

Learning from Speculating and the No Trade Theorem

Dmitry Lubensky*
Kelley School of Business
Indiana University

Doug Smith
Federal Trade Commission[†]

October 15, 2014

Abstract

Due to the no trade theorems of Aumann (1976) and others, models of speculative trade have relied on the presence of irrational noise traders. We present a model in which trade occurs when all agents are fully rational speculators that share a common prior and trade a common value asset based only on private information. The motive that helps agents overcome adverse selection, and hence the no trade result, is experimentation. An agent that engages in trade receives a payoff but also learns about his type. If he learns he has a low skill the agent exits while if he learns he has a high skill he remains and continues to benefit from trade against future cohorts. Younger cohorts appear as noise traders in a single period snapshot since they enter unprofitable trades, thus the model provides a rational foundation for some perceived behavioral biases including overconfidence.

*dlubensk@indiana.edu

[†]dsmith2@ftc.gov.

The views expressed in this article are those of the authors and do not necessarily reflect those of the Federal Trade Commission or any Commissioner.

1 Introduction

An important discovery of economic theory is that asset trade cannot occur solely for informational reasons. This fact was initially conceived by Aumann (1976) and sparked a robust literature which continued to find this result in increasingly complex settings (seminal works include Milgrom and Stokey (1982), Kreps (1977), and Tirole (1982)). As it turns out, regardless of how one models a market, rational agents cannot trade for purely speculative purposes. Tirole states that even in a dynamic setting, “speculation relies on inconsistent plans and is ruled out by rational expectations”.

The idea behind these results is adverse selection. A buyer of an asset is concerned that the seller trades only when he has some private negative information. Since in equilibrium the seller only agrees to trade when it gives him a positive payoff, if trades are zero-sum then the buyer is not willing to trade himself.

To overcome adverse selection, a speculator must believe that his counterparty is willing to lose money. This can occur if the opposing trader has ulterior motives for asset ownership such as the hedging of risk or liquidity concerns, or if the opposing trader is also a speculator but suffers from behavioral biases or a misunderstanding of the market environment. Starting with Kyle (1985) and Glosten and Milgrom (1985), a myriad of asset trading models has emerged which relies on the existence of these traders that are willing to lose money, the so-called noise traders, in order to study the behavior of informed speculators.

In this paper, we demonstrate that the presence of noise traders is not necessary to generate information-based speculative trade. Instead, we present a model in which all agents are rational speculators that trade a purely common value asset based only on private information. The motive that helps agents overcome adverse selection is experimentation. By engaging in a trade, an agent receives a payoff but also learns about his type; if he learns he has a low skill the agent exits the market while if he learns he has a high skill, he may remain and continue to make profitable trades in the future. Adverse selection is still present and agents sometimes enter trades that on average will lose them money, however the value of the information learned from the trade overcomes this loss.

We study a dynamic environment in which in every period a cohort of new agents is born and each agent has the opportunity to transact in an existing pool. Agents have either a high or low skill but do not directly observe this; instead each may learn about his skill through trading. In every period agents receive private signals about the value of an asset, with high skill agents receiving signals that are more accurate. Agents are then randomly pairwise matched and choose whether to buy, sell, or abstain from trading conditional on their private signals. This interaction generates a stage payoff for each

agent and provides information to help them learn their type. For instance, an agent whose pre-trade signal predicts a high asset value but whose realized payoff reveals a low asset value may shift his posterior toward the low-skill type. The period ends with a pool of traders, each holding a belief about their type that depends on their trading history. In the next period, there is an inflow of a new cohort and an outflow of existing traders at an exogenous rate. We solve for the steady state equilibrium of this market.

We consider the case in which assets are purely common value, so that the sum of the payoffs of two matched traders is zero. In a steady state, the pool of active traders is comprised of agents with varying beliefs about their skill, in which the new cohort's belief is the most pessimistic. If agents in this pool played for one period only, the no-trade theorem would arise as agents would be concerned about trading against a more skilled counterparty. However, in the dynamic environment even the most pessimistic traders, i.e. agents in the new cohort, still find it optimal to trade despite facing an adversely selected pool and expecting to lose money on the first trade. Their rationale is that if through trading they learn that they are a high skill type, new agents can continue to trade in future periods and earn a positive continuation value while if they learn that they are a low skill type, they can cut their losses by essentially exiting and abstaining from future trading. We demonstrate that in the zero sum environment, the information value of the first trade exactly offsets the losses from adverse selection and thus an equilibrium is supported in which all new agents as well agents from older cohorts trade.

In a single period snapshot of our dynamic model, new agents willingly lose money on that period's trade and thereby are observationally similar to the noise traders in the classic models such as Kyle (1985) and Glosten and Milgrom (1985). In this sense, one interpretation of our model is that it identifies learning about one's skill as a foundation for the existence of noise traders. Because over their lifetimes our noise traders do not lose money, we avoid the concern that markets ought to weed out ineffective traders. In fact, our model directly describes the way in which this selection occurs. Furthermore, we highlight the fact that it is not only speculators, i.e. the experienced traders within our model that on average make money on this period's trade, but also the noise traders that make decisions based on information about the value of the asset. This is an important distinction when evaluating how well the market aggregates available information, or how changes in the uncertainty over the asset's value affect trading patterns, or a myriad of similar lines of inquiry.

Our model may also provide a rational justification for trading previously explained by behavioral biases. For example, in any given period there is a set of agents that are overconfident. These are traders that have experienced a history of fortuitous outcomes despite having low skill and whose future earnings will end up lower than their current expectation. However, these traders are not irrational – their overconfidence is simply

a result of proper Bayes updating and over time will be corrected. In our model, the number of overconfident traders is not an exogenous parameter but rather emerges from the fundamental information structure of the market.

Several recent empirical studies support the idea that learning about one's skill as a trader provides an important rationale for trading. Barber, Lee, Liu, and Odean (2010) demonstrates that traders expect future performance to be correlated with their past performance and are more likely to exit when initially losing money. Seru, Shumway, and Stoffman (2010) finds a similar result at the individual level and also shows that older cohorts outperform younger cohorts, mostly due to the selection inherent in the exit decision. Some theoretical work has also examined the role of learning in speculative trade. Bond and Eraslan (2010) supports information-based bilateral trade by having the process of trading release information that improves the productive value of the asset for the owner. In our work the process of trade is also necessary to release information, but by contrast the information is not productive as there are no gains from trade in the market. Closer in spirit is Gervais and Odean (2001), in which speculators learn about their skill as they trade in multiple periods. The paper focuses on boundedly rational speculators that over-attribute prior success to skill, and demonstrate that overconfidence is strongest early in a speculator's career and diminishes over time. In our setting a similar dynamic occurs but due to the overlapping generations structure we can support trade without the use of noise traders and generate behavior that is rational in the long run but seemingly overconfident in the short run.

The rest of the paper proceeds as follows. Section 2 presents an illustrative example in which we describe a simple information structure such that purely speculative zero-sum trade is supported in steady state. Section 3 then describes the general model and proves the existence of information-based trade in this substantially broader environment. In general, there will also be a zero-trade equilibrium and section 4 discusses reasons why the equilibrium with trade may be more likely to occur. Section 5 then concludes.

2 Learning about Ability: An Example

In this section we develop an example of a simple overlapping generations game that captures many of the important features of the main model and provides an outline of the arguments used for our main result in the general case. The goal of this section is to provide a closed-form tractable solution that demonstrates (i) that there exists a steady state equilibrium with trade and (ii) that no trade would occur in the single-period snapshot of this environment. Many of our particular assumptions in this section, for instance the distributions of asset values, are made to facilitate exposition and are relaxed in the main model.

Entry and exit. Time flows discretely in periods (... , t-1, t, t+1, ...). Each period a measure one of new agents enter and each agent i independently draws a type

$$\theta_i = \begin{cases} G & \text{with probability } \frac{1}{2} \\ B & \text{with probability } \frac{1}{2} \end{cases}.$$

Types correspond to whether an agent is good (G) or bad (B) at obtaining or interpreting information about the value of an asset. Entering agents do not directly observe their type but they may learn about it by interacting in the market.

Agents live exactly two periods. We refer to agents in their first period as "young" and to agents in their second period as "old". The pool of agents is comprised of two overlapping generations and all agents in the pool are randomly matched, regardless of age.

Asset Value. In any given period, an asset has a value $v = v_1 + v_2$, in which v_1 and v_2 are independent components of the asset. Each component is stochastic and drawn from a distribution with density $\varphi(v_i | \alpha_i)$, which depends on whether it is a high state $\alpha_i = -1$ or a low state $\alpha_i = 1$ for that component, both equally likely. The density in the high state is

$$\varphi(v_i | \alpha_i = 1) = \begin{cases} \frac{2}{5}(1 + v_i) & \text{if } v_i \in [-1, 0) \\ \frac{4}{5} & \text{if } v_i \in [0, 1] \end{cases},$$

and the density in the low state is

$$\varphi(v_i | \alpha_i = -1) = \begin{cases} \frac{4}{5} & \text{if } v_i \in [-1, 0) \\ \frac{2}{5}(1 - v_i) & \text{if } v_i \in [0, 1] \end{cases}.$$

States are symmetric so that $\varphi(v_i | \alpha_i = 1) = \varphi(-v_i | \alpha_i = -1)$, and the distributions satisfy $E[v_i | \alpha_i = -1] < E[v_i] = 0 < E[v_i | \alpha_i = 1]$.

Stage game. In every period agents are randomly matched in pairs and assigned an asset that they may trade. Within a pair, one agent receives a private signal $s_1 \in \{-1, 1\}$ about state α_1 and the other agent a private signal $s_2 \in \{-1, 1\}$ about state α_2 . The quality

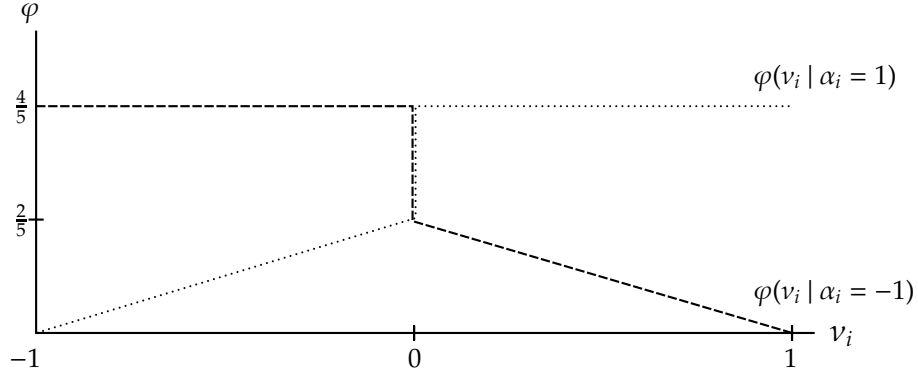


Figure 1: State-dependent densities for component v_i

of the signal depends on the agent's type. Specifically,

$$Pr(s_i = \alpha_i) = \begin{cases} 1 & \text{if } \theta_i = G \\ \frac{1}{2} & \text{if } \theta_i = B \end{cases},$$

so that if agent i is the type G his signal reveals the state α_i with perfect accuracy, while if the agent is type B then his signal is noise.

After each observes their private signals agents simultaneously choose whether to buy, sell, or stay out. Trade occurs if and only if one agent chooses buy and the other agent chooses sell. If a trade occurs then the buyer's payoff is $v_1 + v_2$, the seller's payoff is $-(v_1 + v_2)$, and both observe realizations v_1 and v_2 . If no trade occurs then both agents do not observe the realizations of the components and receive a payoff of zero. In addition, suppose that agent i directly observes agent j 's action whenever agent i chooses "buy" or "sell" but not "stay out".¹

A Trading Equilibrium

We will now show that the following is an equilibrium of the overlapping generations game: all young traders decide to participate in the market, and choose "buy" if their $s_i = 1$ and "sell" if their $s_i = -1$. Old traders participate if and only if they became more optimistic about their skill after their first trade.

¹In the present setting this observation plays no role, however in the more general setup in the following section, the counterparty's action may provide a further source of inference about one's own type.

Proposition 1 *The following is an equilibrium strategy of the above game. Every young agent i chooses buy if $s_i = 1$ and sell if $s_i = -1$. Every old agent i stays out if in the previous period $s_i = 1$ and $v_i < 0$ or $s_i = -1$ and $v_i \geq 0$, else chooses buy if in the current period $s_i = 1$ and sell if in the current period $s_i = -1$.*

We check whether this is an equilibrium in two steps. First we determine the composition of the steady state pool of participating agents if the above strategy is followed. Then we check whether the strategy is a best response when the pool is thus composed.

In the proposed strategy, since all young agents participate (choose buy or sell), the masses of good and of bad young traders are $N(\text{young } G) = N(\text{young } B) = \frac{1}{2}$. To determine the masses of participating old traders of type G and B, we must account for these agents' history from the previous period. Every good old agent must have been born type G, traded as a young agent, and observed a favorable realization after the trade. Therefore, the amount of old good types participating is equal to

$$\begin{aligned} N(\text{old } G) &= N(\text{young } G) \cdot \Pr(\text{trade when young } |G) \cdot \Pr(\text{favorable realization } |G) \\ &= \frac{1}{2} \cdot \left(\frac{1 + N(\text{old } G) + N(\text{old } B)}{2} \cdot \frac{1}{2} \right) \cdot \frac{4}{5} \\ &= \frac{1}{10} (1 + N(\text{old } G) + N(\text{old } B)) \end{aligned}$$

Above, the probability of trading when young equals the probability of being matched with another participating agent, which is $\frac{1}{2}(1 + N(\text{old } G) + N(\text{old } B))$, times the probability of receiving the opposite signal of one's counterparty, which equals $\frac{1}{2}$. The probability of seeing a favorable realization from trading is $\Pr(s_i = 1, v_i \geq 0 |G) + \Pr(s_i = -1, v_i < 0 |G) = \frac{4}{5}$.

There are also old bad agents, that were bad in their first period, traded, and received a favorable realization. The number of these traders is

$$\begin{aligned} N(\text{old } B) &= N(\text{young } B) \cdot \Pr(\text{trade when young } |B) \cdot \Pr(\text{favorable realization } |B) \\ &= \frac{1}{2} \cdot \left(\frac{1 + N(\text{old } G) + N(\text{old } B)}{2} \cdot \frac{1}{2} \right) \cdot \frac{1}{2} \\ &= \frac{1}{16} (1 + N(\text{old } G) + N(\text{old } B)) \end{aligned}$$

The number of bad old agents differs from the number of good old agents only in the final term, the probability of a favorable realization, which for the bad agent equals $\frac{1}{2}$ since his signals about the asset value are uninformative. Solving the system of the two equations above obtains $N(\text{old } G) = \frac{8}{67}$ and $N(\text{old } B) = \frac{5}{67}$. Accordingly, the proportion of agents of

type G in the pool of participants is

$$\begin{aligned}\mu_s &\equiv \frac{N(\text{young } G) + N(\text{old } G)}{N(\text{young } G) + N(\text{old } G) + N(\text{young } B) + N(\text{old } B)} \\ &= \frac{\frac{1}{2} + \frac{8}{67}}{\frac{1}{2} + \frac{8}{67} + \frac{1}{2} + \frac{5}{67}} \\ &= \frac{83}{160}.\end{aligned}$$

We now check whether conditional on this proportion, the strategy is a best response for both young and old traders. We first look at whether old types who participate are best responding. For concreteness we focus on old traders who in the previous period were buyers. Due to the symmetry of our setup, the results translate immediately to old traders who in the previous period were sellers.

First, define

$$\bar{v}_G \equiv E[v_i | \theta_i = G, s_i = 1] = \int_{-1}^1 v_i \varphi(v_i | \alpha_i = 1) dv_i = \frac{1}{3}$$

as the expected value of component i for a buyer of type G and define

$$\bar{v}_B \equiv E[v_i | \theta_i = B, s_i = 1] = \frac{1}{2} \cdot \int_{-1}^1 v_i \varphi(v_i | \alpha_i = 1) dv_i + \frac{1}{2} \cdot \int_{-1}^1 v_i \varphi(v_i | \alpha_i = -1) dv_i = 0$$

as the expected value of component i for a buyer of type B. The former is strictly positive because the agent's signal is informative, while the latter is zero because the signal is noise and the posterior equals the prior. A buyer with belief μ that his type is G has an expected stage payoff conditional on trading given by

$$u_i(\mu, \mu_s) = E[v_i | \mu] + E[v_j | \mu_s] = (\bar{v}_G \cdot \mu + \bar{v}_B \cdot (1 - \mu)) - (\bar{v}_G \cdot \mu_s + \bar{v}_B \cdot (1 - \mu_s)) = \bar{v}_G(\mu - \mu_s).$$

Since old traders maximize their expected stage payoff only, it is sufficient to verify that an old buyer's posterior belief $\mu(v_i) \geq \mu_s$ if and only if $v_i \geq 0$. For an agent that received the buy signal $s_i = 1$, the posterior at $v_i = 0$ is given by

$$\begin{aligned}\mu(v_i = 0, s_i = 1) &= \frac{\varphi(v_i = 0 | \alpha_i = 1)}{\varphi(v_i = 0 | \alpha_i = 1) + \frac{1}{2}(\varphi(v_i = 0 | \alpha_i = 1) + \varphi(v_i = 0 | \alpha_i = -1))} \\ &= \frac{\frac{4}{5}}{\frac{4}{5} + \frac{1}{2}\left(\frac{4}{5} + \frac{2}{5}\right)} \\ &= \frac{4}{7}.\end{aligned}$$

It is easy to check that $\mu(v_i)$ is strictly increasing, thus since $\frac{4}{7} > \mu_s = \frac{83}{160}$, all old agents who as young buyers observed $v_i \geq 0$ best respond by participating. Similarly, due to the discontinuity at $\varphi(v_i = 0)$, it can immediately be seen that the agent's posterior about his type at any $v_i < 0$ is below his prior of $\frac{1}{2}$, since conditional on signal $s_i = 1$ these outcomes are associated more with type B than type G. Thus all old agents who as buyers in the first round observed $v_i < 0$ do not participate. Lastly, to establish that old agents that did not trade in the first period do not participate, note that these traders' belief is $\mu = \frac{1}{2} < \mu_s = \frac{83}{160}$.

It remains now to demonstrate that it is a best response for all young agents to participate. Young buyers face both a stage and a continuation payoff, and their lifetime utility is given by

$$u_0 = \Pr(\text{trade}) \cdot \left(\Pr(G) \left[u_i(1, \mu_s) + \Pr(v_i \geq 0 | G, s_i = 1) \Pr(\text{trade}) u_i(1, \mu_s) \right] \right. \\ \left. + \Pr(B) \left[u_i(0, \mu_s) + \Pr(v_i \geq 0 | B, s_i = 1) \Pr(\text{trade}) u_i(0, \mu_s) \right] \right).$$

A young buyer, if matched with a seller, receives a stage payoff and, conditional on receiving a favorable realization, may trade again in the next period. Recalling that conditional on attempting to buy, the probability of trade equals the probability of being matched with a participating agent times the probability of having opposing signals, that is $\Pr(\text{trade}) = \frac{1}{2} \left(1 + \frac{8}{67} + \frac{5}{67} \right) \cdot \frac{1}{2} = \frac{20}{67}$, the young trader's expected utility evaluates as follows:

$$u_0 = \frac{20}{67} \cdot \left(\frac{1}{2} \left[\bar{v}_G(1 - \mu_s) + \frac{4}{5} \cdot \frac{20}{67} \cdot \bar{v}_G(1 - \mu_s) \right] \right. \\ \left. + \frac{1}{2} \left[\bar{v}_G(0 - \mu_s) + \frac{1}{2} \cdot \frac{20}{67} \cdot \bar{v}_G(0 - \mu_s) \right] \right) \\ = \frac{20}{67} \cdot \frac{1}{2} \bar{v}_G(1 - \mu_s) \left(\left[1 + \frac{4}{5} \cdot \frac{20}{67} \right] - \frac{\mu_s}{1 - \mu_s} \left[1 + \frac{1}{2} \cdot \frac{20}{67} \right] \right).$$

Plugging in $\mu_s = \frac{28}{53}$ obtains $u_0 = 0$, implying that it is a best response for all young traders to participate. We have thus shown that the conjectured trading strategies induce a steady state in which the strategies are best responses, and this concludes the proof. ■

No Trade Theorem

In the previously described equilibrium, the probability that a randomly selected agent trades in a given period is the probability that both the agent and his counterparty choose

to participate times the probability they receive opposing signals, that is

$$\Pr(\text{agent trades}) = \left(\frac{1 + \frac{8}{67} + \frac{5}{67}}{2} \right)^2 \cdot \frac{1}{2} \approx 0.18.$$

Put differently, of the instances in which matched agents observe opposing signals, trade occurs approximately 36% of the time. By contrast, we now demonstrate that in a single snapshot of this game the no-trade theorem would obtain. Specifically, if we take the steady state distribution of agents and allow them to trade for one period only, no trade would occur.

Proposition 2 *Let $Z(\mu)$ denote the probability distribution of the posteriors of participating agents in the steady state equilibrium of Proposition 1. If the stage game were played with no continuation, the probability of trade would be zero.*

To prove this, first note that $Z(\cdot)$ has an atom at $\mu = \frac{1}{2}$ corresponding to young traders and that it has full support on $\mu \in [\frac{1}{2}, 1)$. It is immediate that in any equilibrium of the one shot game, for agents who decide to either buy or sell both buyers and sellers must follow a threshold strategy so that there exists a belief $\bar{\mu}_b$ for buyers and $\bar{\mu}_s$ for sellers so that agents agree to participate only if their belief is higher than their respective threshold. The marginal seller expects to trade with the average buyer and the marginal buyer expects to trade with the average seller, and both are willing to participate only if

$$E[v | \mu_b \geq \bar{\mu}_b, \mu_s = \bar{\mu}_s] \leq 0 \leq E[v | \mu_b = \bar{\mu}_b, \mu_s \geq \bar{\mu}_s].$$

However, since decreasing μ_s and increasing μ_b increases the expected value of the asset, it must be that

$$E[v | \mu_b = \bar{\mu}_b, \mu_s \geq \bar{\mu}_s] \leq E[v | \mu_b = \bar{\mu}_b, \mu_s = \bar{\mu}_s] \leq E[v | \mu_b \geq \bar{\mu}_b, \mu_s = \bar{\mu}_s].$$

Both sets of inequalities hold only if

$$E[v | \mu_b \geq \bar{\mu}_b, \mu_s = \bar{\mu}_s] = E[v | \mu_b = \bar{\mu}_b, \mu_s \geq \bar{\mu}_s],$$

which in turn implies that $\bar{\mu}_b = \bar{\mu}_s = 1$. Thus trade can occur only between the maximally optimistic buyer and the maximally pessimistic seller. In the equilibrium we describe above, the mass of these traders in the steady state pool is zero. ■

The example presented in this section demonstrates that a substantial amount of trade can occur for purely informational purposes, as more than one third of agents with opposing signals engage in trade despite there being no aggregate value from doing so. In light of the robustness of the no trade theorems, this result may be surprising, and our aim is now to explore which assumptions were necessary for it to obtain. As we will demonstrate in

the next section, many of the particulars of the example, such as the functional form of the distributions of asset value, the independence of signals, the length of an agent's life, and even the connection between stage payoffs and inference about skill are irrelevant for the existence of trade. The key feature is that information about one's type may only be obtained by attempting to trade.

3 Learning about Ability: A General Model

Entry and Exit. Time flows discretely in periods $(\dots, t-1, t, t+1, \dots)$. Each period a measure one of agents is born and each agent i independently draws a type

$$\theta_i = \begin{cases} G & \text{with probability } \mu_0 \\ B & \text{with probability } 1 - \mu_0 \end{cases},$$

where $\mu_0 \in (0, 1)$. Entering agents do not observe their type directly, although they may learn about it by interacting in the market. Agents exit the market at an exogenous rate δ per period and live at most T periods. Assume the expected lifetime is finite so that $\min(\frac{1}{\delta}, T) < \infty$.

Stage Game. In every period agents are matched pairwise and have the opportunity to trade an asset with an uncertain common value v , normalized so that $E[v|\theta_i, \theta_j] = 0$. Each agent i receives a private binary signal $s_i \in \{-1, 1\}$ about the realization of v . Signals are generated from a joint distribution $F(v, s_i, s_j|\theta_i, \theta_j)$, and have the following properties. First, a higher signal is associated with a higher expected asset value:

$$E[v|s_i = -1, s_j, \theta_i, \theta_j] \leq E[v|s_i = 1, s_j, \theta_i, \theta_j].$$

Next, signals received by good agents are more meaningful than those received by bad agents, in terms of the expected value of the asset:

$$E[v|s_i = 1, s_j, \theta_i, \theta_j] - E[v|s_i = -1, s_j, \theta_i, \theta_j] \text{ increases in } \theta_i \text{ (from B to G)}.$$

This condition would hold, for instance, if the signals received by bad agents were a garbling of the signals received by good agents, as in the example of the preceding section. Finally, positive and negative signals s_i are of the same strength:

$$E[v|s_i, s_j, \theta_i, \theta_j] = -E[v|-s_i, -s_j, \theta_i, \theta_j].$$

Recalling that the ex-ante expected value is zero, this condition ensures that holding fixed agents' types, whenever a set of signals (s_i, s_j) increases the expected asset value above zero, the set of opposite signals $(-s_i, -s_j)$ reduces the expected value below zero by the

same amount.

Having observed their private signal each agent i simultaneously chooses action $a_i \in \{\text{buy, sell, stay out}\}$. Trade occurs if and only if one agent chooses “buy” and the other agent chooses “sell”. If a trade occurs then the buyer’s payoff is v and the seller’s payoff is $-v$; if no trade occurs both agents receive a payoff of zero.

Inference from Trading. At the end of the stage, if an agent chose $a_i \in \{\text{buy, sell}\}$ then he observes the other player’s action a_j along with an additional signal $\gamma_i \sim \Psi(\gamma_i | a_j, \theta_i, \theta_j, s_i, s_j, v)$ with support on $[0, 1]$, which he may use to draw inference about his type. In the opening example γ_i revealed both components of the asset’s value, however in principle can provide information in a variety of ways, including but not limited to observing the relationship between the realized outcome v and the agent’s own signal. We assume that agent i ’s choice of action – beyond choosing to participate – does not affect the signal γ_i that he receives. This assumption rules out the possibility that trading is inherently more informative than not trading, independent of observed and inferred signals. Such situations introduce potential complications that we do not explore in this paper. To ensure that learning occurs when actively pursuing a trade, the agent does not observe γ_i or a_j when playing “stay out”.

Let $y_i = (s_i, a_i, a_j, \gamma_i)$ be player i ’s stage observation conditional on participating, and let $y_i = (s_i, \text{stay out})$ be the observation when not participating. Let $h = (y_1, \dots, y_t)$ denote an agent’s history with t observations and let H denote the set of all feasible histories.

Steady state equilibrium

A steady state equilibrium is described by two objects: a distribution of the composition of active agents and a strategy σ^* prescribing an agent’s action for each history such that (i) is a best response given the distribution and (ii) keeps the distribution invariant across time. We begin by describing strategies and then demonstrate the existence of a steady state equilibrium in which all new agents participate.

Stage Game. First we argue that the stage game decision boils down to a decision about whether to participate. Due to the fact that the information value of choosing “buy” or “sell” is assumed to be the same, at any history h in which an agent i chooses to participate it is dominant to follow his signal and play

$$a(s_i, h) = a(s_i) = \begin{cases} \text{buy} & \text{if } s_i = 1 \\ \text{sell} & \text{if } s_i = -1 \end{cases}.$$

Furthermore, given that conditional on participating $a(s_i)$ is used, and given the fact that signals are of equal strength, at any history at which it is optimal to play “buy” when $s_i = 1$ it must also be optimal to play “sell” when $s_i = -1$, and vice versa. Therefore we suppress the stage decision and instead let $\sigma(h) = 0$ correspond to playing “stay out”, $\sigma(h) = 1$ correspond to playing $a_i(h)$, and $\sigma(h) \in (0, 1)$ correspond to a mixed strategy.

Agent i 's stage payoff when $\sigma(h) = 0$ is zero. The expected stage payoff when $\sigma(h) = 1$ depends on the type of the agent and that of his counterparty. When types are θ_i and θ_j , agent i 's expected utility from participating is

$$u(\theta_i, \theta_j) \equiv E[u_i | \sigma(h) = 1, \theta_i, \theta_j] = \Pr(s_i = 1, s_j = -1 | \theta_i, \theta_j) E[v | s_i = 1, s_j = -1, \theta_i, \theta_j] \\ - \Pr(s_i = -1, s_j = 1 | \theta_i, \theta_j) E[v | s_i = -1, s_j = 1, \theta_i, \theta_j],$$

in which all probabilities and expectations are derived from $F(v, s_1, s_2 | \theta_1, \theta_2)$. Note that by symmetry among players, $u(B, B) = u(G, G) = 0$ and $u(G, B) = -u(B, G)$. Let $\bar{u} = u(G, B)$ and note that $\bar{u} > 0$ by the assumption that good types get stronger signals. With uncertainty over types, so that $\Pr(\theta_i = G) = \mu_i$ and $\Pr(\theta_j = G) = \mu_j$, agent i 's expected utility from participating is

$$u(\mu_i, \mu_j) \equiv E[u(\theta_i, \theta_j)] = \mu_i \mu_j u(G, G) + (1 - \mu_i)(1 - \mu_j) u(B, B) \\ + \mu_i(1 - \mu_j) u(G, B) + (1 - \mu_i) \mu_j u(B, G) \\ = (\mu_i - \mu_j) \bar{u}.$$

Thus, at any given history agent i 's stage payoff is simply the difference between his expected type and that of his counterparty, scaled by a constant.

Beliefs. To evaluate an agent's strategy at some history $h = (y_1, \dots, y_t)$, it is necessary to compute that agent's expected type $\mu(h)$, which in turn requires describing the inference from each stage observation y_i in that history. We will demonstrate that while in principle inference from trading depends on the equilibrium strategy of other agents, presently it is sufficient to consider the proportion of good traders in the steady state pool.

Consider when the agent participates by playing $a_i \in \{\text{buy, sell}\}$, in which case the observation $y_i = (s_i, a_i, a_j, \gamma_i)$ is generated from distribution $\Psi(\gamma_i | a_j, \theta_i, \theta_j, s_i, s_j, v)$. By observing his counterparty's action a_j , agent i can compute the likelihood $\ell(\gamma_i | \theta_i, \theta_j)$ using distribution Ψ .² If μ_s denotes the proportion of good traders that participate in the pool then

²Given distribution $\Psi(\gamma_i | a_j, \theta_i, \theta_j, s_i, s_j, v)$, forming likelihood $\ell(\gamma_i | \theta_i, \theta_j)$ would be a straightforward statistical exercise if a_j were exogenous. However because a_j is an equilibrium strategy, it must accounted for, which is accomplished by using the fact that if $a_j = \text{“buy”}$ then $s_j = 1$, if $a_j = \text{“sell”}$ then $s_j = -1$, and if $a_j = \text{“stay out”}$ then it is not informative about s_j .

observed outcome y_i has an associated likelihood ratio given by

$$\lambda(y, \mu_s) = \frac{\mu_s \ell(y_i|G, G) + (1 - \mu_s) \ell(y_i|G, B)}{\mu_s \ell(y_i|B, G) + (1 - \mu_s) \ell(y_i|B, B)}.$$

If agent i plays “stay out” then the signal is uninformative and $\lambda(y_i, \mu_s) = 1$.

Let $\lambda_0 = \frac{\mu_0}{1-\mu_0}$ be the prior likelihood of a new agent. An agent that has observed a sequence of outcomes $h = (y_1, \dots, y_\tau)$ has beliefs described by the product of each outcome’s likelihood,

$$\lambda(h, \mu_0, \mu_s) = \lambda_0 \prod_{t=1}^{\tau} \lambda(y_t, \mu_s).$$

Thus, in a steady state with a proportion μ_s of good traders, an agent with prior μ_0 and history h has posterior

$$\mu(h, \mu_0, \mu_s) \equiv \frac{\lambda(h, \mu_0, \mu_s)}{1 + \lambda(h, \mu_0, \mu_s)}.$$

The key observation here is that for any history h an agent’s belief is fully determined by his prior μ_0 and the proportion of good traders μ_s in the steady state pool.

Payoffs in the dynamic game. For a given history $h = (y_1, \dots, y_\tau)$, for $t \leq \tau$ let $h_t = (y_1, \dots, y_t)$ denote the first t observations in that history. When an agent faces a pool with a proportion μ_s good types and follows strategy σ , the chance of observing history h is given by

$$\rho(h, \mu_s, \sigma) = \delta^\tau \prod_{t=1}^{\tau} \left(\mu_0 \sigma(h_t) \ell(y_t|G, \mu_s) + (1 - \mu_0) \sigma(h_t) \ell(y_t|B, \mu_s) + (1 - \sigma(h_t)) \mathcal{I}(y_t = \emptyset) \right),$$

where $\ell(y_t|\theta, \mu_s) = \mu_s \ell(y_t|\theta, G) + (1 - \mu_s) \ell(y_t|\theta, B)$ and $y_t = \emptyset$ is shorthand for the observation y_t when playing “stay out”. The agent’s expected lifetime utility is

$$U(\sigma, \mu_0, \mu_s) = \int_{h \in H} \left[\rho(h, \mu_s, \sigma) \cdot u_i(\mu(h, \mu_0, \mu_s), \mu_s) \cdot \sigma(h) \right],$$

and by substituting for u_i is

$$U(\mu_0, \mu_s, \sigma) = \bar{u} \int_{h \in H} \left[\rho(h, \mu_s, \sigma) \cdot (\mu(h, \mu_0, \mu_s) - \mu_s) \cdot \sigma(h) \right].$$

The agent chooses strategy $\sigma(h)$ to maximize this expected lifetime utility.

Optimal strategy and maximized utility. Here we demonstrate that an agent’s optimization problem when facing a pool with proportion μ_s of good types depends on two state variables: his belief about his own type and the number of periods that remain until T . We explicitly derive the optimal strategy which is a time-dependent threshold strategy in beliefs and show that the value function increases in own belief and in time until T .

To simplify expressions, we drop μ_0 and μ_s from an agent’s optimization problem since he takes these as given. Thus, for example, beliefs are written as $\mu(h) = \mu(h, \mu_0, \mu_s)$.

Lemma 3 *Belief μ and time t are sufficient for optimization. That is, if $\sigma^*(h)$ is optimal then there exists a strategy $\sigma^*(\mu(h), t(h))$ that is also optimal.*

Proof Agents face a dynamic decision problem, for which it is sufficient to focus only on payoff-relevant variables. As demonstrated earlier, the stage payoff $u(\mu(h), \mu_s)$ depends only on the agent’s belief and the agent’s lifetime payoff is integrated over future histories, for which $t(h)$ is relevant to determine the expected number of remaining periods. No other component of the agent’s history, for instance the specific path by which the belief μ was reached, is relevant for computing continuation payoffs. ■

Given belief μ and time t are sufficient, we can define (recursively) the value function

$$V(\mu, t) = \max\left(\delta V(\mu, t + 1), u(\mu) + \delta E[V(\mu', t + 1)]\right).$$

The first term in the max operator corresponds to playing “stay out” in period t , receiving a stage payoff of zero, and keeping the same belief μ in the following period. The second term take corresponds to playing “buy” or “sell” in period t , receiving a stage payoff, and receiving a new posterior in period $t + 1$ after generating an informative signal about one’s type in period t . If $T = \infty$ so that agents are infinitely lived, t is not informative about the expected number of periods remaining and thus $V(\mu, t) = V(\mu)$ for all t , guaranteeing the existence of a solution for the recursive equation above using a contraction mapping argument. If $T < \infty$ and agents are finitely lived, then $V(\mu, t)$ can be constructed by backward induction, starting with $V(\mu, T) = \max(0, u(\mu) - \mu_s)$.

Lemma 4 *The optimal strategy $\sigma^*(h)$ is a time-dependent threshold strategy. That is, there exists a sequence of beliefs $\bar{\mu}_1, \dots, \bar{\mu}_T$ so that $\sigma^*(h) = 1$ iff $\mu(h) \geq \bar{\mu}_{|h|}$. If $T = \infty$ then $\bar{\mu}_t = \bar{\mu}$ for all t . If $T < \infty$ then $\bar{\mu}_T = \mu_s$ and $\bar{\mu}_{t < T}$ is solved by backward induction.*

Proof See Appendix. ■

Characterization of steady state equilibrium

A steady state equilibrium requires that the composition of the pool of agents remains invariant across time and that all agents act optimally conditional on the pool's composition. To solve for a steady state equilibrium, we first conjecture that the steady state proportion of good types in the pool must have a particular value μ_s^* . We find σ^* , the agents' optimal strategy conditional on facing a pool with μ_s^* good types. Then, we demonstrate that this strategy is not only a best response to μ_s^* but in fact maintains μ_s^* . Specifically, we will show that using σ^* we can construct a set of masses $m^*(h)$ of agents at each history so that when agents use σ^* , the masses constitute a steady state and the proportion of good types in this steady state is μ_s^* .

Lemma 5 *There exists a unique μ_s^* so that*

$$\begin{aligned} V(\mu, t = 0 | \mu_s^*) &= 0 \text{ if } \mu \leq \mu_0 \\ &\text{and} \\ V(\mu, t = 0 | \mu_s^*) &> 0 \text{ if } \mu > \mu_0. \end{aligned}$$

Proof First note that $V(\mu, t = 0 | \mu_s^* = 1) = 0$ since at any history h the stage payoff $u_i(\mu(h), 1) \leq 0$ for all $\mu(h) \in [0, 1]$. Similarly, $V(\mu, t = 0 | \mu_s^* = 0) > 0$ since $u_i(\mu_0, 0) > 0$. Finally, note that $V(\mu, t = 0 | \mu_s^*)$ is continuous and monotone non-increasing in μ_s because $u_i(\mu, \mu_s)$ has these properties. Thus μ_s^* must exist and is unique. ■

Lemma 6 *There exists a strategy σ^* that is a best response to μ_s^* with $\sigma^*(h_0) = 1$.*

Proof To see this, suppose that new agents strictly prefer to stay out in the first period. Then, they will stay out in all subsequent periods (see Lemma 11 in the Appendix). By continuity of $V(\mu, t = 0 | \mu_s^*)$ in the first argument, if staying out is strictly preferred by new agents with belief μ_0 , it is also strictly preferred by some hypothetical new agents with belief $\mu > \mu_0$, which implies $V(\mu, t = 0 | \mu_s^*) = 0$, which contradicts the construction of μ_s^* . Thus, it is a best response for new agents to enter. ■

Lemma 7 *The strategy σ^* generates μ_s^* . That is, if $\mu_s(\sigma^*, \mu_s^*)$ is defined as an agent's expected posterior over all possible histories in which he participates, then $\mu_s(\sigma^*, \mu_s^*) = \mu_s^*$.*

Proof Recall that μ_s^* was chosen so that a new agent's expected lifetime payoff is zero. We use this fact and re-arrange terms.

$$\begin{aligned}
V(\mu, t = 0 | \mu_s^*) = 0 &= \int_{h \in H} (\rho(h, \mu_s^*, \sigma^*) u(\mu(h, \mu_0, \mu_s^*), \mu_s^*) \sigma^*(h)) dh \\
&= \bar{u} \int_{h \in H} (\rho(h, \mu_s^*, \sigma^*) (\mu(h, \mu_0, \mu_s^*) - \mu_s^*) \sigma^*(h)) dh \\
&= \bar{u} \left(\int_{h \in H} (\rho(h, \mu_s^*, \sigma^*) \mu(h, \mu_0, \mu_s^*) \sigma^*(h)) dh \right) - \mu_s^* \int_{h \in H} \rho(h, \mu_s^*, \sigma^*) \sigma^*(h) dh \\
\mu_s^* &= \frac{\int_{h \in H} (\rho(h, \mu_s^*, \sigma^*) \mu(h, \mu_0, \mu_s^*) \sigma^*(h))}{\int_{h \in H} \rho(h, \mu_s^*, \sigma^*) \sigma^*(h)} \\
\mu_s^* &= \mu_s(\sigma^*, \mu_s^*)
\end{aligned}$$

■

The proof of Lemma 7 is the key to understanding the paper's main result. In a steady state in which the proportion of good traders is such that a new entrant will exactly break even when playing optimally, the expected type of the new player in all his future histories must be equal to the average type in the pool. Thus his optimal strategy $\sigma^*(\mu_s^*)$ will generate the μ_s^* to which it is best-responding.

Lemma 8 *If the mass of participating agents at every history h is $m^*(h) = \rho(h, \mu_s^*, \sigma^*)$ then $m^*(h)$ constitutes a steady state.*

Proof Suppose $m_0(h) = \rho(h, \mu_s^*, \sigma^*)$ describes the mass of agents at each history in a particular point in time $t = 0$. We need to demonstrate that when all agents play strategy σ^* then $m_1(h) = m_0(h)$. To see this, consider history $h_\tau = (y_1, \dots, y_{\tau-1}, y_\tau)$ and its sub history $h_{\tau-1} = (y_1, \dots, y_{\tau-1})$:

$$\begin{aligned}
m_1(h_\tau) &= m_0(h_{\tau-1}) \cdot Pr(y_\tau | h_{\tau-1}, \sigma^*, \mu_s^*) \\
&= \rho(h_{\tau-1}, \mu_s^*, \sigma^*) \cdot Pr(y_\tau | h_{\tau-1}, \sigma^*, \mu_s^*) \\
&= \rho(h_\tau, \mu_s^*, \sigma^*) \\
&= m_0(h_\tau).
\end{aligned}$$

Thus the mass at every history remains invariant across time and the set of masses $m(h) = \rho(h, \mu_s^*, \sigma^*)$ constitutes a steady state. ■

The stationary equilibrium and the no trade theorem

We note that the no trade theorem would obtain in a single period snapshot of our steady state equilibrium. That is, suppose that the distribution of agents' types is given by the

steady state masses $m(h)$, that agents choose in or out this period, and that the game ends after that. The steady state distribution over histories can be translated into a probability distribution over beliefs, so that

$$Z(\mu) = \frac{\int_{h: \mu(h) \leq \mu} m(h)}{\int_h m(h)}.$$

As we established in the previous section, trade can only occur if both of the agents in a pair have the highest belief in the pool, which has probability 0 if there is no atom at the maximum of the support of distribution $Z(\mu)$. By contrast, in the dynamic steady state environment trade occurs with strictly positive probability, as any two agents for whom $\mu(h) \geq \bar{\mu}_{|h|}$ will trade.

4 Equilibrium Selection

The previous section demonstrates the existence of an equilibrium where all new agents enter and continue to actively trade as long as their belief remains above threshold levels. However, it is immediately clear that initial entry with any probability is an equilibrium, including a probability of zero so that no trade takes place. Multiple equilibria are not uncommon of course, but given that initial entry is a weak best response, one may wonder whether what reason there is to expect the equilibrium above in which trade occurs versus an equilibrium with no trade. To this end we briefly discuss two circumstances under which the trade equilibrium might be “selected” instead of the no-trade equilibrium.

Gains from Trade. Consider an environment with a small number of agents that trade assets for classical reasons such as liquidity or risk-hedging concerns, and a large pool of speculators that are potential entrants. As long as the classical agents make up a positive proportion of the market, every speculator has a strict incentive to enter since they will on average capture some of the gains from trade. As entry persists and the proportion of speculators grows, the gains from trade per speculator diminish toward zero, and in the limit the environment is as the one described in the main model. Thus, while there is a continuum of equilibria at zero gains from trade, the limit of equilibria as gains from trade approach zero is the equilibrium with certain initial entry.

Path to Steady State. An alternative way to select the trading equilibrium is to examine the path to the steady state. Namely, consider a brand new market in which the first cohort of new agents makes up the entire pool. In this initial period the pool is not adversely selected so that in expectation the stage payoff from the first trade is zero and, due to the anticipated entry of new agents in the following period, the information value is strictly positive. Entry is thus a strict best response for this first cohort and, as experienced traders

accumulate and the pool becomes more adversely selected, the payoff to entry remains positive but approaches zero as the pool approaches the steady state mix. Along the way toward steady state, not only do older generations make up a smaller proportion of the pool but also marginally skilled (in expectation) agents who otherwise would have exited in steady state remain while the pool is less selected, thus further slowing the adjustment.

5 Conclusion

Classic no trade theorems imply that a rational speculator is willing to trade based on his private information only if his counterparty is willing to lose money. This observation has led to a class of models, starting with Kyle (1985) and Glosten and Milgrom (1985), in which a proportion of market participants known as noise traders either have preferences that are different from the speculators' and willingly lose money over their lifetime or misunderstand the market and either lose money or accept excessive risk. In this paper, we demonstrate that the presence of noise traders is not necessary to support information-based trade.

In our model all agents have the same preference for the asset and trade based on private information. The adverse selection of the no trade theorems is mitigated by agents' incentive to learn about their ability through participating in the market. An agent who learns that his ability is high remains in the market and continues to make profitable trades, while an agent who learns that his ability is low exits. The option value inherent in this information induces young agents to engage in trades with more experienced counterparties, trades on which young agents lose money on average. The learning motive thus not only rationalizes trade by young traders, but also alleviates adverse selection concerns for older agents who expect to transact with these younger cohorts. The key contrast from standard noise trader models is that while some agents are willing to lose money on a given trade, they are not willing to lose money over their lifetime. Thus, the concern that such traders would be selected out of the market due to their persistent losses does not apply here.

Some parallels exist between our framework and a model of learning by doing. It is straightforward to construct a learning by doing model where older agents that have participated in trade are more skilled than their younger counterparts, and younger agents are willing to accept short term expected losses from trade because they anticipate positive expected payoffs when they are old. The mechanism at work is different from our model, where an individual agent does not improve in ability over time, but agents are able to self select as they receive signals about their types. In this sense, an equilibrium with trade in our setting requires a more nuanced explanation. Exploring the differences and similarities between models where agents learn about their type and models where

agents learn by doing is an area for future research.

An appealing feature of our model is that aside from demonstrating that the no trade theorems do not rule out purely speculative trade, it provides an additional rationale for trading that may bridge the gap between the classic motives associated with asset trading and observed trading volume. As noted in Odean (1999), “Trading volume on the world’s markets seems high, perhaps higher than can be explained by models of rational markets”, in which the author later explains rational markets to be those that attend to “investors’ rebalancing and hedging needs” (p. 1279). In recent years especially, with the proliferation of automated high frequency trading, it is difficult to support the theory that classical motives are behind most trading transactions. With this model we provide an additional rational explanation for trade, and show that a substantial volume may be generated even when there are no gains from trade.

References

- AUMANN, R. J. (1976): “Agreeing to Disagree,” *The Annals of Statistics*, 4(6), pp. 1236–1239.
- BARBER, B., Y.-T. LEE, Y.-J. LIU, AND T. ODEAN (2010): “Do Day Traders Rationally Learn About Their Ability?,” *Working Paper*.
- BOND, P., AND H. ERASLAN (2010): “Information-based trade,” *Journal of Economic Theory*, 145(5), 1675 – 1703.
- DE LONG, J. B., A. SHLEIFER, L. H. SUMMERS, AND R. J. WALDMANN (1990): “Noise trader risk in financial markets,” *Journal of political Economy*, pp. 703–738.
- GERVAIS, S., AND T. ODEAN (2001): “Learning to be overconfident,” *Review of Financial studies*, 14(1), 1–27.
- GLOSTEN, L. R., AND P. R. MILGROM (1985): “Bid, ask and transaction prices in a specialist market with heterogeneously informed traders,” *Journal of financial economics*, 14(1), 71–100.
- KREPS, D. M. (1977): “A note on fulfilled expectations; equilibria,” *Journal of Economic Theory*, 14(1), 32–43.
- KYLE, A. S. (1985): “Continuous Auctions and Insider Trading,” *Econometrica*, 53(6), pp. 1315–1335.
- MILGROM, AND STOKEY (1982): “Information, trade and common knowledge,” *J. Econ. Theory*, 26, 17–27.

ODEAN, T. (1999): "Do Investors Trade Too Much?," *American Economic Review*, 89(5), 1279–1298.

SERU, A., T. SHUMWAY, AND N. STOFFMAN (2010): "Learning by trading," *Review of Financial studies*, 23(2), 705–739.

TIROLE, J. (1982): "On the Possibility of Speculation under Rational Expectations," *Econometrica*, 50(5), pp. 1163–1181.

6 Appendix

Characterization of Optimal Strategy $\sigma^*(h)$ (proof of Lemma 4)

Here we prove that the optimal strategy $\sigma^*(h)$ consists of a sequence of time-dependent thresholds $\bar{\mu}_t$, $t = 1, \dots, T$. In the ensuing analysis the prior μ_0 and steady state pool μ_s are fixed and omitted from the notation.

Lemma 9 $V(0, t) = 0$, $V(1, t) > 0$, and $V_\mu(\mu, t) \geq 0$.

Proof That $V(0, t) = 0$ follows from the fact that every maximized stage payoff is bounded above by zero; that $V(1, t) > 0$ follows from the fact that any given stage payoff $u(1)$ is strictly positive if $\mu_s < 1$. To show that $V_\mu \geq 0$, fix a period t and consider two beliefs $\mu_1 < \mu_2$ with associated likelihood ratios $\lambda_i = \frac{\mu_i}{1-\mu_i}$, so that $\lambda_1 < \lambda_2$. Let $\sigma^*(\mu, t)$ be an optimal strategy. We now define a strategy for an agent with belief μ_2 at time t that imitates an agent with belief μ_1 . Let h_t be the first t observations of a history h (of length at least t). Define imitation likelihood

$$\tilde{\lambda}(h) = \begin{cases} \frac{\lambda_1}{\lambda_2} \cdot \lambda(h) & \text{if } |h| \geq t \text{ and } \lambda(h_t) = \lambda_2 \\ \lambda(h) & \text{otherwise} \end{cases}$$

Let $\tilde{\mu}(h) = \frac{\tilde{\lambda}(h)}{\tilde{\lambda}(h)+1}$ and let the imitation strategy be $\tilde{\sigma}(h) = \sigma^*(\tilde{\mu}(h))$. Thus, the strategy $\tilde{\sigma}$ is constructed so that if an agent reaches belief μ_2 in period t , all his ensuing actions are the same as that of an agent that reaches belief μ_1 in period t and follows strategy σ^* . Also, define $H(\mu, t) \equiv \{h \mid |h| > t \text{ and } \mu(h_t) = \mu\}$ as the set of histories for which the belief after the first t observations is μ .

Now, we demonstrate that the payoff to an agent with belief μ_2 in period t that pursues the imitation strategy is at least as high as the payoff to agent with belief μ_1 in period t that pursues the optimal strategy.

$$\begin{aligned} V(\mu_1, t) &= u(\mu_1) + \mu_1 E_{h \in H(\mu_1, t)}[u(1)\sigma^*(h) \mid \sigma^*, \theta = 1] + (1 - \mu_1) E_{h \in H(\mu_1, t)}[u(0)\sigma^*(h) \mid \sigma^*, \theta = 0] \\ &\leq u(\mu_2) + \mu_2 E_{h \in H(\mu_1, t)}[u(1)\sigma^*(h) \mid \sigma^*, \theta = 1] + (1 - \mu_2) E_{h \in H(\mu_1, t)}[u(0)\sigma^*(h) \mid \sigma^*, \theta = 0] \\ &= u(\mu_2) + \mu_2 E_{h \in H(\mu_2, t)}[u(1)\tilde{\sigma}(h) \mid \tilde{\sigma}, \theta = 1] + (1 - \mu_2) E_{h \in H(\mu_2, t)}[u(0)\tilde{\sigma}(h) \mid \tilde{\sigma}, \theta = 0] \\ &\leq V(\mu_2, t) \end{aligned}$$

In the first line, expectations are taken with respect to the probability distribution over continuation histories, which is determined by the agent's true type and his strategy. The expectation when $\theta = 1$ is weakly positive and when $\theta = 0$ is weakly negative (since all stage payoffs are positive and negative, respectively). Thus in the second line, replacing μ_1 with μ_2 puts more weight on the positive expectation, and since $u(\mu_2) > u(\mu_1)$, the net

effect is an increase in the right hand side. In the third line, the strategy $\tilde{\sigma}(h)$ replaces σ^* , and we look over the set of continuation histories $H(\mu_2, t)$ instead of $H(\mu_1, t)$. This is the key step because the imitation strategy $\tilde{\sigma}$ is constructed so that, conditional on the agent's true type, the distribution over continuation histories is the same, and thus expected continuation payoffs are the same. The third line describes the value for an agent with belief μ_2 that pursues an imitation strategy, which must be weakly worse than pursuing the optimal strategy, leading to the inequality in the the fourth line. ■

Lemma 10 *The value function decreases with respect to time, i.e. $V(\mu, t) \leq V(\mu, t + 1)$.*

Proof This follows from the fact that the availability of extra period is an option. That is, an agent in period t with belief μ can use the strategy he would use if the period were $t + 1$, $\sigma(\mu, t) = \sigma^*(\mu, t + 1)$, receive the same payoffs and then at the end have an additional period, the payoff of which is at least zero. Namely,

$$V(\mu, t) \geq V(\mu, t + 1) + \delta^{T-t} E[\max(0, u(\mu')) \mid \sigma(\mu, t) = \sigma^*(\mu, t + 1)] \geq V(\mu, t + 1).$$

■

Lemma 11 *There exists an optimal strategy $\sigma^*(h)$ so that for any history y_1, \dots, y_t , if $\sigma^*(y_1, \dots, y_t) = 0$ then $\sigma^*(y_1, \dots, y_t, \emptyset) = 0$.*

Proof The idea is that the value of information diminishes as agents have fewer remaining periods. Suppose there is an optimal strategy $\sigma(h)$ in which for some history an agent with belief μ in period t plays "out", and then having the same belief μ in period $t + 1$ plays "in". If the player pursues an alternative strategy σ^* in which he plays "in" in period t , "out" in period $t + 1$, and subsequently follows the imitation strategy $\sigma^*(y_1, \dots, y_{t-1}, y_t, \emptyset, y_{t+2}, \dots, y_T) = \sigma(y_1, \dots, y_{t-1}, \emptyset, y_t, y_{t+1}, \dots, y_T)$ then he receives the same payoff, so $\sigma^*(h)$ is also a best response. ■

Corollary 12 *If $\sigma^*(h) = 0$ then $V(\mu(h), t(h)) = 0$.*

Proof By the previous lemma, an agent that plays out continues to play out, thus his continuation payoff is zero. ■

Lemma 13 *Agents use time-dependent threshold strategies, so that $\sigma^*(\mu, t) = 1$ iff $\mu \geq \bar{\mu}_t$. If $T = \infty$ then $\bar{\mu}_t = \mu$ for all t ; if $T < \infty$ then thresholds are computed by backward induction, with $\bar{\mu}_T = \mu_s$.*

Proof

$$V(\mu, t) = \max \{0, u(\mu) + \delta E[V(\mu', t + 1) \mid \mu]\}$$

The payoff to playing “in” in period t , the second term inside the max operator, is strictly increasing in μ , since the stage payoff is strictly increasing and the continuation payoff is weakly increasing (as shown in the proof of lemma 9). This payoff is also strictly negative when $\mu = 0$ and strictly positive when $\mu = 1$. Thus, there exists some $\bar{\mu}_t$ so that $V(\mu, t) > 0$ iff $\mu > \bar{\mu}_t$, which in turn implies the optimal strategy is $\sigma^*(\mu, t) = 1$ iff $\mu \geq \bar{\mu}_t$.

When $T = \infty$, $V(\mu, t) = V(\mu)$ for all t , thus $\bar{\mu}_t = \bar{\mu}$ for all t . When $T < \infty$ we solve for the time-dependent thresholds using backward induction.

Period $t = T$

$$\bar{\mu}_T : 0 = u(\bar{\mu}_T)$$

$$V(\mu, T) = \begin{cases} u(\mu) & \text{if } \mu \geq \bar{\mu}_T \\ 0 & \text{if } \mu < \bar{\mu}_T \end{cases}$$

Period $t < T$

$$\bar{\mu}_t : 0 = u(\bar{\mu}_t) + \delta E[V(\mu', t + 1)|\bar{\mu}_t]$$

$$V(\mu, T) = \begin{cases} u(\mu) + \delta E[V(\mu', t + 1)|\bar{\mu}_t] & \text{if } \mu \geq \bar{\mu}_t \\ 0 & \text{if } \mu < \bar{\mu}_t \end{cases}$$

It is easily seen that the final period threshold is $\bar{\mu}_T = \mu_s$, and that all preceding thresholds are computed by the backward induction process. ■