Spring 4-25-2013

Essays on Price Dynamics and Market Selection

Filippo Massari
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Economics

Dissertation Examination Committee:

John Nachbar, Chair

Werner Ploberger, Co-Chair

Marcus Berliant

Bart Hamilton

Paulo Natenzon

Maher Said

Essays on Price Dynamics and Market Selection

by

Filippo Massari

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2013

St. Louis, Missouri

Table of Contents

Acknowledgment

This thesis would not have been made possible without the help and guidance of many people. First and foremost, I would like to thank my two thesis co-chairs, Werner Ploberger and John Nachbar. Werner believed in and supported me from the beginning. He carefully nudged me along, patiently answering myriad of questions.The hours we spent talking together will always be one of my best and most useful memories of my PhD. John was instrumental in helping me take the all-important step from student to researcher. He helped me to find the best way to link my results to economics in a clear, interesting and ,most importantly, understandable fashion. He endured my "crazy" talk and helped me to focus my research.

I also owe a great debt of gratitude to the other members of my dissertation committee. I want to thank Marcus Berliant, Bart Hamilton, Paulo Natenzon and Maher Said for their help in polishing my results and presentation skills. Also, I am especially thankful to Larry Blume, Philip Dybvik, David Easley, Alvaro Sandroni, Massimo Marinacci, Fabio Maccheroni, and Simone Cerreia Vioglio for their precious advise and guidance during the market.

Many friends and graduate colleagues were instrumental in helping me to become an economist. I would like to first thank Eleonora Melato, without whom I would have had a much harder time adjusting to St. Louis. In no particular order, I wish to thank the following people: Aleksandr Yankelevich, Nicholas Papageorge, Melanie Blackwell, Rohan Dutta, Taro Kumano, Hassan Faghani Dermi, Carmen Astorne-Figari, Brady Vaughan, Juan Block, Julieta Caunedo, Jan Auerbach, Francesco Carli, Mushegh Harutyunyan, Francesca Picciocchi, Ben Poremba and my aunt Silvia. My most special thanks belongs to Stephanie Heger, for bringing me great happiness in the last two years.

This dissertation could not have been possible without the financial support of the John Stuart Mills Fellowship, my family and the Mortara Fellowship from Banca d'Italia.

This work is dedicated to my Mom, Dad and Olivia.

Chapter 1: Price Dynamic and Market Selection

Abstract

This paper analyzes and characterizes the dynamics of wealth-share and equilibrium price in a stochastic general equilibrium model with heterogeneous consumers. The characterization enables a comparison between probabilistic learning and price evolution, revealing that prices incorporate "sparse" information efficiently. Results on wealth-share are obtained by comparing traders' optimal investment-consumption plans against their prices. This novel approach extends recent results in the literature by providing a condition that is necessary as well as sufficient for a trader to vanish. The results are applied to survival in iid, survival in large economies, and survival of traders that follow strategies commonly observed in real markets.

1 Introduction

The *market selection hypothesis*, henceforth MSH, states that the market selects for traders who produce the most accurate forecasts. This hypothesis, first articulated by Alchian (1950) and Friedman (1953), is one of the supporting arguments for the plausibility of rational expectations. As the wealth-share of traders with rational expectations converges to one, financial markets can be understood, to a large extent, using models with a representative trader who has rational expectations.

The MSH literature typically focuses on the asymptotic characterization of wealth-shares. The problem is reduced to one of identifying the investment strategy that ensures the highest capital accumulation rate, defining an appropriate measure of probabilistic distance and ascertaining the way beliefs, risk attitudes, discount factors and saving rates influence the individual investment strategy.

The baseline theory depends on the seminal work of Kelly (1956). An implication of his results is that the self-financing investment strategy that maximizes capital accumulation corresponds to the optimal investment strategy of an investor with rational expectations and log utility (Kelly rule). Consequently, if all traders have log utility and the same discount factor, the MSH holds because the trader with rational expectation will have the highest capital accumulation rate and will asymptotically dominate the market. The robustness of the MSH to deviations from log utility has been tested. In economies with an exogenously given saving rate, Blume-Easley (1992) shows that the MSH can fail because a trader with rational expectations but non log utility can have lower capital accumulation rate than a trader with incorrect beliefs. The situation is qualitatively different in general equilibrium models. Sandroni (2000) and Blume-Easley (2006), using discrete time models, and Yan (2006), using a continuum time model, show that in a non growing economy, the MSH holds true and traders risk attitude plays no role in their survival.

General Equilibrium models which study the MSH explore market selection from different perspectives but share the same structure: first, construct the investment strategy of each trader, then, proceed with a pairwise comparison of each strategy to determine the asymptotic distribution of wealth. The approach leaves no room to discuss the way equilibrium prices evolve. Fundamental questions regarding the dependence of price evolution on risk attitudes, the informational content of equilibrium prices and the amount of information needed by a trader to survive or beat the market cannot be answered. In order to answer these questions I propose a new approach. My first step is to characterize equilibrium price dynamics. To discuss survival, I compare each trader optimal investment-consumption plan with its price and show that the wealth-share of a trader converges to 0 if his optimal consumption plan becomes too expensive.

1.1 Price dynamics and their informational content

The key step to my approach is to provide a characterization of prices that is general enough to cover a wide spectrum of economies and informative enough to derive meaningful results. The information structure adopted is standard, in market selection literature, and is a key component of the model. Each trader has a personal probabilistic view of the world, his beliefs. I impose no restrictions on beliefs: they can be correct, incorrect, evolve in a way that is consistent with Bayes rule or not. Each trader "agrees to disagree" with the other traders. Unlike in the models in the information exchange literature (Grossmann-Stiglits 1980), each trader believes that there is nothing to learn from prices and is engaged in speculative trading. This modelling assumption is well described by William Feather's words: *One of the funny things about the stock market is that every time one person buys, another sells, and both think they are astute.*

To make things simple, I assume that traders receive no private information which, given the complete freedom in the evolution of traders' beliefs, comports no loss of generality. I define the *sparse information* of an economy to be the set of all traders' beliefs and the sequence of realized states. In this framework, this is all of the available information. Sparse information together with the initial wealth share distribution and traders preferences are the key elements in the determination of equilibrium prices. From the literature, we know that if all traders have the same discount factor, the market selection mechanism implies that the trader with the most accurate beliefs will dominate the market and that prices will reflect his beliefs. A key question is how fast equilibrium prices converge to their asymptotic values? In particular, if we introduce a new trader into the economy and we give him full knowledge of the sparse information, would this trader be able to construct forecasts that are more accurate than the forecasts implicit in market prices and would he be able to make money in the market?

My main contribution is to show that it is impossible to use sparse information to produce forecasts that are more accurate than the forecasts implicit in market prices and that the answer to the last question depends on the size of the economy: if the economy is small, a trader that uses sparse information will always be one of the survivors and never be the only one. If the economy is large, access to sparse information is neither sufficient or necessary to survive.

Blume-Easley (2009) provides a first attempt at characterizing price dynamics. They present an example of an iid Arrow's security economy populated by traders with log-utility to show that the equilibrium price evolves similarly to a Bayesian updating scheme. My paper builds on Blume-Easley's example and characterizes the equilibrium price process in every complete-market, exchange economy.

In order to highlight the similarities between price evolution and probabilistic learning, I mainly focus on the discounted prices of Arrow's securities in an economy with no aggregate risk and a homogeneous discount factor. For the degenerate case of a representative agent economy, it is an easy exercise to show that the discounted price of the contingent claim on a certain period-state coincides with the probability of that period-state. In economies populated by traders with log-utility, henceforth *log-economy*, I show that discounted equilibrium prices define, in each period, a probability measure and that there exists a prior distribution on the beliefs of the traders in the economy (the *market prior*) such that discounted prices coincide, on every path, with the probability obtained via Bayes' rule from the market prior.

The analysis of economies that are not log-economies is complicated by the fact that discounted equilibrium prices are not probabilities. Specifically, the sum of the discounted prices of contingent claims on next-period consumption may differ from 1. There are two ways to address this problem: to normalize the discounted prices to sum to 1 in each period or to directly study the discounted price process. Normalized prices reflect the probability that prices attach to future events (the price's *informational content*), which I will interpret as the "market's beliefs". Discounted prices capture their wealth-share implications (the price's *investment strategy content*), which I will interpret as the "market's investment strategy." My main result is to show that the market's beliefs incorporate sparse information efficiently and that a trader vanishes if his investment strategy is worse than the market's.

For heterogeneous discount factor economies, the equivalence between price evolution and Bayesian updating is generally lost even asymptotically. This is due to the fact that the differences in discount factors can negate the effect of more accurate forecasts. In this setting, unlike in any Bayesian procedure, equilibrium prices need not converge to the beliefs of the trader with the most accurate beliefs.

1.2 Wealth-shares

The asymptotic characterization of the wealth-share of traders operating in complete markets has been studied extensively. Most of the existing models rely on pairwise comparisons of traders' investment strategies. The approach I adopt is different (as the "leading example" in Section 4 shows). I compare traders' investment strategies against equilibrium prices instead of against each other. As Blume-Easley (2009) and Mailath-Sandroni (2003) have pointed out, pairwise investment strategy comparison cannot be used to identify a sufficient condition for a trader to vanish. Their papers reveal that, in order to verify whether a trader vanishes, we must compare his beliefs against all of the other traders' beliefs simultaneously. My approach incorporates this idea, as equilibrium prices are the most natural aggregator of traders' beliefs. Theorem 4 provides a necessary and sufficient condition for a trader to vanish that applies to any complete-market, exchange economy.

Section 6 contains some of the implications of my vanishing condition. First, I discuss survival in small and large economies. In large economies, contrary to the current findings for small economies (Sandroni (2000), Blume-Easley (2006), Kogan-Ross et al.(2006, 2009) and Yan (2006)), I show that risk attitudes can have an effect on survival even if the economy does not grow. This result also has implications on the amount of information needed for a trader to survive. In small economies, a trader that knows the beliefs of all of the other traders is always one of the survivors. But if the economy is large, this result holds true only if all of the traders have log utility. This point illustrates that focusing only on individual characteristics is not enough to determine the asymptotic fate of a trader.

Second, I apply my vanishing condition to discuss the merits of investment strategies that are commonly adopted in real world markets. I consider two strategies: the diversified investment strategy, which is equivalent to holding a long position on a market index, and the copy-trader strategy, which corresponds to the strategy of always investing all money in the hedge fund that performed the best in the past (this strategy is commonly advertised by providers of internet-trading like www.etoro.com and www.currensee.com). My results provide theoretical support in favor of the diversified portfolio strategy and against the copy-trader strategy.

Third, I complete Blume-Easley's discussion on the necessary and sufficient condition to survive in iid economies.

My approach relies on fewer assumptions and incorporates, in a unifying framework, results that were previously treated as exceptions. These results include survival in an economy with some Bayesian learners (Blume-Easley(2006) Section 3.3) and the "almost" necessary and sufficient condition for iid economies of Blume-Easley (2009).

The remainder of my paper proceeds as follows: Section 2 outlines the economic environment and introduces the relevant characteristics and definitions of the model. Section 3 presents the notions of informational content and investment strategy content of equilibrium prices and shows that normalized prices incorporate sparse information efficiently. Section 4 illustrates, by means of an example, my approach and compares it with the existing approach. Sections 5 discusses the small and large sample properties of equilibrium prices and their dependence on the size of the economy. Sections 6 presents my results about survival in small and large economies. Section 7 concludes. In Appendix A, I compare the notion of distance that I adopt with the distances adopted in Sandroni (2000) and Blume-Easley(2006). The other appendices contain the proofs.

2 The model

The model is an infinite horizon Arrow-Debreu exchange economy with complete set of claims to consumption. Time is discrete and begins at date 0. At each date there is a finite set of states $\mathcal{S} \equiv \{1, ..., s, ..., S\}$ with cardinality $|\mathcal{S}|$=S. The set of all infinite sequences of states is $\mathcal{S}^\infty$ with representative sequence of realizations $\sigma = (\sigma_1, ...)$. Let $\sigma^t = (\sigma_1, ..., \sigma_t)$ denote the partial history through date t of path σ and $\mathcal{S}^t$ be the set containing all the different sequences of length t. The partial history of length t with last element s is denoted by: $\sigma^{t,s}(\sigma_1, ..., \sigma_{t-1}, s)$. Σ is the product σ-algebra generated by the sequences in $\mathcal{S}^\infty$. Let P be the true probability measure on Σ. For any probability measure p on Σ, $p(\sigma^t)$ is the marginal probability of the partial history σ^t; that is, $p(\sigma^t) = p(\{\sigma_1 \times ... \times \sigma_t\} \times S \times S \times ...)$. The information set at time t is $\mathcal{F}_t$. WLOG, given the absolute freedom I allow on the determination of the probability measures on Σ, $\mathcal{F}_t \equiv \sigma^t$: all that is learned from the past is the sequence of realized states. The conditional probability of a generic state σ_t and of the specific state s in period t, given the information set $\mathcal{F}_{t-1}$ can therefore be defined as $p(\sigma|\mathcal{F}_{t-1}) = p(\sigma|\sigma^{t-1}) = \frac{p(\sigma^t)}{p(\sigma^{t-1})}$ and $p(s|\mathcal{F}_{t-1}) = p(s|\sigma^{t-1}) = \frac{p(\sigma^{t,s})}{p(\sigma^{t-1})}$, respectively. Finally, the expectation operator without a subscript is the expectation with respect to the true measure P at time 0 and E_t is the expectation according to true probability at time t.

The economy contains a set of traders $\mathcal{I}$. If the number of traders is finite then the economy is called small. If there is a continuum of traders, the economy is large. Each trader has consumption set $\mathbb{R}_+$. A consumption plan $c : S^\infty \to \prod_{t=0}^{\infty} \mathbb{R}_+$ is a sequence of $\mathbb{R}_+$- valued functions $\{c_t(\sigma)\}_{t=0}^{\infty}$ in which each $c_t(\sigma)$ is $\mathcal{F}_t$-measurable. Each trader i is endowed with a particular consumption plan, called the endowment stream and denoted by $e_t^i(\sigma)$. Each trader i is characterized by a payoff function $u^i : \mathbb{R}_+ \to [-\infty, +\infty]$ over consumption, a discount factor $\beta_i \in (0, 1)$ and a subjective probability p^i on Σ, his beliefs. The words "distribution" and "probability measure" are used interchangeably and are to be understood in their most general sense: "distribution" means every $\mathcal{F}_t$ measurable function that generates a probability measure on Σ. There are no restrictions on the type of processes generating the sequence of realizations, nor on the way traders' beliefs evolve. Traders can have ideosincratic iid beliefs, be Bayesian learners or use any other rule that maps information from the past to produce forecasts about the future.

Trader i's utility for a consumption plan c is:

$$U^i(c) = E_{p^i} \sum_{t=0}^{\infty} \beta_i^t u_i(c_t(\sigma)).$$

A present value price system is formally defined as follows:

Definition 1. *A present value price system is a sequence of $\mathcal{F}_{t-1}$-measurable functions $\{q(\sigma^t)\}_{t=0}^{\infty}$.*

In terms of asymptotic wealth-shares, traders can either vanish or survive:

Definition 2. *Trader i vanishes on path σ if $\lim_{t\to\infty} c_t(\sigma) \to 0$. He survives on path σ if $\limsup_t c_t(\sigma) > 0$.*

It is important to stress that survival is not a normative concept. A trader that vanishes is acting optimally, given his preferences, discount factor and beliefs. My discussion about the copy-trader strategy and the diversified portfolio strategy does not constitute an exception to this point of view. My critique is on the way the traders use the same information to produce forecasts. A trader that adopts the copy-trader strategy is "criticized" because he uses a suboptimal updating scheme, not because his consumption-investment plan leads him to extinction.

2.1 Assumptions

A competitive equilibrium is a present value price system and, for each trader, a consumption plan that is affordable, preference maximal on the budget set and mutually feasible. I make use of a minimal set of assumptions to ensure the existence and uniqueness of the competitive equilibrium (see Peleg-Yaeri (1970)). These are the assumptions I will refer to:

- A1: The payoff functions $u_i : \mathbb{R}_+ \to [-\infty, +\infty]$ are C^1, concave and strictly increasing and satisfy the Inada condition at 0; that is, $u_i^{'}(c) \to \infty$ as $c \searrow 0$.

- A2: There are numbers $0 < f \leq F < +\infty$ such that for each trader i, all dates t and all paths σ, $f \leq \inf_{\sigma^t} \sum_{i \in \mathcal{I}} e_t^i(\sigma) \leq \sup_{\sigma^t} \sum_{i \in \mathcal{I}} e_t^i(\sigma) \leq F$.

- A3: For all traders i, all dates t and all paths σ, $p^i(\sigma^t) > 0 \Leftrightarrow P(\sigma^t) > 0$.

- A4: Each trader has iid beliefs and the union of traders' beliefs $\left(\mathcal{P} = \bigcup_{i \in \mathcal{I}} p^i\right)$ coincides with the S-dimensional-simplex: $\mathcal{P} = \Delta^S$.

Assumption A1 is a collection of standard properties that the payoff functions must satisfy, the Inada condition at 0 ensures that equilibrium prices are uniquely identified.[1] Assumption A2 uniformly bounds the aggregate endowment above and away from 0. Assumption A3 is needed in order to exclude pathological economies in which the equilibrium does not exist (for example economies in which all of the traders attach 0 probability to an event that occurs with positive probability). I will make use of Assumption A4 only in large economies to ensure that an appropriate prior distribution (in the Bayesian sense) can be define on the space of beliefs.

In order to discuss price evolution in detail, I will need a more restrictive environment:

Definition 3. *A CRRA economy is an economy in which all of the traders have the same CRRA payoff function:* $u(c) = \frac{c^{1-\gamma}-1}{1-\gamma}$.

I will also need a technical assumption on the prior distributions of learning traders.

[1] If we are only interested in the necessary condition to survive, then we can dispose of the Inada condition following the same reasoning of Section 4 in Blume-Easley (2006).

Definition 4. *A prior probability distribution $w(i)$is smooth if it is three times differentiable and strictly positive at the index corresponding to the model with maximum likelihood $\hat{\imath}$.*[2]

3 Equilibrium prices

3.1 Equilibrium prices: Characterization

I study a complete market economy with time 0 trading of a complete set of claims to consumption contingent on histories σ^t. The price of a claim that pays a unit of consumption at the end of σ^t is given by $q(\sigma^t)$. As usual, it is possible to support a Pareto optimal allocation with an appropriate initial distribution of wealth, sequential trading of one period Arrow's security and a set of "natural" limits on the quantities of Arrow's securities that can be issued in each history, date pair; for example, see Ljungqvist-Sargent (2004, ch.8). The equilibrium conditions imply that the price of an Arrow's security that pays a unit of consumption in state s, period t, after history σ^{t-1} is $q(s|\sigma^{t-1}) = \frac{q(\sigma^{t,s})}{q(\sigma^{t-1})}$. For the case of homogeneous discount factor economies, particular attention will be given to time 0 normalized prices: $q^n(\sigma^t) = \frac{q(\sigma^t)}{\sum_{\hat{\sigma}^t \in \mathcal{S}^t} q(\hat{\sigma}^t)}$ and to time 0 discounted prices: $q^d(\sigma^t) = \frac{q(\sigma^t)}{\beta^t}$, as they capture the informational content and the investment strategy content of equilibrium prices, respectively.

From the equilibrium condition we obtain that the "risk adjusted" wealth shares capture the weight that each trader has in the determination of equilibrium prices. By construction, their sum equals 1 and can be interpreted as a probability distribution on the set of beliefs in the economy: the *market prior* distribution.

Definition 5. *In a small economy, the market prior distribution is given by:*

$$w_0(i) = \left[\frac{\frac{1}{u_1'(c_0^1)}}{\sum_{i \in \mathcal{I}} \frac{1}{u_i'(c_0^i)}}, ..., \frac{\frac{1}{u_I'(c_0^I)}}{\sum_{i \in \mathcal{I}} \frac{1}{u_i'(c_0^i)}} \right]$$

In large economy, the market prior distribution is given by a function $w_0(i)$ that is smooth and

[2] A tighter set of assumptions can be found in Clarke-Barron (1990), Phillips-Ploberger (2003) or Grünwald (2007).

satisfies:

$$\int_{\mathcal{I}} w_0(i)d\theta(i) = 1$$

Where $\theta(i)$ is the Lebesgue measure on $\mathcal{I}$.

The smoothness assumption is a natural one for a large economy as it is equivalent to require that no trader has positive wealth-share or price impact.

3.2 Equilibrium prices: Informational content

In order to discuss the informational content of equilibrium prices I need to provide a definition of efficient forecasting that is independent of the true probability and to define a benchmark against which to compare the probabilities implicit in the equilibrium prices.

In an economy that is not a log-economy, normalized equilibrium prices do not evolve according to Bayes' rule. Nevertheless, via the equilibrium conditions, normalized prices map sparse information into future probabilities. I define a function with this property to be a *forecasting schemes*. It should be noted that forecasting schemes always implicitly depend on a model class (a prior support, in Bayesian terminology). A comparison between two different forecasting schemes can therefore be meaningful only if both schemes use the same sparse information: the same model class and the same sequence of realizations.

The benchmark to which equilibriums prices are compared is Bayesian updating. This choice depends on the fact that Bayesian updating produces consistent forecasts with optimal convergence rate.

Definition 6.

Let $\mathcal{P}$ be a set of probabilistic models on Σ and $p^B(\sigma^t)$ be the likelihood attached by a Bayesian model obtained with uniform prior on the models in $\mathcal{P}$ (if well defined). A probability p on Σ is informationally efficient with respect to $\mathcal{P}$, if:

$$\forall \sigma \in \Sigma, \lim_{t\to\infty} \log \frac{p^B(\sigma^t)}{p(\sigma^t)} < \infty,$$

Definition 6 states that a probability measure p on Σ is efficient if, in every sequence of realizations, the log-likelihood ratio between p and the likelihood obtained via Bayes' rule from a uniform prior on $\mathcal{P}$ is bounded above. In other words, on every path, σ, there is no statistical test based on the log likelihood principle that can claim that p^B is a better description of the sequence of realizations than p.

Theorem 1.

- *In a small homogeneous discount factor economy that satisfies A1-A3, normalized equilibrium prices are informationally efficient with respect to* $\mathcal{P} = \{p^i : i \in \mathcal{I}\}$.
- *In a large CRRA homogeneous discount factor economy that satisfies A1-A4, normalized equilibrium prices are informationally efficient with respect to* $\mathcal{P} = \{p^i : i \in \mathcal{I}\}$.[3]

Proof. For small economies, see Appendix B. For large economies, it follows from Theorem 6 in Appendix C □

Theorem 1 tells us that normalized equilibrium prices produce forecasts that are as accurate as the forecasts that a Bayesian with the same information would produce. Nevertheless, unless all of the traders in the economy have log utility, the way in which normalized prices evolve is not consistent with Bayes' rule. Therefore, normalized prices can be seen as a generalization of the Bayesian paradigm, which cannot be judged as irrational.

It is easy to show that Theorem 1 does not apply to economies with heterogeneous discount factors. This is because the convergence to a specific model is determined by its likelihoods and its associated discount factor. If the differences in the discount factors overcome the effect of more accurate forecasts, normalized prices, unlike Bayesian updating, will not converge to the bast model in its prior support.

3.3 Equilibrium prices: Investment strategy content

In order to discuss the asymptotic fate of the traders in the economy; I make use of a generalization of a standard measure of distance between probability distributions, the Kullback-Leibler divergence

[3]The existence of a well defined prior on $\mathcal{P} = \{p^i : i \in \mathcal{I}\}$ is ensured by assumption $A4$.

(K-L*).

Definition 7.

i) The market investment strategy's doubling rate is: $D_t^*(P||q) \equiv E\left(log\frac{P(\sigma^t)}{q(\sigma^t)}\right)$.

ii) The trader i's investment strategy doubling rate is: $D_t^*(P||\beta_i^t p^i) \equiv E\left(log\frac{P(\sigma^t)}{\beta_i^t p^i(\sigma^t)w_0(i)}\right)$.

The K-L* used for this definition is more general than the standard one, as it does not require q to be a probability measure. In particular, it is well defined for equilibrium price, $q(\sigma^t)$, and for discounted probabilities, $\beta_i^t p^i(\sigma^t)$. This technical modification allows for a direct comparison of traders' investment strategies against prices. In my model, traders are not betting against each other, but against a "fictitious" trader: the market. Following this interpretation, equilibrium prices can be thought of as the investment strategy of the market. This way of thinking about prices provides a good interpretation for discounted prices that do not sum to 1. They represent investment strategies that are not self-financing. The periods in which they sum to more (less) than 1 can be interpreted as periods in which some external funding increases (decreases) the capital invested. Clearly, the use of external funding has a negative effect on the rate at which the initial capital is lost (hence "negative distance").

In the next sections I make use of the following definitions:

Definition 8.

Trader i's investement strategy is as good as the market's investment strategy if

$$\lim_{t\to\infty} \left|D_t^*(P||q) - D_t^*(P||p^i) + t\log\beta_i\right| < \infty.$$

Trader i's investment strategy is worse than the market's investment strategy if

$$\lim_{t\to\infty} D_t^*(P||q) - D_t^*(P||p^i) + t\log\beta_i = -\infty.$$

Using these definitions, I will show that a trader survives if his investment strategy is as good as the market's strategy.

Definition 8 reveals that the performance of the investment strategy of a trader depends on two factors. The first one is trader's ability to produce accurate forecasts and is captured by the "standard" K-L* distance between his beliefs and the true distribution: $D^*(P||p^i)$. The second factor is trader's discount factor: β_i. The discount factors, can be thought of as the fraction of capital that is reinvested in every period by trader i using the strategy p^i. Clearly, higher β, corresponds to a greater capital accumulation rate. Moreover, it shows that risk attitudes play no role in traders' investment strategies. Different risk attitudes can have an effect on survival only through their effect on equilibrium prices.

4 The leading example

In this Section I present an example to illustrates the main features of my approach and why it is more informative than trader by trader pairwise comparison. Consider an iid Arrow's security economy with S states of the world distributed according to P. There are I traders with log utility functions, identical discount factors, the same initial wealth and iid beliefs p^i. There is a trader with correct beliefs. The solution is interior and the FOC are: $\frac{\beta^t p^i(\sigma^t)}{c_t^i(\sigma)} = \frac{1}{c_0^i} q(\sigma^t)$.

My approach:

Using the FOC, we see that normalized prices coincide with discounted equilibrium prices which are given by:

$$q^n(\sigma^{t,s}) = q^d(\sigma^{t,s}) = \sum_{i\in\mathcal{I}} p^i(\sigma^{t,s})w_0(i) \quad ; \quad q^d(s|\sigma^{t-1}) = \frac{\sum_{i\in\mathcal{I}} p^i(\sigma^{t,s})w_0(i)}{\sum_{j\in\mathcal{I}} p^j\left(\sigma^{t-1}\right)w_0(j)}.$$

Price dynamics:

On every path σ, the probability that the Bayesian posterior probability mass function, obtained from a prior $w_0(i)$ on the model class $\mathcal{P} = \{i \in \mathcal{I}\}$, attaches to state s at time t is given by:

$$W(s|\sigma^{t-1}) = \frac{\sum_{i\in\mathcal{I}} p^i(\sigma^{t,s})w_0(i)}{\sum_{j\in\mathcal{I}} p^j\left(\sigma^{t-1}\right)w_0(j)} = \sum_{i\in\mathcal{I}} p^i(s)w(i|\sigma^{t-1})$$

where $w(i|\sigma^{t-1}) = \frac{p^i(\sigma^{t-1})w_0(i)}{\sum_{j\in\mathcal{I}} p^j(\sigma^{t-1})w_0(j)}$ is the Bayesian prior probability of model i, given the sequence of realizations σ^{t-1} and $w_0(i)$. Clearly $W(s|\sigma^{t-1}) = q^d(s|\sigma^{t-1})$ on every path. Hence, the discounted price evolution coincides with Bayesian updating, price converges to the true model exponentially fast and normalized prices are informationally efficient.

Market selection:

Substituting the price equation in the FOCs I obtain:

$$c_t^i(\sigma) = \frac{p^i(\sigma^t)}{q^d(\sigma^t)} = \frac{p^i(\sigma^t)w_0(i)}{\sum_{j\in\mathcal{I}} p^j(\sigma^t)w_0(j)} = w(i|\sigma^t) = e^{\ln\frac{P(\sigma^t)}{q^d(\sigma^t)} - \ln\frac{P(\sigma^t)}{p(\sigma^t)w_0(i)}} \approx e^{D_t^*(P||q^d) - D_t^*(P||p^i)} \tag{1}$$

Equation 1 has four interpretations:

- Economic interpretation: $c_t^i(\sigma) \to 0 \Leftrightarrow \frac{p^i(\sigma^t)w_0(i)}{\sum_{i\in\mathcal{I}} p^i(\sigma^t)w_0(i)} \to 0$. Trader i vanishes if and only if his effect on next period prices vanishes.

- Probabilistic interpretation: $c_t^i(\sigma) \to 0 \Leftrightarrow w(i|\sigma^t) \to 0$. Trader i vanishes if and only if the measure associated with model i converges to 0. By assumption there is a trader with correct beliefs, hence, by consistency $w(i|\sigma^{t-1}) \to^{P\text{-a.s.}} 0$ if and only if his beliefs are not correct.

- Forecasting interpretation: $c_t^i \to 0 \Leftrightarrow e^{D_t^*(P||q^n) - D_t^*(P|p^i)} \to 0$. Trader i vanishes in probability if and only if his forecasts are worse than the market's.

- Investment strategy interpretation: $c_t^i \to 0 \Leftrightarrow e^{D_t^*(P||q) - D_t^*(P||\beta_i^t p^i)} \to 0$. Trader i vanishes in probability if and only if his investment strategy is worse than the market's.

The existing approach:

The standard approach focuses on the log ratios of the FOCs of different traders:

$$\log\frac{c_t^j(\sigma)}{c_t^i(\sigma)} = \log(\frac{c_0^j}{c_0^i}) + \log\frac{p_t^j(\sigma)}{p^i(\sigma^t)} = \log(\frac{c_0^j}{c_0^i}) + \sum_{\tau=1}^{t}\sum_{s\in\mathcal{S}} I_{\sigma=s}^{\tau}(\sigma)\log\frac{p^j(s)}{p^i(s)}.$$

As an implication of the Strong Law of Large Numbers:

$$\frac{1}{t}\log\frac{c_t^j(\sigma)}{c_t^i(\sigma)} \approx \frac{1}{t}\log(\frac{c_0^j}{c_0^i}) + \sum_{s\in\mathcal{S}} P(s)\log\frac{p^j(s)}{p^i(s)} \approx D_t^*(P||p^i) - D_t^*(P||p^j).$$

Pairwise comparison of the consumption ratio of each couple of traders leads to the conclusion that only the trader with lowest K-L* distance (hence correct beliefs, because $(D^*(p^i||P) = 0 <=> p^i = P)$ survives. Hence, trader i vanishes if there is another trader with a better investment strategy. By standard economic arguments this implies that equilibrium prices converges to P. The approach is silent about price dynamics.

The main theme of the paper is to generalize the example to the case of different payoff functions, non-iid (and possibly incorrect) beliefs, different discount factors and large economies.

5 Equilibrium price dynamics

5.1 Equilibrium prices: Small sample properties

In this Section I discuss the evolution of discounted equilibrium prices. My reference model will be the evolution of the posterior probabilities obtained via Bayes' rule from the market prior. In order to make the comparison meaningful I will focus on economies with no aggregate risk. Parallels between the Bayesian methodology and the evolution of market prices are informally argued by Blume-Easly (2009) for an iid log-economy. The next Proposition shows that, on every path σ, an exact equivalence holds independent of the beliefs of the traders.

Proposition 1. *In a homogeneous discount factor log-economy (that satisfies A4, if large) and with no aggregate risk, discounted price follows a Bayesian mixture model.*[4]

- $q^d(\sigma^t) = \int_{\mathcal{I}} p^i(\sigma^t) w_0(i) d\theta(i)$.
- $q^d(s|\mathcal{F}_{t-1}) = \int_{\mathcal{I}} p^i(s|\mathcal{F}_{t-1}) \frac{p^i(\sigma^{t-1})w_0(i)}{\int_{\mathcal{I}} p^j(\sigma^{t-1})w_0(j)d\theta(j)} d\theta(i) = \int_{\mathcal{I}} p^i(s|\mathcal{F}_{t-1}) w(i|\sigma^{t-1}) d\theta(i)$.

With $w(i|\sigma^{t-1}) = \frac{p^i(\sigma^{t-1})w_0(i)}{\int_{\mathcal{I}} p^j(\sigma^{t-1})w_0(j)d\theta(j)}$ *representing the prior at t-1.*

[4]The result holds in both large and small economies. The integrals have to be interpreted as Lebesgue integrals.

Proof. See Appendix B, Lemma 2. □

The generalization of Proposition 1 to different risk attitudes is complicated for two reasons. First, even for the iid case, discounted prices are not probabilities. Second, as Blume-Easley (2006) point out, risk attitudes have an effect on the rate of convergence of wealth-shares. Defining $w(ML|\sigma^t)$ to be the wealth share of the agent with maximal likelihood on σ^t. Their observation suggests the following Proposition:

Proposition 2. *Given two CRRA, homogeneous discount factor economies with no aggregate risk that only differ in the IES parameter's value (that satisfy A4, if large), then* $\forall t, \forall \sigma \in \Sigma:$

$$
\begin{aligned}
&\textit{i)}\ 0<\gamma'\leq 1\leq\gamma''<\infty \Rightarrow \sum_{s\in\mathcal{S}} q^d_{\gamma'}(s|\sigma^{t-1}) \geq 1 \geq \sum_{s\in\mathcal{S}} q^d_{\gamma''}(s|\sigma^{t-1});\\
&\textit{ii)}\ 0<\gamma'<\gamma''<\infty \qquad \Rightarrow q_{\gamma'}(\sigma^t) \geq q_{\gamma''}(\sigma^t);\\
&\textit{iii)}\ 0<\gamma'<\gamma''<\infty \qquad \Rightarrow w_{\gamma'}(ML|\sigma^t) \geq w_{\gamma''}(ML|\sigma^t).
\end{aligned}
$$

Proof. See Appendix B. □

Proposition 2 describes the small sample properties of discounted prices. The first two lines generalize Proposition 1 to CRRA-economies that are not log. They show that there is a negative relation between the value of γ and the price level and that log-economies constitute the knife's edge case in which discounted prices determine a probability measure on Σ. The economic intuition is as follows. Each trader subjectively believes that equilibrium prices are incorrect, hence the no trade theorems do not apply and each trader sees the market not only as an instrument to transfer wealth, but also as a "favorable" gamble. Depending on the risk attitudes, traders may be eager or reluctant to exploit this opportunity and the economy generates discounted prices with a positive (negative) risk premium if $\gamma > 1 (< 1)$. The last line shows that the value of γ is also inversely related to the wealth share of the trader with beliefs of maximal likelihood. The reason is that a lower γ implies more extreme investment strategies and hence faster accumulation of capital for the trader with beliefs with maximal likelihood. The persistence of these effects and their asymptotic implications are discussed in the next section.

5.2 Equilibrium prices: Asymptotic properties

The results of this section differs depending on the size of the economy. The following definition will play a fundamental role. What it captures is the notion of two functions that convergence to 0 or diverge to infinity at the same rate.

Definition 9. *Given two functions $f(x)$ and $g(x)$, $f(x) \doteq g(x)$ if and only if* $\exists 0 < \underline{a} \leq \bar{a} < \infty$: $\forall x,\ f(x) \in [\underline{a}g(x), \bar{a}g(x)]$.

All of the results that follows can be seen as implications of this technical Theorem:

Theorem 2. *In an economy that satisfies A1-A3:*

$$i)\ \text{If the economy is small: } \forall \sigma \in \Sigma,\ q(\sigma^t) \doteq \sum_{i \in \mathcal{I}} \beta_i^t p^i(\sigma^t) w_0(i);$$

$$ii)\ \text{In some large economies: } \forall \sigma \in \Sigma,\ q(\sigma^t) \not\doteq \int_{\mathcal{I}} \beta_i^t p^i(\sigma^t) w_0(i) d\theta(i).$$

Proof. $i)$ See Appendix B. $ii)$ Implied by Corollary 4 in Appendix C, comparing the case of $\gamma = 1$ with $\gamma \neq 1$. □

For small economies, Theorem 2 provides an extremely strong characterization of the rate of convergence of equilibrium prices. The constant factor implicit in the $\doteq$ notation has no asymptotic relevance. It should also be noted that there is no reference to the true probability distribution. This property of equilibrium prices is universal: it holds on every path of realization. The existence and the properties of a true distribution are orthogonal to this characterization. Theorem 2 shows that small economies and large economies are qualitatively different as the relationship between prices, beliefs and discount factors of the traders in the economy are not the same.

5.2.1 Prices in small economies

Theorem 2 implies that if the economy is small and irrespectively from the risk attitudes of the traders then, WLOG, we can use $\sum_{i \in \mathcal{I}} \beta_i^t p^i(\sigma^t)$ to determine the asymptotic properties of equilibrium prices. More specifically, equilibrium price satisfy these two conditions:

Corollary 1. *In a small economy that satisfies A1-A3:*

$$i)\ \forall \sigma \in \Sigma, \ \limsup_t \left| \frac{\max_{i \in \mathcal{I}} \beta_i^t p^i(\sigma^t) w_0(i)}{q(\sigma^t)} \right| < +\infty$$
$$ii)\ \ \forall \sigma \in \Sigma, \ \liminf_t \left| \frac{\max_{i \in \mathcal{I}} \beta_i^t p^i(\sigma^t) w_0(i)}{q(\sigma^t)} \right| > 0$$

Proof. Follows from Theorem 2 □

The first condition states that the discounted beliefs of each trader are more than absolutely continuous with respect to equilibrium prices (this condition is stronger than absolute continuity as it holds on every path and not only for sets of positive measure). The second condition states that equilibrium prices do not attach too much "weight" to any sequence. This ensures that the doubling rate of market's investment strategy is, on every path, bounded above by the doubling rate the trader with the best investment strategy in the market. As an application, the following Proposition holds:

Proposition 3. *In a small economy that satisfies A1-A3, the market investment strategy is as good as the best investment strategy in the economy.*

Proof. See Appendix B. □

These results depend crucially on the assumption about the cardinality of $\mathcal{I}$. The following example shows that Proposition 3 is incorrect in a large economy:

Example: Consider the following economy: the true data generating process is Bernoulli, iid $\theta \in (0,1)$. We have a continuum of traders with identical discount factors, log utility (by Proposition 1, discounted prices are probabilities) and iid beliefs. The beliefs of the traders and the initial wealth-shares are such that the market prior is smooth on the interval $(0,1)$. Therefore there is a trader with rational expectations. If Proposition 3 were correct, $D^*(P||q^d) - D^*(P||P) < \infty$. This is to say, discounted equilibrium prices merge with the true probability. This is impossible as it would imply that discounted equilibrium prices converge to the true distribution faster than the Bayesian estimator with the same prior, contradicting the complete class theorem.

For homogeneous discount factor economies, Theorem 2 allows for a precise characterization of the rate of convergence of next period discounted prices. The next Proposition shows that in small economies risk attitudes have no effect on the rate of convergence of next period discounted prices.

Proposition 4. *In a homogeneous discount factor small economy with constant aggregate endowment that satisfies A1, if there is a unique trader with beliefs with the highest likelihood, next period discounted prices converge exponentially fast to the beliefs of the traders with the most accurate beliefs.*

Proof. By Lemma 4 in Appendix B the convergence rate is the same as the convergence rate we would have in a log-economy, hence, by Proposition 1 the same as Bayesian learning. Standard results in bayesian learning (see Marinacci-Massari) ensures that the convergence rate is exponential. □

Proposition 4 tells us that, in a small economy, equilibrium price converges exponentially fast, in contrast with the finding of Yan (2008), it suggests that market selection provides a valid theoretical argument in favor of the rational expectations assumption.

5.2.2 Prices in large economies

In large economies, the effect of risk attitudes can persist asymptotically. The reason is that discounted prices do not converge to probabilities fast enough to overcome the effect of the risk premium discussed in Proposition 2. Unlike the small economy case, the investment strategy content and the informational content of equilibrium prices are not asymptotically equivalent. For the case of iid CRRA homogeneous discount factor economies the following Theorem holds:

Theorem 3. *Given two large CRRA, homogeneous discount factor economies with no aggregate risk that only differ in the IES parameter's value and satisfy A4, then*

$$0 < \gamma' < \gamma'' < \infty \Rightarrow D_t^*(P||q^d_{\gamma'}) - D_t^*(P||q^d_{\gamma''}) \to -\infty.$$

Proof. See Appendix C, Corollary 4. □

Theorem 3 tells us that in large economies risk attitudes have an asymptotic effect on equilibrium prices. Specifically, it shows that the equilibrium price in the economy with the lower value of γ generate a better doubling rate (because it diverges to $+\infty$ at a slower rate). As it is shown by Proposition 2, the higher doubling rate is due the fact that the more risk loving the traders, the higher is the price level. The comparison between Theorems 1 and 3 shows that discounted prices reflect both the risk attitudes of the traders in the market and the aggregate information. In sharp contrast with the small economy case, the effect of risk attitudes does not fade away fast enough and it affects the asymptotic properties of equilibrium prices. In the next Section I discuss how this effects survival.

6 Wealth-shares Dynamics

In this Section I use my characterization of equilibrium prices to discuss wealth-share dynamics. My approach addresses the problem of wealth-share convergence via a direct comparison between equilibrium prices, trader beliefs and discount factors. The restriction to Arrow's securities economy is done WLOG as the price of more complex assets in a complete market can always be obtained via no arbitrage pricing. The following Proposition links the standard definition of survival to my setting:

Proposition 5. *In an economy (either small or large) that satisfies A1-A3, trader i vanishes on path σ if and only if:*

$$\forall s \in S \quad \lim_{t\to\infty} \frac{\beta_i^t p^i(\sigma^{t,s}) w_0(i)}{q(\sigma^{t,s})} \to 0$$

Proof. See Appendix B □

In small economies $\frac{\beta_i^t p^i(\sigma^{t,s}) w_0(i)}{q(\sigma^{t,s})} \mathrel{\mathring{=}} \frac{\beta_i^t p^i(\sigma^t) w_0(i)}{\sum_{\mathcal{I}} \beta_j^t p_t^j(\sigma) w_0(j)}$ therefore the wealth of trader i is proportional to the weight his beliefs have in the determination of next period prices and the following interpretation holds. If trader i's consumption goes to 0 on σ, then trader i does not use the market to transfer consumption on σ, hence his beliefs have no weight on the determination of the equilibrium prices. Conversely, if trader i's beliefs have no weight on the determination of the equilibrium prices on σ, then he is not moving consumption on that path, which is to say that his consumption is 0 on that path. Given Proposition 5, I can apply my results in price dynamic to derive asymptotic results in wealth shares. I begin with a general necessary condition to vanish:

Proposition 6. *In an economy (either small or large) that satisfies A1-A3,*

- *Trader i vanishes in probability if his investment strategy is worse than the market's strategy.*
- *Trader i survives in probability if his investment strategy is as good as the market's strategy.*

Proof. see Appendix B. □

My condition to vanish is weaker than Sandroni's condition, as it only implies that a trader vanishes in probability, not P-a.s.. Nevertheless, it is more general: it can be used to discriminate between learning traders, it relies on fewer assumptions and it can be applied to large economies. In particular, Proposition 7 and 11 cannot be proved using either Sandroni's or Blume-Easley's measure of distance (an equivalent result to Propositions 7 is proven in Blume-Easley (2006) as a special case). Proposition 6 cannot be used to prove P-a.s convergence because it is a result on the asymptotic value of consumption, not on the number of times consumption is positive. Consumption can converge to 0 in expectation and yet be positive sporadically but infinitely often. In order to capture this second aspect, a more technical condition must be imposed:

Theorem 4. *In an economy (either small or large) that satisfies A1-A3:*

$$E\left(\limsup \frac{\beta_i^t p^i(\sigma^t) w_0(i)}{q(\sigma^t)}\right) = 0 \Leftrightarrow \text{ Trader } i \text{ vanishes } P\text{-a.s.}$$

Proof. See Appendix B □

Theorem 4 suggests that in order to capture the sufficient condition to vanish we need to control for two things. First, the definition of survival states that a trader survives if his consumption is positive infinitely often, not that his expected asymptotic consumption must be positive. The difference is subtle, but Blume-Easley (2009) show that is possible for a trader to have positive wealth-share so infrequently that his expected consumption goes to 0. Second, we need to keep track of the geometry of the beliefs in the economy: a trader that makes more accurate forecasts than every other trader in the market and with a discount factor that exactly compensates for his better forecasting skills can vanish if the set of sequences in which he has positive wealth has measure 0. A detailed discussion about conditions under which this can occur can be found in Blume-Easley(2009).

Proposition 6 can be used to discriminate between Bayesian learners with smooth priors and different cardinality in their prior support. As already observed by Blume-Easley (2006), the general result is that in the presence of multiple Bayesian traders with nested prior supports, a trader vanishes if there is another trader whose prior support has lower dimensionality and contains the true model. The result captures an intuitive idea: more parameters to be estimate correspond to slower convergence.This observation is captured in the following Proposition:

Proposition 7. *In a small homogeneous discount factor economy that satisfies A1-A3, with two Bayesian traders i and j with smooth priors, nested prior supports and such that the dimensionality of j's prior support is higher than the dimensionality of i's prior support. If the true model belongs to i's prior support, trader j vanishes in probability.*

Proof. Follows from Proposition 6 and Corollary 4 in Appendix C (with $\gamma = 1$). □

In the next section I consider some relevant applications of Theorem 4 that are new to the literature.

6.1 Learning, risk attitudes and the number of traders

In this section I discuss the amount of information needed to survive in an economy and its relation with the size of the economy and risk attitudes. For small economies, the first result I present is standard in the market selection literature. It states that if a Bayesian trader puts positive weight on the true model, then that trader survives P-a.s..

Corollary 2. *In a small homogeneous discount factor economy that satisfies A1-A3, a Bayesian trader (B) with positive wealth share and whose prior attaches positive probability to the true model survives P-a.s..*

Proof. See Appendix B. □

An equivalent result is also true for large economies:

Proposition 8. *In a homogeneous discount factor CRRA economy that satisfies A2-A3, populated by a continuum of traders whose beliefs satisfy A4 and by a Bayesian trader (B) whose prior attaches positive probability to the true model; the Bayesian trader survives P-a.s..*

Proof. See Appendix B. □

Therefore, a trader that attaches positive probability to P always survives.
The problem with these results is that they depend on ex-ante knowledge of the true distribution. The next results discuss survival without referring to the true distribution. Proposition 2 tells us that risk attitudes have an effect on the rate of convergence of discounted prices. If the economy is small, Theorem 2 implies that this effect has no asymptotic relevance as discounted prices merge with the beliefs of the trader with the most accurate beliefs. Therefore the following Corollary holds.

Corollary 3. *In a small, homogeneous discount factor economy that satisfies A1-A3, a Bayesian trader survives on every path if and only if he has access to sparse information.*

Proof. Follows from Theorem 2 and Proposition 5. □

The picture is different in large economies. Theorem 3 tells us that risk attitudes can have an asymptotic effect on price dynamics. The next Proposition discusses its wealth-shares implications.

Proposition 9. *In a large, homogeneous discount factor, iid, CRRA economy that satisfies A2-A3, populated by a continuum of traders $\mathcal{I}$ with beliefs that smoothly cover the simplex and by a Bayesian trader with positive initial wealth-share with a smooth prior on the union of the beliefs of* *the other traders; the asymptotic wealth-share of the Bayesian trader (B) depends on risk attitudes as follows:*

i) $\gamma \in (0, 1) \Leftrightarrow$ *trader B vanishes P-a.s.*

ii) $\gamma = 1 \Leftrightarrow$ *trader B survives and his wealth-share is less than 1 P-a.s.*

iii) $\gamma > 1 \Leftrightarrow$ *trader B is the only survivor P-a.s.*

Proof. See Appendix B. □

The intuition goes as follows. By assumption, we have a continuum of traders with beliefs that smoothly cover the simplex, therefore there is a trader in the economy with rational expectations and an infinitesimal fraction of wealth-share. We want to compare the rate at which his wealth-share grows and the rate at which the Bayesian learns. As Proposition 2 shows, the parameter γ regulates the rate at which wealth moves between traders. $\gamma = 1$ constitutes the knife's edge in which this rate coincides with the learning rate of the Bayesian trader, consequently this is the only case in which the Bayesian trader survives without dominating. For $\gamma < 1$, wealth-shares move faster than the log case and the learning rate of the Bayesian is not fast enough to compensate for the informational disadvantage. The case $\gamma > 1$ is specular, and the Bayesian trader learns fast enough to become the only survivor. [5]

It is easy to find arguments in favor of the hypothesis that real world markets are populated by traders that are more risk loving than log. For example, the bias in the incentives of portfolio managers due to the fact that they invest other people's money and participate in the gains without (directly) participating in the losses. Proposition 9 is good news for the theory of market selection because it implies that the selection mechanism becomes faster the more risk loving are the traders in the economy. On the other side, it is bad news for the investors. It shows that the more risk loving are the traders in the economy the more information is needed to survive. In particular, if $\gamma < 1$, to know the beliefs of all of the traders in the market is not enough to survive. The only traders that survive in a large risk loving market are the lucky ones and those with insider

[5] To be precise, the above argument should be made in terms of neighborhoods of the true parameters (in the Fisher information metrics) with positive Lebesgue measure (i.e. positive wealth share) and not in terms of a trader with infinitesimal wealth share.

information.

The following stronger result shows that the γ parameter and the dimensionality of the market prior have comparable effects on the rate of convergence.

Proposition 10. *In a large homogeneous discount factor economy, to have more information than the market is not necessary nor sufficient to vanish.*

Proof. Examples. $\square$

The first example illustrates the case in which the Bayesian learner knows more than the market and yet vanishes. There are two parameter to be estimated and the Bayesian learner knows one of them. The second example discusses the case in which the Bayesian learner knows less than the market and yet become the only survivor. There are two parameters to be estimated and everybody but the Bayesian knows one of them.

Example 1 Consider a homogeneous discount factor economy with three states $\mathcal{S} = \{a, b, c\}$ with iid true probabilities $P = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$. Trader B is a Bayesian who knows that the probability of state a, has a uniform prior on $(0, \frac{2}{3})$ on the probabilities of the other two states and holds $\frac{1}{2}$ of the wealth of the economy. The other half of the wealth is evenly distributed to a continuum of traders I_c with iid beliefs that cover the tridimensional simplex. All traders have a CRRA payoff functions with the same parameter γ. The discounted equilibrium prices are given by:

$$
\begin{aligned}
q^d(\sigma^t) &= \left(\frac{1}{2} p_t^B(\sigma)^{\frac{1}{\gamma}} + \frac{1}{2} \int \int p(\sigma^t | \theta 1, \theta_2)^{\frac{1}{\gamma}} d\theta_1 d\theta_2 \right)^{\gamma} \\
&= \left(\frac{1}{2} \left(\int_0^{\frac{2}{3}} p(\sigma^t | \theta_2) d\theta_2 \right)^{\frac{1}{\gamma}} + \frac{1}{2} \int \int p(\sigma^t | \theta 1, \theta_2)^{\frac{1}{\gamma}} d\theta_1 d\theta_2 \right)^{\gamma}
\end{aligned}
$$

The FOCs implies:

$$c^B(\sigma^t) = c_0^B \left(\frac{p^B(\sigma^t)}{q^d(\sigma^t)}\right)^{\frac{1}{\gamma}} = \frac{c_0^B\left(\int_0^{\frac{2}{3}} p(\sigma^t|\theta_2)d\theta_2\right)^{\frac{1}{\gamma}}}{\left(c_0^B\int_0^{\frac{2}{3}} p(\sigma^t|\theta_2)d\theta_2\right)^{\frac{1}{\gamma}} + \int\int p(\sigma^t|\theta_1,\theta_2)^{\frac{1}{\gamma}}d\theta_1 d\theta_2} = \frac{c_0^B}{c_0^B + \frac{\int\int p(\sigma^t|\theta_1,\theta_2)^{\frac{1}{\gamma}}d\theta_1 d\theta_2}{\left(\int_0^{\frac{2}{3}} p(\sigma^t|\theta_2)d\theta_2\right)^{\frac{1}{\gamma}}}}$$

$$\overset{\circ}{=}_{By\ Cor.4, Appendix\ C} \frac{c_0^B}{c_0^B + \frac{e^{\frac{1}{\gamma}\left(\ln P(\sigma^t) - \frac{2\gamma}{2}\ln t\right)}}{e^{\frac{1}{\gamma}\left(\ln P(\sigma^t) - \frac{1}{2}\ln t\right)}}} \quad P\text{-a.s.}$$

Therefore the Bayesian trader vanishes P-a.s. if and only if $\gamma < \frac{1}{2}$

Example 2 Consider a homogeneous discount factor economy with three states $\mathcal{S} = \{a, b, c\}$ with iid true probabilities $P = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$. Trader B is a Bayesian with uniform prior on the tridimensional simplex and holds $\frac{1}{2}$ of the wealth of the economy. $(0, \frac{2}{3})$ on the probabilities of the other two states . The other half of the wealth is evenly distributed to a continuum of traders I_c with iid beliefs so that all of the traders agree on the probability of the first state, and their beliefs cover the bidimensional simplex. All traders have a CRRA payoff functions with the same parameter γ. Following the same logic of the previous example the FOCs implies:

$$c^B(\sigma^t) = c_0^B \left(\frac{p^B(\sigma^t)}{q^d(\sigma^t)}\right)^{\frac{1}{\gamma}} \qquad \overset{\circ}{=}_{By\ Cor.4, Appendix\ C} \frac{c_0^B}{c_0^B + \frac{e^{\frac{1}{\gamma}\left(\ln P(\sigma^t) - \frac{\gamma}{2}\ln t\right)}}{e^{\frac{1}{\gamma}\left(\ln P(\sigma^t) - \frac{2}{2}\ln t\right)}}} \quad P\text{-a.s.}$$

Therefore the Bayesian trader is the only survivor P-a.s. if and only if $\gamma > 2$

6.2 Real world investment strategies

This Section describes the merit of investment strategies that are commonly adopted in real world markets. I consider two strategies: the diversified investment strategy and the copy-trader strategy. Given a set of investment opportunities (for example the strategies of the other traders, different edge fonds or different stocks) the diversified investment strategy consist on investing a fraction of the initial capital in each investment in the first period and hold the position (this strategy correspond to holding a long position on a market index). The copy-trader strategy is to invest, in each period, all of the capital according to the investment that had the highest cumulated wealth up to the previous period (to copy the strategy of the trader that did best in the past). This strategy is intuitive and is commonly advertised by providers of internet-trading (for example, www.etoro.com, www.currensee.com). My results provide theoretical support for the diversified portfolio strategy and against the copy-trader one. If the two strategies consider the same set of possible investments, the diversified portfolio strategy always guaranties survival while the copy-trader strategy leads to extinction if the best strategy (the leader's strategy) changes often. This result does not depends on the presence of frictions in the market, the copy-trader strategy is suboptimal because produces lower capital accumulation rate, not because there is a cost in moving the capital between investment strategies.

Proposition 11. *In small economy that satisfies A1-A3, in which there are two traders d, c that consider the same set of traders and adopt respectively the diversified investment strategy and the copy-trader strategy, then:*

- *Trader d survives on every path.*
- *Trader s survives if and only if the leaders do not alternate infinitely often.*

Proof. See Massari (2012a) □

The following example illustrates the content of Proposition 11:

Example Consider a log-economy with two states $\mathcal{S} = \{a, b\}$, iid true probability $P := [\frac{1}{2}, \frac{1}{2}]$, four traders (1,2,3,4) with log utilities, the same discount factors and $e_t^i = \frac{1}{4}\ \forall t,\ i = 1, 2, 3, 4$. Traders 1 and 2 have iid beliefs: $p_1 := [\frac{2}{3}, \frac{1}{3}], p_2 := [\frac{1}{3}, \frac{2}{3}]$. Traders 3 follows the diversified investment strategy and trader 4 follows the copy-trader strategy:

$$p^3(a|\sigma^{t-1}) := p^1(a)\frac{\frac{1}{2}p^1(\sigma^{t-1})}{\frac{1}{2}p^1(\sigma^{t-1}) + \frac{1}{2}p^2(\sigma^{t-1})} + p^2(a)\frac{\frac{1}{2}p^2(\sigma^{t-1})}{\frac{1}{2}p^1(\sigma^{t-1}) + \frac{1}{2}p^2(\sigma^{t-1})}$$

$$p^4(a|\sigma^{t-1}) := \begin{cases} p^i(a) : i = \arg\max\left\{p^1(\sigma^{t-1}), p^2(\sigma^{t-1})\right\} \\ \text{random } 1/2p^1(a), 1/2p^2(a) \text{ if ties occur.} \end{cases}$$

Proposition 11 implies that trader 4 vanishes P-a.s.

More formally, the two strategies described above correspond to investing using two bayesian forecasting procedures: Bayesian Mixture (Bernardo-Smith 1994) and Bayesian Factor Model (Kass-Raftery 1995), respectively. The surprising result is that if the posterior does not converge, Bayesian Mixture forecasts outperform Bayesian Factor forecasts in terms of likelihood.[6]

6.3 Survival in iid economies

Precise results on who survives in iid economies with heterogeneous discount factors have been proved surprisingly hard to obtain. In this section I conclude the project of Blume-Easley (2009) about survival in the iid setting. For an easier comparison, I will adopt their notation and approach. Blume-Easley (2009) present an almost tight geometric condition to verify if a trader survives in an iid economy. They first apply the necessary condition for extinction from Blume-Easley (2006) to all of the traders in the economy in order to identify the set of potential survivors (SUR). Then, they apply their geometric condition to further refine this set and identify the actual survivors. More formally,

Definition 10. *Trader's i survival index: $s_i = \log \beta_i - E \log \frac{P}{p^i}$*

Let SUR denote the set of traders with maximal survival index (in my setting this is set of traders with the best investment strategies). Blume-Easley show that the fate of a trader in SUR is determined by how his beliefs are positioned relative to the beliefs of the other traders in SUR. Let $lo(p^j) = \left(\log\left(\frac{p(s)}{p(S)}\right)\right)_{s=1}^{S-1}$ and $C\left\{lo(p^j)\right\}_{j\in SUR}$ be the closed convex cone generated by the log-probability vectors of the traders.

Definition 11. *Trader i is interior if $lo(p^j)$ is in the relative interior of $C\left\{lo(p^j)\right\}_{j\in SUR}$. He is extremal if $lo(p^j)$ is an extreme point; that is, not a non-negative linear combination of the other $lo(p^j)$.*

[6] This result is consistent but not overlapping with the finding in Grünwald-Rooij (2005). In particular, it conforms to the general observation that forecasting schemes that can produce forecasts that correspond to models outside the model prior support perform better, in terms of sum squared error and log loss function, than forecasting schemes that only formulates forecasts that corresponds to model that belong to the model prior support.

Blume-Easley (2009) shows that if $|\mathcal{S}| \leq 3$, a trader survives if and only if he is in SUR. If $|\mathcal{S}| > 3$, interior traders vanish and extremal traders survive. Nothing is said about the asymptotic fate of traders on the boundaries of $C\left\{lo(p^j)\right\}_{j\in SUR}$.

Geometrically, the fate of an internal (extremal) trader is determined by the time spent by a discrete $\mathcal{S}-1$ dimensional Random Walk in compact sets (on $\mathcal{S}-1$ dimensional half-spaces). These results are readily available and have been used by Blume-Easley (2009). In order to determine the fate of the traders whose beliefs lie on the boundaries of $C\left\{lo(p^j)\right\}_{j\in SUR}$ it is necessary to have results on the time spent by a discrete $\mathcal{S}-1$ dimensional Random Walk on $\mathcal{S}-r$ conical subspaces. In Appendix C, I prove the needed results. The next Theorem completes their Theorem including these cases:

Definition 12. *Let K^2 be the set of traders in SUR such that $lo(p^j)$ is a non-negative linear combination of $lo(p^i)_{j\neq i\in SUR}$ in at most two dimensions.*

Theorem 5. *Assume A1-A3, in a iid economy*

$$Trader\ i\ survives \Leftrightarrow i \in K^2$$

Proof. See Appendix D and Blume-Easley (2009) (Theorems 2 and 3). □

7 Conclusion

This paper introduces a new approach to the market selection mechanism. By interpreting prices as investment schemes, I am able to compare price evolution and probabilistic learning, even though prices are not necessarily probabilities and their evolution is not obtained via Bayes' rule. My characterization reveals a new implication of the market selection hypothesis: prices convey all of the information available in the market as efficiently as a Bayesian with the same information. In line with the results of Sandroni (2000), Blume-Easley (2006) and Yan (2008) for non-growing economies, I find that, if the economy is small, different risk attitudes have no asymptotic effect on wealth-share. Nevertheless, I show that the result does not apply to large economies. The amount of information needed by a trader to survive also depends on the size of the economy. In small economies, a trader that knows the beliefs of all of the other traders is always among the survivors. In large economies, if the traders are more risk loving than log, a trader survives only if he has more information than the market; if the traders are less risk loving than log, a trader with less information than the market can be the only survivor. Log-economies constitute the knife's edge case: knowledge of all traders' beliefs ensures survival. In addition, this paper extends existing results by providing a necessary and sufficient condition for a trader to vanish that applies to general complete market economies. On the way, it discusses survival in iid economies, which concludes the analysis of Blume-Easley (2006), and the wealth-shares implications of investment strategies that are commonly observed in real markets. My results provide theoretical support in favor of the diversified portfolio strategy and against the copy-trader strategy.

Appendices

A Appendix

In this section I discuss the relationship between the notion of distance I adopt ($D^*_t(P||p^i)$) and the distances adopted by Sandroni (Entropy:$\mathcal{E}_{i,t}$) and Blume-Easley (Sum of Conditional Relative Entropies $D_t(P||p^i)$).

Definition 13. *Entropy of trader i:*

$$\mathcal{E}_{i,t} \equiv -\frac{1}{t}\sum_{\tau=1}^{t} E_\tau\left(\log\frac{P_\tau(\sigma|\mathcal{F}_{\tau-1})}{p_\tau(\sigma|\mathcal{F}_{\tau-1})}\right) = -\frac{1}{t}\sum_{\tau=1}^{t} d(P_\tau||p^i_\tau)$$

Definition 14. *Sum of conditional relative entropies of trader i:*

$$D_t(P||p^i) \equiv \sum_{\tau=1}^{t} E_\tau\left(\log\frac{P_\tau(\sigma|\mathcal{F}_{\tau-1})}{p(\sigma|\mathcal{F}_{\tau-1})}\right) = \sum_{\tau=1}^{t} d(P_\tau||p^i_\tau)$$

The K-L and $\mathcal{E}$ distances are approximations of the K-L* distance I am using. Most of the times these distances provide equivalent implications, nevertheless examples in which this is not the case can be constructed.

The scope of Assumption A5 in Blume-Easly's (2006) and Sandroni's (2000) Definition 8 (equivalent to A5) is to make these distances equivalent.

- A5: There is a $\delta > 0$ such that for all paths σ, all dates t, $s \in S$, and traders i, $P(s|\mathcal{F}_{t-1})(\sigma) > 0 \Rightarrow p(s|\mathcal{F}_{t-1})(\sigma) > \delta$ and $p^i(s|\mathcal{F}_{t-1})(\sigma) > \delta$.

An assumption similar to A5 is needed in terms of $\mathcal{E}$ because the sum of conditional relative entropies can diverge on a set of measure 0 fast enough to invalidate the Strong Law of Large Numbers. This implies that it is possible to construct an economy (an example can be found in Blume-Easley (2006)) with beliefs and true probabilities such that $\mathcal{E}$ and K-L fail to correctly identify the limit of the summations. An intuition of it is given:

example 1.

Let: $s_t = \begin{cases} t^2 - 1 \quad w.p. \ \frac{1}{t^2} \\ -1 \quad w.p. \ 1 - \frac{1}{t^2} \end{cases}$

Then:

$$-\infty =^a \lim_{t\to\infty} \sum_{\tau=1}^{t} s_\tau = \lim_{t\to\infty} E\left(\sum_{\tau=1}^{t} s_\tau\right) \neq \lim_{t\to\infty} \sum_{\tau=1}^{t} E_t(s_\tau) = \lim_{t\to\infty} \sum_{\tau=1}^{r} 1 - \frac{1}{\tau^2} - (1 - \frac{1}{\tau^2}) = 0$$

$$-1 =^a \lim_{t\to\infty} \frac{1}{t} \sum_{\tau=1}^{t} s_\tau = \lim_{t\to\infty} E\left(\frac{1}{t} \sum_{\tau=1}^{t} s_\tau\right) \neq \lim_{t\to\infty} \frac{1}{t} \sum_{\tau=1}^{t} E_t(s_\tau) = \lim_{t\to\infty} \frac{1}{t} \sum_{\tau=1}^{t} 1 - \frac{1}{\tau^2} - (1 - \frac{1}{\tau^2}) = 0$$

a) By Borel-Cantelli Lemma

Nevertheless Massari (2012) shows that even assuming A5 the use of the distance proposed by Blume-Easley can lead to incorrect results. This finding has serious implications as it shows that the $D(.||.)$ distance cannot be used to identify the necessary condition for extinction for a non-iid economy.

On the other hand, the measure adopted by Sandroni is to coarse to discriminate between learning rates. If two traders learns at different rates, they would both pass Sandroni's criterion. In particular the measure of distance he adopt cannot be used to discriminate between Bayesian learner with different dimensionality in the prior support because the differences in the likelihoods are "killed" by the $\frac{1}{t}$ term.

These approximations do not make the model more tractable and can lead to mistakes (as Massari (2012) shows). Therefore, nothing is lost by using a more abstract, but correct, notion of distance. The convergence Theorems are less demanding in terms of assumptions (A5 becomes superfluous) and special cases, instead of being ruled out, can be treated and solved individually.

B Appendix

Lemma 1. *In a small economy that satisfies A1-A3:* $\forall t, \forall \sigma \in \Sigma, \exists a, b \in \mathbb{R}:$

$$0 < a < \sum_{i \in \mathcal{I}} \frac{1}{u_i'(c_t^i(\sigma))} < b < \infty \tag{2}$$

Proof. *i*): $\sum_{i \in \mathcal{I}} \frac{1}{u_i'(c_t^i(\sigma))} < \infty$:
$\forall \sigma \in \Sigma,\ \forall i \in \mathcal{I}, u_i'(c_t^i(\sigma)) > 0$ as the total endowment is finite (by A2), the payoff functions are monotone and strictly concave with positive derivative at 0 (by A1).
ii): $0 < \sum_{i \in \mathcal{I}} \frac{1}{u_i'(c_t^i(\sigma))}$:
$\sum_{i \in \mathcal{I}} \frac{1}{u_i'(c_t^i(\sigma))} = 0 \Leftrightarrow \forall i \in \mathcal{I}, u_i'(c_t^i(\sigma) = \infty$ which is true iif all the traders have 0 consumption and satisfy the Inada condition at 0. The first requirements is impossible as it violates the market clearing condition: $\sum_{i \in \mathcal{I}} c_t^i = \sum_{i \in \mathcal{I}} e_t^i > 0$.

$\square$

Proof of Theorem 1 (small economies)

Since the bayesian model cannot do better, in terms of likelihood, than the best model in its prior support, the desired result can be proven by showing that under A1-A3, if all of the traders have the same discount factor: $\forall i \in I, \exists v_0^i > 0: \quad \log \frac{1}{v_0^i} > \log \frac{p^i(\sigma^t)}{q^n(\sigma^t)} \quad \forall \sigma^t \in \Sigma.$

Proof. Rearranging the FOCs, normalized equilibrium prices are:

$$q^n(\sigma^t) = \frac{q(\sigma^t)}{\sum_{\sigma^t} q(\sigma^t)} = \frac{\frac{\beta^t \sum_{i\in I} p^i(\sigma^t) \frac{1}{u_i'(c_0)}}{\sum_{i\in I} \frac{1}{u_i'(c_t(\sigma))}}}{\beta^t \sum_{\sigma^t\in\mathcal{S}^t} \left(\frac{\sum_{i\in I} p^i(\sigma^t)\frac{1}{u_i'(c_0)}}{\sum_{i\in I}\frac{1}{u_i'(c_t(\sigma))}} \right)} \tag{3}$$

By Lemma 1 $\exists a, b : 0 < a \leq b < +\infty$

$$q^n(\sigma^t) \in [\frac{\frac{\beta^t \sum_{i\in I} p^i(\sigma^t)\frac{1}{u_i'(c_0)}}{b}}{\beta^t \sum_{\sigma\in\mathcal{S}^t}\left(\frac{\sum_{i\in I} p^i(\sigma^t)\frac{1}{u_i'(c_0)}}{a}\right)}, \frac{\frac{\beta^t \sum_{i\in I} p^i(\sigma^t)}{a}}{\beta^t \sum_{\sigma\in\mathcal{S}^t}\left(\frac{\sum_{i\in I} p^i(\sigma^t)\frac{1}{u_i'(c_0)}}{b}\right)}]$$

$$=> q^n(\sigma^t) \in [\frac{a}{b}\sum_{i\in I} p^i(\sigma^t)\frac{1}{u_i'(c_0)}, \frac{b}{a}\sum_{i\in I} p^i(\sigma^t)\frac{1}{u_i'(c_0)}]$$

$$=> \forall i \in I, \forall\sigma \in \Sigma, \forall t,\ q^n(\sigma^t) > \frac{a}{b} p^i(\sigma^t)\frac{1}{u_i'(c_0)}$$

And the result follows rearranging, noticing that in equilibrium $\forall i \in \mathcal{I}$, $+\infty > \frac{1}{u_i'(c_0)} > 0$ and taking log. □

Proof of Theorem 2, (small economies)

Proof. The equilibrium conditions implies $\forall t$, $\frac{q(\sigma^t)}{p^i(\sigma^t)} = \frac{\beta^t u_i'(c_t)}{u_i'(c_0)}$

Rearranging and summing over traders:

$$q(\sigma^t) = \frac{\sum_{i\in\mathcal{I}} \beta_i^t \frac{p^i(\sigma^t)}{u_i'(c_0)}}{\sum_{j\in\mathcal{I}} \frac{1}{u_j'(c^j(\sigma^t))}} \tag{4}$$

Lemma 1 implies $\forall t, \forall \sigma \in \Sigma$:

$$\frac{\sum_{i\in\mathcal{I}} \frac{\beta_i^t p^i(\sigma^t)}{u_i'(c_0)}}{b} \leq q(\sigma^t) \leq \frac{\sum_{i\in\mathcal{I}} \frac{\beta_i^t p^i(\sigma^t)}{u_i'(c_0)}}{a}$$

Which is to say $\forall i, \forall t, \forall \sigma \in \Sigma$:

$$q(\sigma^t) \doteq \sum_{i\in\mathcal{I}} \frac{\beta_i^t p^i(\sigma^t)}{u_i'(c_0)} \doteq^a \sum_{i\in\mathcal{I}} \beta_i^t p^i w_0^i(\sigma^t)$$

a :The result follows noticing that in equilibrium $\forall i \in \mathcal{I}$, $+\infty > \frac{1}{u_i'(c_0)} > 0$.

□

Proof of Proposition 3

Proof. I have to show two things:

i) $\lim_{t\to\infty} D^*(P||q^d) - \min_{i\in\mathcal{I}}(D^*(P||\beta_i^t p^i) < +\infty$

Suppose not. This implies that $\exists\sigma \in \Sigma, \exists i \in \mathcal{I}:\ \log\frac{\beta_i^t p^i(\sigma^t)}{q^d(\sigma^t)} \to +\infty$.

Contradicting Theorem 2.

ii) $\lim_{t\to\infty} D^*(P||q^d) - \min_{i\in\mathcal{I}}(D^*(P||\beta_i^t p^i) > -\infty$

Suppose not. This implies that $\exists\sigma \in \Sigma : \forall i \in \mathcal{I},\ \frac{\beta_i^t p^i(\sigma^t)}{q^d(\sigma^t)} \to -\infty$

Contradicting Theorem 2. □

Lemma 2. *Proof.* Standard calculations show that for a homogeneous discount factors CRRA economy ($\forall i \in \mathcal{I},\ u^i(c) = \frac{c^{1-\gamma}-1}{1-\gamma}$) with constant aggregate endowment ($\forall t, \sum_{i\in\mathcal{I}} e_t^i = e$):

$$q_\gamma^d(\sigma^t) = \left(\sum_{i\in\mathcal{I}} p^i(\sigma^t)^{\frac{1}{\gamma}} w_0(i)\right)^\gamma \ ;\ q_\gamma^d(\sigma|\mathcal{F}_{t-1}) = \frac{\left(\sum_{i\in\mathcal{I}} p^i(\sigma^t)^{\frac{1}{\gamma}} w_0(i)\right)^\gamma}{\left(\sum_{i\in\mathcal{I}} p^i\left(\sigma^{t-1}\right)^{\frac{1}{\gamma}} w_0(i)\right)^\gamma} \tag{5}$$

Log economy: $\lim\gamma \to 1$

$$\hat{q}_\gamma(\sigma^t) = \sum_{i\in\mathcal{I}} p^i(\sigma^t) w_0(i)$$

$$q^d(\sigma|\mathcal{F}_{t-1}) = \sum_{i\in\mathcal{I}} \frac{p^i(\sigma^t|\mathcal{F}_{t-1})p^i\left(\sigma^{t-1}\right) w_0(i)}{\sum_{\mathcal{I}} p^i\left(\sigma^{t-1}\right) w_0(j)} = \sum_{i\in\mathcal{I}} p_t^i(\sigma|\mathcal{F}_{t-1}) w(i|\sigma^{t-1})$$

□

Lemma 3. *In a small CRRA economy with constant aggregate endowment, if all of the traders have identical discount factors, these two conditions are equivalent: i) : $\exists A : P(A) = 1$ and $\forall\sigma \in A$ $\forall i \in \mathcal{I}, \limsup \frac{p^i(\sigma^t)}{p^{ML}(\sigma^t)} = \liminf \frac{p^i(\sigma^t)}{p^{ML}(\sigma^t)}$*
ii) : $\forall\gamma_1, \gamma_2 \in [0, +\infty),\ \frac{q_{\gamma_1}(\sigma|\mathcal{F}_{t-1})}{q\gamma_2(\sigma|\mathcal{F}_{t-1})} \to 1$ P-a.s.

Proof. i) ⇒ ii)

By Lemma 2 (defining $p_0^i(\sigma) = (w_0(i))^\gamma$):

$$
\begin{aligned}
\frac{q_{\gamma_1}(\sigma|\mathcal{F}_{t-1})}{q_{\gamma_2}(\sigma|\mathcal{F}_{t-1})} &= \frac{\left(\sum_{i\in\mathcal{I}} p^i(\sigma^t)^{\frac{1}{\gamma_1}}\right)^{\gamma_1}}{\left(\sum_{i\in\mathcal{I}} p^i\left(\sigma^{t-1}\right)^{\frac{1}{\gamma_1}}\right)^{\gamma_1}} \frac{\left(\sum_{i\in\mathcal{I}} p^i\left(\sigma^{t-1}\right)^{\frac{1}{\gamma_2}}\right)^{\gamma_2}}{\left(\sum_{i\in\mathcal{I}} p^i(\sigma^t)^{\frac{1}{\gamma_2}}\right)^{\gamma_2}} \\
&=^a \frac{\left(1+\sum_{i\neq ML}\left(\frac{p^i(\sigma^t)}{p^{ML}(\sigma^t)}\right)^{\frac{1}{\gamma_1}}\right)^{\gamma_1}}{\left(1+\sum_{i\neq ML}\left(\frac{p^i(\sigma^{t-1})}{p^{ML}(\sigma^{t-1})}\right)^{\frac{1}{\gamma_1}}\right)^{\gamma_1}} \frac{\left(1+\sum_{i\neq ML}\left(\frac{p^i\left(\sigma^{t-1}\right)}{p^{ML}(\sigma^{t-1})}\right)^{\frac{1}{\gamma_2}}\right)^{\gamma_2}}{\left(1+\sum_{i\neq ML}\left(\frac{p^i(\sigma^t)}{p^{ML}(\sigma^t)}\right)^{\frac{1}{\gamma_2}}\right)^{\gamma_2}} \\
&\rightarrow^b \frac{(1+K)^{\gamma_1}}{(1+K)^{\gamma_1}} \frac{(1+K)^{\gamma_2}}{(1+K)^{\gamma_2}} = 1
\end{aligned}
$$

a : Dividing and multiplying by $\frac{p^{ML}(\sigma^t)}{p^{ML}(\sigma^{t-1})}$.

b : $i)$ implies $\forall\sigma\in A, \forall i\in\mathcal{I}, \left(\frac{p^i(\sigma^t)}{p^{ML}(\sigma^t)}\right)^{\gamma} \rightarrow \left(\frac{p^i\left(\sigma^{t-1}\right)}{p^{ML}_{t-1}\sigma}\right)^{\gamma}$

$ii) \Rightarrow i)$: By contradiction: $\neg i) \Rightarrow \left(\frac{p^i(\sigma^t)}{p^{ML}(\sigma^t)}\right)^{\gamma} \not\rightarrow \left(\frac{p^i\left(\sigma^{t-1}\right)}{p^{ML}_{t-1}\sigma}\right)^{\gamma}$ hence the last implication (b) does not hold.

□

Lemma 4. *In a small economy with constant aggregate endowment that satisfies A1, if all of the traders have identical discount factors, these two conditions are equivalent:*

$i)$: $\exists A : P(A) = 1$ *and* $\forall\sigma\in A\ \forall i\in\mathcal{I}, \limsup \frac{p^i(\sigma^t)}{p^{ML}(\sigma^t)} = \liminf \frac{p^i(\sigma^t)}{p^{ML}(\sigma^t)}$

$$
ii) : \begin{cases} \frac{q^d_{gen}(\sigma|\mathcal{F}_{t-1})}{q^d_{\log}(\sigma|\mathcal{F}_{t-1})} \rightarrow 1 \ P\text{-}a.s. \\ \frac{q^d_{\log}(\sigma|\mathcal{F}_{t-1})}{q^d_{gen}(\sigma|\mathcal{F}_{t-1})} \rightarrow 1 \ P\text{-}a.s. \end{cases}
$$

Proof. The result follows from Lemma 3, The convergence rate of a general economy can be sandwiched between the linear case $\lim\gamma\rightarrow 0$ and an arbitrarily large γ. □

Proof of Proposition 2

Proof. For small economies, integrals have to be understood in the Lebesgue sense.

$i)$ By Equation 5, $q^d_\gamma(\sigma|\mathcal{F}_{t-1}) = \frac{\left(\int_{\mathcal{I}} p^i(\sigma^t)^{\frac{1}{\gamma}} w_0(i)d\theta(i)\right)^\gamma}{\left(\int_{\mathcal{I}} p^j(\sigma^{t-1})^{\frac{1}{\gamma}} w_0(j)d\theta(j)\right)^\gamma}$. For $\gamma \leq 1$, the result is an implication of the Minkowsky 's inequality for integrals (Folland (1999) pg. 194): let $r = \frac{1}{\gamma} \geq 1$, then,

$$\begin{aligned}
\sum_{s\in\mathcal{S}}\left(\int_{\mathcal{I}} p^i(s|\mathcal{F}_{t-1})^r \frac{\left(p^i\left(\sigma^{t-1}\right)(w_0(i))^{\frac{1}{r}}\right)^r}{\int_{\mathcal{I}}\left(p^j\left(\sigma^{t-1}\right)(w_0(j))^{\frac{1}{r}}\right)^r d\theta(j)} d\theta(i)\right)^{\frac{1}{r}} &\geq \left(\int_{\mathcal{I}}\left(\sum_{s\in\mathcal{S}} p^i(s|\mathcal{F}_{t-1}) \frac{p^i\left(\sigma^{t-1}\right)(w_0(i))^{\frac{1}{r}}}{\left(\int_{\mathcal{I}}\left(p^j\left(\sigma^{t-1}\right)(w_0(j))^{\frac{1}{r}}\right)^r d\theta(j)\right)^{\frac{1}{r}}}\right)^r d\theta(i)\right)^{\frac{1}{r}} \\
&= \left(\int_{\mathcal{I}} \frac{\left(p^i\left(\sigma^{t-1}\right)(w_0(i))^{\frac{1}{r}}\right)^r}{\int_{\mathcal{I}}\left(p^j\left(\sigma^{t-1}\right)(w_0(j))^{\frac{1}{r}}\right)^r d\theta(j)} d\theta(i)\right)^{\frac{1}{r}} = 1
\end{aligned}$$

For $\gamma > 1$ we have that $r < 1$ hence the quantities of interests are not norms (fail to satisfy the triangular inequality). The result holds as we can use a "reversed" Minkowsky's inequality for integrals. The inequality is obtained by plugging the "reversed" Hölder's Inequality in the proof of Minkowski's inequality for integrals at page 194 of Folland (1999).

$$\begin{aligned}
\sum_{s\in\mathcal{S}}\left(\int_{\mathcal{I}} p^i(s|\mathcal{F}_{t-1})^r \frac{\left(p^i\left(\sigma^{t-1}\right)(w_0(i))^{\frac{1}{r}}\right)^r}{\int_{\mathcal{I}}\left(p^j\left(\sigma^{t-1}\right)(w_0(j))^{\frac{1}{r}}\right)^r d\theta(j)} d\theta(i)\right)^{\frac{1}{r}} &\leq \left(\int_{\mathcal{I}}\left(\sum_{s\in\mathcal{S}} p^i(s|\mathcal{F}_{t-1}) \frac{p^i\left(\sigma^{t-1}\right)(w_0(i))^{\frac{1}{r}}}{\left(\int_{\mathcal{I}}\left(p^j\left(\sigma^{t-1}\right)(w_0(j))^{\frac{1}{r}}\right)^r d\theta(j)\right)^{\frac{1}{r}}}\right)^r d\theta(i)\right)^{\frac{1}{r}} \\
&= \left(\int_{\mathcal{I}} \frac{\left(p^i\left(\sigma^{t-1}\right)(w_0(i))^{\frac{1}{r}}\right)^r}{\int_{\mathcal{I}}\left(p^j\left(\sigma^{t-1}\right)(w_0(j))^{\frac{1}{r}}\right)^r d\theta(j)} d\theta(i)\right)^{\frac{1}{r}} = 1
\end{aligned}$$

$ii)$ The result follows from proposition 6.12 pg 186, Folland (1999) that states:

If $P(X) < \infty$ and $0 < p < q \leq \infty$, then $L^p(P) \subset L^q(P)$ and $||f||_p \leq ||f||_q P(X)^{\frac{1}{p}-\frac{1}{q}}$.

The result apply to this setting as follows:

By Equation 5, $P(X) = \int_{\mathcal{I}} w^i_0 = 1$ and $q^d_\gamma(\sigma^t) = \left(\int_{\mathcal{I}}\left(p^i(\sigma^t)\right)^{\frac{1}{\gamma}} w_0(i)\right)^\gamma = ||p||_{\frac{1}{\gamma}}$.

Therefore for $0 \leq \gamma' < \gamma'' < \infty$, $q^d_{\gamma''} = ||p||_{\frac{1}{\gamma''}} \leq ||p||_{\frac{1}{\gamma'}} = q^d_{\gamma'}$.

$iii)$ For large economies that satisfy A4, the result is trivially true as no trader ever has positive wealth share. For small economies, define $\forall i \in \mathcal{I}, p^i_0(\sigma) = (w_0(i))^\gamma$. Using the FOC and Equation 5 we find that in a CRRA economy with parameter γ, the wealth share of trader i in the partial history σ^t is given by $w(i|\sigma^t) = \frac{p^i(\sigma^t)^{\frac{1}{\gamma}}}{\sum_{j\in\mathcal{I}} p^j(\sigma^t)^{\frac{1}{\gamma}}}$. Therefore, the wealth share of the trader with beliefs with maximal likelihood (p^{ML}) is given by: $w(ML|\sigma^t) = \frac{p^{ML}(\sigma^t)^{\frac{1}{\gamma}}}{\sum_{j\in\mathcal{I}} p^j(\sigma^t)^{\frac{1}{\gamma}}} = \frac{1}{1+\sum_{j\neq ML}\left(\frac{p^j(\sigma^t)}{p^{ML}(\sigma^t)}\right)^{\frac{1}{\gamma}}}$ and the result follows because

$$\frac{\partial w(ML|\sigma^t)}{\partial \gamma} = -\frac{\frac{1}{\gamma^2}\left(-\sum_{j\neq ML}\left(\frac{p^j(\sigma^t)}{p^{ML}(\sigma^t)}\right)^{\frac{1}{\gamma}}\ln\left(\frac{p^j(\sigma^t)}{p^{ML}(\sigma^t)}\right)\right)}{1+\sum_{j\neq ML}\left(\frac{p^j(\sigma^t)}{p^{ML}(\sigma^t)}\right)^{\frac{1}{\gamma}}} \leq^a 0$$

$a)$: $\forall j \neq ML, \left(\frac{p^j(\sigma^t)}{p^{ML}(\sigma^t)}\right) < 1$ by construction with equality if we have multiple models with maximal likelihood.

□

Proof of Proposition 5:

Proof. By the FOCs: $\frac{u_i'(c_0)}{u_i'(c(\sigma^{t,s}))} = \frac{\beta_i^t p^i(\sigma^{t,s})}{q(\sigma^{t,s})}$. Therefore, $\forall \sigma^{t,s} \in \Sigma$

$$\frac{\beta_i^t p^i(\sigma^{t,s})}{q(\sigma^{t,s})} \to 0 \Leftrightarrow \frac{u_i'(c_0)}{u_i'(c^i(\sigma^{t,s}))} \to 0 \Leftrightarrow^a u_i'(c^i(\sigma^{t,s})) \to +\infty \Leftrightarrow^b c^i \to 0$$

a : By A2: $c_0^i < +\infty \Rightarrow^{by\ A1} u_i'(c_0) > 0$.

b : By A1 $\lim_{c\to 0} u_i'(c) = +\infty$ and the payoff functions are strictly concave, increasing. □

Proof of Proposition 6

Proof. From the equilibrium condition $\forall \sigma^t, \log c_t^i(\sigma) = \log \frac{\beta^t p^i(\sigma^t)}{q(\sigma^t)}$. Therefore $E\left(\log c_t^i(\sigma)\right) = E\left(\log \frac{\beta^t p^i(\sigma^t)}{q(\sigma^t)}\right)$. The result follows because convergence in expectation implies convergence in probability. □

Proof of Theorem 4:

Proof. In an economy that satisfies A1-A3 Note that:

$$E\left(\limsup \frac{\beta_i^t p_t^i(\sigma)\, w_0(i)}{\int_{\mathcal{I}} \beta_j^t p_t^j(\sigma)\, w_0(j) d\theta(j)}\right) = 0 \Leftrightarrow \forall \epsilon > 0, E\left(I_{\limsup \frac{\beta_i^t p_t^i(\sigma) w_0(i)}{\int_{\mathcal{I}} \beta_j^t p_t^j(\sigma) w_0(j) d\theta(j)} > \epsilon}\right) = 0$$

$$\Leftrightarrow \forall \epsilon > 0, P\left\{\limsup \frac{\beta_i^t p_t^i(\sigma)\, w_0(i)}{\int_{\mathcal{I}} \beta_j^t p_t^j(\sigma)\, w_0(j) d\theta(j)} > \epsilon\right\} = 0$$

$$\Leftrightarrow^{\text{by Prop. 5}} \forall \epsilon_1 > 0, P\left\{\limsup c_t^i > \epsilon_1\right\} = 0$$

$$\Leftrightarrow \forall \epsilon_1 > 0, P\left\{c_t^i > \epsilon_1 \ i.o.\right\} = 0$$

$$\Leftrightarrow^a c_t^i \to 0 \ P\text{-a.s.}$$

$$\Leftrightarrow \text{Trader } i \text{ vanishes } p\text{-a.s.}$$

a: the <= implication is obvious, the => implication is an application of Borel-Cantelli Lemma (see exercise EA13.1 C in Williamson)

□

Proof of Corollary 2 and Proposition 8:

Proof. Let ν be the probability that the Bayesian attaches to the true model. Then, $\forall \sigma \in \Sigma, \forall t, p^B(\sigma^t) \geq \nu P(\sigma^t)$. Moreover, since P is the data generating proces, $\exists A \in \Sigma : P(A) = 1$ and $\forall \sigma \in A, \forall p$ on $\Sigma, \lim_{t\to\infty} \frac{p^i(\sigma^t)}{P(\sigma^t)} < \infty$.

Small economies:

By Theorem 2, $q(\sigma^t) \doteq \sum_{i\in\mathcal{I}} \beta_i^t p^i(\sigma^t) w_0(i)$. Therefore $\forall \sigma \in A, \forall t, \exists b < \infty :$

$$\frac{p^B(\sigma^t)}{q(\sigma^t)} \geq \frac{\frac{w_0(B)\nu P(\sigma^t)}{P(\sigma^t)}}{b\frac{\sum_{i\in\mathcal{I}} p^i(\sigma^t)}{P(\sigma^t)}} = \frac{w_0(B)v}{b\sum_i \frac{p^i(\sigma^t)}{P(\sigma^t)}} > 0.$$

The result follows because by Proposition 5, a trader survive if $\lim_{t\to\infty} \frac{p^B(\sigma^t) w_0(B)}{q(\sigma^t)} > 0$ infinitely often.

Large economies:

The FOC implies that $\forall \sigma \in A$:

$$c^B(\sigma^t) = \frac{p^B(\sigma^t)^{\frac{1}{\gamma}} w_0(B)}{w_0(B)p^B(\sigma^t)^{\frac{1}{\gamma}} + \int_{\mathcal{I}\setminus\mathcal{B}} p^i(\sigma^t)^{\frac{1}{\gamma}} w_0(i)d\theta(i)} \geq \frac{w_0(B)\nu^{\frac{1}{\gamma}}}{\frac{p^B(\sigma^t)^{\frac{1}{\gamma}} w_0(B) + \int_{\mathcal{I}\setminus\mathcal{B}} p^i(\sigma^t)^{\frac{1}{\gamma}} w_0(i)d\theta(i)}{P(\sigma^t)^{\frac{1}{\gamma}}}}$$

$$\mathrel{\overset{\circ}{=}} \frac{1}{1 + \left(\frac{\left(\int_{\mathcal{I}\setminus\mathcal{B}} p^i(\sigma^t)^{\frac{1}{\gamma}} w_0(i)d\theta(i) \right)^{\gamma}}{P(\sigma^t)} \right)^{\frac{1}{\gamma}}} \mathrel{\overset{\circ}{=}}^{By Cor.4} \frac{1}{1 + \left(\frac{P(\sigma^t)}{t^{\frac{\gamma k}{2}} P(\sigma^t)} \right)^{\frac{1}{\gamma}}} \quad P\text{-a.s.} > 0$$

and the result follows from Proposition 5.

$\square$

Proof of Proposition 9:

Proof. Let $g(i)$ be the smooth prior of the Bayesian trader and remember that $\theta(i)$ has an atom at B and $w_0(i)$ is smooth for $i \in \mathcal{I}_c$. The FOCs imply:

$$c^B(\sigma^t) = w_0(B) \left(\frac{p^B(\sigma^t)}{q(\sigma^t)} \right)^{\frac{1}{\gamma}}$$

$$= \frac{w_0(B) \left(\int_{i\in\mathcal{I}_c} p^i(\sigma^t)g(i)d\theta(i) \right)^{\frac{1}{\gamma}}}{w_0(B) \left(\int_{i\in\mathcal{I}_c} p^i(\sigma^t)g(i)d\theta(i) \right)^{\frac{1}{\gamma}} + \int_{i\in\mathcal{I}_c} p^i(\sigma^t)^{\frac{1}{\gamma}} w_0(i)d\theta(i)}$$

$$\mathrel{\overset{\circ}{=}}^{\text{by Cor 4}} \frac{e^{\frac{1}{\gamma}\left(\ln P(\sigma^t) - \frac{k}{2}\ln t\right)}}{e^{\frac{1}{\gamma}\left(\ln P(\sigma^t) - \frac{k}{2}\ln t\right)} + e^{\frac{1}{\gamma}\left(\ln P(\sigma^t) - \frac{\gamma k}{2}\ln t\right)}} \quad P\text{-a.s.}$$

Note that for $\gamma < 1$ the RHS converges to 0, for $\gamma = 1$ the RHS converges to a constant $\in (0, 1)$ and for $\gamma > 1$ the RHS converges to 1. Therefore the result follows from Proposition 5. $\square$

C Appendix

In this appendix I make use of the notation O(.) and o(.) with the following meanings. The big-O notation, $f(x) = O(g(x))$, means $\limsup_{x\to\infty} \frac{|f(x)|}{|g(x)|} < \infty$. The little-o notation, $f(x) = o(g(x))$, abbreviates $\lim_{x\to\infty} \frac{f(x)}{g(x)} \to 0$.

Using Lemma 2 and letting the number of traders go to infinity we obtain the following equation for the normalized equilibrium prices in a CRRA economy with a continuum of traders:

$$q^n_\gamma(\sigma^t) = \frac{\left(\int_{\mathcal{I}} p^i(\sigma^t)^{\frac{1}{\gamma}} w(i) d\theta(i)\right)^\gamma}{\sum_{\sigma^t} \left(\int_{\mathcal{I}} p^j(\sigma^t)^{\frac{1}{\gamma}} w(j) d\theta(j)\right)^\gamma}$$

The proof of Theorem 3 relies on a modification of standard proofs on the asymptotic convergence rate of Bayesian posterior (Clarke-Barron (1990), Phillips-Ploberger (2003)) ($\gamma = 1$)with smooth positive Lebesgue prior together with the proof of the asymptotic minimum regret property of the Normalized Maximum Likelihood distribution ($\gamma = 0$) in Grünwald (2007).

Theorem 6. *For all* $\gamma \in [0, +\infty)$, *if* $\int_{\mathcal{I}} det I(p^i)^{-\frac{1+\gamma}{2}} < \infty$[7] *and the market prior is smooth:*

$$\left|\ln q^n_\gamma(\sigma^t) - \ln q^n_1(\sigma^t)\right| = \left| e^{\ln p^{\hat{i}(\sigma^t)}(\sigma^t) - \frac{k}{2}\ln\frac{t}{2\pi} + O(1)} - e^{\ln p^{\hat{i}(\sigma^t)}(\sigma^t) - \frac{k}{2}\ln\frac{t}{2\pi} + O(1)} \right| = O(1) \quad P\text{-}a.s.$$

Proof. I focus on the case $k = 1$, which is to say on the Bernoulli family: $S = \{a, b\}$. The generalization to the multinomial family is straightforward. Let $I(p^i)$ be the Fisher information, $\hat{i}(\sigma^t)$ be the model with Maximum likelihood on σ^t, which is to say the model such that $p^{\hat{i}(\sigma^t)}(a) = \frac{\sum_{\nu=1}^t I_{\sigma_n = a}}{t}$.

The priors are smooth, therefore satisfy the standard regularity conditions which make the error in the third and higher order Taylor expansion around $\hat{i}$ be $o(1)$.

[7]This assumption comport no loss of generality because it is always true for the multinomial model with probabilities strictly bounded away from the boundaries of the simplex.

$$
\begin{aligned}
q^n_\gamma(\sigma^t) &= \frac{\left(\int_{\mathcal{I}} p^i(\sigma^t)^{\frac{1}{\gamma}} w(i)d\theta(i)\right)^\gamma}{\sum_{\sigma^t}\left(\int_{\mathcal{I}} p^i(\sigma^t)^{\frac{1}{\gamma}} w(i)d\theta(i)\right)^\gamma} \\
&\approx^{a,b} \frac{e^{\ln p^{\hat{i}(\sigma^t)}(\sigma^t)+\gamma\ln\sqrt{\gamma}+\gamma\ln w(\hat{i})-\gamma\frac{1}{2}\ln\frac{t}{2\pi}-\gamma\ln\sqrt{\det I(p^{\hat{i}})}+o(1)}}{\sum_{\tau=0}^t\binom{t}{\tau}\left(\int_{\mathcal{I}} p^i(\sigma^t_\tau)^{\frac{1}{\gamma}} w(i)d\theta(i)\right)^\gamma} \\
&\approx^{c} \frac{e^{\ln p^{\hat{i}(\sigma^t)}(\sigma^t)+\gamma\ln\sqrt{\gamma}+\gamma\ln w(\hat{i})-\gamma\frac{1}{2}\ln\frac{t}{2\pi}-\gamma\ln\sqrt{\det I(p^{\hat{i}})}+o(1)}}{\sum_{\tau=0}^t e^{\left(-\ln p^{\hat{i}^\tau_t}(\sigma^t_\tau)-\frac{1}{2}\ln 2\pi t-\ln\sqrt{\det I(\hat{i}^\tau_t)}+o(1)\right)} e^{\ln p^{\hat{i}^\tau_t}(\sigma^t_\tau)+\gamma\ln\sqrt{\gamma}+\gamma\ln w(\hat{i}^\tau_t)-\gamma\frac{1}{2}\ln\frac{t}{2\pi}-\gamma\ln\sqrt{\det I(\hat{i}^\tau_t)}+o(1)}} \\
&\approx \frac{e^{\ln p^{\hat{i}(\sigma^t)}(\sigma^t)+\gamma\ln w(\hat{i})-\gamma\frac{1}{2}\ln\frac{t}{2\pi}-\gamma\ln\sqrt{\det I(p^{\hat{i}})}+o(1)}}{\sum_{\hat{i}^\tau_t=\{0,\frac{1}{t},\dots,1\}} e^{\gamma\ln w(\hat{i}^\tau_t)-(1+\gamma)\ln\sqrt{t\det I(\hat{i}^\tau_t)}+\frac{1-\gamma}{2}\ln 2\pi+o(1)}} \\
&\approx \frac{e^{\ln p^{\hat{i}(\sigma^t)}(\sigma^t)+\gamma\ln w(\hat{i})-\gamma\frac{1}{2}\ln\frac{t}{2\pi}-\gamma\ln\sqrt{\det I(p^{\hat{i}})}+o(1)}}{\frac{t}{\sqrt{(2\pi)^{1-\gamma}t^{1+\gamma}}}\int_0^1 det I(\hat{i}^\tau_t)^{-\frac{1+\gamma}{2}}\ w(\hat{i}^\tau_t)d\hat{i}^\tau_t} \\
&= e^{\ln p^{\hat{i}(\sigma^t)}(\sigma^t)-\frac{1}{2}\ln\frac{t}{2\pi}+O(1)}
\end{aligned}
$$

Which prove the claim as $e^{\ln p^{\hat{i}(\sigma^t)}(\sigma^t)-\frac{1}{2}\ln\frac{t}{2\pi}+O(1)}$ does not depend on γ.

a : numerator: by Lemma 6

b : With σ^t_τ denoting the sequences of length t in which Maximum likelihood estimator for the probability of state a is given by $p^{\hat{i}^\tau_t}(a)=\frac{\tau}{t}$ (which is to say, such that $\sum_{\nu=1}^t I_{\sigma_\nu=a}=\tau$).

c : Using Stirling's approximation of the factorial and Lemma 6

$\square$

Lemma 5. *Let $\mathcal{M}$ be an exponential family of distributions parametrized by $\mathcal{I}$. Let σ^t be any sequence with $p^{\hat{i}(\sigma)}\in\mathcal{I}$ and let w be an arbitrary prior distribution. Then:*

$$\gamma\ln\int_{\mathcal{I}} p^i(\sigma^t)^{\frac{1}{\gamma}}w(i)d\theta(i)=\gamma\ln\int_{\mathcal{I}} e^{-\frac{t}{\gamma}D(p^{\hat{i}(\sigma^1)}||p^i)}w(i)d\theta(i)+\ln p^{\hat{i}(\sigma^t)}(\sigma^t)$$

Proof.

$$\begin{aligned}
\gamma \ln \int_{\mathcal{I}} p^i(\sigma^t)^{\frac{1}{\gamma}} w(i) d\theta(i) &= \gamma \ln \int_{\mathcal{I}} p^i(\sigma^t)^{\frac{1}{\gamma}} w(i) d\theta(i) + \gamma \ln p^{\hat{i}(\sigma^t)}(\sigma^t)^{\frac{1}{\gamma}} - \gamma \ln p^{\hat{i}(\sigma^t)}(\sigma^t)^{\frac{1}{\gamma}} \\
&= \gamma \ln \int_{\mathcal{I}} \frac{p^i(\sigma^t)}{p^{\hat{i}(\sigma^t)}(\sigma^t)}^{\frac{1}{\gamma}} w(i) d\theta(i) + \gamma \ln p^{\hat{i}(\sigma^t)}(\sigma^t)^{\frac{1}{\gamma}} \\
&= \gamma \ln \int_{\mathcal{I}} e^{\frac{1}{\gamma}\left(\ln p^i(\sigma^t) - \ln p^{\hat{i}(\sigma^t)}(\sigma^t)\right)} w(i) d\theta(i) + \gamma \ln p^{\hat{i}(\sigma^t)}(\sigma^t)^{\frac{1}{\gamma}} \\
&=^a \gamma \ln \int_{\mathcal{I}} e^{-\frac{t}{\gamma} D(p^{\hat{i}(\sigma^1)} || p^i)} w(i) d\theta(i) + \ln p^{\hat{i}(\sigma^t)}(\sigma^t)
\end{aligned}$$

a: as $\mathcal{M}$ is an exponential family □

Lemma 6. *Let $\mathcal{M}$ be an exponential family parametrized by $\mathcal{I}$ (in the canonical parametrization of distributions). Let $\mathcal{I}_0$ be non empty compact set in the interior of $\mathcal{I}$. Let σ^t be any sequence with $p^{\hat{i}(\sigma)} \in \mathcal{I}_0$ and let w be a prior distribution on $\mathcal{I}$ which is continuous and strictly positive on $\mathcal{I}_0$. Then*

$$\gamma \ln \int_{\mathcal{I}} p^i(\sigma^t)^{\frac{1}{\gamma}} w(i) d\theta(i) = \ln p^{\hat{i}(\sigma^t)}(\sigma^t) + \gamma \ln \sqrt{\gamma} + \gamma \ln w(\hat{i}) - \gamma \frac{1}{2} \ln \frac{t}{2\pi} - \gamma \ln \sqrt{\det I(p^{\hat{i}})} + o(1)$$

Proof. The proof uses standard arguments in Bayesian statistics, and is an adaptation (the only difference is that I have the γ term) of the proof give by Grüwald (pg 248). By Lemma 5

$$\gamma \ln \int_{\mathcal{I}} p^i(\sigma^t)^{\frac{1}{\gamma}} w(i) d\theta(i) = \gamma \ln \int_{\mathcal{I}} e^{-\frac{t}{\gamma} D(p^{\hat{i}(\sigma^1)} || p^i)} w(i) d\theta(i) + \ln p^{\hat{i}(\sigma^t)}(\sigma^t)$$

To proof the claim I will approximate the integral using a second order Taylor series expansion and proving that the expansion does not loose relevant information.

$$D(p^{\hat{i}(\sigma^1)} || p^i) \approx \frac{1}{2}\left(p^{\hat{i}(\sigma^1)}(a) - p^i(a)\right)^2 I(p^{\hat{i}}) + o\left(\left(p^{\hat{i}(\sigma^1)}(a) - p^i(a)\right)^2\right) \tag{6}$$

for $0 < \alpha < \frac{1}{2}$ let $B_t = [p^{\hat{i}(\sigma^1)}(a) - t^{-\frac{1}{2}+\alpha}, p^{\hat{i}(\sigma^1)}(a) + t^{-\frac{1}{2}+\alpha}]$.

Since $w(i)$ is continuous on $\mathcal{I}$ and strictly positive at $\hat{i}$ there is a T, such that $\forall t > T$ $w(i) > 0, \forall \theta \in B_t$. In what follows I always assume $t > T$.

Note that:

$$\int_{\mathcal{I}} e^{-\frac{t}{\gamma}D(p^{\hat{i}(\sigma^1)}||p^i)}w(i)d\theta(i) = \int_{i\in\mathcal{I}\setminus B_t} e^{-\frac{t}{\gamma}D(p^{\hat{i}(\sigma^1)}||p^i)}w(i)d\theta(i) + \int_{i\in B_t} e^{-\frac{t}{\gamma}D(p^{\hat{i}(\sigma^1)}||p^i)}w(i)d\theta(i)$$

I need to bound the two integrals:

First integral: $\mathcal{I}_1 = \int_{i\in\mathcal{I}\setminus B_t} e^{-\frac{t}{\gamma}D(p^{\hat{i}(\sigma^1)}||p^i)}w(i)d\theta(i)$

Remember that $D(p^{\hat{i}(\sigma^1)}||p^i)$ as a function of p^i is strictly convex and has a minimum at $p^i = p^{\hat{i}(\sigma^1)}$. $D(p^{\hat{i}(\sigma^1)}||p^i)$ is increasing in $|p^i(a) - p^{\hat{i}(\sigma^1)}(a)|$, so that:

$$0 < \int_{\theta\in\mathcal{I}\setminus B_t} e^{-\frac{t}{\gamma}\min_{\theta\in\mathcal{I}\setminus B_t} D(p^{\hat{i}(\sigma^1)}||p^i)}w(i)d\theta(i)$$

By Equation 6

$$\min_{\theta\in\mathcal{I}\setminus B_t} D(p^{\hat{i}(\sigma^1)}||p^i) \geq \frac{1}{2}t^{-1+2\alpha}I(p^{\hat{i}})$$

Therefore, since $I(p^i)$ is continuous and > 0 and also $\int_{\mathcal{I}\setminus B_t} w(i)d\theta(i) < \infty$.

$$0 < \int_{\theta\in\mathcal{I}\setminus B_t} e^{-\frac{t}{\gamma}D(p^{\hat{i}(\sigma^1)}||p^i)}w(i)d\theta(i) < \int_{\theta\in\mathcal{I}\setminus B_t} e^{-\left(n\frac{1}{2\gamma}t^{-1+2\alpha}I(p^{\hat{i}})\right)}w(i)d\theta(i) < ke^{-ct^{2\alpha}}$$

For some $c > 0, k < \infty$.

Second integral: $\mathcal{I}_2 = \int_{\theta\in B_t} e^{-\frac{t}{\gamma}D(p^{\hat{i}(\sigma^1)}||p^i)}w(i)d\theta(i)$

Let

$$I_t^- := \inf_{i'\in B_t} I(i'),\quad I_t^+ := \sup_{i'\in B_t} I(i'),\quad w_t^- := \inf_{i'\in B_t} w(i'),\quad w_t^+ := \sup_{i'\in B_t} w(i')$$

By Equation 6

$$\mathcal{I}_2 \approx \int_{\theta\in B_t} e^{-\frac{t}{\gamma}(p^{\hat{i}(\sigma^1)}(a)-p^i(a))^2 I(i')}w(i)d\theta(i)$$

Where i' depends on θ. using the definitions above, we get

$$w_t^- \int_{\theta\in B_t} e^{-\frac{t}{\gamma}(p^{\hat{i}(\sigma^1)}(a)-p^i(a))^2 I_t^+} d\theta(i) \leq \mathcal{I}_2 \leq w_t^+ \int_{\theta\in B_t} e^{-\frac{t}{\gamma}(p^{\hat{i}(\sigma^1)}(a)-p^i(a))^2 I_t^-} d\theta(i)$$

Perform the substitution $z := (p^{\hat{i}(\sigma^1)}(a) - p^i(a))\sqrt{t\frac{I_t^+}{\gamma}}$ on the left integral and $z := (p^{\hat{i}(\sigma^1)}(a) - p^i(a))\sqrt{t\frac{I_t^-}{\gamma}}$ in the right integral to get

$$\frac{w_t^-}{\sqrt{t\frac{I_t^+}{\gamma}}} \int_{|z|<t^\alpha\sqrt{I_t^-}} e^{-\frac{1}{2}z^2} d\theta(i) \leq \mathcal{I}_2 \leq \frac{w_t^+}{\sqrt{t\frac{I_t^-}{\gamma}}} \int_{|z|<t^\alpha\sqrt{I_t^-}} e^{-\frac{1}{2}z^2} d\theta(i)(i)$$

Both integral are standard Gaussian. Since, as $n \to \infty$ $I_t^- \to I(p^{\hat{i}})$ and $I_t^+ \to I(p^{\hat{i}})$ the domain of integration tend to infinity for both integral so that they both converge to $\sqrt{2\pi}$. Since $w_t^+ \to w(\hat{i})$ and $w_t^- \to w(\hat{i})$ the constant in both integral converges to $\frac{w(\hat{i})}{\sqrt{t\frac{I(p^{\hat{i}})}{\gamma}}}$ and we get

$$\mathcal{I}_2 \to \frac{\sqrt{2\pi}w(\hat{i})}{\sqrt{t\frac{I(p^{\hat{i}})}{\gamma}}}$$

Therefore

$$\mathcal{I}_1 + \mathcal{I}_2 \approx \frac{\sqrt{2\pi}w(\hat{i})}{\sqrt{t\frac{I(p^{\hat{i}})}{\gamma}}}$$

Hence

$$\gamma \ln \int_{\mathcal{I}} p^i(\sigma^t)^{\frac{1}{\gamma}} w(i) d\theta(i) = \gamma \ln(\mathcal{I}_1 + \mathcal{I}_2) + \ln p^{\hat{i}(\sigma^t)}(\sigma^t)$$

$$\to \ln p^{\hat{i}(\sigma^t)}(\sigma^t) + \gamma \ln\sqrt{\gamma} + \gamma \ln w(\hat{i}) - \gamma\frac{1}{2}\ln\frac{t}{2\pi} - \gamma \ln\sqrt{\det I(p^{\hat{i}})} + o(1)$$

Note that the convergence is uniform because w and I are continuous functions over a compact set: $\mathcal{I}_0$. □

Corollary 4. *Under the assumption of Lemma 6:*

$$q_\gamma^d(\sigma^t) \approx e^{\ln p^{\hat{i}(\sigma^t)}(\sigma^t) - \gamma\frac{k}{2}\ln t + O(1)} \mathrel{\overset{\circ}{=}} \frac{p^{\hat{i}(\sigma^t)}}{t^{\frac{\gamma k}{2}}}$$

Moreover, if $P \in \mathcal{I}$, then $p^{\hat{i}(\sigma^t)} \to^{P\text{-}a.s.} P$.

Proof. The first point follows from Lemma 6. The second from the fact that if the data generating process is P, then P-a.s. $p^{\hat{i}(\sigma^t)} \equiv P$. □

D Appendix

In this section I show that a s-dimensional random walk generated by the sum of normalized extractions from a multinomial with $s+1$ states, lies infinitely often in a s-dimensional neighborhood of convex cones, whose dimensionality is at least s-2 and that it does not lies infinitely often in the s dimensional neighborhood of any cone whose dimensionality is less then s-2. With this result, the proof in Blume-Easleys (2009) can be extended to traders whose beliefs lies on the boundary of $C(lo(p^j))_{j \in SUR}$.

D.1 The proof

My strategy is to use the multinomial process that generates the Random walk to obtain the unconditional probabilities that the Random walk is in the desired set in each period. Then I use the first Borel-Cantelli lemma and a generalization of the Second Borel-Cantelli Lemma to identify the sets that are reached infinitely often.

First Borel-Cantelli Lemma 1. *Suppose E_1, E_2, . . . are events with*

$$\sum_{n=1}^{\infty} P(E_n) < \infty$$

Then $P\{E_n i.o.\} = 0$

Kochen-Stone Lemma 2. *Suppose E_1, E_2, . . . are events with*

$$\sum_{n=1}^{\infty} P(E_n) = \infty \text{ and } \liminf_{t \to \infty} \frac{\sum_{m=1}^{t} \sum_{n=1}^{t} P(E_m \cap E_n)}{\left(\sum_{n=1}^{t} P(E_n)\right)^2} < \infty$$

Then $P\{E_n i.o.\} = 1$

Proof. For references see pg. 74 of "Brownian motion" Peter Mörters and Yuval Peres ☐

I make use of the notation O(.), o(.) and $\circeq$ with the following meanings. The big-O notation, $f(x) = O(g(x))$, means $\limsup_{x\to\infty} \frac{|f(x)|}{|g(x)|} < \infty$. The little-o notation, $f(x) = o(g(x))$, abbreviates $\lim_{x\to\infty} \frac{f(x)}{g(x)} \to 0$.

Theorem 7. *Let B_t be a s-dimensiona random walk of the form:*
$\forall t B_t = [\sum_{\tau=1}^{t} I_{x=1} - P_1, ..., \sum_{\tau=1}^{t} I_{x=s+1} - P_{s+1}]'$ *with $X \approx f(x_1, ...x_{s+1}, 1)$ a multinomial with cell probabilities: $[P_1, ..., P_{s+1}]$.*
Let K^{s-r} a convex cones in $s-r$ dimensions and C^s a s-dimensional compact set: $[-1,1]^s \subset C^s$.
Then:
i)$r < 3 \Rightarrow P\{B_t \in K^{s-r} + C^s \text{ i.o.}\} = 1$
ii)$r \geq 3 \Rightarrow P\{B_t \in K^{s-r} + C^s \text{ i.o.}\} = 0$

Proof. The cases $r = 0$ and $r = 1$ are implied by the case $r = 2$. Therefore I will focus WLOG on the cases $r = 2$ and $r > 2$.
Let $E_n = \{B_t \in K^{s-r} + C^s\}$.

$r = 2$. From Lemma 8, $P(E_\tau) \circeq \frac{1}{\tau}$:

$$\lim_{t\to\infty} \sum_{\tau=1}^{t} P(E_\tau) = \lim_{t\to\infty} \sum_{\tau=1}^{t} \frac{1}{\tau} = \lim_{t\to\infty} \log t = \infty$$

To prove the result I need to show that the conditions for Kochen-Stone's lemma are satisfied:

$$\liminf_{t\to\infty} \frac{\sum_{m=1}^{t} \sum_{n=1}^{t} P(E_m \cap E_n)}{\left(\sum_{n=1}^{t} P(E_n)\right)^2} < \infty$$

Numerator:

$$\begin{aligned}
\sum_{m=1}^{t}\sum_{n=1}^{t}P(E_m\cap E_n) &= \sum_{m=n}^{t}P(E_m)+2\sum_{m=1}^{t-1}\sum_{n>m}^{t}P(E_m)P(E_{n-m}) \\
&\leq^{By\ Lem\ 9}\sum_{m=n}^{t}P(E_m)+2A\sum_{m=1}^{t-1}\sum_{n>m}^{t}P(E_m)P(E_{n-m}) \\
&\mathrel{\mathring{=}}^{By\ Lem\ 8}\sum_{\tau=1}^{t}\frac{1}{\tau}+2A\sum_{m=1}^{t-1}\sum_{n>m}^{t}\frac{1}{m}\frac{1}{n-m} \\
&\mathrel{\mathring{=}}^{a}\int_1^t\frac{1}{\tau}d\tau+2A\int_1^{t-1}\frac{1}{m}\int_{n=m+1}^{t}\frac{1}{n-m}dndm \\
&\mathrel{\mathring{=}}\log t+2A\int_1^{t-1}\frac{1}{m}\log{(n-m)}\mid_{m+1}^{t}dm \\
&=\log t+2A\int_1^{t-1}\frac{1}{m}\log{(t-m)}\,dm \\
&\leq\log t+2A\int_1^{t-1}\frac{1}{m}\log{(t)}\,dm=\log t+2A\,(\log t)^2
\end{aligned}$$

$a)$: I can approximate the sum with an integral because the elements are positive.

Denominator:

$$\left(\sum_{\tau=1}^{t}P(E_\tau)\right)^2\mathrel{\mathring{=}}\left(\sum_{\tau=1}^{t}\frac{1}{\tau}\right)^2\mathrel{\mathring{=}}(\log t)^2$$

Therefore

$$\liminf_{t\to\infty}\frac{\sum_{m=1}^{t}\sum_{n=1}^{t}P(E_m\cap E_n)}{\left(\sum_{n=1}^{\infty}P(E_n)\right)^2}\leq\liminf_{t\to\infty}\frac{(\log t)^2}{(\log t)^2}<\infty$$

and by Kochen-Stone Lemma $P\left\{E_n\ i.o.\right\}=1$.

$r\geq 3$. From Lemma 8, $\exists\eta>0:$

$$\lim_{t\to\infty}\sum_{\tau=1}^{t}P\left(E_\tau\right)=\lim_{t\to\infty}\sum_{\tau=1}^{t}\left(\frac{1}{\sqrt{\tau}}\right)^r>\lim_{t\to\infty}\eta\sum_{\tau=1}^{t}\left(\frac{1}{\sqrt{\tau}}\right)^3<\infty$$

Therefore by the first Borel-Cantelli Lemma $P\left\{E_n\ i.o.\right\}=0$.

$\square$

Lemma 7. *Let $\boldsymbol{k}_t$ follows a multinomial with probabilities $\boldsymbol{m} = [P_1, ..., P_{s+1}]$ and t extractions.*

$$P\left(\boldsymbol{k}_t = \lceil t\boldsymbol{m} + \boldsymbol{i} \rceil\right) = \binom{t}{k_1, ..., k_{s+1}} \prod_{i=1}^{s+1} P_i^{k_i} \triangleq \left(\frac{1}{\sqrt{t}}\right)^s e^{-\sum_{j=1}^{s+1} \frac{(i_j^2)}{2P_j t}} \quad \textit{for } i < \sqrt{t \log t}$$

Proof.

$$\binom{t}{k_1, ..., k_{s+1}} \prod_{j=1}^{s+1} P_j^{k_j} = \frac{t!}{\prod_{j=1}^{s+1} k_j!} \prod_{j=1}^{s+1} P_j^{k_j} = e^{\ln t! + \sum_{j=1}^{s+1} k_j \ln P_j - \sum_{j=1}^{s+1} \ln k_j!}$$

I focus on the exponent:

$$
\begin{aligned}
&\ln t! + \sum_{j=1}^{s+1} k_j \ln P_j - \sum_{j=1}^{s+1} \ln k_j! \\
&=^a -t + \left(t + \frac{1}{2}\right) \ln t + \sum_{j=1}^{s+1} k_j \ln P_j - \left(\sum_{j=1}^{s+1} \left(-k_j + \left(k_j + \frac{1}{2}\right) \ln k_j \right) \right) + O(1) \\
&=^b -t + \left(t + \frac{1}{2}\right) \ln t + \sum_{j=1}^{s+1} k_j \ln P_j - \left(-t + \sum_{j=1}^{s+1} \left(k_j + \frac{1}{2}\right) \ln k_j \right) + O(1) \\
&= t \ln t + \frac{1}{2} \ln t + \sum_{j=1}^{s+1} k_j \ln \frac{P_j}{k_j} - \sum_{j=1}^{s+1} \frac{1}{2} \ln k_j + O(1) \\
&= t \ln t + \frac{1}{2} \ln t + \sum_{j=1}^{s+1} k_j \ln \frac{tP_j}{k_j} - \sum_{j=1}^{s+1} k_j \ln t - \sum_{j=1}^{s+1} \frac{1}{2} \ln k_j + O(1) \\
&=^b t \ln t + \frac{1}{2} \ln t + \sum_{j=1}^{s+1} k_j \ln \frac{tP_j}{k_j} - t \ln t - \sum_{j=1}^{s+1} \frac{1}{2} \ln k_j + O(1) \\
&= \frac{1}{2} \ln t + \sum_{j=1}^{s+1} k_j \ln \frac{tP_j}{k_j} - \sum_{j=1}^{s+1} \frac{1}{2} \ln k_j + O(1) \\
&= \frac{1}{2} \ln t - \sum_{j=1}^{s+1} k_j \ln \frac{k_j}{tP_j} - \sum_{j=1}^{s+1} \frac{1}{2} \ln \frac{tk}{t} + O(1) \\
&= \frac{1}{2} \ln t - \sum_{j=1}^{s+1} k_j \ln \frac{k_j}{tP_j} - \sum_{j=1}^{s+1} \frac{1}{2} \ln t - \sum_{j=1}^{s+1} \frac{1}{2} \ln \left(P_j + \frac{i_j}{t}\right) + O(1) \\
&= -\frac{s}{2} \ln t - \sum_{j=1}^{s+1} (tP_j + i_j) \ln \left(1 + \frac{i_j}{tP_j}\right) - \sum_{j=1}^{s+1} \frac{1}{2} \ln \left(P_j + \frac{i_j}{t}\right) + O(1) \\
&=^c -\frac{s}{2} \ln t - \sum_{j=1}^{s+1} \left((tP_j + i_j) \left(\left(\frac{i_j}{tP_j}\right) - \frac{1}{2} \left(\frac{i_j}{tP_j}\right)^2 + o\left(\frac{i_j}{tP_j}\right)^3 \right) \right) - \frac{1}{2} \sum_{j=1}^{s+1} \ln \left(P_j + \frac{i_j}{t}\right) + O(1) \\
&=^d -\frac{s}{2} \ln t - \sum_{j=1}^{s+1} i_j - \sum_{j=1}^{s+1} \left(\frac{i_j^2}{tP_j} - \frac{1}{2} \frac{i_j^2}{tP_j} - \frac{1}{2} \frac{i_j^3}{(tP_j)^2} + o\left(\frac{i_j^3}{t^2}\right) \right) - O(1) + O(1) \\
&=^b -\frac{s}{2} \ln t - 0 - \sum_{j=1}^{s+1} \frac{(i_j^2)}{2P_j t} + O\left(\frac{i_j^3}{t^2}\right) + O(1) \\
&\overset{\circ}{=}^e e^{\ln\left(\left(\frac{1}{\sqrt{t}}\right)^s\right) - \sum_{j=1}^{s+1} \frac{(i_j^2)}{2P_j t}}
\end{aligned}
$$

$a)$: By Sterling formula: $\ln n! \in \left[-n + \left(n + \frac{1}{2}\right) \ln n + 1, -n + \left(n + \frac{1}{2}\right) \ln n + \ln \sqrt{2\pi}\right]$

$b)$: By construction: $\sum_{i=1}^{s+1} k_j = t$ and $\sum_{i=1}^{s+1} i_j = 0$

$c)$: Taylor series approximation. The approximation can be done because $\frac{\sqrt{t \log t}}{t} \to 0$

$d)$: $\frac{\sqrt{t \log t}}{t} \to 0 \Rightarrow \ln\left(P_j + \frac{i_j}{t}\right) \to \ln(P_j) \in (-\infty, +\infty)$.

$e)$: $\frac{\sqrt{t \log t}}{t} \to 0 \Rightarrow \left(\frac{i_j^3}{t^2}\right) \to 0$

$\square$

Lemma 8. *Let $X_t \approx f(x_1, ..., x_{s+1}, t)$ a multinomial distribution with cell probabilities $\boldsymbol{m} = [P_1, ..., P_{s+1}]$.Let $\hat{X}_t^s$ be the first s coordinates of the vector $[x_1 - tP_1, ..., x_{s+1} - tP_{s+1}]$, C^s be a compact subset of $\mathbb{R}^s$: $[-1,1]^s \subset C^s$ and K^{s-r} be a (s-r)-dimensional closed convex cone in $\mathbb{R}^s$. Then*

i) $P(\hat{X}_t^s \in C^s) \doteq \left(\frac{1}{\sqrt{t}}\right)^s$

ii) $P(\hat{X}_t^s \in K^{s-r} + C^s) \doteq \left(\frac{1}{\sqrt{t}}\right)^r$

Proof. $i)$: Let $k^s \in \mathbb{Z}^s$ be the first s coordinates of the vector $[k_1, ..., k_s, k_{s+1} = -\sum_{j=1}^{s} k_j]$.

$$P(\hat{X}_t^s \in C^s) = P(X_t^s \in C^s + t\mathbf{m}) = \sum_{k^s \in (C^s + t\mathbf{m}) \cap \mathbb{Z}^s}^{t} \binom{t}{k_1, ..., k_{s+1}} \prod_{i=1}^{s+1} P_i^{k_i}$$

Note that:

a) Let $\mathcal{A} := \{k^s \in (C^s + t\mathbf{m}) \cap \mathbb{Z}^s\}$. C^s compact $\Rightarrow |(C^s + t\mathbf{m}) \cap \mathbb{Z}^s| < \infty \Rightarrow \exists B < \infty : |\mathcal{A}| = B$

b) By Lemma 7 $\forall k^s \in (C^s + t\mathbf{m}) \cap \mathbb{Z}^s$, $P(k^s) \doteq \left(\frac{1}{\sqrt{t}}\right)^s$. Because C^scompact $\Rightarrow \max_j \frac{i_j^2}{t} \to 0$

Therefore

$$\min_{k^s \in (C^s + t\mathbf{m}) \cap \mathbb{Z}^s} P(k_t^s) \leq P(\hat{X}_t^s \in C^s) \leq B \max_{k^s \in (C^s + t\mathbf{m}) \cap \mathbb{Z}^s} P(k_t^s)$$

which implies $P(\hat{X}_t^s \in C^s) \doteq \left(\frac{1}{\sqrt{t}}\right)^s$

ii): Let D_i be the extreme points of a section of K (points that are not a non-negative linear combination of other points in K).

Let start with $s - r = 1$

$$
\begin{aligned}
P(\hat{X}_t^s \in K^1 + C^s) &= P\left(X_t^s \in (t\mathbf{m}^s + C^s + \tau D) \cap \mathbb{Z}^s\right) \quad for\ 0 \le \tau \le t \\
&\doteq^a \sum_{\tau=1}^{t} P\left(X_t^s \in (t\mathbf{m}^s + C^s + \tau D) \cap \mathbb{Z}^s\right) \\
&\doteq^b \sum_{\tau=1}^{\sqrt{t \log t}} \frac{1}{\sqrt{t^s}} e^{-\frac{(\tau)^2}{t}} \\
&\doteq^c \int_{\tau=1}^{\sqrt{t \log t}} \frac{1}{\sqrt{t^s}} e^{-\frac{(\tau)^2}{t}} d\tau \\
&\doteq \frac{1}{\sqrt{t^s}} \sqrt{\frac{t\pi}{2}} \\
&\doteq \frac{1}{\sqrt{t^r}}
\end{aligned}
$$

$a)$: $[-1,1]^s \subset C^s \Rightarrow \forall \tau, \tau D + C^s \cap Z^s \neq \emptyset$

$b)$: By lemma 7. This is done WLOG because the Law of Iterated Logarithm implies: $\max \tau = O(\sqrt{t \log \log t})$ P-a.s.

$c)$: The approximation is done WLOG because each element of the sum is positive.

For $s-r=2$

$$
\begin{aligned}
P(\hat{X}_t^s \in K^2 + C^s) &= P\left(X_t^s \in (t\mathbf{m}^s + C^s + \tau Chull(D_1, D_2)) \cap \mathbb{Z}^s\right) \quad for\ 0 \le \tau \le t \\
&\doteq \sum_{\tau_1, \tau_2 \in (\tau Chull(D_1, D_2))}^{t} P\left(X_t^s \in (t\mathbf{m}^s + C^s + \tau_1 D_1 + \tau_2 D_2) \cap \mathbb{Z}^s\right) \\
&\doteq \sum_{\tau_1, \tau_2 \in (\tau Chull(D_1, D_2))}^{\sqrt{t \log t}} \frac{1}{\sqrt{t^s}} e^{\frac{(\tau_1)^2 + (\tau_2)^2}{t}} \\
&\doteq \int_{\tau_1, \tau_2 \in (\tau Chull(D_1, D_2))}^{\sqrt{t \log t}} \frac{1}{\sqrt{t^s}} e^{\frac{(\tau_1)^2 + (\tau_2)^2}{t}} d\tau_1 d\tau_2 \\
&\doteq \frac{1}{\sqrt{t^s}} \frac{\sqrt{t^2}\pi}{2} \\
&\doteq \frac{1}{\sqrt{t^r}}
\end{aligned}
$$

The same argument generalizes to arbitrary values of $s-r$. $\square$

Lemma 9. *Let* $E_n = \{B_t \in K^{s-r} + C^s\} = \left\{\hat{X}_n \in K^{s-r} + C^s\right\}$.

$$
for\ n > m, \exists A < +\infty : P\{E_m \cap E_n\} \le AP\{E_m\}P\{E_{n-m}\}
$$

Proof.

$$\begin{aligned}
P\{E_m \cap E_n\} &= \sum_{X_m} P\{E_m \cap E_n | X_m\} P(X_m) \\
&=^a \sum_{X_m} P\{E_m | X_m\} P\{E_n | X_m\} P(X_m) \\
&= \sum_{X_m} I_{E_m} P\{E_n | X_m\} P(X_m) \\
&=^b \sum_{X_m \in E_m \cap \mathbb{Z}^s} P\{\hat{E}_{n-m}(X_m) | X_m \in E_m \cap \mathbb{Z}^s\} P(X_m) \\
&\leq^c P\{E_m\} \max_{k^r \in C^r} P\{\bar{E}_{n-m}(k^r)\} \\
&\circeq^d P\{E_m\} \frac{1}{\sqrt{t-m}^r} = P\{E_m\} P\{E_{n-m}\}
\end{aligned}$$

$a)$: The process is Markov and X_m is in the past of E_n

$b)$: Letting $\hat{E}_i(X_m) = \{B_i^s \in (K^{s-r} + C^s - X_m^s)\}$.

$c)$: Letting $\bar{E}_{n-m}(k^r) = \left\{B_{n-m}^r \in (C^r - k^r)\right\}$.

Note that:

$\max_{k^r \in C^r} P\{\bar{E}_{n-m}(k^r)\}$ exists because C^r is compact.

Moreover

$$\begin{aligned}
\max_{k^r \in C^r} P\{\bar{E}_{n-m}(k^r)\} &= \max_{k^s \in C^s + \mathbb{R}^{s-r}} P\{\hat{E}_{n-m}(k^s)\} \\
&\geq \max_{k^s \in C^s + K^{s-r}} P\{\hat{E}_{n-m}(k^s)\} \\
&\geq \max_{k^s \in (C^s + K^{s-r}) \cap \mathbb{Z}^s} P\{\hat{E}_{n-m}(k^s)\} \\
&= \max_{X_m \in E_m} P\{\hat{E}_{n-m}(X_m)\}
\end{aligned}$$

$d)$: Follows from Lemma 8 as $C^r - k^r$ is compact $\square$

8 References

Alchian A., 1950. Uncertainty, Evolution and Economic Theory, Journal of Political Economy, 58, 211221

Ash R., 1999 , Probability and Measure Theory. Academic Press

Borovička J. 2010 Survival and stationary equilibria with heterogeneous beliefs under recursive preferences, Job Market Paper

Bernardo, J.M. and Smith,A.F.M 1994. Bayesian Theory, Chirchester: Wiley

Blume L, Easley D. 1992. Evolution and market behavior. J. Econ. Theory 58:940

Blume L, Easley D. 2002. Optimality and natural selection in markets. J. Econ. Theory 107:95135

Blume L, Easley D. 2006. If you're so smart, why aren't you rich? Belief selection in complete and incomplete markets. Econometrica 74:92966

Blume L, Easley D. 2009a. The market organism: long-run survival in markets with heterogeneous traders. J. Econ. Dyn. Control 5:102335

Blume L, Easley D. 2010. Heterogeneity,selection and wealth dynamics. Annu. Rev. Econ. 2010. 2:42550

Clarke B.S, Barron A.R. 1990. Information theoretic asymptotics of Bayes methods. IEEE Trans. Inform. Theory 36:45371

Cover T. M.,J. A.Thomas. 1991, Elements of Information Theory. Wiley and Son,Inc

Cvitanic J., Malamud 2011, Price impact and portfolio impact. Journal of Financial Economics 100, 201-225

De Finetti, B. 1974. Theory of Probability. A critical introductory treatment. London:John Wiley and Sons.

De Long, J. B., A. Shleifer, L. Summers, R. Waldmann (1990): Noise Trader Risk in Financial Markets, Journal of Political Economy, 98, 703738.[930,957]

De Long, J. B., A. Shleifer, L. Summers, R. Waldmann (1991): The Survival of Noise Traders in Financial Markets, Journal of Business, 64, 119.[930,931,957]

Erdös P. "A problem about prime numbers and random walks" Illinois J. Math., vol. 5 (1961), pp. 352-353

Fama E. 1965. The behavior of stock market prices. J. Bus. 38:34105

Folland G.B. 1999, Real Analysis: Modern Techniques and Their Applications. Wiley New York ISBN-10: 0471317160, 2nd edition.

Friedman, M. (1953): Essays in Positive Economics. Chicago: University of Chicago Press.[930,932]

Grossman S. and Stiglitz, J. 1980. On the impossibility of informationally efficient markets. American Economic Review 70, 393408

Grünwald P. 2007, The Minimum Description Length Principle, MIT Press ISBN-10: 0262072815

Grünwald P. and Rooij S. 2005 ”Asymptotic Log-loss of Prequential Maximum Likelihood Codes”

Hutter M. 2004 Universal Artificial Intelligence: Sequential Decisions Based On Algorithmic Probability ISBN 3-540-22139-5

Kalai E, Lehrer E. 1993. Rational learning leads to Nash equilibrium. Econometrica 61:101945

Kass R.E., A.E. Raftery 1995. Bayes factors. Journal of the American Statistical. Association 90: 773-795.

Kelly J. 1965. A new interpretation of information rate. Bell Sys. Tech. Journal, 35: 917-926.

Kochen S., Charles S. A note On the Borel-Cantelli Lemma. Cornell Univerity

Kogan, L., S. Ross, J. Wang, M. Westerfield. 2006. The price impact and survival of irrational traders. J. Finance 61 195229.

Kogan, Leonid, Stephen A. Ross, Jiang Wang, and Mark M. Westerfield, 2009. Market selection.

Ljungqvist, L., Sargent, T.J. 2004. Recursive Macroeconomic Theory, MIT press.

Mailath G.,Sandroni A. 2003. Market Selection and Asymmetric Information. Review of Economic Studies. 70(2): 343-369.

Marinacci M., Massari F. Learning in Ambiguous settings. Working Paper

Massari F. 2012. Comment on If you're so smart, why aren't you rich? Belief selection in complete and incomplete markets. Forthcoming, Econometrica

Massari F. 2012a Bayesian updating and bounded memory. Working paper.

Mörters P., Peres Y. 2010, Brownian motion. Cambridge Univ. Pr.

Phillips P.C.B., Ploberger W. 2003. Empirical limits for time series econometric models. Econometrica 71:62773

Rissanen, J. J. 1986. Stochastic Complexity and Modeling, Annals of Statistics, 14, 1080-1100

Rissanen, J. J.1987 Stochastic Complexity (with discussion), Journal of the Royal Statistical Society, 49,223-239, and 252-265

Sandroni A. 2000. Do markets favor agents able to make accurate predictions? Econometrica 68:130342

Surowiecki J. 2004. The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations. New York: Random House

Yan H. 2008. Natural Selection in Financial Markets: Does It Work?. Management Science. 54 (11): 19351950.

Williamas, D. 1991, Probability with Martingales, Cambridge Univ. Pr. ISBN 9780521406055.

Chapter 2: On survival of traders in complete markets

Abstract

Market Selection studies the evolution of wealth shares among traders with heterogeneous beliefs operating in a market. For the case of homogeneous discount factors, the sufficient condition proposed in Blume-Easley (2006) for the wealth share of a trader to converge to 0 (vanish) implies that a trader vanishes whenever there is another trader in the economy whose beliefs are, in every period, closer to the true distribution. Surprisingly, this intuitive result can fail to be correct. In this paper I present a non pathological counter-example to Blume-Easley's (2006) result and discuss its implications.

1 Introduction

Market Selection studies the evolution of market prices and wealth shares in economies where traders have heterogeneous beliefs. One focus of this literature has been on establishing sufficient conditions for the wealth share of a trader to converge to 0 (vanish). Theorem 8 in Blume and Easley (2006) gives a sufficient condition for the case in which the true probability need not be absolutely continuous with respect to survivors (i.e. non-vanishing traders) beliefs. For the case in which players have identical discount factors, the condition of Theorem 8 implies that a trader vanishes whenever there is another trader in the economy whose beliefs are, in every period, closer to the true distribution. This paper provides a non-pathological counter example. The example also provides insight into equilibrium prices in market selection models.

In Sections 2 and 3 I introduce the notation and the sufficient condition of Blume-Easley (2006). Sections 4, 4.1 and 5 are dedicated to the counter-example, comments and a remark on the proof of Theorem 8.

2 The model

I follow the notation of Blume-Easley (2006). The model is a discrete time, infinite horizon, exchange economy that allocates a single commodity good. Time is discrete and begins at date 0. At each date t there is a finite set of states $\mathcal{S} \equiv \{1, ..., S\}$ with cardinality $|\mathcal{S}|$=S. The set of all infinite sequences of states is Σ with representative sequence of realizations $\sigma = (\sigma_0, ...)$. Let $\sigma^t = (\sigma_0, ..., \sigma_t)$ denote the partial history through date t of the path σ, and $I_t^s(\sigma)$ is the indicator function (defined on path σ) that takes the value 1 if the t-th element of path σ takes the value s ($\sigma_t = s$) and 0 otherwise. The set Σ, with its σ-field, is the relevant measurable space.The information partition at period-event σ_t is $\mathcal{F}_t$. Let p be the "true" probability measure on Σ. For any probability measure q on Σ, $q_t(\sigma)$ is the marginal probability of the partial history $(\sigma_0, ..., \sigma_t)$; that is, $q_t(\sigma) = q(\{\sigma_0, ...\sigma_t\} \times S \times S \times ...)$.The expectation operator without a subscript is the expectation with respect to the true measure p. The little-o notation, $f(x) = o(g(x))$, abbreviates $\lim_{x\to\infty} \frac{f(x)}{g(x)} \to 0$.

2.1 Traders

The economy contains a finite number of traders I, each with consumption set $\mathbb{R}_+$. A consumption plan $c : \Sigma \to \prod_{t=0}^{\infty} \mathbb{R}_+$ is a sequence of $\mathbb{R}_+$- valued functions $\{c_t(\sigma)\}_{t=0}^{\infty}$ in which each c_t is $\mathcal{F}_t$-measurable. Each trader i is endowed with a particular consumption plan, called the endowment stream and denoted by e^i. Trader i has a utility function $U_i(c) : c \to [-\infty, +\infty)$ which is the expected present discounted value of some payoff stream with respect to some beliefs. The beliefs of trader i are represented by a probability distribution p^i on Σ. Each trader also has a payoff function $u^i : \mathbb{R}_+ \to [-\infty, +\infty)$ over consumption and a discount factor $\beta_i \in (0, 1)$. Trader i's consumption plan c utility is:

$$U^i(c) = E_{p^i} \sum_{t=0}^{\infty} \beta_i^t u_i(c_t(\sigma)).$$

In terms of asymptotic consumption, traders can either vanish or survive:

Definition 1.

Trader i vanishes on path σ if $\lim_{t\to\infty} c_t(\sigma) \to 0$.

He survives on path σ if $\limsup_t c_t(\sigma) > 0$.

In the notation of Blume-Easley (2006), these are the axioms I will refer to:

- A1: The payoff functions $u_i : \mathbb{R}_+ \to [-\infty, +\infty)$ are C^1, strictly concave, strictly monotonic and satisfy the Inada condition at 0.
- A2: The endowment processes e_t^i are bounded above and away from 0.
- A3: For each trader i, all dates t, and all paths σ if $p_t(\sigma) > 0$ then $p_t^i(\sigma) > 0$.
- A5: There is a $\delta > 0$ such that for all paths σ, all dates t, $s \in S$, and traders i, $p(s|\mathcal{F}_{t-1})(\sigma) > 0 => p(s|\mathcal{F}_{t-1})(\sigma) > \delta$ and $p^i(s|\mathcal{F}_{t-1})(\sigma) > \delta$.

Axioms A1 and A2 are standard in General Equilibrium models. A3 and A5 are technical axioms introduced to avoid pathological behavior of log likelihood ratios on the boundaries of the simplex. A4 is not stated because it plays no role in the discussion.

3 Blume-Easley's sufficient conditions to vanish

Blume-Easley characterize the long run fate of the traders in the economy in terms of Pareto optimality of the equilibrium allocation. In particular the economy is organized by a Social Planner that solves the following maximization problem:

$$\max_{c^i} \sum_{i \in I} \lambda_i E_{p_i} \sum_{t=1}^{\infty} \beta^t u_i(c_t^i(\sigma))$$

$$\text{such that} \quad \sum_{i \in I} c_t^i - e_t^i = 0$$

$$\forall t, \sigma \quad c_t^i(\sigma) \geq 0$$

The first order conditions (A1 guaranties an interior solution) implies that $\forall i, j \in I$, $\forall t$, and $\forall \sigma \in \Sigma$, the optimal allocation of resources satisfies:

$$\frac{u^{i\prime}(c_t^i(\sigma))}{u^{j\prime}(c_t^j(\sigma))} = \frac{\lambda_j \beta_j^t p_t^j(\sigma)}{\lambda_i \beta_i^t p_t^i(\sigma)} = \frac{\lambda_j \beta_j^t \prod_{s \in S} p_\tau^j(s|\mathcal{F}_{\tau-1})^{\sum_{\tau=1}^t I_\tau^s(\sigma)}}{\lambda_i \beta_i^t \prod_{s \in S} p_\tau^i(s|\mathcal{F}_{\tau-1})^{\sum_{\tau=1}^t I_\tau^s(\sigma)}} \tag{1}$$

If we define the two random variables:

$$Z_t^k := -\sum_{s\in S} I_t^s(\sigma)\log p^k(s|\mathcal{F}_{t-1}) \quad \text{and} \quad \bar{Z}_t^k(\sigma^{t-1}) := \sum_{\tau=1}^{t} E\{Z_\tau^k|\mathcal{F}_{\tau-1}\}$$

Equation 1 can be rewritten as:

$$\log\frac{u^{i'}(c_t^i(\sigma))}{u^{j'}(c_t^j(\sigma))} = \log\frac{\lambda_j}{\lambda_i} + t\log\frac{\beta_j}{\beta_i} + \sum_{\tau=1}^{t}(Z_\tau^i - Z_\tau^j) \tag{2}$$

B.-E. and S. sufficient condition to vanish 1.

Assume Axioms 1-3 and 5. On the event:

$$\lim_{t\to\infty}\left[t\log\frac{\beta_j}{\beta_i} + \bar{Z}_t^i(\sigma^{t-1}) - \bar{Z}_t^j(\sigma^{t-1})\right] = +\infty$$

$c^i(\sigma^t) \to 0\ p - a.s.$

It is instructive to compare the implications of Theorem 8, Blume-Easley (2006), and Proposition 3, Sandroni (2000), in a homogeneous discount factor economy ($t\log\frac{\beta_j}{\beta_i} = 0$).
Using Equation 2 we can show that Theorem 8 implies:

$$\lim_{t\to\infty}\left[\bar{Z}_t^i(\sigma^{t-1}) - \bar{Z}_t^j(\sigma^{t-1})\right] = +\infty => \lim_{t\to\infty}\sum_{\tau=1}^{t}(Z_\tau^i - Z_\tau^j) = +\infty$$

while Proposition 3 only implies:

$$\lim_{t\to\infty}\frac{1}{t}\left[\bar{Z}_t^i(\sigma^{t-1}) - \bar{Z}_t^j(\sigma^{t-1})\right] > 0 => \lim_{t\to\infty}\sum_{\tau=1}^{t}(Z_\tau^i - Z_\tau^j) = +\infty.$$

The two implications are not equivalent. The $\frac{1}{t}$ term “kills” all of the deviations of order $o(t)$. In the next Section I show that there are simple economies that satisfy A1-A3, A5 in which these deviations occur and invalidate Theorem 8.

4 The counter-example

Consider an economy such that the distribution of states is given by iid draws on $\mathcal{S} = \{a, b\}$ with $p(a) = \frac{1}{2} = p(b)$. There are three traders $\{1, 2, 3\}$ operating in the market with log utility (A1 is satisfied[8]) and homogeneous discount factors. Endowments are positive and the aggregate endowment is constant in each period (A2 is satisfied).

The beliefs are as follows: trader 1 and 2 have i.i.d. beliefs with $p_1(a){=}p_2(b){=}\frac{1}{3}$; trader 3 is a Bayesian learner whose prior support consists of singletons $\{p_1, p_2\}$ to which he assigns equal probability: $\frac{1}{2}$. With these specifications, for all t, the conditional probabilities that traders 1,2,3 attach to the realization a at period t, given the information set $\mathcal{F}_{t-1}$ are respectively: $p_{1,t}(a|\mathcal{F}_{t-1}) = 1 - p_{2,t}(a|\mathcal{F}_{t-1}) = \frac{1}{3}$ and

$$p_{3,t}(a|\mathcal{F}_{t-1}) = \frac{\frac{1}{2}p_1(a|\mathcal{F}_{t-1})p_{1,t-1}(\sigma) + \frac{1}{2}p_2(a|\mathcal{F}_{t-1})p_{2,t-1}(\sigma)}{\frac{1}{2}p_{1,t-1}(\sigma) + \frac{1}{2}p_{2,t-1}(\sigma)} \tag{3}$$

The conditional probabilities of state b at time t are $1 - p_{i,t}(a|\mathcal{F}_{t-1})$.
Therefore, in each period and in each path, the beliefs of all traders belong to the interval $\left[\frac{1}{3}, \frac{2}{3}\right]$ and A3-A5 are satisfied.

Claim : Trader 1 and 2 cannot both vanish p-a.s..
Here I prove the stronger result that there is no path in which both traders vanish.
Suppose, by contradiction, that there is a path $\sigma \in \Sigma$ in which the consumption of traders 1 and 2 goes to 0. Then, the FOC (Equation 1) for this economy would imply that the following conditions hold simultaneously:

- $\frac{\lambda_3 p_{3,t}(\sigma)}{\lambda_1 p_{1,t}(\sigma)} = \frac{c_t^3(\sigma)}{c_t^1(\sigma)} \to \infty$
- $\frac{\lambda_3 p_{3,t}(\sigma)}{\lambda_2 p_{2,t}(\sigma)} = \frac{c_t^3(\sigma)}{c_t^2(\sigma)} \to \infty$

This is impossible. Solving Equation 3 recursively, we see that:$\forall \sigma \in \Sigma,\ \forall t$

$$p_{3,t}(\sigma) = \frac{1}{2}p_{1,t}(\sigma) + \frac{1}{2}p_{2,t}(\sigma) < \max(p_{1,t}(\sigma), p_{2,t}(\sigma))$$

[8] If we define $\log 0 = -\infty$.

Therefore the left-hand-side of at least one of the FOC must be less than ∞.

This result contradicts Theorem 8. As Lemma 10 (see Appendix) shows, the beliefs of trader 1 and 2 satisfy Blume-Easley's sufficient condition to vanish.

4.1 Comments

- The sufficient condition in Sandroni (2000) and the necessary and sufficient condition of Massari (2011) both correctly identify the set of traders that vanishes.

- $\forall t$, $p_{3,t} \in (p_{1,t}, p_{2,t})$, and $p(a) = \frac{1}{2}$, therefore trader 3 makes better forecasts about the future than trader 1 and 2 in every period. Yet his wealth share is strictly bounded away from 1. In particular, if all traders start with the same wealth shares (Pareto weights), the FOCs imply that in every path trader 3's wealth share is always between the wealth shares of the other two traders.

- The counter-example does not depend on log preferences, on the symmetry of the problem around $p = \frac{1}{2}$, or on the dimensionality of the state space. It can be shown that Theorem 8 does not apply to any homogeneous discount factors economy such that there is a subset $\hat{I}$ of the survivors[9] with asymptotically fixed beliefs and a Bayesian learner whose prior contains the beliefs of some of the survivors in $\hat{I}$.

- It is easy to show that if all the traders have the same initial wealth shares, trader 3's beliefs coincide with the normalized equilibrium prices. The counter-example provides us an interesting insight about the argument in the book: "The Wisdom of Crowds" (Surowiecki 2004). He claims that equilibrium prices can produce more accurate forecasts than the beliefs of all traders in the market. The example shows that this intuition is true with a grain of salt. On one hand, we have that according to any probabilistic criterion (for example in terms of KullbackLeibler distance and expected Euclidean distance), normalized equilibrium prices are

[9] The condition of Massari (2011) can be used to identify the survivors.

the best forecasts for the next period realizations. On the other hand, these better forecasts cannot be used by any trader in the market to increase his wealth share.[10]

- If Theorem 8 of Blume-Easley (2006) were correct it would lead to absurd conclusions. Consider the case of a financial market in which there are two portfolios of assets that, on average, perform equally well and in which traders are allowed to invest in both portfolios. Suppose Blume-Easley's sufficient condition were correct. Then, to invest half of the initial capital in each portfolios would ensure (p-a.s.) a rate of return that is higher than the rate of return of the portfolio that performs best. This conclusion is simply too good to be true.

5 A comment on the proof of Theorem 8

In this Section, I show where the proof given by Blume-Easley (2006) of Theorem 8 (page 963) breaks down. The key step of the proof is to show that p-a.s.:

$$\lim_{t\to\infty}[t\log\frac{\beta_j}{\beta_i}+\bar{Z}_t^i(\sigma^{t-1})-\bar{Z}_t^j(\sigma^{t-1})]=+\infty \Rightarrow \lim_{t\to\infty}[t\log\frac{\beta_j}{\beta_i}+\sum_{\tau=1}^{t}(Z_\tau^i-Z_\tau^j)]=+\infty.$$

Blume-Easley use a version of the Strong Law of Large Numbers (specifically Freedman (1973)) to show that there are random variables $A_t^k(\sigma^t)=\frac{\sum_{\tau=1}^t Z_\tau^k}{\bar{Z}_t^k(\sigma^{t-1})}$ that converge to 1 p-a.s. such that:

$$t\log\frac{\beta_j}{\beta_i}+\sum_{\tau=1}^{t}(Z_\tau^i-Z_\tau^j)=t\log\frac{\beta_j}{\beta_i}+A_t^i\bar{Z}_t^i(\sigma^{t-1})-A_t^j\bar{Z}_t^j(\sigma^{t-1}).$$

The problem is that the above equation does not imply, as they claim, that:

$$\begin{aligned}&\lim_{t\to\infty}[t\log\frac{\beta_j}{\beta_i}+\bar{Z}_t^i(\sigma^{t-1})-\bar{Z}_t^j(\sigma^{t-1})]=+\infty\\ &\Rightarrow \lim_{t\to\infty}[t\log\frac{\beta_j}{\beta_i}+A_t^i\bar{Z}_t^i(\sigma^{t-1})-A_t^j\bar{Z}_t^j(\sigma^{t-1})]=+\infty \quad p-a.s.\end{aligned}$$

This implication is not true because the rate of convergence to 1 of the random variables A_t^i and A_t^j can be different!

[10]In Massari (2011) it is shown that this property of market prices holds in any economy that satisfies A1-A2.

For an intuition, consider the case in which $\beta_i = \beta_j$ and the random variables $A^i_t(\sigma^t)$ and $A^j_t(\sigma^t)$ are such that:

$$A^i_t(\sigma^t) = 1 - \frac{1}{t} \quad \text{and} \quad A^j_t(\sigma^t) = 1 - \frac{1}{\log t}.$$

Then

$$\lim_{t\to\infty}[\bar{Z}^i_t(\sigma^{t-1}) - \bar{Z}^j_t(\sigma^{t-1})] = \lim_{t\to\infty}\sum_{\tau=1}^{t} E\log\Big(\frac{p^j(\sigma|\mathcal{F}_{\tau-1})}{p^i(\sigma|\mathcal{F}_{\tau-1})}\Big) = \infty$$
$$\neq> \lim_{t\to\infty} A^i_t\bar{Z}^i_t(\sigma^{t-1}) - A^j_t\bar{Z}^j_t(\sigma^{t-1}) = \lim_{t\to\infty}\sum_{\tau=1}^{t} E\log\Big(\frac{p^j(\sigma|\mathcal{F}_{\tau-1})(p^i(\sigma|\mathcal{F}_{\tau-1}))^{\frac{1}{t}}}{p^i(\sigma|\mathcal{F}_{\tau-1})(p^j(\sigma|\mathcal{F}_{\tau-1}))^{\frac{1}{\log t}}}\Big)$$
$$= \lim_{t\to\infty}[\sum_{\tau=1}^{t} E\log\frac{p^j(\sigma|\mathcal{F}_{\tau-1})}{p^i(\sigma|\mathcal{F}_{\tau-1})} + E\log\Big(\frac{(p^i(\sigma|\mathcal{F}_{\tau-1}))^{\frac{1}{t}}}{(p^j(\sigma|\mathcal{F}_{\tau-1}))^{\frac{1}{\log t}}}\Big)] = \infty.$$

The implication is not true as it is possible that $\lim_{t\to\infty} E\Big[\log\frac{(p^i(\sigma|\mathcal{F}_{\tau-1}))^{\frac{1}{t}}}{(p^j(\sigma|\mathcal{F}_{\tau-1}))^{\frac{1}{\log t}}}\Big] \to -\infty$ fast enough. In the counter-example, trader 1 and 2's beliefs are fixed while trader 3's beliefs do not converge, therefore trader 1 and 2's rate of convergence to the average payoff is faster than trader 3's. This is to say that A^1_t and A^2_t converge to 1 faster than A^3_t. Moreover, it can be shown that $A^3_t \to 1$ from below. This is enough to invalidate Theorem 8.

A Appendix

Lemma 10. *In the setting of the counterexample, for i=1,2:*

$$\lim_{t\to\infty}[t\log\frac{\beta_3}{\beta_i} + \bar{Z}^i_t(\sigma^{t-1}) - \bar{Z}^3_t(\sigma^{t-1})] = +\infty \quad p-a.s. \tag{4}$$

Proof.

First note that the first element of Equation 4 is 0 as all the traders in the market have the same discount factor and that condition 4 can be conveniently rewritten as:

$$\lim_{t\to\infty}[\sum_{\tau=1}^{t} E\log\frac{p_\tau(s|\mathcal{F}_{\tau-1})}{p_{i,\tau}(s|\mathcal{F}_{\tau-1})} - \sum_{\tau=1}^{t} E\log\frac{p_\tau(s|\mathcal{F}_{\tau-1})}{p_{3,\tau}(s|\mathcal{F}_{\tau-1})}] = \infty \quad p-a.s.$$

Note that:

- *i)* For i=1,2 $\forall t$, $E\log\frac{p_t(\sigma|\mathcal{F}_{t-1})}{p_{i,t}(\sigma|\mathcal{F}_{t-1})}=\frac{1}{2}\log\frac{\frac{1}{2}}{\frac{1}{3}}+\frac{1}{2}\log\frac{\frac{1}{2}}{\frac{2}{3}}=\frac{1}{2}\log\frac{9}{8}>0$ Because the two traders do not update their beliefs and the true probability is iid.

- *ii)* $\forall t, p_{3,t}(\sigma|\mathcal{F}_{t-1})\in\left[\frac{1}{3},\frac{2}{3}\right]$ therefore $\forall t$ $E\log\frac{p_t(\sigma|\mathcal{F}_{t-1})}{p_{3,t}(\sigma|\mathcal{F}_{t-1})}\leq E\log\frac{p_t(\sigma|\mathcal{F}_{t-1})}{p_{i,t}(\sigma|\mathcal{F}_{t-1})}$.

- *iii)* $\sum_{\tau=1}^{t} I_{\tau}^{a}=\frac{t}{2} => p_{3,t+1}(\sigma|\mathcal{F}_t)=\frac{1}{2}$ This is because:

$$\begin{aligned}
p_{3,t+1}(a|\mathcal{F}_t) &= \frac{\frac{1}{2}p_{1,t+1}(a|\mathcal{F}_t)p_{1,t}(\sigma)+\frac{1}{2}p_{2,t+1}(a|\mathcal{F}_t)p_{2,t}(\sigma)}{\frac{1}{2}p_{1,t}(\sigma)+\frac{1}{2}p_{2,t}(\sigma)}\\
&=^{a} \frac{\frac{1}{2}(p_{1,t+1}(a|\mathcal{F}_t)p_{1,t}(\sigma)+p_{2,t+1}(a|\mathcal{F}_t)p_{1,t}(\sigma))}{\frac{1}{2}p_{1,t}(\sigma)+\frac{1}{2}p_{1,t}(\sigma)}\\
&=^{b} \frac{\frac{1}{2}p_{1,t}(\sigma)}{p_{1,t}(\sigma)}=\frac{1}{2}
\end{aligned}$$

  Where in (a) I used the the symmetry of $p_{1,t}(\sigma)$ and $p_{2,t}(\sigma)$ around $\frac{1}{2}$ to conclude that $\sum_{\tau=1}^{t} I_{\tau}^{a}=\frac{t}{2} => p_{1,t}(\sigma)=p_{2,t}(\sigma)$. And (b) is a consequence of the fact that in each period $p_2(a|\mathcal{F}_{t-1})+p_1(a|\mathcal{F}_{t-1})=1$.

- *iv)* $\sum_{\tau=1}^{t} I_a=\frac{t}{2}$ occurs infinitely often p-a.s.. This is because:

  Let $y_\tau=\begin{cases} 1\ if\ s_\tau=a\\ -1\ if\ s_\tau=b\end{cases}$

  The Law of Iterated Logarithm (see D.Williams page 208 for reference) implies:

$$\begin{cases}\limsup_t\sum_{\tau=1}^{t}\frac{y_\tau}{\sqrt{2t\log\log t}}=1\ p-a.s.\\ \liminf_t\sum_{\tau=1}^{t}\frac{y_\tau}{\sqrt{2t\log\log t}}=-1\ p-a.s.\end{cases}$$

Therefore $p\{t : \sum_{\tau=1}^{t} I_\tau^a = \frac{t}{2}\ i.o.\} = 1$

Consequently on a set of p measure 1

$$\lim_{t\to\infty}\left[\sum_{\tau=1}^{t} E\log\frac{p_\tau(\sigma|\mathcal{F}_{t-1})}{p_{i,\tau}(\sigma|\mathcal{F}_{\tau-1})} - \sum_{\tau=1}^{t} E\log\frac{p_\tau(\sigma|\mathcal{F}_{\tau-1})}{p_{3,\tau}(\sigma|\mathcal{F}_{\tau-1})}\right]$$
$$\geq^a \lim_{t\to\infty}\left[\sum_{K<t:\sum_{\tau=1}^{K-1} I_\tau^a=\frac{K}{2}} E\log\frac{p_K(\sigma|\mathcal{F}_{K-1})}{p_{i,K}(\sigma|\mathcal{F}_{K-1})} - \sum_{K<t:\sum_{\tau=1}^{K-1} I_\tau^a=\frac{K}{2}} E\log\frac{p_K(\sigma|\mathcal{F}_{K-1})}{p_{3,K}(\sigma|\mathcal{F}_{K-1})}\right]$$
$$=^b \lim_{K\to\infty}\left[K(\frac{1}{2}\log\frac{9}{8} - \frac{1}{2}\log\frac{\frac{1}{2}}{\frac{1}{2}})\right] = +\infty$$

$ii) => (a)$, $iii)$ implies the second term of (b) and $iv) => K \to^{a.s} \infty$. □

B References

Blume L, Easley D. 1992. Evolution and market behavior. J. Econ. Theory 58:940

Blume LE, Easley D. 2002. Optimality and natural selection in markets. J. Econ. Theory 107:95135

Blume LE, Easley D. 2006. If youre so smart, why arent you rich? Belief selection in complete and incomplete markets. Econometrica 74:92966

Blume LE, Easley D. 2009a. The market organism: long-run survival in markets with heterogeneous traders. J. Econ. Dyn. Control 5:102335

Blume LE, Easley D. 2010. Heterogeneity,selection and wealth dynamics. Annu. Rev. Econ. 2010. 2:42550

De Finetti, B. 1974. Theory of Probability. A critical introductory treatment. London:John Wiley and Sons.

Freedman, D. A. 1973. Another Note on the BorelCantelli Lemma and the Strong Law, with the Poisson Approximation as a By-Product. The Annals of Probability, 1, 910925. [963]

Hutter M. 2004 Universal Artificial Intelligence: Sequential Decisions Based On Algorithmic Probability ISBN 3-540-22139-5

Massari. F, Working Paper, Price Dynamics and Market Selection (2011).

T.M.Cover,J.A.Thomas. 1991, Elements of Information Theory. Wiley and Son, Inc ISBN 0-471-06259-6

Sandroni A. 2000. Do markets favor agents able to make accurate predictions? Econometrica 68:130342

Surowiecki J. 2004. The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations. New York: Random House

Williams D. 1991 Probability with Martingales. Cambridge University Press ISBN 978-0-521-40605-5