ *are endogenous coefficients of price function* $P_1, P_2$, *i.e.*$P_1(\theta, x_1) = A_1(\lambda_1, \lambda_2) + B_1(\lambda_1, \lambda_2)\theta + C_1(\lambda_1, \lambda_2)x_1$; $P_2(\theta, x_1) = A_2(\lambda_2) + B_2(\lambda_2)\theta + C_2(\lambda_2)x_2$;

**Proof.** See appendix. ■

The lemma is straightforward to understand. The expression of $P_2 + D_2$ is given by:

$$\begin{aligned} P_2 + D_2 &= \underbrace{A_2 + B_2\theta + C_2x_2}_{P_2} + \underbrace{\theta + \varepsilon_2}_{D_2} \\ &= A_2 + (B_2 + 1)\theta + C_2x_2 + \varepsilon_2 \end{aligned}$$

Thus *unconditional* variance $Var(P_2 + D_2)$ is given by: $(B_2+1)^2\sigma_\theta^2 + C_2^2\sigma_x^2 + \sigma_\varepsilon^2$. For informed agents, they observe $\theta$ perfectly, so what remains uncertain to them is just $x_2$ and $\varepsilon_2$. Thus $Var(P_2 + D_2|P_1, \theta) = C_2^2\sigma_x^2 + \sigma_\varepsilon^2$ (Note that $P_1$ is not useful in predicting $x_2$ or $\varepsilon_2$).

To the extent that $P_1$ is only a noisy signal of $\theta$, uncertainty in $\theta$ cannot be completely wiped out by observing just $P_1$. Specifically, the term $\Omega$ captures how useful $P_1$ is in reducing the uncertainty in $\theta$. $\Omega$ is just equal to *the fraction of variance* of current price $P_1 = A_1 + B_1\theta + C_1x_1$ due to fundamental $\theta$:

$$\Omega = \frac{B_1^2\sigma_\theta^2}{B_1^2\sigma_\theta^2 + C_1^2\sigma_x^2}$$

To understand this, consider extreme cases. Suppose in equilibrium $B_1 = 0, C_1 \neq 0$ . Then $P_1 = A_1 + C_1x_1$. All the variations in $P_1$ is due to noise $x$, thus $\Omega = 0$. According to lemma 4.5, $Var(P_2 + D_2|P_1) = (B_2+1)^2\sigma_\theta^2 + C_2^2\sigma_x^2 + \sigma_\varepsilon^2$, the unconditional variance. If in equilibrium $B_1 \neq 0, C_1 = 0$, then $P_1 = A_1 + B_1\theta$: all the variation in $P_1$ is due to $\theta$. This makes $P_1$ a clean signal of the fundamental, thus $\Omega = 1$. In this case uninformed agents, as informed agents, know perfect the value of $\theta$. Hence as expected $Var(P_2 + D_2|P_1) = C_2^2\sigma_x^2 + \sigma_\varepsilon^2$. More generally, the higher the variance of $P_1$ due to $\theta$, the more informative $P_1$ is, hence the lower the stock payoff uncertainty faced by the uninformed agents. The crucial thing to notice that $\Omega$ is related to the *variance* of $P_1$.

Combining lemma 4.2 and lemma 4.5, we obtain an expression for $\pi_d$:

$$\begin{aligned}
\pi_d(\lambda_1,\lambda_2) &= \sqrt{\frac{Var(D_2+P_2|P_1)}{Var(D_2+P_2|P_1,\theta)}} \\
&= \sqrt{\frac{(B_2(\lambda_2)+1)^2\sigma_\theta^2 + C_2(\lambda_2)^2\sigma_x^2 + \sigma_\varepsilon^2 - \Omega(\lambda_1,\lambda_2)[(B_2(\lambda_2)+1)^2\sigma_\theta^2]}{C_2(\lambda_2)^2\sigma_x^2+\sigma_\varepsilon^2}} \\
&= \sqrt{1+\frac{(1-\Omega(\lambda_1,\lambda_2))[(B_2(\lambda_2)+1)^2\sigma_\theta^2]}{C_2(\lambda_2)^2\sigma_x^2+\sigma_\varepsilon^2}} \\
&= \sqrt{1+\frac{(1-\frac{B_1(\lambda_1,\lambda_2)^2\sigma_\theta^2}{B_1(\lambda_1,\lambda_2)^2\sigma_\theta^2+C_1(\lambda_1,\lambda_2)^2\sigma_x^2})[(B_2(\lambda_2)+1)^2\sigma_\theta^2]}{C_2(\lambda_2)^2\sigma_x^2+\sigma_\varepsilon^2}}
\end{aligned} \tag{4.2}$$

Where the first equality follows from lemma 4.2. Second equality follows from 4.5. Third and last equality follows from simplification.

**Proposition 4.2** *there exists some function* $\Pi : \mathbb{R}^4 \to \mathbb{R}$ *such that:*

$$\pi_d(\lambda_1,\lambda_2) = \Pi(B_1^2, (B_2+1)^2, C_1^2, C_2^2) \tag{4.3}$$

*Where* $B_1 = B_1(\lambda_1,\lambda_2), B_2 = B_2(\lambda_2), C_1 = C_1(\lambda_1,\lambda_2), C_2 = C_2(\lambda_2)$ *are all endogenous coefficients on the price function* $P_1$ *and* $P_2$

**Proof.** Directly follows from the last equality of equation 4.2. ■

Proposition 4.2 lies in the heart of the intuition: value of information is a quadratic function of equilibrium price coefficients $B_1, B_2$, as both substitutability and complementarity work through variance (of stock price or stock payoff). Note that $\lambda_1$ mainly work through $B_1$, the loading coefficient of current price $P_1$, whereas $\lambda_2$ mainly work through $B_2$, the loading coefficient of future price $P_2$. To see how change in $\lambda_1$ and $\lambda_2$ affects value of information, take derivatives (by applying the chain rule) and evaluate the derivative at $\lambda = 0$, focusing on its effect through coefficient $B_1$

and $B_2$[6]:

$$\frac{\partial \pi_d(\lambda,\lambda)}{\partial \lambda_1} \approx \underbrace{\Pi_1}_{-} \underbrace{2B_1}_{=0 \text{ when } \lambda=0} \underbrace{\frac{\partial B_1}{\partial \lambda_1}}_{+} = 0 \tag{4.4}$$

Where $H_1$ is just derivative with respect to its first argument. It is negative capturing the substitutability: when $B_1$ increases, current price becomes more sensitive to $\theta$, value of information decreases. Note that $B_1(\lambda,\lambda) = 0$ when $\lambda = 0$. That is, if there are no informed agents today ($\lambda = 0$), no one knows the true value of $\theta$. Thus equilibrium price will be insensitive to $\theta$: $B_1 = 0$. Thus equation 4.4 makes it clear that since $\pi_d$ is a quadratic function of $B_1$ *passing through the origin*, marginal increase of $B_1$ from 0 has no effect on the value of information.

In contrast, as $\pi_d$ is a quadratic function of $B_1$ *not* passing through the origin, marginal increase of $B_1$ from 0 has nontrivial effect on the value of information.

$$\frac{\partial \pi_d(\lambda,\lambda)}{\partial \lambda_2} \approx \underbrace{\Pi_2}_{+} ( \underbrace{2B_2}_{=0 \text{ when } \lambda=0} +2) \underbrace{\frac{\partial B_2}{\partial \lambda_2}}_{+} > 0 \tag{4.5}$$

Where $H_2$ is just derivative with respect to its second argument. It is positive capturing the complementarity. Why does $(B_2+1)^2$ enter into the value of information instead of $B_2^2$? This is because agents not only care about future stock price, but also future dividends. In other words, they care about total stock payoff $P_2 + D_2 = A_2 + (B_2+1)\theta + C_2 x_2 + \varepsilon_2$. In contrast the current price is just $P_1 = A_1 + B_1\theta + C_1 x_1$. Thus stock payoff has higher loading on $\theta$ than current stock price. This, combined with the fact that variance is a convex function of the loading coefficients $B_1$ and $B_2$, implies that perturbing future price's loading on fundamental $B_2$ has larger impact on the value of information than perturbing current price's loading on fundamental $B_1$. As a result, complementarity may dominate substitutability [7].

[6] For illustrative purpose we ignore the fact that $C_1, C_2$ are also functions of $\lambda$. In the formal proof of proposition 4.1, this is guaranteed by the regularity condition B.16 that the derivative with respect to $C_1$ and $C_2$ are sufficiently small.

[7] One may wonder why substitutability dominates complementarity when $\lambda$ is large. It is due to other terms of the derivative in equation 4.4 and 4.5. In particular, term $\Pi_1$ may be greater than $\Pi_2$, so $\frac{\partial \pi_d(\lambda,\lambda)}{\partial \lambda_1}$ may be greater than $\frac{\partial \pi_d(\lambda,\lambda)}{\partial \lambda_2}$. This effect does not play a major role when $\lambda$ is small, so that $\frac{\partial \pi_d(\lambda,\lambda)}{\partial \lambda_1}$ is of an order of magnitude smaller than $\frac{\partial \pi_d(\lambda,\lambda)}{\partial \lambda_2}$.

# 5 Dynamics

In this section I depart from the steady state analysis and study the dynamics of the model. In the long run, there exists a unique low-volatility steady state to which nonstationary equilibria may converge. At that steady state information acquisitions are complements ($\pi'(\lambda) > 0$). Given the highly nonlinear structure of the model, theoretically I can only characterize the limiting case where $\sigma_\theta$ is taken to some extremely small number. I also verify that the results carry over to more general set of parameters, by conducting numerical experiments. I also numerically characterize the set of nonstaionary equilibria.

The following lemma states that for all equilibria, the (potentially time-varying) coefficients on the equilibrium price function must satisfy a first order difference equation:

**Lemma 5.1** *Suppose* $\{P_t(s_t), \lambda_t, \{s_{ti}(s_t), b_{ti}(s_t)\}_{i=U,I}\}_{t=1}^{\infty}$ *is an equilibrium in the infinite horizon economy. Then the coefficients of the price function* $P_t(s_t) = A_t + B_t\theta(s_t) + C_t x_t(s_t)$ *satisfy a first-order difference equation:*

$$(B_t, C_t) \quad = \quad F(B_{t+1}, C_{t+1})$$

*for some function* $F$*.*

**Proof.** See appendix. ■

Intuitively, as agents only live for two periods, they only care about today and tomorrow's price. Hence given tomorrow's price function $P_{t+1}(\theta, x) = A_{t+1} + B_{t+1}\theta + C_{t+1}x.$, agents optimally chooses whether to become informed or not today. Thus $\lambda_t$ is a function of $B_{t+1}, C_{t+1}$: $\lambda_t = \lambda_t(B_{t+1}, C_{t+1})$. To the extent that today's aggregate demand of stock is a function of $\lambda_t, B_{t+1}, C_{t+1}$, today's equilibrium price is also a function of these variables.That is:

$$(B_t, C_t) = F_2(\lambda_t(B_{t+1}, C_{t+1}), B_{t+1}, C_{t+1})$$

This implicitly defines a transition function from $B_{t+1}, C_{t+1}$ to $B_t, C_t$. Thus we obtain $F$. Given the transition function $F$, one is able to examine stability of different steady states. The notion of stability I adopt is:

**Definition 5.1** *A steady state with price function $P = \tilde{A} + \tilde{B}\theta + \tilde{C}x$ is:*

1. *stable if there exists a two-dimensional neighborhood of $(\tilde{B}, \tilde{C})$ such that for any sequence $\{(B_t, C_t)\}_{t=0}^{\infty}$ that starts in that neighborhood, it converges to $(\tilde{B}, \tilde{C})$ as $t \to \infty$.*
2. *saddle-path stable if there exists a single-dimensional neighborhood (stable manifold) of $(\tilde{B}, \tilde{C})$ such that for any sequence $\{(B_t, C_t)\}_{t=0}^{\infty}$ that starts in that neighborhood, it converges to $(\tilde{B}, \tilde{C})$ as $t \to \infty$.*
3. *unstable if it is not stable nor saddle-path stable.*

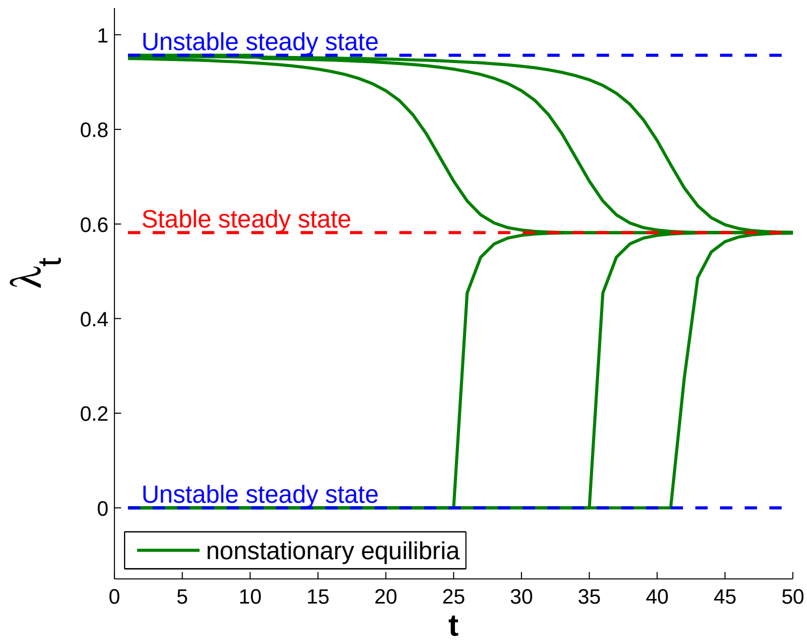
The next theorem shows the stability of different steady states:

**Theorem 2** *Suppose $\sigma_\theta$ is sufficiently small and strictly positive. Then:*

1. *There exists $\chi$ such that three steady states exist with different level of $\lambda$ : $\lambda_1 = 0$,$\lambda_2 \in (0, 1)$,$\lambda_3 = 1$.*
2. *The steady states associated with $\lambda_1$ and $\lambda_3$ are unstable.*
3. *The steady state associated with $\lambda_2$ is either stable or saddle-path stable with $\pi'(\lambda_2) > 0$.*

The theorem is proved under the limiting case where $\sigma_\theta$ tends to 0. The result of the theorem, however, holds under more general parameterizations. In general, if there exists 3 steady states, the steady state with intermediate level $\lambda$ is the only steady state to which nonstationary equilibria may converge (figure 4)[8].

[8] The crucial step of solving for nonstationary equilibria is to approximate the stable manifold $\phi(B_t, C_t)$ around the intermediate steady state using eigenvalues of the Jacobian matrix of the transition function evaluated at the steady state. Details are available upon request.




$\sigma_\theta^2 = 0.01, \sigma_\varepsilon^2 = 0.4, \sigma_x^2 = 0.3, \chi^U = 1.135$

Figure 4: steady states (blue) and nonstationary equilibria(red)

# 6 Applications

## 6.1 The Persistence of Uncertainty after the 2008 Crisis

The dynamic repeated-information-decision framework allows the model to speak to the issue of persistence of information. In particular, the theory provides a novel explanation for the persistently high uncertainty after the 2008 recession based on expectation: if people expects high uncertainty tomorrow (no one acquires information tomorrow), their incentive to learn today is also reduced. Thus the economy may be "trapped" in a high-uncertainty steady state due to pessimistic beliefs of agents about the future. This creates room for policy intervention.

To illustrate the idea, I introduce unexpected belief shocks into the model. The belief shock works as follows. Suppose the economy operates at a steady state where $\lambda = 1$. The steady state is supported by investors' belief that $\lambda = 1$ in the future. If, however, at some date there is a shock to investors' belief of how many informed investors there are in the future, then the economy will

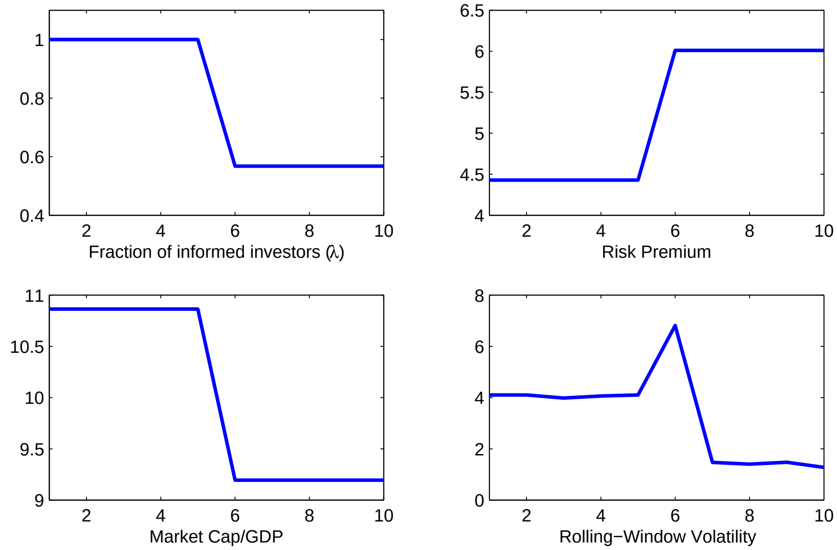



Figure 5: Impulse response to belief shock in period 6

suddenly shift to another steady state with fewer informed investors. As there are less information in this economy, investors in aggregate face more uncertainty. This resembles "uncertainty shock" which drives up risk premium and depresses asset price, leading to a stock market crash. Note that the economy would be trapped at the low-$\lambda$ steady state absent any belief shocks thereafter.

Interestingly, average stock market volatility increases on impact due to the large drop of asset price, but ends up lower than pre-crisis level even though uncertainty remains high. This echoes Fajgelbaum et al. (2014)'s point that stock market volatility might be a poor measure of uncertainty. That is, stock market volatility can be low whereas uncertainty is high.

## 6.2 The decrease of information cost and the growth of passive investing

The past 20 years witnesses substantial transformation of information technology and rapid development of the Internet. This makes more data easily accessible to investors. According to the conventional wisdom (and prediction of the static Grossman and Stiglitz (1980) model), this should lead to more investors acquiring information in the stock market, as information becomes less costly

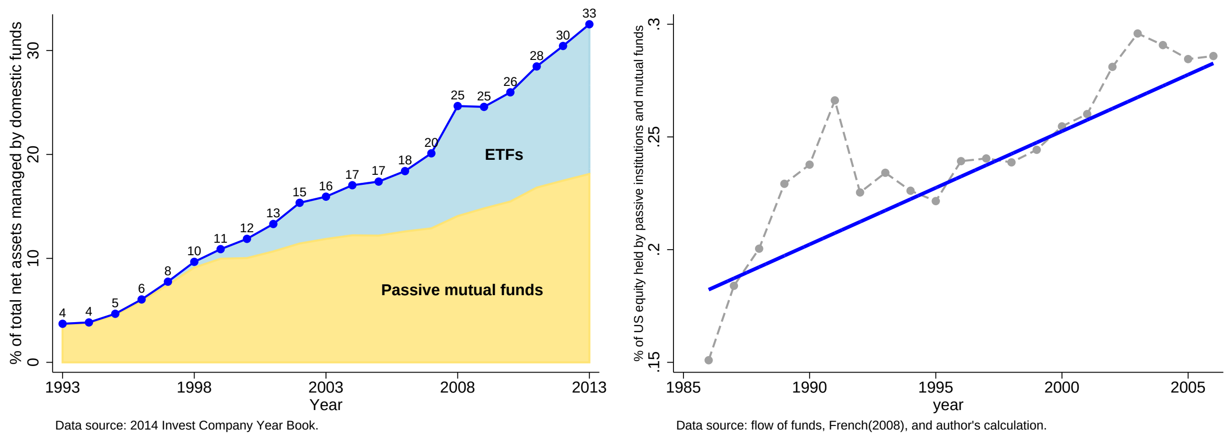



Figure 6: The left panel displays measure 1, total net assets held by ETFs and passive mutual funds as percentage of total net asset held by all domestic mutual funds and ETFs. The right panel displays measure 2, US equity held by passive institutional investors and mutual funds as percentage of US equity held by all institutions and mutual funds.

to acquire.

This prediction is at odds with the data. To show this, I use passive investing v.s. active investing as a proxy for the fraction of informed and uninformed investors (Garleanu. et al. (2015)). I collect data from Invest Company Year Book, Flow of Funds, and French (2008) and obtain two measures of popularity of passive investing: first, total net assets held by ETFs and passive mutual funds as a percentage of total net asset held by all domestic mutual funds; second, US equity held by passive institutional and mutual funds as a percentage of US equity held by all institutions and mutual funds. Both statistics show that passive investing grows relative to active investing (figure 6), implying that agents are acquiring less information despite the fact that information collection has become less costly.

The prediction from the dynamic model is consistent with such observation. As shown in figure 7, a drop of the information cost reduces the number of informed agents at the unique low-volatility stable steady state. This is because value of information $\pi(\lambda)$ is locally increasing at the stable steady state[9].

[9]Here I restrict my attention to low-volatility steady states. One can also examine the stability of different high-volatility steady states. Turns out there exists a unique stable high-volatility steady state. At that high-volatility steady state, a drop of the information cost reduces the number of informed agents. Thus, results are qualitatively unchanged if one studies high-volatility steady states.

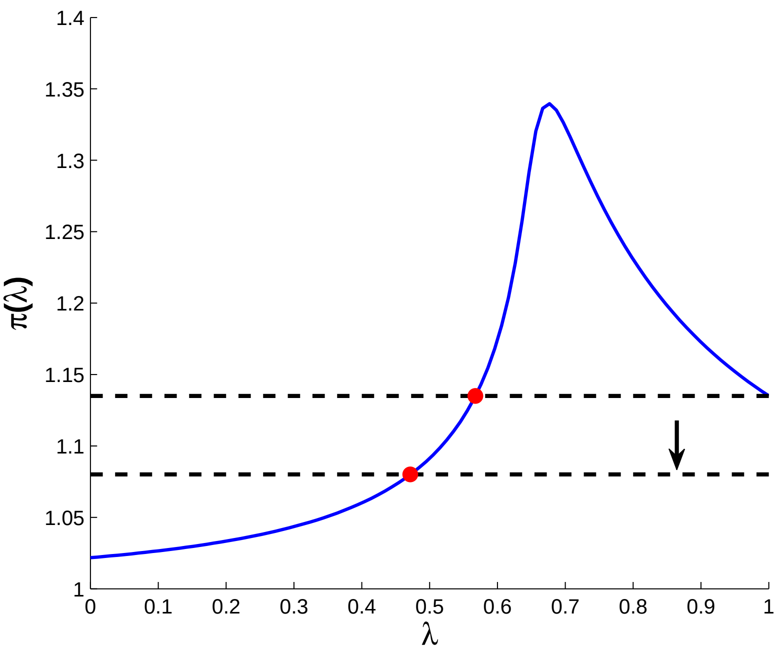



Figure 7: As information cost drops, less agents choose to aquire information.

I further conduct numerical experiment to examine the quantitative potential of the model. To do so, I first need to calibrate the model to some key moments of the US stock market. I set risk free rate $R$ to 1.03 as it is an annual model. I set risk aversion coefficient $\alpha$ to 1.5, as in Veldkamp (2006a). I set variance of noise supply $\sigma_x^2$ to 0.25 as in Easley et al. (2015). There are three parameters left to estimate: prior uncertainty $\sigma_\theta^2$, idiosyncratic noise in dividend payout $\sigma_\varepsilon^2$, and information cost $\chi$. I estimate the three parameters to match: equity premium 6.5% (Campbell and Cochrane, 1999; Bansal and Yaron, 2004), average price dividend ratio 21.1 (Campbell and Cochrane, 1999), and 82% informed investors as in the data(trend) of 1986. The resulting parameters are $\sigma_\theta^2 = 0.002$, $\sigma_\varepsilon^2 = 0.308$, and $\chi_{1986} = 0.00564$, where $\chi_{1986}$ is the (indirectly inferred) information cost in 1986. Note that the model produces a sharpe ratio of 0.89, in the data it is 0.5.

Next, I assume that the information cost decreases at constant annual rate $\mu_\chi$. That is, $\chi_t = (1-\mu_\chi)\chi_{t-1}$. I set $\mu_\chi$ such that the model produces the same fraction of informed investors as in the data(trend) in 1995. The resulting $\mu_\chi$ is 6.35% . Given $\mu_\chi$, I am able to back out values of $\chi_t$ for each $t$. Then I solve for the long-run steady state fraction of informed agents for each $\chi_t$.

| Exogenously Determined | Value | Source |
|---|---|---|
| Risk Free Rate | $R = 1.03$ | Annual Model |
| Risk Aversion | $\alpha = 1.5$ | Veldkamp (2006a) |
| Volatility of Noise Trade | $\sigma_x^2 = 0.25$ | Easley et al. (2015) |
| **Endogenously Determined** | **Value** | **Source** |
| Prior Uncertainty | $\sigma_\theta^2 = 2.00 \times 10^{-3}$ | Mean Price/Dividend Ratio 21.1 |
| Volatility of dividend noise | $\sigma_\varepsilon^2 = 3.08 \times 10^{-1}$ | Equity Premium 6.5% |
| Information cost | $\chi_{1986} = 5.64 \times 10^{-2}$ | 82% active investors |
| **Untargeted Moment** | | |
| Sharpe ratio | 0.89 (model) | 0.5(data) |

Table 1: Parameterization

Figure 8 displays the model-generated growth of passive investing, against the data[10].

I also compute other statistics associated with different information costs(figure 9). As information cost drops, equity premium drops, market capitalization increases. This is because as stock price becomes less sensitive to fundamental, stocks become a safer asset for uninformed agents. As a result, they demand lower equity premium and push up stock price. These predictions are consistent with the data. The model predicts that stock return variance is decreasing. This is hard to identify in the data. The model also predicts that turnover stays roughly constant whereas in the data it increases.

The key channel through which the model successfully replicate the decreasing equity premium and increasing market capitalization as in the data is through decreasing informativeness of stock price. Do we observe that in the data? The literature uses firm-specific return variation as a proxy for informativeness of stock price (Roll (1988); Durnev et al. (2003)). One measure of firm-specific return variation is stock market synchronizability (Morck et al. (1999)). Here I compute average

[10] For simplicity here I do an exercise of comparing across steady states associated with different information cost. It is certainly more desirable to do a full-blown quantitative exercise and solve for a transition path where agents take into account the fact that the information cost is dropping. I left this to future research. Yet the result presented here should be informative of what to expect in the full-blown quantitative exercise.

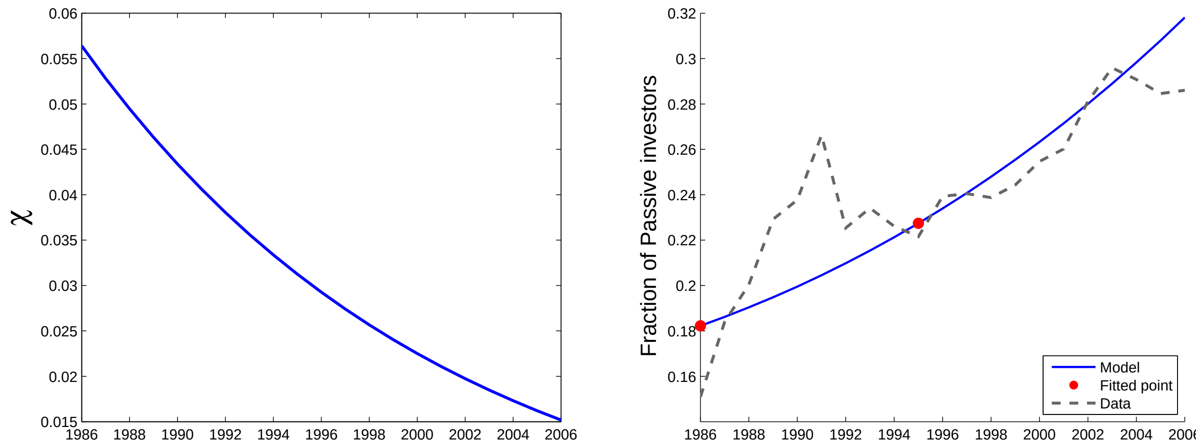



Figure 8: Left panel: implied cost of acquiring information; Right panel: Model-generated growth of passive investors;

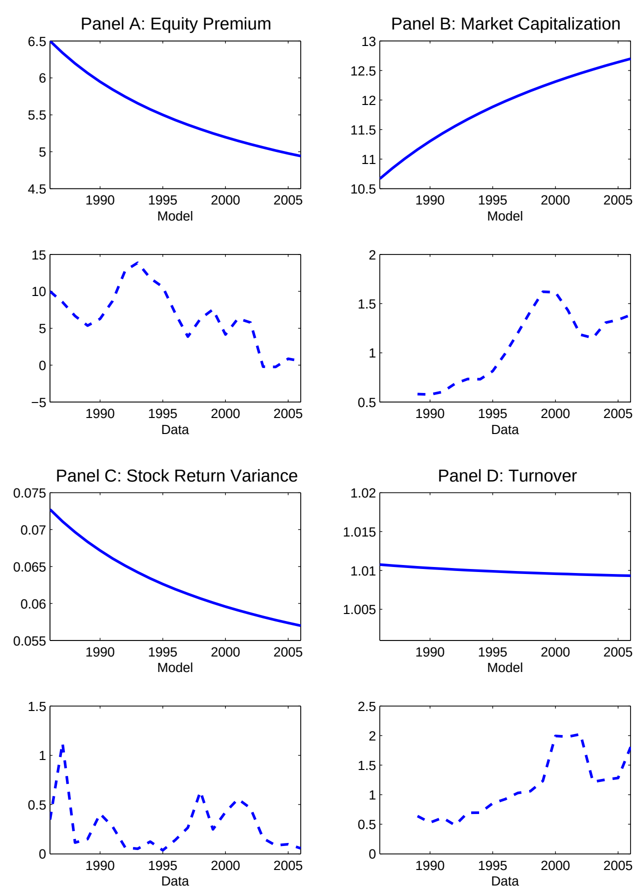



Figure 9

stock return correlation with market return as a measure of stock market synchronizability. A higher stock return correlation implies higher stock market synchronizability. This in turn implies lower firm-specific return variation and hence lower informativeness of stock price. Figure 10 depicts average stock return correlation of individual stocks in SP500 with the index return. Consistent with the finding in Morck et al. (1999), prior to 1990s stock return correlation was decreasing. After the mid-1990s, however, the decreasing trend was muted and finally reverted and stock return correlation started to increase. This provides some direct evidence that stock price did become less informative after the 1990s.

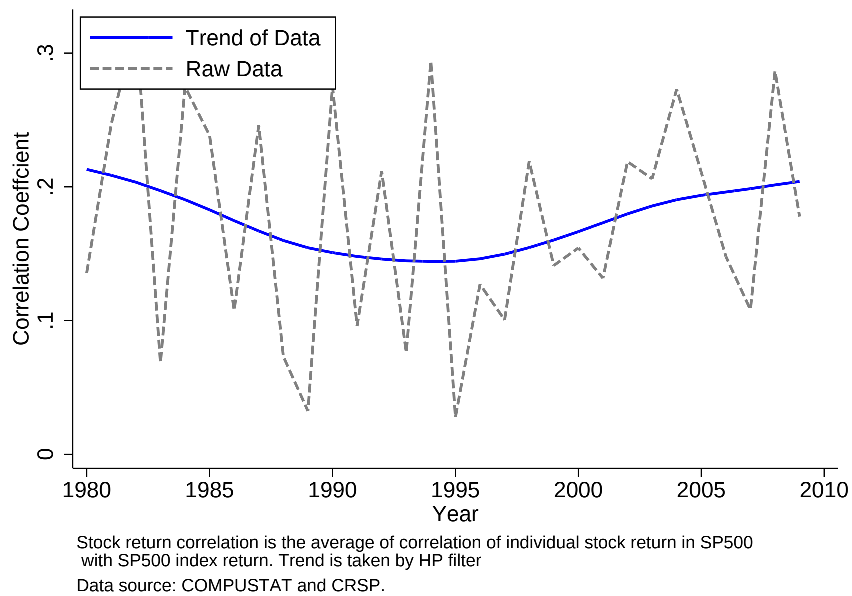



Figure 10: Stock price informativeness (proxied by stock return correlation) decreased since mid-1990s

# 7 Robustness

## 7.1 Model Assumptions

In the baseline model, I assume that the noise supply shock follows an *i.i.d.* process; fundamental is time-invariant; and uninformed agents are born without observing past prices and dividends. All the assumptions are made to keep the analysis simple. Neither of these results is essential. To make the point in a unified framework, in appendix A I describe an economy with random walk noisy supply, mean-reverting fundamental, and uninformed agents endowed with information about the history. I solve the model numerically and show that multiplicity remains in that framework.

**Noise supply shock** Avdis (2014) shows that when the noise supply follows a mean-reverting process, there exists multiplicity. Moreover, there is no multiplicity when the noise supply shock follows a random walk. Similarly, Zhang (2012) only considers *i.i.d.* noise supply. Unlike Avdis (2014) and Zhang (2012), the structure of noise supply does not alter the multiplicity result in this paper. In particular, multiplicity result holds when I replace the assumption of *i.i.d* noise supply with random walk supply [11]. The key difference, is that there is interim dividend payout in my model, which naturally introduces additional loading on fundamental into the stock payoff, making stock payoff more sensitive to fundamental than current stock price. As variance of stock payoff and stock price are convex functions of the loading coefficients, complementarity (which works through stock payoff) may dominate substitutability (which works through current stock price).

**Stochastic fundamental** In the main analysis, fundamental $\theta$ is assumed to be time-invariant, *i.e.* fundamental is extremely persistent. If fundamental is *i.i.d.*, then there is no multiplicity because acquiring information today is not helpful in predicting future price, which is only a function of future fundamental. The magnitude of persistence, however, does not seem to affect the main result. For instance, in appendix A, I set the persistence of fundamental to be as low as 0.1 and show that there still exists multiplicity.

**Information structure** That uninformed agents observe nothing about history is a strong assumption. In fact, the literature that studies infinite-horizon models with information asymmetry typically assumes that uninformed agents observe the past realization of dividends and prices to infer the value of current fundamental (Wang (1993, 1994); Spiegel (1998); Watanabe (2008), among others). Does the multiplicity result follows from the perhaps strong assumption that uninformed agents observe nothing about history? The answer is no. To illustrate, I deliberately endow unin-

[11] One can apply almost the same technique as in the paper to solve for a model with random walk noise supply. The only problem is, now the noise supply shock is a random walk, aggregate stock supply is a nonstationary process and does not have a well-defined unconditional variance. Therefore what one need to do is to follow Watanabe (2008) and assumes that agents observe the last period aggregate supply of stocks. Results are available upon request.

formed agents with much better knowledge of history than what the literature typically assumes. In particular, I assume that uninformed agents know *perfectly* the past realizations of fundamental.

## 7.2 High volatility equilibria

So far we have only considered low-volatility equilibria and show that (theorem 1) there exists some level of information cost $\chi$ such that multiple low-volatility steady states exist. This multiplicity result carries over to high-volatility equilibria. To show this, first define the class of high-volatility exogenous-information steady state:

**Definition 7.1** *A high-volatility exogenous-information steady state* $\Phi^H(\lambda)$ *given* $\lambda$ *is* $\{P(s), \lambda, \{s_i(s), b_i(s\}_{i=U,I}\}$ *such that*

1. *It is an exogenous-information steady state. In particular, there exists* $\{A, B, C\}$ *such that*

   $P(s) = A + B\theta(s) + Cx(s)$.

2. $$\begin{aligned} C &= \frac{R+\sqrt{R^2-4\alpha^2\sigma_x^2(\sigma_\varepsilon^2+\sigma_\theta^2)}}{2\alpha\sigma_x^2} && \text{if} \quad \lambda = 0 \\ C &= \frac{\lambda(B+1)/B+\sqrt{[\lambda(B+1)/B]^2-4\alpha^2\sigma_x^2\sigma_\varepsilon^2}}{2\alpha\sigma_x^2} && \text{if} \quad \lambda \in (0,1] \end{aligned}$$

Note that now $C$ is the positive root to a quadratic equation. Now one can define steady-state value of acquiring information at high-volatility equilibria:

**Definition 7.2** *Given* $\Phi^H(\lambda)$*, denote the expected utility of the informed* $W^{I,H}$ *and uninformed* $W^{U,H}$*. Define:*

$$\pi^H(\lambda) = W^{U,H}/W^{I,H}$$

*Where* $W^{i,H}, i = I, U$ *are given by:*

$$\begin{aligned} W^{i,H} &= \int_{\theta,x} W^{i,H}(P(\theta,x))df(\theta,x) \\ W^{i,H}(P) &= \max_{e,b,c(\theta,\varepsilon',x')} \int_{\theta,\varepsilon',x'} U(c(\theta,\varepsilon',x'))dh(\theta,\varepsilon',x'|I^i) \\ eP+b &\leq w \\ c(\theta,\varepsilon',x') &\leq (D(\theta,\varepsilon')+P(\theta,x')e+Rb) \end{aligned}$$

*Where* $I^U = \{P\}, I^I = \{P, \theta\}$.

**Lemma 7.1** *Under some regularity conditions,*

$$\frac{\partial \pi^H(\lambda)}{\partial \lambda} > 0,$$

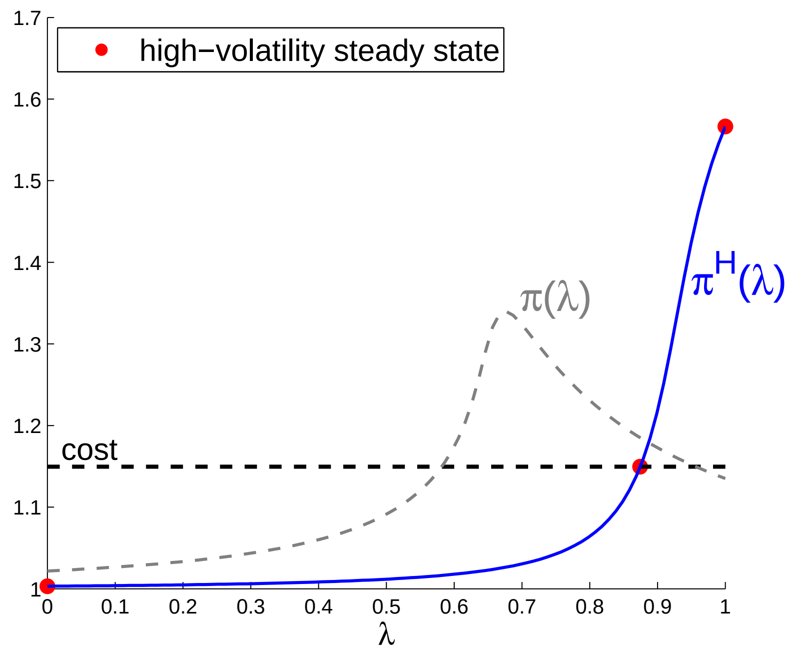
*for* $\lambda$ *sufficiently small.*

**Proof.** The proof closely mirrors the prove of lemma 4.3. There is a slight change of the regularity condition B.16. In particular, instead of plugging in $C = \frac{r - \sqrt{r^2 - 4\alpha^2\sigma_x^2\sigma_\varepsilon^2}}{2\alpha\sigma_x^2}$ into the condition, plug in $C = \frac{r + \sqrt{r^2 - 4\alpha^2\sigma_x^2\sigma_\varepsilon^2}}{2\alpha\sigma_x^2}$. ■

**Theorem 3** *Under some regularity condition, there exists* $\chi$ *such that multiple steady states exist.*

**Proof.** Identical to the proof of theorem 1 given lemma 7.1. ■

To illustrate, I plot $\pi^H(\lambda)$ together with $\pi(\lambda)$. A common feature is that both curves are upward sloping when $\lambda$ is sufficiently small. Thus there exist a level of information cost $\chi$ such that multiple steady states exist.




*Parameter values:* $\sigma_\theta^2 = 0.01, \sigma_\varepsilon^2 = 0.3, \sigma_x^2 = 0.4, \chi^U = 1.01.$

*Note that three steady states exist:* $\lambda = 0, 0.74, 1$

Figure 11: Value of information at high-volatility steady sate $\pi^H(\lambda)$

What are the stability of different high-volatility steady states? One can numerically verify that the middle steady states is stable whereas the other two boundary steady states are saddle-path stable (*i.e.* there exists a single-dimension stable manifold around the steady state).

# 8 Conclusion

In a dynamic environment information acquisitions are not only static substitutes, but also dynamic complements. In this paper I study the dynamic complementarity in information acquisition in an overlapping-generation noisy rational expectation model with endogenous information acquisition. In this environment multiple steady states arise associated with different fraction of informed investors. The crucial step of the proof is to show that the value of acquiring information increases with the mass of the informed when there are few or no informed agents. The multiplicity makes the economy susceptible to belief shocks, potentially leading to huge and persistent spike of uncertainty.

Generally three steady states exist and only the middle steady state is stable. At the middle steady state strategic complementarity dominates substitutability, thus the net effect is that the gain from acquiring information increases with the mass of informed.This implies that Grossman and Stiglitz (1980)'s substitutability result may not be robust to an extension to the dynamic environment. As the model admits a unique stable steady state, there are sharp and interesting long-run implications. In particular, it yields comparative statics results in sharp contrast to that of a static model(Grossman and Stiglitz (1980)), due to the (locally) increasing value of information. I numerically solve for nonstationary equilibria and illustrate that the transition can be used to understand the recent growth of index-related investing (e.g. ETFs). The transition also qualitatively matches several recent trends in equity premium, market capitalization and stock price correlation.

There are a couple of extensions of the model that one may consider. The overlapping generation framework is desirable as it kills heterogeneity across agents. With CARA preference wealth heterogeneity does not matter a lot but information heterogeneity may as the conditional variance enters into agents' asset demand. Thus it is challenging to consider information choice with infinitely-lived agents. As an intermediate step it might be interesting to consider generalized overlapping-generation structures where agents live for more than two periods. We leave it to future research.

# References


Elias Albagli. Investment horizons and asset prices under asymmetric information. *Journal of Economic Theory*, (0):–, 2015. ISSN 0022-0531.

Franklin Allen, Stephen Morris, and Hyun Song Shin. Beauty Contests and Iterated Expectations in Asset Markets. *Review of Financial Studies*, 19(3):719–752, 2006.

Efstathios Avdis. Essays on information acquisition. *Dissertations*, 2012.

Efstathios Avdis. Information trade-offs in dynamic financial markets. *Working Paper*, 2014.

Philippe Bacchetta and Eric Van Wincoop. Can information heterogeneity explain the exchange rate determination puzzle? *American Economic Review*, 96(3):552–576, 2006.

Philippe Bacchetta, CÃ©dric Tille, and Eric van Wincoop. Self-Fulfilling Risk Panics. *American Economic Review*, 102(7):3674–3700, December 2012.

Ravi Bansal and Amir Yaron. Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. *Journal of Finance*, 59(4):1481–1509, 08 2004. URL http://ideas.repec.org/a/bla/jfinan/v59y2004i4p1481-1509.html.

Nicholas Barberis, Andrei Shleifer, and Jeffrey Wurgler. Comovement. *Journal of Financial Economics*, 75(2):283–317, February 2005.

Gadi Barlevy and Pietro Veronesi. Information acquisition in financial markets: a correction. 2007.

Bruno Biais, Peter Bossaerts, and Chester Spatt. Equilibrium Asset Pricing and Portfolio Choice Under Asymmetric Information. *Review of Financial Studies*, 23(4):1503–1543, April 2010.

Nicholas Bloom. The Impact of Uncertainty Shocks. *Econometrica*, 77(3):623–685, 05 2009.

DP Brown and RH Jennings. On technical analysis. *Review of Financial Studies*, 2(4):527–551, 1989.

John Y. Campbell and John Cochrane. Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy*, 107(2):205–251, April 1999.

John Y Campbell and Albert S Kyle. Smart Money, Noise Trading and Stock Price Behaviour. *Review of Economic Studies*, 60(1):1–34, January 1993.

Giovanni Cespa and Xavier Vives. Dynamic trading and asset prices: Keynes vs. hayek. 79(2): 539–580, 2012.

Giovanni Cespa and Xavier Vives. Expectations, liquidity and, short-term trading. *CESifo Working Paper Series No. 3390*, 2014.

Christophe Chamley. Complementarities in information acquisition with short-term trades. *Theoretical Economics*, 2(4), December 2007.

Da and Shive. When the bellwether dances to noise: Evidence from exchange-traded funds. *mimeo*, 2013.

James Dow and Gary Gorton. Arbitrage Chains. *Journal of Finance*, 49(3):819–49, July 1994. URL https://ideas.repec.org/a/bla/jfinan/v49y1994i3p819-49.html.

Artyom Durnev, Randall Morck, Bernard Yeung, and Paul Zarowin. Does Greater Firm-Specific Return Variation Mean More or Less Informed Stock Pricing? *Journal of Accounting Research*, 41(5):797–836, December 2003.

David A. Easley, Maureen O'Hara, and Liyan Yang. Differential access to price information in financial markets. *Journal of Financial and Quantitative Analysis*, 2015.

Pablo Fajgelbaum, Edouard Schaal, and Mathieu Taschereau-Dumouchel. Uncertainty Traps. NBER Working Papers 19973, National Bureau of Economic Research, Inc, March 2014.

Kenneth R. French. Presidential Address: The Cost of Active Investing. *Journal of Finance*, 63 (4):1537–1573, 08 2008.

Kenneth A. Froot, David S. Scharfstein, and Jeremy C. Stein. Herd on the street: Informational inefficiencies in a market with short-term speculation. *The Journal of Finance*, 47(4):1461–1484, 1992. ISSN 1540-6261.

Jayant Vivek Ganguli and Liyan Yang. Complementarities, Multiplicity, and Supply Information. *Journal of the European Economic Association*, 7(1):90–115, 03 2009.

Pingyang Gao. Keynesian Beauty Contest, Accounting Disclosure, and Market Efficiency. *Journal of Accounting Research*, 46(4):785–807, 09 2008.

Nicolae Garleanu., Stavros Panageas, and Jianfeng Yu. Impediments to financial trade: Theory and measurement. *Working Paper*, 2015.

Sanford J Grossman and Joseph E Stiglitz. On the impossibility of informationally efficient markets. *American Economic Review*, 70(3):393–408, June 1980.

Carolina Manzano and Xavier Vives. Public and private learning from prices, strategic substitutability and complementarity, and equilibrium multiplicity. *Journal of Mathematical Economics*, 47 (3):346 – 369, 2011. Mathematical Economics {II} : Special Issue in honour of Andreu Mas-Colell.

Randall Morck, Bernard Yeung, and Wayne Yu. The Information Content of Stock Markets: Why Do Emerging Markets Have Synchronous Stock Price Movements? 1999.

Richard Roll. R2. *The Journal of Finance*, 43(3):541–566, 1988. ISSN 1540-6261. doi: 10.1111/j.1540-6261.1988.tb04591.x. URL http://dx.doi.org/10.1111/j.1540-6261.1988.tb04591.x.

Matthew Spiegel. Stock Price Volatility in a Multiple Security Overlapping Generations Model. *Review of Financial Studies*, 11(2):419–47, 1998.

Robert F. Stambaugh. Presidential Address: Investment Noise and Trends. *Journal of Finance*, 69(4):1415–1453, 08 2014.

Nancy L. Stokey, Robert E. Lucas, Jr., and Edward C. Prescott. *Recursive Methods in Economic Dynamics*. Harvard University Press, Cambridge, MA, 1989.

Laura L. Veldkamp. Information Markets and the Comovement of Asset Prices. *Review of Economic Studies*, 73(3):823–845, 2006a.

Laura L. Veldkamp. Media frenzies in markets for financial information. *American Economic Review*, 96(3):577–601, 2006b. doi: 10.1257/aer.96.3.577.

Robert E Verrecchia. Information Acquisition in a Noisy Rational Expectations Economy. *Econometrica*, 50(6):1415–30, November 1982.

Jiang Wang. A Model of Intertemporal Asset Prices under Asymmetric Information. *Review of Economic Studies*, 60(2):249–82, April 1993.

Jiang Wang. A Model of Competitive Stock Trading Volume. *Journal of Political Economy*, 102 (1):127–68, February 1994.

Masahiro Watanabe. Price Volatility and Investor Behavior in an Overlapping Generations Model with Information Asymmetry. *Journal of Finance*, 63(1):229–272, 02 2008.

Jeffrey Wurgler. On the economic consequences of index-linked investing. *NBER working paper*, 2010.

Jinfan Zhang. Learning, short horizon, and asset pricing. *Working Paper*, 2012.

# A Appendix: An economy with mean-reverting fundamental, random walk noise supply, and uninformed observing the history

In this section I present a model that relaxes a number of assumptions in the baseline model. I show that in such an environment, multiplicity remains. Unfortunately, the model is hard to characterize theoretically due to its complexity. Therefore I comply with the literature and solve for the steady states numerically.

Time is discrete and runs from negative infinity to positive infinity. As in the baseline model, there is a single consumption goods that agents treat as numeraire. There are two assets. The bond is in perfect elastic supply and the stock pays out dividends every period:

$$D_t = \theta_t + \varepsilon_t$$

, where $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$.

$\theta_t$ follows:

$$\theta_{t+1} = \rho\theta_t + \varepsilon_t^\theta$$

where $\varepsilon_t^\theta \sim \mathcal{N}(0, \sigma_\theta^2)$. I follow Watanabe (2008) and Spiegel (1998) and assume that the aggregate supply of stock follows a random walk process:

$$N_t = N_{t-1} + \eta_t$$

One interpretation is that every period there are noise traders that demand random amount of stock $\eta_t$. There are overlapping-generations of investors who acquires information and trade. Specifically, in period $t$:

1. $\varepsilon_t$ is realized. Dividend $D_t = \theta_t + \varepsilon_t$ is paid.
2. $\varepsilon_t^\theta$ is realized. Fundamental $\theta_{t+1} = \rho\theta_t + \varepsilon_t^\theta$ is also realized.
3. Generation-$t$ agents are born. They are endowed with $w$ units of bonds, CARA utility, and information set $\{D_t\} \cup \{\theta_{t-j}, N_{t-j}\}_{\infty}^{j=1}$
4. Generation-$t$ agents choose whether to acquire information. At cost $\chi$, they also observe $\{\theta_t, \varepsilon_t^\theta\}$
5. Market opens, generation-$t$, $t-1$ agents and the noise traders trade on the financial market. Everyone observes price $P_t$.
6. Generation $t-1$ agents exit and consume their wealth.

Note that in the model, fundamental is stochastic and noise supply follows random walk. Uninformed agents are allowed to observe historical realizations of $\theta_t$. Given the concern that the main result in the baseline model may be driven by the perhaps strong assumption that uninformed agents observe nothing about the history, here I deliberately endow them with much better knowledge of the history than what the literature typically assumes. The usual assumption made in the literature is that agents observe dividends and prices to infer historical realizations of fundamental.

Here I endow the uninformed agents more information by letting them know the true historical realizations of fundamental.

As in the baseline model, each generation of agents make two choices sequentially. First, they make their information acquisition choice:

$$V_t \;=\; \max\{V_t^I, V_t^U\}$$

Where $V_t^I$ denotes expected utility for agents acquiring information (later I call them informed agents); $V_t^U$ denotes expected utility for agents not acquiring information (later I call them uninformed agents). $V_t^I$ and $V_t^U$ are determined by agents' portfolio choice:

$$\begin{aligned}
V_t^i &= \textstyle\int_P V_t^i(P)dF_t(P) \\
V_t^i(P) &= \max_{e_t,b_t,c(P_{t+1},\theta_{t+1},\varepsilon_{t+1})} \textstyle\int_{P_{t+1},\theta_{t+1},\varepsilon_{t+1}} U(c(P_{t+1},\theta_t,\varepsilon_{t+1}))dH_t(P_{t+1},\theta_{t+1},\varepsilon_{t+1}|I_t^i) \\
e_tP_t + b_t &\leq w - \mathbb{1}\{i = I\}\chi \\
c(P_{t+1},\theta,\varepsilon') &\leq (D(\theta_{t+1},\varepsilon_{t+1}) + P_{t+1})e + Rb_t
\end{aligned}$$

Where $I_t^U = \{P_t, D_t\} \cup \{\theta_{t-j}, N_{t-j}\}_{\infty}^{j=1}$, $I_t^I = \{\theta_t, \varepsilon_t^\theta\} \cup I_t^U$, $U(c) = -\exp(-\alpha c)$, $\alpha$ is the risk averse parameter. $D(\theta_{t+1}, \varepsilon_{t+1}) = \theta_{t+1} + \varepsilon_{t+1}$, $F_t, H_t$ are equilibrium objects as price is determined in general equilibrium.

I take a guess and verify strategy and conjecture that the equilibrium price function takes the form:

$$P_t = a + b_1\theta_t + b_2\theta_{t-1} + b_3\varepsilon_t^\theta + cD_t + d\eta_t + eN_{t-1}$$

Where $a, b_1, b_2, b_3, c, d, e$ are scalars. Thus one can define a steady state in this environment:

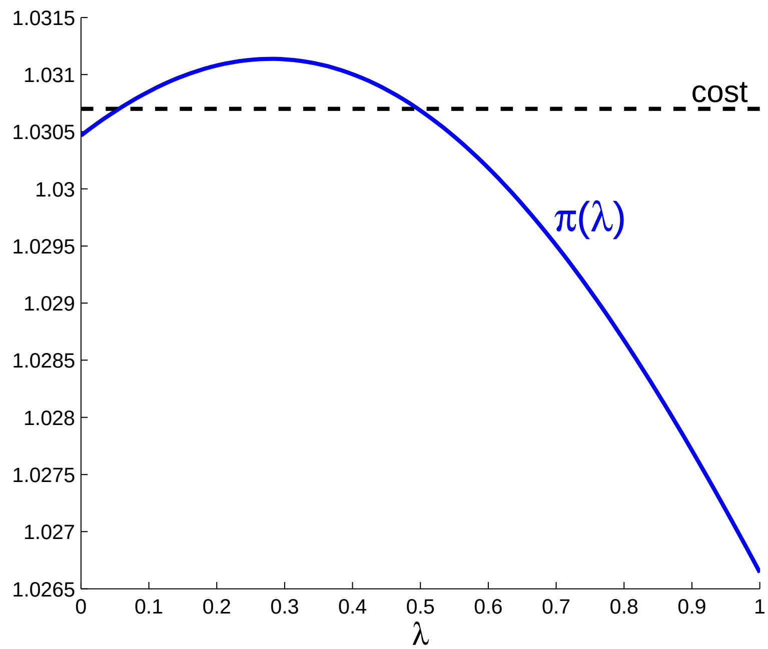
**Definition A.1** *Denote* $s = \{\theta_t, \theta_{t-1}, \varepsilon_t^\theta, \eta_t, N_{t-1}\}$*, A steady state is* $\{P(s), \lambda, \{s_i(s), b_i(s)\}_{i=U,I}\}$ *s.t:*

1. $e_i(s), b_i(s)$ *solves uninformed and informed agents' problem given* $P(s)$*.*
2. *Market clears:* $\lambda e_I(s) + (1-\lambda)e_U(s) = N_{t-1} + \eta_t, \forall s_t, t$ *.*
3. $V_U = V_I$ *if* $\lambda \in (0,1)$*; if* $\lambda = 0$*,* $V_U \geq V_I$*; if* $\lambda = 1$*,* $V_U \leq V_I$

Where $V_U$ and $V_I$ solves the uninformed and informed agents' problem respectively. The individual agents' problems are analogous to that stated previously.

To solve for steady state, I take the approach described in section 4.1 and solve for the value of information $\pi(\lambda)$ for each $\lambda$. I set the persistence parameter $\rho$ to be as low as 0.1 to show that the multiplicity result is not sensitive to the choice of persistence parameter $\rho$. Figure 12 displays the result.

The intuition is that more informed agents tomorrow increases the predictive power of $\varepsilon_t^\theta$. To see this, note that tomorrow's informed agents know perfectly the value of $\theta_{t+1}$ whereas uninformed agents only observes some noisy signal to infer the value of $\theta_{t+1}$. Thus, when there are more informed agents tomorrow, tomorrow's price $P_{t+1}$ will be more heavily loaded on $\theta_{t+1}$, and thus on $\varepsilon_t^\theta$. This increase agents' incentive to learn $\varepsilon_t^\theta$ today.




*Parameter values:* $\sigma_\theta^2 = 0.01, \rho = 0.1, \sigma_\eta^2 = 0.3, \sigma_\varepsilon^2 = 0.3$

Figure 12: Value of information in the extended model: $\pi(\lambda)$

# B Appendix: Proofs

## Useful Result:

**Proposition B.1** *At exogenous-information steady state $\phi(\lambda)$, the value of information*

$$\pi(\lambda) = \frac{W^U}{W^I} = \sqrt{\frac{Var(P(s') + D(s')|P)}{Var(P(s') + D(s')|P, \theta)}}$$

**Proof.** This is an extension of the Theorem 2 in Grossman and Stiglitz (1980). It states that the value of information, defined as ratio of expected utility for the uninformed and informed, is the relative stock payoff uncertainty of the uninformed and informed agents. Note that $P(s')$ denotes next period equilibrium price whereas $P$ denotes current price.

Suppose equilibrium price $P(s) = A + B\theta(s) + Cx(s)$. Simplify the budget constraint: $c(s') = (D(s') + P'(s') - RP)e$. Plug into utility function, we obtain that the expected utility of each type of agents after market opens:

$$W^i(P) = \max_e \int_{s'} U((D(s') + P'(s') - RP)e)dH(s'|I^i)$$

Given CARA utility:

$$\begin{aligned} W^i(P) &= \max_e \int_{s'} U((D(s')+P'(s')-RP)e)dH(s'|I^i) \\ &= \max_e \int_{s'} -e^{(-\alpha((D(s')+P'(s')-RP)e))}dH(s'|I^i) \\ &= \max_e -\exp[E[-\alpha((D(s')+P'(s')-RP)e)|I^i] + \tfrac{1}{2}Var(-\alpha((D(s')+P'(s')-RP)e)|I^i)] \\ &= \max_e -\exp[-\alpha(E[D(s')+P'(s')-RP|I^i]e - \tfrac{1}{2}\alpha e^2 Var(D(s')+P'(s')-RP|I^i))] \end{aligned} \tag{B.1}$$

hence maximizing over the objective function is equivalent to maximizing:

$$\max_e E[D(s')+P'(s')-RP|I^i]e - \frac{1}{2}\alpha e^2 Var(D(s')+P'(s')-RP|I^i)$$

Solve for optimal $s^*$:

$$e^* = \frac{E[D(s')+P'(s')-RP|I^i]}{\alpha Var(D(s')+P'(s')-RP|I^i)}$$

Plug back to the original objective function:

$$\begin{aligned} W^i(P) &= -\exp[-\tfrac{1}{2}\alpha\tfrac{(E[D(s')+P'(s')-RP|I^i])^2}{\alpha Var(D(s')+P'(s')-RP|I^i)}] \\ &= -\exp[-\tfrac{1}{2}\tfrac{(E[D(s')+P'(s')|I^i]-RP)^2}{Var(D(s')+P'(s')|I^i)}] \end{aligned} \tag{B.2}$$

Where the second equation follows because $P$ is realized at this stage. Let:

$$h = Var(D(s')+P'(s')|I^U) - Var(D(s')+P'(s')|I^I) > 0$$

The reason why it is greater than 0 is that uninformed has residual uncertainty over $\theta$ whereas the informed are perfectly informed about $\theta$. Taking conditional expectation of the informed $W_I(P)$ of the uninformed agents' information set:

$$\begin{aligned} E[W^i(P)|I^U] &= E[-e^{-\frac{1}{2}\frac{(E[D(s')+P'(s')|I^I]-RP)^2}{Var(D(s')+P'(s')|I^I)}}|I^U] \\ &= E[-e^{-\frac{1}{2}\frac{(E[D(s')+P'(s')|I^I]-RP)^2}{h}\frac{h}{Var(D(s')+P'(s')|I^I)}}|I^U] \\ &= E[-e^{-\frac{1}{2}\frac{h}{Var(D(s')+P'(s')|I^I)}z^2}|I^U] \end{aligned} \tag{B.3}$$

Where $z = \frac{(E[D(s')+P'(s')|I^I]-RP)}{\sqrt{h}}$.

Thus by moment generating function of a non-central chi-squared distribution (formula $A21$ of Grossman and Stiglitz (1980)):

$$\begin{aligned} E[W^i(P)|I^U] &= \frac{1}{\sqrt{1+\frac{h}{Var(D(s')+P'(s')|I^I)}}}\exp(\frac{-E[z|I^U]^2\frac{1}{2}\frac{h}{Var(D(s')+P'(s')|I^I)}}{1+\frac{h}{Var(D(s')+P'(s')|I^I)}}) \\ &= \sqrt{\frac{Var(D(s')+P'(s')|I^I)}{Var(D(s')+P'(s')|I^U)}}\exp(\frac{-E[z|I^U]^2\frac{1}{2}\frac{h}{Var(D(s')+P'(s')|I^i)}}{1+\frac{h}{Var(D(s')+P'(s')|I^i)}}) \\ &= \sqrt{\frac{Var(D(s')+P'(s')|I^I)}{Var(D(s')+P'(s')|I^U)}}\exp(\frac{-\frac{1}{2}(E[D(s')+P'(s')|I^U]-RP)^2}{Var(D(s')+P'(s')|I^U)}) \\ &= \sqrt{\frac{Var(D(s')+P'(s')|I^I)}{Var(D(s')+P'(s')|I^U)}}W_U(P) \end{aligned} \tag{B.4}$$

Integrate on both sides with respect to current state $s$, one get:

$$W_I = \sqrt{\frac{Var(D(s')+P'(s')|I^I)}{Var(D(s')+P'(s')|I^U)}} W_U \tag{B.5}$$

As $I^I = \{P, \theta\}, I^U = \{P\}$,

$$\pi(\lambda) = \frac{W_U}{W_I} = \sqrt{\frac{Var(D(s') + P(s')|P)}{Var(D(s') + P(s')|P, \theta)}}$$

■

## Proof of lemma 4.1

It suffices to show that coefficients $B, C$ of the price function exist when $\lambda$ is sufficiently small, as all the other objects can be easily constructed. As shown in lemma 4.3, the coefficients satisfies the following system of equation:

$$\begin{bmatrix} B \\ C \end{bmatrix} - \begin{bmatrix} \frac{1}{R}\lambda (B+1) L \\ \frac{1}{R}\alpha \left[C^2\sigma_x^2 + \sigma_\varepsilon^2\right] L \end{bmatrix} = 0$$

Where functon $L$ is defined below. When $\lambda \to 0$, $L \to 1 + \frac{(B+1)^2\sigma_\theta^2}{C^2\sigma_x^2+\sigma_\varepsilon^2} > 0$, thus the system of equation converge to:

$$\begin{bmatrix} B \\ C \end{bmatrix} - \begin{bmatrix} 0 \\ \frac{1}{R}\alpha \left[C^2\sigma_x^2 + \sigma_\varepsilon^2\right] [1 + \frac{(B+1)^2\sigma_\theta^2}{C^2\sigma_x^2+\sigma_\varepsilon^2}] \end{bmatrix} = 0$$

Thus

$$\begin{aligned} B &\to 0 \\ C &\to \frac{R - \sqrt{[R]^2 - 4\alpha^2\sigma_x^2(\sigma_\varepsilon^2+\sigma_\theta^2)}}{2\alpha\sigma_x^2} \end{aligned}$$

Thus under assumption 1, B,C are well defined for $\lambda$ sufficiently small.

## Proof of lemma 4.3

Period t investors' portfolio choice problem is given by:

$$\max_e E\left[U\left(e\left(D_{t+1} + P_{t+1} - RP_t\right)\right)|I\right]$$

Where $I$ denotes his information set.

Given CARA utility function, this is equivalent to maximizing:

$$\max_e E\left[\left(D_{t+1} + P_{t+1} - RP_t\right)|I\right] e - \frac{1}{\alpha} e^2 Var\left[D_{t+1} + P_{t+1} - RP_t | I\right]$$

Hence agents demand functions are given by:

$$e_t^* = \frac{E\left[\left(D_{t+1} + P_{t+1} - RP_t\right)|I\right]}{\alpha Var\left[D_{t+1} + p_{t+1} - RP_t|I\right]} = \frac{E\left[\left(D_{t+1} + P_{t+1}\right)|I\right] - RP_t}{\alpha Var\left[D_{t+1} + p_{t+1}|I\right]}$$

Write out explicitly the formula for tomorrow's dividend and stock price:

$$\begin{aligned} D_{t+1} + P_{t+1} &= \theta + \varepsilon_{t+1} + A_{t+1} + B_{t+1}\theta + C_{t+1}x_{t+1} \\ &= A_{t+1} + \left(B_{t+1} + 1\right)\theta + \varepsilon_{t+1} + C_{t+1}x_{t+1} \end{aligned}$$

We first look for demand for the informed agents $e_{tI}$. Given his information set $I^I = \{P_t, \theta\}$, the conditional mean and variance of stock payoff $P_{t+1} + D_{t+1}$ is given by:

$$\begin{aligned} E\left[\left(D_{t+1} + P_{t+1}\right)|I^I\right] &= A_{t+1} + \left(B_{t+1} + 1\right)\theta \\ Var\left[\left(D_{t+1} + P_{t+1}\right)|I^I\right] &= \sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2 \end{aligned}$$

Plug in the formula into informed demand, we get:

$$e_{tI} = \frac{A_{t+1} + \left(B_{t+1} + 1\right)\theta - RP_t}{\alpha\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]}$$

From the market clearing condition:

$$\lambda_t e_{tI} + \left(1 - \lambda_t\right)e_{tU} + x_t = 1$$

Plug in

$$e_{tI} = \frac{A_{t+1} + \left(B_{t+1} + 1\right)\theta - RP_t}{\alpha\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]}$$

we get

$$\lambda_t \frac{A_{t+1} + \left(B_{t+1} + 1\right)\theta - RP_t}{\alpha\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]} + \left(1 - \lambda_t\right)e_{tU} + x_t = 1$$

Rearrange

$$\lambda_t\left(B_{t+1} + 1\right)\theta + \alpha\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]x_t = \alpha\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right] - \alpha\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]\left(1 - \lambda_t\right)e_{tU} - \lambda_t\left(A_{t+1} + RP_t\right)$$

uninformed agents understands the structure of the equilibrium, thus he knows $\lambda_t$ and $s_{tU}$. They also observe $P_t$. Thus knowing $P_t$ is equivalent to knowing the right hand side of the equation, equivalent to knowing the left hand side of the equation, which serves as a noisy signal of $\theta, s\left(P_t\right)$ :

$$s\left(P_t\right) = \lambda_t\left(B_{t+1} + 1\right)\theta + \alpha\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]x_t$$

Given this signal, uninformed agents do Bayesian updating as follows. Random variable $D_{t+1}+P_{t+1}$

and $s(P_t)$ are jointly normal with the following mean and variance matrix:

$$\left[\begin{array}{c} D_{t+1}+P_{t+1} \\ s(P_t) \end{array}\right] = \left[\begin{array}{c} A_{t+1}+(B_{t+1}+1)\theta+\varepsilon_{t+1}+C_{t+1}x_{t+1} \\ \lambda_t(B_{t+1}+1)\theta+\alpha\left[\sigma_\varepsilon^2+C_{t+1}^2\sigma_x^2\right]x_t \end{array}\right]$$
$$\sim \mathbb{N}\left(\mu^U,\Sigma^U\right)$$

Where:

$$\mu^U = \left[\begin{array}{c} A_{t+1}+(B_{t+1}+1)\mu \\ \lambda_t(B_{t+1}+1)\mu \end{array}\right]$$
$$\Sigma^U = \left[\begin{array}{cc} (B_{t+1}+1)^2\sigma_\theta^2+\sigma_\varepsilon^2+C_{t+1}^2\sigma_x^2 & \lambda_t[B_{t+1}+1]^2\sigma_\theta^2 \\ \lambda_t[B_{t+1}+1]^2\sigma_\theta^2 & \lambda_t^2(B_{t+1}+1)^2\sigma_\theta^2+\alpha^2\left[\sigma_\varepsilon^2+C_{t+1}^2\sigma_x^2\right]^2\sigma_x^2 \end{array}\right]$$

Thus from conditional expectation formula for normal variables, the conditional mean of $P_{t+1}+D_{t+1}$ is given by:

$$E[D_{t+1}+P_{t+1}|P_t] = A_{t+1}+(B_{t+1}+1)\mu+M^U\left[\lambda_t(B_{t+1}+1)\theta+\alpha\left[\sigma_\varepsilon^2+C_{t+1}^2\sigma_x^2\right]x_t\right] \quad \text{(B.6)}$$

Where $M^U = \frac{\lambda_t[B_{t+1}+1]^2\sigma_\theta^2}{\lambda_t^2(B_{t+1}+1)^2\sigma_\theta^2+\alpha^2\left[\sigma_\varepsilon^2+C_{t+1}^2\sigma_x^2\right]^2\sigma_x^2}$

The conditional variance of $P_{t+1}+D_{t+1}$ is given by:

$$Var[D_{t+1}+P_{t+1}|P_t] = (B_{t+1}+1)^2\sigma_\theta^2+\sigma_\varepsilon^2+C_{t+1}^2\sigma_x^2-\frac{\left[\lambda_t[B_{t+1}+1]^2\sigma_\theta^2\right]^2}{\lambda_t^2(B_{t+1}+1)^2\sigma_\theta^2+\alpha^2\left[\sigma_\varepsilon^2+C_{t+1}^2\sigma_x^2\right]^2\sigma_x^2} \quad \text{(B.7)}$$

Thus we get the demand of the uninformed:

$$s_{tU} = \frac{E[D_{t+1}+P_{t+1}|P_t]-RP_t}{\alpha Var[D_{t+1}+P_{t+1}|P_t]}$$

Plug the demand functions into the market clearing condition:

$$\lambda_t\frac{A_{t+1}+(B_{t+1}+1)\theta-RP_t}{\alpha\left[\sigma_\varepsilon^2+C_{t+1}^2\sigma_x^2\right]}+(1-\lambda_t)\frac{E[D_{t+1}+P_{t+1}|P_t]-RP_t}{\alpha Var[D_{t+1}+P_{t+1}|P_t]}+x_t=1$$

Where $E[D_{t+1}+P_{t+1}|P_t]$ and $Var[D_{t+1}+P_{t+1}|P_t]$ are given by B.6 and B.7 respectively. Rearrange this expression, we get a linear expression for price functin $P_t$:

$$P_t = A_t+B_t\theta+C_tx_t$$

Where

$$A_t = v^{-1}\left[\lambda_t \frac{A_{t+1}}{\alpha\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]} + (1-\lambda_t)\frac{A_{t+1} + (B_{t+1}+1)\mu}{\alpha V_t^I} - 1\right] \tag{B.8}$$

$$B_t = v^{-1}\left[\lambda_t \frac{(B_{t+1}+1)}{\alpha\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]} + (1-\lambda_t)\frac{M^U\left[\lambda_t(B_{t+1}+1)\right]}{\alpha V_t^I}\right] \tag{B.9}$$

$$C_t = v^{-1}\left[(1-\lambda_t)\frac{M^U\left[\alpha\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]\right]}{\alpha V_t^I} + 1\right] \tag{B.10}$$

Where:

$$V_t^U = (B_{t+1}+1)^2\sigma_\theta^2 + \sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2 - \frac{\left[\lambda_t\left[B_{t+1}+1\right]^2\sigma_\theta^2\right]^2}{\lambda_t^2(B_{t+1}+1)^2\sigma_\theta^2 + \alpha^2\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]^2\sigma_x^2}$$

$$M_t^U = \frac{\lambda_t\left[B_{t+1}+1\right]^2\sigma_\theta^2}{\lambda_t^2(B_{t+1}+1)^2\sigma_\theta^2 + \alpha^2\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]^2\sigma_x^2}$$

$$v = R\left(\frac{\lambda_t}{\alpha\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]} + \frac{(1-\lambda_t)}{\alpha V_t^I}\right)$$

Define the relative uncertainty of uninformed and informed:

$$\begin{aligned}
X_t &= \frac{Var(P_{t+1}+D_{t+1}|P_t)}{Var(P_{t+1}+D_{t+1}|P_t,\theta)} \\
&= \frac{(B_{t+1}+1)^2\sigma_\theta^2 + \sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2 - \frac{\left[\lambda_t[B_{t+1}+1]^2\sigma_\theta^2\right]^2}{\lambda_t^2(B_{t+1}+1)^2\sigma_\theta^2+\alpha^2\left[\sigma_\varepsilon^2+C_{t+1}^2\sigma_x^2\right]^2\sigma_x^2}}{\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]} \\
&= \frac{\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2 + \left[1 - \frac{\lambda_t^2[B_{t+1}+1]^2\sigma_\theta^2}{\lambda_t^2(B_{t+1}+1)^2\sigma_\theta^2+\alpha^2\left[\sigma_\varepsilon^2+C_{t+1}^2\sigma_x^2\right]^2\sigma_x^2}\right](B_{t+1}+1)^2\sigma_\theta^2}{\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]} \\
&= \frac{\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2 + \left[\frac{\alpha^2\left[\sigma_\varepsilon^2+C_{t+1}^2\sigma_x^2\right]^2\sigma_x^2}{\lambda_t^2(B_{t+1}+1)^2\sigma_\theta^2+\alpha^2\left[\sigma_\varepsilon^2+C_{t+1}^2\sigma_x^2\right]^2\sigma_x^2}\right](B_{t+1}+1)^2\sigma_\theta^2}{\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]} \\
&= 1 + \frac{\alpha^2\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]\sigma_x^2}{\lambda_t^2(B_{t+1}+1)^2\sigma_\theta^2 + \alpha^2\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]^2\sigma_x^2}(B_{t+1}+1)^2\sigma_\theta^2
\end{aligned}$$

Plug $V_t^U = X_t\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]$ into B.9 and B.10, rearrange, one can get:

$$\begin{bmatrix} B_t \\ C_t \end{bmatrix} = \begin{bmatrix} \frac{1}{R}\lambda_t(B_{t+1}+1)\frac{\left[X_t+(1-\lambda_t)M_t^U\right]}{(\lambda_t X_t + (1-\lambda_t))} \\ \frac{1}{R}\alpha\left[C_{t+1}^2\sigma_x^2 + \sigma_\varepsilon^2\right]\frac{\left[(1-\lambda_t)M_t^U + X_t\right]}{(\lambda_t X_t+(1-\lambda_t))} \end{bmatrix}$$

Define $L_t = \frac{\left[X_t+(1-\lambda_t)M_t^U\right]}{(\lambda_t X_t+(1-\lambda_t))}$. Then the transition function is given by:

$$\left[\begin{array}{c} B_t \\ C_t \end{array}\right] - \left[\begin{array}{c} \frac{1}{R}\lambda_t\left(B_{t+1}+1\right)L_t \\ \frac{1}{R}\alpha\left[C_{t+1}^2\sigma_x^2+\sigma_\varepsilon^2\right]L_t \end{array}\right] = 0$$

Where

$$\begin{aligned}
L_t &= \frac{\left[X_t+(1-\lambda_t)M_t^U\right]}{(\lambda_t X_t+(1-\lambda_t))} \\
X_t &= 1+\frac{\alpha^2\left[\sigma_\varepsilon^2+C_{t+1}^2\sigma_x^2\right]\sigma_x^2}{\lambda_t^2\left(B_{t+1}+1\right)^2\sigma_\theta^2+\alpha^2\left[\sigma_\varepsilon^2+C_{t+1}^2\sigma_x^2\right]^2\sigma_x^2}\left(B_{t+1}+1\right)^2\sigma_\theta^2 \\
M_t^U &= \frac{\lambda_t\left[B_{t+1}+1\right]^2\sigma_\theta^2}{\lambda_t^2\left(B_{t+1}+1\right)^2\sigma_\theta^2+\alpha^2\left[\sigma_\varepsilon^2+C_{t+1}^2\sigma_x^2\right]^2\sigma_x^2}
\end{aligned}$$

Next we turn to steady state. At steady state, we can get rid of the $t$ subscripts. Thus we arrive at an expression for the value of information:

$$\begin{aligned}
\pi(\lambda) &= \frac{W^U}{W^I} && \text{(B.11)} \\
&= \sqrt{\frac{V^U}{V^I}} && \text{(B.12)} \\
&= \sqrt{X(\lambda, B, C)} && \text{(B.13)} \\
&= \sqrt{1+\frac{\alpha^2\left[\sigma_\varepsilon^2+C^2\sigma_x^2\right]\sigma_x^2}{\lambda^2\left(B+1\right)^2\sigma_\theta^2+\alpha^2\left[\sigma_\varepsilon^2+C^2\sigma_x^2\right]^2\sigma_x^2}\left(B+1\right)^2\sigma_\theta^2} && \text{(B.14)}
\end{aligned}$$

(The first equality follows from definition. The second equality follows from proposition B.1. The third equality is by definition of function $X$. )

Subject to:

$$\left[\begin{array}{c} B \\ C \end{array}\right] = \left[\begin{array}{c} \frac{1}{R}\lambda\left(B+1\right)L(B,C,\lambda) \\ \frac{1}{R}\alpha\left[C^2\sigma_x^2+\sigma_\varepsilon^2\right]L(B,C,\lambda) \end{array}\right] \quad \text{(B.15)}$$

Where

$$\begin{aligned}
L(B,C,\lambda) &= \frac{\left[X(B,C,\lambda)+(1-\lambda_t)M^U(B,C,\lambda)\right]}{(\lambda X(B,C,\lambda)+(1-\lambda))} \\
X(B,C,\lambda) &= 1+\frac{\alpha^2\left[\sigma_\varepsilon^2+C^2\sigma_x^2\right]\sigma_x^2}{\lambda_t^2\left(B+1\right)^2\sigma_\theta^2+\alpha^2\left[\sigma_\varepsilon^2+C^2\sigma_x^2\right]^2\sigma_x^2}\left(B+1\right)^2\sigma_\theta^2 \\
M^U(B,C,\lambda) &= \frac{\lambda\left[B+1\right]^2\sigma_\theta^2}{\lambda^2\left(B+1\right)^2\sigma_\theta^2+\alpha^2\left[\sigma_\varepsilon^2+C^2\sigma_x^2\right]^2\sigma_x^2}
\end{aligned}$$

Essentially, the constraints B.15 define a set of implicite function $B(\lambda), C(\lambda)$.

Thus

$$\pi'(\lambda) = \frac{1}{2}X^{-\frac{1}{2}}\left(X_\lambda+X_B B_\lambda+X_C C_\lambda\right)$$

Where $B_\lambda, C_\lambda$ are total differentiation of the constraint:

$$\begin{bmatrix} B \\ C \end{bmatrix} - \begin{bmatrix} \frac{1}{R}\lambda (B+1) L(B,C,\lambda) \\ \frac{1}{R}\alpha \left[C^2\sigma_x^2 + \sigma_\varepsilon^2\right] L(B,C,\lambda) \end{bmatrix} = 0$$

Define $G(\lambda, B, C) = RB - \lambda (B+1) L(B,C,\lambda); H(\lambda, B, C) = RC - \alpha \left[C^2\sigma_x^2 + \sigma_\varepsilon^2\right] L(B,C,\lambda)$, Then the following must hold:

$$\begin{aligned} G(\lambda, B, C) &= 0 \\ H(\lambda, B, C) &= 0 \end{aligned}$$

Total differentiation:

$$\begin{bmatrix} G_B & G_C \\ H_B & H_C \end{bmatrix} \begin{bmatrix} B_\lambda \\ C_\lambda \end{bmatrix} = - \begin{bmatrix} G_\lambda \\ H_\lambda \end{bmatrix}$$

$$\begin{bmatrix} R - \lambda L - \lambda (B+1) L_B & -\lambda (B+1) L_C \\ -\alpha \left[C^2\sigma_x^2 + \sigma_\varepsilon^2\right] L_B & R - 2\alpha C\sigma_x^2 L - \alpha \left[C^2\sigma_x + \sigma_\varepsilon\right] L_C \end{bmatrix} \begin{bmatrix} B_\lambda \\ C_\lambda \end{bmatrix} = - \begin{bmatrix} -(B+1) L - \lambda (B+1) L_\lambda \\ -\alpha \left[C^2\sigma_x^2 + \sigma_\varepsilon^2\right] L_\lambda \end{bmatrix}$$

$$\begin{bmatrix} B_\lambda \\ C_\lambda \end{bmatrix} = \begin{bmatrix} R - \lambda L - \lambda (B+1) L_B & -\lambda (B+1) L_C \\ -\alpha \left[C^2\sigma_x^2 + \sigma_\varepsilon^2\right] L_B & R - 2\alpha C\sigma_x^2 L - \alpha \left[C^2\sigma_x + \sigma_\varepsilon\right] L_C \end{bmatrix}^{-1} \begin{bmatrix} (B+1) L + \lambda (B+1) L_\lambda \\ \alpha \left[C^2\sigma_x^2 + \sigma_\varepsilon^2\right] L_\lambda \end{bmatrix}$$

Now we need to check the value of each term when $\lambda = 0$

$$\begin{aligned}
X &= \frac{(B+1)^2 \sigma_\theta^2 \alpha^2 V^I \sigma_x^2}{\lambda^2 (B+1)^2 \sigma_\theta^2 + \alpha^2 \left(C^2\sigma_x^2 + \sigma_\varepsilon^2\right)^2 \sigma_x^2} + 1 = \frac{\sigma_\theta^2}{C^2\sigma_x^2 + \sigma_\varepsilon^2} + 1 \\
\frac{\partial X}{\partial B} &= \frac{(B+1)^2 \sigma_\theta^2 \alpha^2 V^I \sigma_x^2 \alpha^2 \left(C^2\sigma_x^2 + \sigma_\varepsilon^2\right)^2 \sigma_x^2}{\left[\lambda^2 (B+1)^2 \sigma_\theta^2 + \alpha^2 \left(C^2\sigma_x^2 + \sigma_\varepsilon^2\right)^2 \sigma_x^2\right]^2} = \frac{\sigma_\theta^2}{\left(C^2\sigma_x^2 + \sigma_\varepsilon^2\right)} \\
\frac{\partial X}{\partial C} &= \frac{(B+1)^2 \sigma_\theta^2 \alpha^2 \sigma_x^2 \left[\lambda^2 (B+1)^2 \sigma_\theta^2 - \alpha^2 \left(C^2\sigma_x^2 + \sigma_\varepsilon^2\right)^2 \sigma_x^2\right] 2C\sigma_x^2}{\left[\lambda^2 (B+1)^2 \sigma_\theta^2 + \alpha^2 \left(C^2\sigma_x^2 + \sigma_\varepsilon^2\right)^2 \sigma_x^2\right]^2} = -\frac{\sigma_\theta^2 \sigma_x^2 2C}{\left(C^2\sigma_x^2 + \sigma_\varepsilon^2\right)^2} \\
\frac{\partial X}{\partial \lambda} &= \frac{-(B+1)^2 \sigma_\theta^2 \alpha^2 V^I \sigma_x 2\lambda (B+1)^2 \sigma_\theta^2}{\left[\lambda^2 (B+1)^2 \sigma_\theta^2 + \alpha^2 \left(C^2\sigma_x^2 + \sigma_\varepsilon^2\right)^2 \sigma_x^2\right]^2} = 0
\end{aligned}$$

$$
\begin{aligned}
M^U &= \frac{\lambda (B+1)^2 \sigma_\theta^2}{\lambda^2 (B+1)^2 \sigma_\theta^2 + \alpha^2 \left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2 \sigma_x^2} = 0 \\
\frac{\partial M^U}{\partial B} &= \frac{2\lambda (B+1) \sigma_\theta^2 \alpha^2 \left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2 \sigma_x}{\left[\lambda^2 (B+1)^2 \sigma_\theta^2 + \alpha^2 \left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2 \sigma_x^2\right]^2} = 0 \\
\frac{\partial M^U}{\partial C} &= -\frac{\lambda (B+1)^2 \sigma_\theta^2}{\left[\lambda^2 (B+1)^2 \sigma_\theta^2 + \alpha^2 \left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2 \sigma_x^2\right]^2} 2\alpha^2 \left(C^2\sigma_x^2+\sigma_\varepsilon^2\right) \sigma_x^2 2C\sigma_x^2 = 0 \\
\frac{\partial M^U}{\partial \lambda} &= \frac{(B+1)^2 \sigma_\theta^2 \left[\lambda^2 (B+1)^2 \sigma_\theta^2 + \alpha^2 \left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2 \sigma_x^2\right] - \lambda (B+1)^2 \sigma_\theta^2 \left[2\lambda (B+1)^2 \sigma_\theta^2\right]}{\left[\lambda^2 (B+1)^2 \sigma_\theta^2 + \alpha^2 \left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2 \sigma_x^2\right]^2} \\
&\to 0
\end{aligned}
$$

Thus:

$$
\begin{aligned}
L &= \frac{\left[X + (1-\lambda) M^U\right]}{(\lambda X + (1-\lambda))} = X + M^U = \frac{\sigma_\theta^2}{C^2\sigma_x^2+\sigma_\varepsilon^2} + 1 \\
L_B &= \frac{\left[X + (1-\lambda) M^U\right]' (\lambda X + (1-\lambda)) - \left[X + (1-\lambda) M^U\right] (\lambda X + (1-\lambda))'}{(\lambda X + (1-\lambda))^2} \\
&= \frac{\left[X_B + (1-\lambda) M_B^U\right] (\lambda X + (1-\lambda)) - \left[X + (1-\lambda) M^U\right] (\lambda X_B)}{(\lambda X + (1-\lambda))^2} = \frac{\sigma_\theta^2}{C^2\sigma_x^2+\sigma_\varepsilon^2} \\
L_C &= \frac{\left[X + (1-\lambda) M^U\right]' (\lambda X + (1-\lambda)) - \left[X + (1-\lambda) M^U\right] (\lambda X + (1-\lambda))'}{(\lambda X + (1-\lambda))^2} \\
&= \frac{\left[X_C + (1-\lambda) M_C^U\right] (\lambda X + (1-\lambda)) - \left[X + (1-\lambda) M^U\right] (\lambda X_C)}{(\lambda X + (1-\lambda))^2} = -\frac{\sigma_\theta^2 \sigma_x^2}{\left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2} 2C \\
L_\lambda &= \frac{\left[X + (1-\lambda) M^U\right]' (\lambda X + (1-\lambda)) - \left[X + (1-\lambda) M^U\right] (\lambda X + (1-\lambda))'}{(\lambda X + (1-\lambda))^2} \\
&= \frac{\left[X_\lambda - M^U + (1-\lambda) M_\lambda^U\right] (\lambda X + (1-\lambda)) - \left[X + (1-\lambda) M^U\right] (X + \lambda X_\lambda - 1)}{(\lambda X + (1-\lambda))^2} \\
&= \frac{\sigma_\theta^2}{\alpha^2 \left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2 \sigma_x^2} - \left[1 + \frac{\sigma_\theta^2}{C^2\sigma_x^2+\sigma_\varepsilon^2}\right] \left(\frac{\sigma_\theta^2}{C^2\sigma_x^2+\sigma_\varepsilon^2}\right)
\end{aligned}
$$

$$
\begin{aligned}
\begin{bmatrix} B_\lambda \\ C_\lambda \end{bmatrix} &= \begin{bmatrix} r - \lambda L - \lambda (B+1) L_B & -\lambda (B+1) L_C \\ -\alpha \left[C^2\sigma_x + \sigma_\varepsilon\right] L_B & r - 2\alpha C \sigma_x L - \alpha \left[C^2\sigma_x + \sigma_\varepsilon\right] L_C \end{bmatrix}^{-1} \begin{bmatrix} (B+1) L + \lambda (B+1) L_\lambda \\ \alpha \left[C^2\sigma_x + \sigma_\varepsilon\right] L_\lambda \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{R}\left(1 + \frac{\sigma_\theta^2}{C^2\sigma_x^2+\sigma_\varepsilon^2}\right) \\ \frac{\alpha\sigma_\theta^2\left(1+\frac{\sigma_\theta^2}{C^2\sigma_x^2+\sigma_\varepsilon^2}\right) + R\alpha\sigma_\theta^2\left(\frac{1}{\alpha^2(C^2\sigma_x^2+\sigma_\varepsilon^2)\sigma_x} - 1 - \frac{\sigma_\theta^2}{C^2\sigma_x^2+\sigma_\varepsilon^2}\right)}{R(R-2\alpha C\sigma_x^2)} \end{bmatrix} \text{ when } \lambda = 0
\end{aligned}
$$

Thus

$$\left.\frac{d\pi}{d\lambda}\right|_{\lambda=0} = \frac{dX\left(B\left(\lambda\right), C\left(\lambda\right), \lambda\right)}{d\lambda} = \frac{\partial X}{\partial B}\frac{\partial B}{\partial \lambda} + \frac{\partial X}{\partial C}\frac{\partial C}{\partial \lambda} + \frac{\partial X}{\partial \lambda}$$

$$= \frac{\sigma_\theta^2}{(C^2\sigma_x^2+\sigma_\varepsilon^2)}\frac{1}{R}\left(1+\frac{\sigma_\theta^2}{C^2\sigma_x^2+\sigma_\varepsilon^2}\right)$$

$$-\frac{\sigma_\theta^2\sigma_x^2 2C}{(C^2\sigma_x^2+\sigma_\varepsilon^2)^2}\frac{\alpha\sigma_\theta^2\left(1+\frac{\sigma_\theta^2}{C^2\sigma_x^2+\sigma_\varepsilon^2}\right)+R\alpha\sigma_\theta^2\left(\frac{1}{\alpha^2(C^2\sigma_x^2+\sigma_\varepsilon^2)\sigma_x}-1-\frac{\sigma_\theta^2}{C^2\sigma_x^2+\sigma_\varepsilon^2}\right)}{R\left(R-2\alpha C\sigma_x^2\right)} > 0$$

Simplify, one get:

$$1+\frac{\sigma_\theta^2}{C^2\sigma_x^2+\sigma_\varepsilon^2}-\frac{\alpha\sigma_\theta^2\sigma_x^2 2C}{C^2\sigma_x^2+\sigma_\varepsilon^2}\frac{\left(1+\frac{\sigma_\theta^2}{C^2\sigma_x^2+\sigma_\varepsilon^2}\right)+R\left(\frac{1}{\alpha^2(C^2\sigma_x^2+\sigma_\varepsilon^2)\sigma_x}-1-\frac{\sigma_\theta^2}{C^2\sigma_x^2+\sigma_\varepsilon^2}\right)}{(R-2\alpha C\sigma_x^2)} > 0$$

In the low volatility economy:

$$C = \left(\frac{R-\sqrt{R^2-4\alpha^2\sigma_x^2(\sigma_\varepsilon^2+\sigma_\theta^2)}}{2\alpha\sigma_x^2}\right)$$

Thus, if

$$1+\frac{\sigma_\theta^2}{C^2\sigma_x^2+\sigma_\varepsilon^2}-\frac{\alpha\sigma_\theta^2\sigma_x^2 2C}{C^2\sigma_x^2+\sigma_\varepsilon^2}\frac{\left(1+\frac{\sigma_\theta^2}{C^2\sigma_x^2+\sigma_\varepsilon^2}\right)+R\left(\frac{1}{\alpha^2(C^2\sigma_x^2+\sigma_\varepsilon^2)\sigma_x}-1-\frac{\sigma_\theta^2}{C^2\sigma_x^2+\sigma_\varepsilon^2}\right)}{(R-2\alpha C\sigma_x^2)} > 0 \qquad \text{(B.16)}$$

with $C = \left(\frac{R-\sqrt{R^2-4\alpha^2\sigma_x^2(\sigma_\varepsilon^2+\sigma_\theta^2)}}{2\alpha\sigma_x^2}\right)$ holds, then $\left.\frac{d\pi}{d\lambda}\right|_{\lambda=0} > 0$

## Proof of theorem 1

The proof follows directly from the continuity of $\pi(\lambda)$ intermediate value theorem. $\pi(\lambda)$ is differentiable, hence continuous. Given that $\pi'(\lambda) > 0$ for $\lambda$ sufficiently small, pick $\chi$ such that $e^{\alpha\chi} = \pi(\lambda_1)$ for some $\lambda_1$ sufficiently small but strictly positive. Then we know that $\lambda_1$ is a steady state. Also, we know that $\pi(0) < \pi(\lambda_1) = e^{\alpha\chi}$. Thus we know that $\lambda_0 = 0$ is another steady state as no one is informed and the gain from acquiring information is less than the cost. Thus we find multiple steady states.

## Proof of proposition 4.1

By proof of lemma 4.3, The coefficients $\{A_2, B_2, C_2\}$ on the price function $P_2 = A_2 + B_2\theta + C_2x_2$ must satisfy:

$$A_2 = v^{-1}\left[\lambda_2 \frac{A_2}{\alpha\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]} + (1-\lambda_2)\frac{A_2 + (B_2+1)\mu}{\alpha V_2^U} - 1\right] \quad \text{(B.17)}$$

$$B_2 = v^{-1}\left[\lambda_2 \frac{(B_2+1)}{\alpha\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]} + (1-\lambda_2)\frac{M_2^U\left[\lambda_2(B_2+1)\right]}{\alpha V_2^U}\right] \quad \text{(B.18)}$$

$$C_2 = v^{-1}\left[(1-\lambda_2)\frac{M_2^U\left[\alpha\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]\right]}{\alpha V_2^U} + 1\right] \quad \text{(B.19)}$$

Where:

$$V_2^U = (B_2+1)^2\sigma_\theta^2 + \sigma_\varepsilon^2 + C_2^2\sigma_x^2 - \frac{\left[\lambda_2\left[B_2+1\right]^2\sigma_\theta^2\right]^2}{\lambda_2^2(B_2+1)^2\sigma_\theta^2 + \alpha^2\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]^2\sigma_x^2}$$

$$M_2^U = \frac{\lambda_2\left[B_2+1\right]^2\sigma_\theta^2}{\lambda_2^2(B_2+1)^2\sigma_\theta^2 + \alpha^2\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]^2\sigma_x^2}$$

$$v = R\left(\frac{\lambda_2}{\alpha\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]} + \frac{(1-\lambda_2)}{\alpha V_2^U}\right)$$

That is, $\{A_2, B_2, C_2\}$ are all functions of $\lambda_2$.

Likewise the coefficients $\{A_1, B_1, C_1\}$ on the price function $P_1 = A_1 + B_1\theta + C_1x_1$ must satisfy:

$$A_1 = v^{-1}\left[\lambda_1 \frac{A_2}{\alpha\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]} + (1-\lambda_1)\frac{A_2 + (B_2+1)\mu}{\alpha V_1^U} - 1\right]$$

$$B_1 = v^{-1}\left[\lambda_1 \frac{(B_2+1)}{\alpha\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]} + (1-\lambda_2)\frac{M_1^U\left[\lambda_1(B_2+1)\right]}{\alpha V_1^U}\right]$$

$$C_1 = v^{-1}\left[(1-\lambda_1)\frac{M_1^U\left[\alpha\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]\right]}{\alpha V_1^U} + 1\right]$$

Where:

$$V_1^U = (B_2+1)^2\sigma_\theta^2 + \sigma_\varepsilon^2 + C_2^2\sigma_x^2 - \frac{\left[\lambda_1\left[B_2+1\right]^2\sigma_\theta^2\right]^2}{\lambda_1^2(B_2+1)^2\sigma_\theta^2 + \alpha^2\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]^2\sigma_x^2}$$

$$M_1^U = \frac{\lambda_1\left[B_2+1\right]^2\sigma_\theta^2}{\lambda_1^2(B_2+1)^2\sigma_\theta^2 + \alpha^2\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]^2\sigma_x^2}$$

$$v = R\left(\frac{\lambda_1}{\alpha\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]} + \frac{(1-\lambda_1)}{\alpha V_1^U}\right)$$

As $\{A_1, B_1, C_1\}$ are functions of $\lambda_1$ as well as $\{A_2, B_2, C_2\}$, which are in turn all functions of $\lambda_2$, $\{A_1, B_1, C_1\}$ are functions of both $\lambda_1$ and $\lambda_2$.

Given $\{A_1, B_1, C_1\}$ and $\{A_2, B_2, C_2\}$, we can compute the value of information for period-1 agents, which is:

$$\begin{aligned}
\pi_d(\lambda_1, \lambda_2) &= \sqrt{\frac{Var(D_2 + P_2|P_1)}{Var(D_2 + P_2|P_1, \theta)}} \\
&= \sqrt{1 + \frac{\alpha^2\left[\sigma_\varepsilon^2 + C_2(\lambda_2)^2\sigma_x^2\right]\sigma_x^2}{\lambda_1^2\left(B_2(\lambda_2)+1\right)^2\sigma_\theta^2 + \alpha^2\left[\sigma_\varepsilon^2 + C_2(\lambda_2)^2\sigma_x^2\right]^2\sigma_x^2}\left(B_2(\lambda_2)+1\right)^2\sigma_\theta^2}
\end{aligned}$$

Where $B_2(\lambda_2)$ and $C_2(\lambda_2)$ are implicit functions defined by B.21 and B.22. The first equality follows from lemma 4.2 and the second equality follows from 4.5.

Simplify B.21 and B.22, $B_2, C_2$ must satisfy:

$$\begin{bmatrix} B_2 \\ C_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{R}\lambda_2\left(B_2+1\right)L(B_2, C_2, \lambda_2) \\ \frac{1}{R}\alpha\left[C_2^2\sigma_x^2 + \sigma_\varepsilon^2\right]L(B_2, C_2, \lambda_2) \end{bmatrix}$$

Where

$$\begin{aligned}
L(B, C, \lambda) &= \frac{\left[X(B, C, \lambda) + (1-\lambda)\,M^U(B, C, \lambda\right]}{\left(\lambda X(B, C, \lambda) + (1-\lambda)\right)} \\
X(B, C, \lambda) &= 1 + \frac{\alpha^2\left[\sigma_\varepsilon^2 + C^2\sigma_x^2\right]\sigma_x^2}{\lambda_t^2\left(B+1\right)^2\sigma_\theta^2 + \alpha^2\left[\sigma_\varepsilon^2 + C^2\sigma_x^2\right]^2\sigma_x^2}\left(B+1\right)^2\sigma_\theta^2 \\
M^U(B, C, \lambda) &= \frac{\lambda\left[B+1\right]^2\sigma_\theta^2}{\lambda^2\left(B+1\right)^2\sigma_\theta^2 + \alpha^2\left[\sigma_\varepsilon^2 + C^2\sigma_x^2\right]^2\sigma_x^2}
\end{aligned}$$

One can take derivatives with respect to $\lambda_1$:

$$\frac{\partial \pi_d}{\partial \lambda_1} = \pi_d^{-1}\frac{-\alpha^2\left[\sigma_\varepsilon^2 + C_2(\lambda_2)^2\sigma_x^2\right]\sigma_x^2}{\left(\lambda_1^2\left(B_2(\lambda_2)+1\right)^2\sigma_\theta^2 + \alpha^2\left[\sigma_\varepsilon^2 + C_2(\lambda_2)^2\sigma_x^2\right]^2\sigma_x^2\right)^2}\left(\left(B_2(\lambda_2)+1\right)^2\sigma_\theta^2\right)^2\lambda_1$$

Thus when $\lambda_1 = 0$, $\frac{\partial \pi_d}{\partial \lambda_1} = 0$.

To examine the derivative of $\frac{\partial \pi_d}{\partial \lambda_2}$, one can perform a similar total differentiation task as in the proof of lemma 4.3 and easily verify that if B.16 holds, then $\frac{\partial \pi_d}{\partial \lambda_2} > 0$. The proof is available upon request.

## Proof of lemma 4.2

The proof of the lemma closely mirrors the proof of proposition B.1, hence is omitted here.

## Proof of lemma 4.5

By proof of lemma 4.3, The coefficients $\{A_2, B_2, C_2\}$ on the price function $P_2 = A_2 + B_2\theta + C_2x_2$ must satisfy:

$$A_2 = v^{-1}\left[\lambda_2 \frac{A_2}{\alpha\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]} + (1-\lambda_2)\frac{A_2 + (B_2+1)\mu}{\alpha V_2^U} - 1\right] \tag{B.20}$$

$$B_2 = v^{-1}\left[\lambda_2 \frac{(B_2+1)}{\alpha\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]} + (1-\lambda_2)\frac{M_2^U\left[\lambda_2(B_2+1)\right]}{\alpha V_2^U}\right] \tag{B.21}$$

$$C_2 = v^{-1}\left[(1-\lambda_2)\frac{M_2^U\left[\alpha\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]\right]}{\alpha V_2^U} + 1\right] \tag{B.22}$$

Where:

$$\begin{aligned}
V_2^U &= (B_2+1)^2\sigma_\theta^2 + \sigma_\varepsilon^2 + C_2^2\sigma_x^2 - \frac{\left[\lambda_2\left[B_2+1\right]^2\sigma_\theta^2\right]^2}{\lambda_2^2(B_2+1)^2\sigma_\theta^2 + \alpha^2\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]^2\sigma_x^2} \\
M_2^U &= \frac{\lambda_2\left[B_2+1\right]^2\sigma_\theta^2}{\lambda_2^2(B_2+1)^2\sigma_\theta^2 + \alpha^2\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]^2\sigma_x^2} \\
v &= R\left(\frac{\lambda_2}{\alpha\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]} + \frac{(1-\lambda_2)}{\alpha V_2^U}\right)
\end{aligned}$$

That is, $\{A_2, B_2, C_2\}$ are all functions of $\lambda_2$.

Likewise the coefficients $\{A_1, B_1, C_1\}$ on the price function $P_1 = A_1 + B_1\theta + C_1x_1$ must satisfy:

$$\begin{aligned}
A_1 &= v^{-1}\left[\lambda_1 \frac{A_2}{\alpha\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]} + (1-\lambda_1)\frac{A_2 + (B_2+1)\mu}{\alpha V_1^U} - 1\right] \\
B_1 &= v^{-1}\left[\lambda_1 \frac{(B_2+1)}{\alpha\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]} + (1-\lambda_2)\frac{M_1^U\left[\lambda_1(B_2+1)\right]}{\alpha V_1^U}\right] \\
C_1 &= v^{-1}\left[(1-\lambda_1)\frac{M_1^U\left[\alpha\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]\right]}{\alpha V_1^U} + 1\right]
\end{aligned}$$

Where:

$$\begin{aligned}
V_1^U &= (B_2+1)^2\sigma_\theta^2 + \sigma_\varepsilon^2 + C_2^2\sigma_x^2 - \frac{\left[\lambda_1\left[B_2+1\right]^2\sigma_\theta^2\right]^2}{\lambda_1^2(B_2+1)^2\sigma_\theta^2 + \alpha^2\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]^2\sigma_x^2} \\
M_1^U &= \frac{\lambda_1\left[B_2+1\right]^2\sigma_\theta^2}{\lambda_1^2(B_2+1)^2\sigma_\theta^2 + \alpha^2\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]^2\sigma_x^2} \\
v &= R\left(\frac{\lambda_1}{\alpha\left[\sigma_\varepsilon^2 + C_2^2\sigma_x^2\right]} + \frac{(1-\lambda_1)}{\alpha V_1^U}\right)
\end{aligned}$$

As $\{A_1, B_1, C_1\}$ are functions of $\lambda_1$ as well as $\{A_2, B_2, C_2\}$, which are in turn all functions of $\lambda_2$, $\{A_1, B_1, C_1\}$ are functions of both $\lambda_1$ and $\lambda_2$.

To ease notation, below I omit the dependence of coefficients on $\lambda_1$ and $\lambda_2$.

Then

$$\begin{aligned} D_2 + P_2 &= \theta + \varepsilon' + A_2 + B_2\theta + C_2 x' \\ &= A_2 + (B_2+1)\theta + \varepsilon' + C_2 x' \end{aligned}$$

Informed investors know perfectly the value of $\theta$, hence:

$$Var(D_2 + P_2|P_1, \theta) = C_2^2\sigma_x^2 + \sigma_\varepsilon^2$$

Now turn to the uncertainty faced by the uninformed: $Var(D_2 + P_2|P_1)$ Given that $P_1 = A_1 + B_1\theta + C_1 x_1$ is a normally distributed noisy signal about $\theta$, we can apply updating formula for normal variables:

$$\begin{aligned} Var(D_2 + P_2|P_1) &= Var(D_2 + P_2) - \tfrac{[Cov(D_2+P_2,P_1)]^2}{Var(P_1)} \\ &= (B_2+1)^2\sigma_\theta^2 + C_2^2\sigma_x^2 + \sigma_\varepsilon^2 - \tfrac{(B_1(B_2+1)\sigma_\theta^2)^2}{B_1^2\sigma_\theta^2 + C_1^2\sigma_x^2} \\ &= (B_2+1)^2\sigma_\theta^2 + C^2\sigma_x^2 + \sigma_\varepsilon^2 - \Omega(\lambda_1,\lambda_2)(B+1)^2\sigma_\theta^2 \end{aligned}$$

Where $\Omega(\lambda_1, \lambda_2) = \dfrac{B_1(\lambda_1,\lambda_2)^2\sigma_\theta^2}{B_1(\lambda_1,\lambda_2)^2\sigma_\theta^2 + C_1(\lambda_1,\lambda_2)^2\sigma_x^2}$.

## Proof of lemma 5.1

From the proof of lemma 4.3, there is a set of equation (B.9 and B.10) that $\{B_t, C_t, \lambda_t, B_{t+1}, C_{t+1}\}$ must satisfy. Denote it:

$$(B_t, C_t) = \Upsilon(\lambda_t, B_{t+1}, C_{t+1}) \tag{B.23}$$

It suffices to show that when information choice is endogenous, there exist a function $F_1$ such that:

$$\lambda_t = F_1(B_{t+1}, C_{t+1}) \tag{B.24}$$

Plug this back into B.23, we get:

$$(B_t, C_t) = \Upsilon(F_1(B_{t+1}, C_{t+1}), B_{t+1}, C_{t+1}) \tag{B.25}$$

Thus we obtain $F$ which is a composite functin of $\Upsilon$ and $F_1$.

Thus it suffices to find $F_1$. Intuitively $F_1$ is given by the information choice optimality condition, by equating value of information to some transformation of the cost of collecting information, potentially accounting for the boundary condition.

Note that the information choice optimality condition is given by (for interior $\lambda$):

$$\begin{aligned} \pi(\lambda) &= \sqrt{X_t} \\ &= \exp(\alpha\chi) \end{aligned} \tag{B.26}$$

Sqare both sides, an interior $\lambda_t$ is solved by equating:

$$X_t = \frac{(B_{t+1}+1)^2 \sigma_\theta^2 \alpha^2 \left(C_{t+1}^2 \sigma_x^2 + \sigma_\varepsilon^2\right) \sigma_x^2}{\lambda_t^2 (B_{t+1}+1)^2 \sigma_\theta^2 + \alpha^2 \left(C_{t+1}^2 \sigma_x^2 + \sigma_\varepsilon^2\right)^2 \sigma_x^2} + 1 = e^{2\alpha\chi}$$

Thus one can solve for $\lambda_t$, if it is at an interior.

$$\lambda_t = \sqrt{\frac{(B_{t+1}+1)^2 \sigma_\theta^2 \alpha^2 \left(C_{t+1}^2 \sigma_x^2 + \sigma_\varepsilon^2\right) \sigma_x^2 - \left[e^{2\alpha\chi} - 1\right] \alpha^2 \left(C_{t+1}^2 \sigma_x^2 + \sigma_\varepsilon^2\right)^2 \sigma_x^2}{(B_{t+1}+1)^2 \sigma_\theta^2}}$$

One still need to account for boundary conditions:

$$\begin{array}{ll} \lambda_t = 1 \quad \text{if} & \sqrt{\frac{(B_{t+1}+1)^2 \sigma_\theta^2 \alpha^2 \left(C_{t+1}^2 \sigma_x^2 + \sigma_\varepsilon^2\right) \sigma_x^2 - \left[e^{2\alpha\chi} - 1\right] \alpha^2 \left(C_{t+1}^2 \sigma_x^2 + \sigma_\varepsilon^2\right)^2 \sigma_x^2}{(B_{t+1}+1)^2 \sigma_\theta^2}} > 1 \\ \lambda_t = 0 \quad \text{if} & \sqrt{\frac{(B_{t+1}+1)^2 \sigma_\theta^2 \alpha^2 \left(C_{t+1}^2 \sigma_x^2 + \sigma_\varepsilon^2\right) \sigma_x^2 - \left[e^{2\alpha\chi} - 1\right] \alpha^2 \left(C_{t+1}^2 \sigma_x^2 + \sigma_\varepsilon^2\right)^2 \sigma_x^2}{(B_{t+1}+1)^2 \sigma_\theta^2}} < 0 \end{array} \tag{B.27}$$

Thus:

$$\begin{aligned} \lambda_t &= F_1 (B_{t+1}, C_{t+1}) \\ &= \max\left(\min\left(\sqrt{\tfrac{(B_{t+1}+1)^2 \sigma_\theta^2 \alpha^2 \left(C_{t+1}^2 \sigma_x^2 + \sigma_\varepsilon^2\right) \sigma_x^2 - \left[e^{2\alpha\chi} - 1\right] \alpha^2 \left(C_{t+1}^2 \sigma_x^2 + \sigma_\varepsilon^2\right)^2 \sigma_x^2}{(B_{t+1}+1)^2 \sigma_\theta^2}}, 1\right), 0\right) \end{aligned}$$

This concludes the proof.

## Proof of theorem 2

### Auxilary Result: for $\sigma_\theta$ sufficiently small, $\pi(\lambda) > 0, \forall \lambda$

Proof: Following the notation in the proof of lemma 4.3:

$$\begin{aligned} \pi(\lambda) &= \frac{W^U}{W^I} \\ &= \sqrt{X(\lambda, B, C)} \\ &= 1 + \frac{\alpha^2 \left[\sigma_\varepsilon^2 + C^2 \sigma_x^2\right] \sigma_x^2}{\lambda^2 (B+1)^2 \sigma_\theta^2 + \alpha^2 \left[\sigma_\varepsilon^2 + C^2 \sigma_x^2\right]^2 \sigma_x^2} (B+1)^2 \sigma_\theta^2 \end{aligned}$$

Subject to:

$$\begin{bmatrix} B \\ C \end{bmatrix} - \begin{bmatrix} \frac{1}{R} \lambda (B+1) L \\ \frac{1}{R} \alpha \left[C^2 \sigma_x^2 + \sigma_\varepsilon^2\right] L \end{bmatrix} = 0$$

Thus

$$\pi'(\lambda) = \frac{1}{2} X^{-\frac{1}{2}} (X_\lambda + X_B B_\lambda + X_C C_\lambda)$$

Define $G(\lambda, B, C) = RB - \lambda(B+1)L; H(\lambda, B, C) = RC - \alpha\left[C^2\sigma_x^2 + \sigma_\varepsilon^2\right]L$, Then the following must hold:

$$\begin{aligned} G(\lambda, B, C) &= 0 \\ H(\lambda, B, C) &= 0 \end{aligned}$$

Total differentiation:

$$\begin{bmatrix} G_B & G_C \\ H_B & H_C \end{bmatrix} \begin{bmatrix} B_\lambda \\ C_\lambda \end{bmatrix} = - \begin{bmatrix} G_\lambda \\ H_\lambda \end{bmatrix}$$

$$\begin{bmatrix} R - \lambda L - \lambda(B+1)L_B & -\lambda(B+1)L_C \\ -\alpha\left[C^2\sigma_x^2 + \sigma_\varepsilon^2\right]L_B & R - 2\alpha C\sigma_x^2 L - \alpha\left[C^2\sigma_x + \sigma_\varepsilon\right]L_C \end{bmatrix} \begin{bmatrix} B_\lambda \\ C_\lambda \end{bmatrix} = - \begin{bmatrix} -(B+1)L - \lambda(B+1)L_\lambda \\ -\alpha\left[C^2\sigma_x^2 + \sigma_\varepsilon^2\right]L_\lambda \end{bmatrix}$$

$$\begin{matrix} B_\lambda \\ C_\lambda \end{matrix} = \begin{bmatrix} R - \lambda L - \lambda(B+1)L_B & -\lambda(B+1)L_C \\ -\alpha\left[C^2\sigma_x^2 + \sigma_\varepsilon^2\right]L_B & R - 2\alpha C\sigma_x^2 L - \alpha\left[C^2\sigma_x + \sigma_\varepsilon\right]L_C \end{bmatrix}^{-1} \begin{bmatrix} (B+1)L + \lambda(B+1)L_\lambda \\ \alpha\left[C^2\sigma_x^2 + \sigma_\varepsilon^2\right]L_\lambda \end{bmatrix}$$

Now we need to check the value of each term when $\sigma_\theta \to 0$.

$$\begin{aligned} X &= \frac{(B+1)^2\sigma_\theta^2\alpha^2 V^I\sigma_x^2}{\lambda^2(B+1)^2\sigma_\theta^2 + \alpha^2\left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2\sigma_x^2} + 1 \to 1 \\ \frac{\partial X}{\partial B} &= \frac{(B+1)^2\sigma_\theta^2\alpha^2 V^I\sigma_x^2\alpha^2\left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2\sigma_x^2}{\left[\lambda^2(B+1)^2\sigma_\theta^2 + \alpha^2\left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2\sigma_x^2\right]^2} \to 0 \\ \frac{\partial X}{\partial C} &= \frac{(B+1)^2\sigma_\theta^2\alpha^2\sigma_x^2\left[\lambda^2(B+1)^2\sigma_\theta^2 - \alpha^2\left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2\sigma_x^2\right]2C\sigma_x^2}{\left[\lambda^2(B+1)^2\sigma_\theta^2 + \alpha^2\left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2\sigma_x^2\right]^2} \to 0 \\ \frac{\partial X}{\partial \lambda} &= \frac{-(B+1)^2\sigma_\theta^2\alpha^2 V^I\sigma_x 2\lambda(B+1)^2\sigma_\theta^2}{\left[\lambda^2(B+1)^2\sigma_\theta^2 + \alpha^2\left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2\sigma_x^2\right]^2} \to 0 \end{aligned}$$

$$\begin{aligned} M^U &= \frac{\lambda(B+1)^2\sigma_\theta^2}{\lambda^2(B+1)^2\sigma_\theta^2 + \alpha^2\left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2\sigma_x^2} \to 0 \\ \frac{\partial M^U}{\partial B} &= \frac{2\lambda(B+1)\sigma_\theta^2\alpha^2\left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2\sigma_x}{\left[\lambda^2(B+1)^2\sigma_\theta^2 + \alpha^2\left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2\sigma_x^2\right]^2} \to 0 \\ \frac{\partial M^U}{\partial C} &= -\frac{\lambda(B+1)^2\sigma_\theta^2}{\left[\lambda^2(B+1)^2\sigma_\theta^2 + \alpha^2\left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2\sigma_x^2\right]^2} 2\alpha^2\left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)\sigma_x^2 2C\sigma_x^2 \to 0 \\ \frac{\partial M^U}{\partial \lambda} &= \frac{(B+1)^2\sigma_\theta^2\left[\lambda^2(B+1)^2\sigma_\theta^2 + \alpha^2\left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2\sigma_x^2\right] - \lambda(B+1)^2\sigma_\theta^2\left[2\lambda(B+1)^2\sigma_\theta^2\right]}{\left[\lambda^2(B+1)^2\sigma_\theta^2 + \alpha^2\left(C^2\sigma_x^2+\sigma_\varepsilon^2\right)^2\sigma_x^2\right]^2} \\ &\to 0 \end{aligned}$$

Thus:

$$
\begin{aligned}
L &= \frac{\left[X+(1-\lambda)M^U\right]}{(\lambda X+(1-\lambda))} \to 1 \\
L_B &= \frac{\left[X+(1-\lambda)M^U\right]'(\lambda X+(1-\lambda)) - \left[X+(1-\lambda)M^U\right](\lambda X+(1-\lambda))'}{(\lambda X+(1-\lambda))^2} \\
&= \frac{\left[X_B+(1-\lambda)M_B^U\right](\lambda X+(1-\lambda)) - \left[X+(1-\lambda)M^U\right](\lambda X_B)}{(\lambda X+(1-\lambda))^2} \to 0 \\
L_C &= \frac{\left[X+(1-\lambda)M^U\right]'(\lambda X+(1-\lambda)) - \left[X+(1-\lambda)M^U\right](\lambda X+(1-\lambda))'}{(\lambda X+(1-\lambda))^2} \\
&= \frac{\left[X_C+(1-\lambda)M_C^U\right](\lambda X+(1-\lambda)) - \left[X+(1-\lambda)M^U\right](\lambda X_C)}{(\lambda X+(1-\lambda))^2} \to 0 \\
L_\lambda &= \frac{\left[X+(1-\lambda)M^U\right]'(\lambda X+(1-\lambda)) - \left[X+(1-\lambda)M^U\right](\lambda X+(1-\lambda))'}{(\lambda X+(1-\lambda))^2} \\
&= \frac{\left[X_\lambda - M^U+(1-\lambda)M_\lambda^U\right](\lambda X+(1-\lambda)) - \left[X+(1-\lambda)M^U\right](X+\lambda X_\lambda - 1)}{(\lambda X+(1-\lambda))^2} \\
&\to 0
\end{aligned}
$$

Also, from $\begin{bmatrix} B \\ C \end{bmatrix} - \begin{bmatrix} \frac{1}{R}\lambda(B+1)L \\ \frac{1}{R}\alpha\left[C^2\sigma_x^2+\sigma_\varepsilon^2\right]L \end{bmatrix} = 0$, one can show that as $\sigma_\theta \to 0$, the equations converge to:

$$
\begin{bmatrix} B \\ C \end{bmatrix} - \begin{bmatrix} \frac{1}{R}\lambda(B+1) \\ \frac{1}{R}\alpha\left[C^2\sigma_x^2+\sigma_\varepsilon^2\right] \end{bmatrix} = 0
$$

Thus:

$$
\begin{aligned}
B &= \frac{\lambda}{R-\lambda} \\
C &= \frac{R \pm \sqrt{R^2 - 4\alpha^2\sigma_x\sigma_\varepsilon}}{2\alpha\sigma_x}
\end{aligned}
$$

Thus under assumption 1, one can always solve for $B, C$ as well as all other equilibrium objects. Thus for any $\lambda$, an exogenous information linear steady state $\Phi(\lambda)$ always exists.

Thus:

$$
\begin{aligned}
\begin{bmatrix} B_\lambda \\ C_\lambda \end{bmatrix} &\to \begin{bmatrix} R-\lambda & 0 \\ 0 & R-2\alpha C\sigma_x^2 \end{bmatrix}^{-1} \begin{bmatrix} (B+1) \\ 0 \end{bmatrix} \\
&= \frac{1}{(R-\lambda)(R-2\alpha C\sigma_x)} \begin{bmatrix} R-2\alpha C\sigma_x & 0 \\ 0 & R-\lambda \end{bmatrix} \begin{bmatrix} (B+1) \\ 0 \end{bmatrix} \\
&= \frac{1}{(R-\lambda)(R-2\alpha C\sigma_x)} \begin{bmatrix} (R-2\alpha C\sigma_x)(B+1) \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} \frac{(B+1)}{R-\lambda} \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} R \\ 0 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
\lim_{\sigma_\theta \to 0} \frac{\pi'(\lambda)}{\sigma_\theta^2} &= \lim_{\sigma_\theta \to 0} \frac{1}{2} X^{-\frac{1}{2}} \frac{(X_\lambda + X_B B_\lambda + X_C C_\lambda)}{\sigma_\theta^2} \\
&= \lim_{\sigma_\theta \to 0} \frac{1}{2} \frac{(X_\lambda + X_B B_\lambda + X_C C_\lambda)}{\sigma_\theta^2} \\
&= \frac{1}{2} \frac{(B+1)^2 \alpha^4 \left(C^2\sigma_x^2 + \sigma_\varepsilon^2\right)^3 \sigma_x^4}{\left[\lambda^2 (B+1)^2 \sigma_\theta^2 + \alpha^2 \left(C^2\sigma_x^2 + \sigma_\varepsilon^2\right)^2 \sigma_x^2\right]^2} R > 0
\end{aligned}
$$

Because:

$$
\lim_{\sigma_\theta \to 0} \frac{X_\lambda}{\sigma_\theta^2} = \lim_{\sigma_\theta \to 0} \frac{-(B+1)^2 \alpha^2 V^I \sigma_x 2\lambda (B+1)^2 \sigma_\theta^2}{\left[\lambda^2 (B+1)^2 \sigma_\theta^2 + \alpha^2 \left(C^2\sigma_x^2 + \sigma_\varepsilon^2\right)^2 \sigma_x^2\right]^2} = 0;
$$

$$
\lim_{\sigma_\theta \to 0} \frac{X_C}{\sigma_\theta^2} = -\frac{(B+1)^2 \alpha^2 \sigma_x^2 \alpha^2 \left(C^2\sigma_x^2 + \sigma_\varepsilon^2\right)^2 \sigma_x^2 2C\sigma_x^2}{\left[\lambda^2 (B+1)^2 \sigma_\theta^2 + \alpha^2 \left(C^2\sigma_x^2 + \sigma_\varepsilon^2\right)^2 \sigma_x^2\right]^2}
$$

is a finite number, hence

$$
\lim_{\sigma_\theta \to 0} \frac{X_C}{\sigma_\theta^2} C_\lambda = -\frac{(B+1)^2 \alpha^2 \sigma_x^2 \alpha^2 \left(C^2\sigma_x^2 + \sigma_\varepsilon^2\right)^2 \sigma_x^2 2C\sigma_x^2}{\left[\lambda^2 (B+1)^2 \sigma_\theta^2 + \alpha^2 \left(C^2\sigma_x^2 + \sigma_\varepsilon^2\right)^2 \sigma_x^2\right]^2} 0 = 0 .\text{as } \lim_{\sigma_\theta \to 0} C_\lambda = 0
$$

Thus for $\sigma_\theta$ sufficiently small but strictly positive, $\pi'(\lambda) > 0, \forall \lambda$

## Claim 1

Claim 2 follows directly from the auxilary result and intermediate value theorem. Just pick any $\chi$ such that $e^{\alpha\chi} \in (\pi(0), \pi(1))$. Then there exist an $\lambda_2$ such that $\pi(\lambda_2) = e^{\alpha\chi}$. $\lambda_2$ is then a steady state. It can be trivially verified that $\lambda_1 = 0$ and $\lambda_3 = 1$ are both steady states.

**Claim 2**

We first show that the boundary steady states are both unstable. Consider the steady state with $\lambda = 0$. Due to strict monotonicity of $\pi(\lambda)$ and $\pi(\lambda') = e^{\alpha\chi}$ for some $\lambda' > 0, \pi(0) < e^{\alpha\chi}$. Hence any perturbation of $(B, C)$ would still make $\pi(0) < e^{\alpha\chi}$, thus the value of $\lambda$ is unchanged and is equal to 0.note that the (inverse) transition matrix is:

$$\begin{bmatrix} B_t \\ C_t \end{bmatrix} = \begin{bmatrix} \frac{1}{R}\lambda_t (B_{t+1}+1) L_t \\ \frac{1}{R}\alpha \left[C_{t+1}^2\sigma_x^2 + \sigma_\varepsilon^2\right] L_t \end{bmatrix}$$

Hence the Jocobian matrix is:

$$\begin{bmatrix} \frac{1}{r}\lambda F + \frac{1}{r}(B_{t+1}+1)\lambda F_B & \frac{1}{r}(B_{t+1}+1)\lambda F_C \\ \frac{1}{r}\alpha\left[C_{t+1}^2\sigma_x^2+\sigma_\varepsilon^2\right] F_B & \frac{1}{r}\alpha\left[2C_{t+1}\sigma_x^2\right] F + \frac{1}{r}\alpha\left[C_{t+1}^2\sigma_x^2+\sigma_\varepsilon^2\right] F_C \end{bmatrix}$$

Which is equal to

$$\begin{bmatrix} 0 & 0 \\ \frac{1}{r}\alpha\left[C_{t+1}^2\sigma_x^2+\sigma_\varepsilon^2\right] F_B & \frac{1}{r}\alpha\left[2C_{t+1}\sigma_x^2\right] F + \frac{1}{r}\alpha\left[C_{t+1}^2\sigma_x^2+\sigma_\varepsilon^2\right] F_C \end{bmatrix}$$

when $\lambda = 0$.

Taking $\sigma_\theta \to 0$, the jocobian matrix is:

$$\begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{r}\alpha\left[2C_{t+1}\sigma_x^2\right] \end{bmatrix}$$

Given $C = \frac{R - \sqrt{R^2 - 4\alpha^2\sigma_x^2\sigma_\varepsilon^2}}{2\alpha\sigma_x^2}$, the matrix becomes:

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 - \sqrt{1 - \frac{4\alpha^2\sigma_x^2\sigma_\varepsilon^2}{R^2}} \end{bmatrix}$$

Thus locally the transition function becomes:

$$\begin{bmatrix} B_t - B^{ss} \\ C_t - C^{ss} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 - \sqrt{1 - \frac{4\alpha^2\sigma_x^2\sigma_\varepsilon^2}{R^2}} \end{bmatrix} \begin{bmatrix} B_{t+1} - B^{ss} \\ C_{t+1} - C^{ss} \end{bmatrix}$$

Or:

$$\begin{bmatrix} B_t - B^{ss} \\ C_t - C^{ss} \end{bmatrix} = \begin{bmatrix} 0 \\ \left(1 - \sqrt{1 - \frac{4\alpha^2\sigma_x^2\sigma_\varepsilon^2}{R^2}}\right)\left[C_{t+1} - C^{ss}\right] \end{bmatrix}$$

Or

$$\begin{aligned} C_t - C^{ss} &= \left(1 - \sqrt{1 - \frac{4\alpha^2\sigma_x^2\sigma_\varepsilon^2}{R^2}}\right)[C_{t+1} - C^{ss}] \\ [C_{t+1} - C^{ss}] &= \frac{1}{\left(1 - \sqrt{1 - \frac{4\alpha^2\sigma_x^2\sigma_\varepsilon^2}{R^2}}\right)}[C_t - C^{ss}] \end{aligned}$$

As $\frac{1}{\left(1-\sqrt{1-\frac{4\alpha^2\sigma_x^2\sigma_\varepsilon^2}{R^2}}\right)} > 1$, the system diverges, in the sense that any perturbation of $C$ would drive $C_t$ away. Thus $\lambda = 0$ is not stable.

One can follow a similar step and show that the Jocobian matrix of the INVERSE transition function evaluated at $\lambda = 1$, when $\sigma_\theta \to 0$, is:

$$\begin{matrix} \frac{1}{R} & 0 \\ 0 & \left(1 - \sqrt{1 - \frac{4\alpha^2\sigma_x^2\sigma_\varepsilon^2}{R^2}}\right) \end{matrix}$$

With eigenvalues $\frac{1}{R}$ and $\left(1 - \sqrt{1 - \frac{4\alpha^2\sigma_x^2\sigma_\varepsilon^2}{R^2}}\right)$

Thus the eigenvalues of the Jocobian matrix of the transition function is just the inverse of these eigenvalues:

$$\begin{aligned} R &> 1 \\ \frac{1}{\left(1 - \sqrt{1 - \frac{4\alpha^2\sigma_x^2\sigma_\varepsilon^2}{R^2}}\right)} &> 1 \end{aligned}$$

Thus $\lambda = 1$ is also unstable.

## Claim 3

We next show that the interior steady state is saddle. At the interior steady state $\lambda_t$ may vary with perturbation to $(B_t, C_t)$ .To ease exposition, let us denote the dynamic system in the following way:

$$\begin{aligned} X(\lambda_t, S_{t+1}) &= e^{2\alpha c} \\ S_t &= Y(\lambda_t, S_{t+1}) \end{aligned}$$

Where $S_t = (B_t, C_t)$ is the state.

$$Y(\lambda_t, S_{t+1}) = \begin{bmatrix} \frac{1}{R}\lambda_t (B_{t+1} + 1) L_t \\ \frac{1}{R}\alpha \left[C_{t+1}^2\sigma_x^2 + \sigma_\varepsilon^2\right] L_t \end{bmatrix}$$

$$X\left(\lambda_t, S_{t+1}\right) = 1 + \frac{\alpha^2\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]\sigma_x^2}{\lambda_t^2\left(B_{t+1}+1\right)^2\sigma_\theta^2 + \alpha^2\left[\sigma_\varepsilon^2 + C_{t+1}^2\sigma_x^2\right]^2\sigma_x^2}\left(B_{t+1}+1\right)^2\sigma_\theta^2$$

The goal is to solve for the quadratic equation of

$$Q(u) = \begin{vmatrix} u - Y_{1B}\left(\lambda_t\left(S_{t+1}\right), S_{t+1}\right) & -Y_{1C}\left(\lambda_t\left(S_{t+1}\right), S_{t+1}\right) \\ -Y_{2B}\left(\lambda_t\left(S_{t+1}\right), S_{t+1}\right) & u - Y_{2C}\left(\lambda_t\left(S_{t+1}\right), S_{t+1}\right) \end{vmatrix}$$

Where $Y_{1B}\left(\lambda_t\left(S_{t+1}\right), S_{t+1}\right)$ is the derivative of first argument of $Y$ with respect to $B$.where $\lambda_t\left(S_{t+1}\right)$ is implicitly defined by $X\left(\lambda_t, S_{t+1}\right) = e^{2\alpha c}$.

The goal is to show that $Q(1) < 0$ so there is a route of the backward transition function that is larger than 1. So that there is a route of the forward transition function that is smaller than 1 in absolute value. This establishes claim 3. To show this, we need the following lemma

**Lemma B.1**

$$\pi'(\lambda) = \frac{X_\lambda}{\det\left[I - Y_S\right]} Q(1)$$

***Proof.*** *Denote $I$ the identity matrix. $\pi(\lambda)$ is given by:*

$$\pi(\lambda) = X(\lambda, S)$$

*subject to*

$$S = Y(\lambda, S)$$

*Thus*

$$\pi'(\lambda) = X_\lambda + X_s\left[I - Y_s\right]^{-1} Y_\lambda$$

*Note that $S$ is double dimensional, we can write out each elements in the matrix:*

$$\begin{aligned} X_\lambda &= \textit{some scalar} \\ X_S &= \begin{bmatrix} \pi_1 \\ \pi_2 \end{bmatrix}' \\ Y_\lambda &= \begin{bmatrix} a \\ b \end{bmatrix} \\ Y_S &= \begin{bmatrix} c & d \\ e & f \end{bmatrix} \end{aligned}$$

*Use the above notation:*

$$\begin{aligned} \pi'(\lambda) &= X_\lambda + X_s\left[I - Y_s\right]^{-1} Y_\lambda \\ &= X_\lambda + \begin{bmatrix} \pi_1 & \pi_2 \end{bmatrix} \begin{bmatrix} 1-c & -d \\ -e & 1-f \end{bmatrix}^{-1} \begin{bmatrix} a \\ b \end{bmatrix} \end{aligned}$$

*Note that:*

$$\begin{bmatrix} 1-c & -d \\ -e & 1-f \end{bmatrix}^{-1} = \frac{1}{\det\left[I - Y_S\right]} \begin{bmatrix} 1-f & d \\ e & 1-c \end{bmatrix}$$

*Hence:*

$$\begin{aligned}
\pi'(\lambda) &= X_\lambda + \frac{1}{\det[I-Y_S]}\begin{bmatrix}\pi_1 & \pi_2\end{bmatrix}\begin{bmatrix}1-f & d\\ e & 1-c\end{bmatrix}\begin{bmatrix}a\\ b\end{bmatrix}\\
&= \frac{X_\lambda}{\det[I-Y_S]}\left[\det\begin{bmatrix}1-f & d\\ e & 1-c\end{bmatrix} + \frac{1}{X_\lambda}\begin{bmatrix}\pi_1 & \pi_2\end{bmatrix}\begin{bmatrix}1-f & d\\ e & 1-c\end{bmatrix}\begin{bmatrix}a\\ b\end{bmatrix}\right]
\end{aligned}$$

*Now*

$$\begin{aligned}
Q(u) &= \begin{bmatrix} u - Y_{1B}(\lambda_t(S_{t+1}), S_{t+1}) & -Y_{1C}(\lambda_t(S_{t+1}), S_{t+1})\\ -Y_{2B}(\lambda_t(S_{t+1}), S_{t+1}) & u - Y_{2C}(\lambda_t(S_{t+1}), S_{t+1})\end{bmatrix}\\
&= \begin{bmatrix}u & 0\\ 0 & u\end{bmatrix} - \begin{bmatrix} Y_{1B}(\lambda_t(S_{t+1}), S_{t+1}) & Y_{1C}(\lambda_t(S_{t+1}), S_{t+1})\\ Y_{2B}(\lambda_t(S_{t+1}), S_{t+1}) & Y_{2C}(\lambda_t(S_{t+1}), S_{t+1})\end{bmatrix}\\
&= uI - (Y_s + Y_\lambda \lambda_s)
\end{aligned}$$

*Where* $\lambda_s$ *is implicitly defined by* $X(\lambda_t, S_{t+1}) = e^{2\alpha c}$*.Thus* $\lambda_s = -\frac{X_S}{X_\lambda}$*. Hence:*

$$\begin{aligned}
Q(1) &= \det\left[I - Y_S + Y_\lambda \frac{X_S}{X_\lambda}\right] = \det\left[I - \frac{Y_S X_\lambda - Y_\lambda X_S}{X_\lambda}\right]\\
&= \det\left[I - \frac{\begin{bmatrix}c & d\\ e & f\end{bmatrix} X_\lambda - \begin{bmatrix}a\\ b\end{bmatrix}\begin{bmatrix}\pi_1 & \pi_2\end{bmatrix}}{X_\lambda}\right]\\
&= \det\left[I - \begin{bmatrix}c & d\\ e & f\end{bmatrix} + \frac{1}{X_\lambda}\begin{bmatrix}a\pi_1 & a\pi_2\\ b\pi_1 & b\pi_2\end{bmatrix}\right]\\
&= \det\left[I - \begin{bmatrix} c - \frac{1}{\pi_\lambda} a\pi_1 & d - \frac{1}{\pi_\lambda} a\pi_2\\ e - \frac{1}{\pi_\lambda} b\pi_1 & f - \frac{1}{\pi_\lambda} b\pi_2\end{bmatrix}\right]
\end{aligned}$$

$$
\begin{aligned}
\det\left[I-\left[\begin{array}{cc} c-\frac{1}{\pi_\lambda}a\pi_1 & d-\frac{1}{\pi_\lambda}a\pi_2 \\ e-\frac{1}{\pi_\lambda}b\pi_1 & f-\frac{1}{\pi_\lambda}b\pi_2 \end{array}\right]\right] &= 1-\left[c-\frac{1}{\pi_\lambda}a\pi_1+f-\frac{1}{\pi_\lambda}b\pi_2\right]+\left[c-\frac{1}{\pi_\lambda}a\pi_1\right]\left[f-\frac{1}{\pi_\lambda}b\pi_2\right] \\
&\quad -\left[e-\frac{1}{\pi_\lambda}b\pi_1\right]\left[d-\frac{1}{\pi_\lambda}a\pi_2\right] \\
&= 1-\left[c-\frac{1}{\pi_\lambda}a\pi_1+f-\frac{1}{\pi_\lambda}b\pi_2\right]+\left[c-\frac{1}{\pi_\lambda}a\pi_1\right]\left[f-\frac{1}{\pi_\lambda}b\pi_2\right] \\
&\quad -\left[e-\frac{1}{\pi_\lambda}b\pi_1\right]\left[d-\frac{1}{\pi_\lambda}a\pi_2\right] \\
&= 1-\left[c-\frac{1}{\pi_\lambda}a\pi_1+f-\frac{1}{\pi_\lambda}b\pi_2\right]+cf-c\frac{1}{\pi_\lambda}b\pi_2-f\frac{1}{\pi_\lambda}a\pi_1-ed \\
&\quad +e\frac{1}{\pi_\lambda}a\pi_2+d\frac{1}{\pi_\lambda}b\pi_1 \\
&= 1-f-c+fc-de+\frac{1}{\pi_\lambda}a\pi_1+\frac{1}{\pi_\lambda}b\pi_2 \\
&\quad -c\frac{1}{\pi_\lambda}b\pi_2-f\frac{1}{\pi_\lambda}a\pi_1+e\frac{1}{\pi_\lambda}a\pi_2+d\frac{1}{\pi_\lambda}b\pi_1 \\
&= \det\left[\begin{array}{cc} 1-f & d \\ e & 1-c \end{array}\right]+ \\
&\quad \frac{1}{\pi_\lambda}\left[a\left(\pi_1-f\pi_1+e\pi_2\right)+b\left(\pi_2-c\pi_2+d\pi_1\right)\right] \\
&= \det\left[\begin{array}{cc} 1-f & d \\ e & 1-c \end{array}\right]+ \\
&\quad \frac{1}{\pi_\lambda}\left[\begin{array}{cc} \pi_1 & \pi_2 \end{array}\right]\left[\begin{array}{cc} 1-f & d \\ e & 1-c \end{array}\right]\left[\begin{array}{c} a \\ b \end{array}\right]
\end{aligned}
$$

*Hence:* $\pi'(\lambda) = \frac{X_\lambda}{\det[I-Y_S]} Q(1)$ ■

Also, as $\sigma_\theta \to 0$,

$$[I-Y_S] \to \left[\begin{array}{cc} 1-\frac{1}{R}\lambda & 0 \\ 0 & 1-\frac{1}{R}\alpha\left[2C^{ss}\sigma_x^2\right] \end{array}\right]$$

Hence

$$\det[I-Y_S] = \left[1-\frac{1}{R}\lambda\right]\left[1-\frac{1}{R}\alpha\left[2C^{ss}\sigma_x^2\right]\right] > 0$$

In view that $\pi'(\lambda) > 0, \det[I-Y_S] > 0, X_\lambda < 0$ and the lemma, we get:

$$Q(1) < 0$$

Thus there is 1 eigenvalue larger than 1, denote it $u^*$. By property of the inverse Jacobian matrix, there exists an eigenvalue that is the reciprocal of $u^* : \frac{1}{u^*} < 1$. Hence there exists a stable manifold $\phi(B, C)$ around the steady state such that if an equilibrium starts on that manifold, it will converge to the steady state. Thus the middle steady state can be either stable or saddle-path stable.