Learning in Crowded Markets *

[PRELIMINARY DRAFT]

Péter Kondor† Adam Zawadowski‡

Central European University Boston University

February 28, 2015

Abstract

We show that in markets which are prone to crowding, traders might devote excessive resources to learning about the size of the crowd. While the presence of more sophisticated traders in a market often leads to more efficient markets, it is nevertheless detrimental to welfare because of costly learning. The welfare loss due to over-learning is typically larger in markets with (i) weaker anchors, (ii) which load on crash-risk and (iii) where earlier entrants benefit more from later entrants. We heuristically classify popular hedge fund strategies along these dimensions. We also analyze when more sophisticated traders lead to over-entry or under-entry in equilibrium. Finally, we connect our result with the recent surge in exclusive trading venues such as dark-pools.

*For comments and suggestions, we are grateful to Patrick Bolton and seminar participants at Boston University, Central European University, ESSET 2014 (Gerzensee), Brigham Young University. We thank Pellumb Reshidi for his excellent research assistance. The support of the European Research Council Starting Grant #336585 is acknowledged.

†Central European University, Department of Economics, Nádor u. 9., Budapest, H-1051, Hungary, e-mail: kondorp@ceu.hu, http://www.personal.ceu.hu/staff/Peter_Kondor/index.php

‡Boston University School of Management, 595 Commonwealth Avenue, Boston, MA 02215, e-mail: zawa@bu.edu, http://www.people.bu.edu/zawa

1 Introduction

Global markets are increasingly dominated by hedge funds, a group often thought of as sophisticated traders. A fundamental question in finance is whether the larger share of sophisticated traders make markets more efficient and increase welfare. As Stein (2009) points out in his AFA Presidential Address, one caveat is that a broad class of quantitative trading strategies followed by these funds are "unanchored", i.e., a given manager cannot learn from prices in real time exactly how many others are using the same model and taking the same position. Indeed, there is a growing concern among analysts and practitioners that this can create a coordination problem, often labeled as "crowding", if managers following similar strategies inflict negative externalities on others. The August 2007 crash of quant funds is commonly attributed to over-crowded strategies (Khandani and Lo, 2011).[1] Furthermore, Sun et al. (2012) show that trades followed by many hedge fund managers do not perform as well as ones that are not imitated that much.

Our starting point in this paper is that sophisticated investors must be able to learn about how crowded the markets are, even though this learning might be very costly. Our main observation is that this leads to over-learning in general. Therefore, while the presence of more sophisticated traders in a market often leads to more efficient markets, it is nevertheless detrimental to welfare because of costly learning. The welfare loss due to over-learning is typically larger in markets with (i) weaker anchors, (ii) which are more prone to crash-risk and (iii) where earlier entrants benefit more from later entrants. We heuristically classify popular hedge fund strategies along these dimensions. We connect our result with the recent surge in exclusive trading venues such as dark-pools.

We present a novel model of entry with externalities and learning. In our model, early entrants gain from late entrants, while late entrants suffer from early entrants. Our main application is a micro-founded model of capital reallocation. Capital is inefficiently distributed over two islands and traders can ship capital from one location to the other and reap the gains of transferring the capital. Early entrants bid up the purchase price for late entrants, while late entrants increase the exit price

[1] In an early paper, MacKenzie (2003) also attributes the 1998 LTCM episode to crowded trades.

for early entrants who have to exit due to an idiosyncratic liquidity shock. The order in which the trader learns the existence of this trading opportunity is her type. Thus early entrants (higher types) can profit the most, while late entrants (lower types) might even lose on the trade. The market is also subject to an aggregate liquidity shock: if this materializes, all traders have to liquidate in a fire sale against a steep demand curve. Thus the expected loss in a crash is larger in a more crowded market. Before entry, the trader can spend resources to learn about her type, which is analogous to learning about how many traders have already entered (transferred capital), which in turn determines the attractiveness of the opportunity. However, such learning is costly. We assume that making the entry decision more state-contingent on her type (i.e. increasing its variance around unconditional entry level) has a constant marginal cost. This cost shows the strength of the price anchor in the given market. With a low cost to learn the price anchor is strong, while with a high cost the price anchor is weak.

Our main observation is that in markets which are subject to crowding, sophisticated traders engage in socially wasteful learning about the size of the crowd. While every trader is individually motivated to know whether she is ahead of the others, this is irrelevant for the planner. As a result, welfare and market efficiency will be affected very differently by more sophisticated traders. In particular, the mass of sophisticated traders needed to equalize return on capital across markets is too high from a welfare point of view. Market efficiency is achieved at the expense of excessive resources burnt in inefficient learning.

We also show that the welfare loss due to over-learning is larger in markets with weaker anchors, which are more prone to crash-risk and where late entrants are more important for profitability. We heuristically classify popular hedge fund strategies along these dimensions, arguing that carry trades might be among strategies most prone to over-learning.

We also show that more sophisticated traders can lead both to over-entry and under-entry in equilibrium depending on the characteristics of the market. Let us take a market which is prone to crowding and where late entrants are important in increasing profits of early entrants (e.g. carry trade or momentum strategies). In such markets, if the trade is unanchored, there is over-entry by the standard "tragedy of commons" argument. When there is a strong anchor, there is under-entry, because late entrants do not internalize that their presence would benefit early entrants. In our specification, as the mass of sophisticated investors grow, the effective per capita cost of learning decreases. Thus, the initial over-entry turns to under-entry.

Turning to exclusive markets, we asses how welfare changes with the mass of sophisticated traders. The critical observation is simple: as long as the mass of traders is small, they decide to enter without spending much resources on learning the size of the crowd. As a result, welfare is often larger with fewer sophisticated traders, even though the small number of sophisticated traders implies that prices are further from fundamentals. An interpretation of this result provides a rationale for the increasing popularity of exclusive trading venues such as black pools. The exclusivity can reduce participants' concern of crowding and lead to a drastic reduction in the resources traders spend in these markets to learn about the mass and timing of other entrants.

We show that our general model allows for a wide range of other applications differing in the source of the externalities and their welfare implications. For example, firms in industries with knowledge spillovers benefit from others following their location choices. Academics benefit from entering into fields early which later become very popular and thus increasing the number of their citations. Welfare consequences of crowding can be very different across these applications. In our baseline application of capital reallocation, consumers' welfare in local markets is not affected by crowding. However, in most other applications there is a group of traders whose welfare increases due to too much entry. For example, the over-entry of firms in an industry with knowledge spillover might decrease consumer prices in that industry below marginal cost. This harms firms but helps consumers. In our analysis, this effect makes markets with a weaker anchor more attractive from a welfare point of view.

Our paper is connected to various branches of literature. Stein (2009) introduces a simple model of crowding but leaves the effect of learning in this model for future research. Several papers deal with strategic entry in the face of learning: Acemoglu et al. (2011), Bolton and Farrell (1990), Fudenberg and Tirole (1985), Fudenberg and Tirole (1986), Hopenhayn and Squintani (2011), Myatt and Wallace (2012). These papers focus on the timing of entry and not the efficiency of the learning decision. Games of entry (Moinas and Pouget, 2011) and exit (Abreu and Brunnermeier, 2003) in bubbles analyze the effect of no price anchor on the development and sustainability of bubbles. Some papers focus on learning about the capacity of the whole market, such as Zeira (1994), Zeira (1999), and Rob (1991). These papers do not focus on the effect of relative position of entrants and how this interacts with the learning and entry decisions. Glode et al. (2012) focuses on inefficient over-learning and how it may lead to market breakdowns in the face of asymmetric information. Learning in the context of beauty contests is explored in Myatt and Wallace (2012) and Hellwig and Veldkamp (2009). In our model, like in the model of Hellwig and Veldkamp (2009), state-contingency in entry decisions are strategic complements, so learning decisions are also complements. In early papers, Tullock (1967) and Krueger (1974) show the high social costs of competition in rent seeking.

2 Learning and investing in crowded markets

In this part we describe our set up. We first present the reduced form setup we use in the paper and then describe a micro-foundation.

2.1 Payoffs

The heart of our model is an entry game with a continuum (mass M) agents, referred to as traders, each with a type $\theta \in [0,1]$. Each trader can decide to take an action: whether to enter in a market or not. θ is interpreted as the time when agent θ can make this decision. The utility gain (or loss, if negative) from entry is given by

$$\Delta u(\theta) = 1 + \alpha \cdot a(\theta) - \beta \cdot b(\theta) \tag{1}$$

where α and β are constants, $a(\theta)$ is the mass of entrants whose type is higher than θ (the agents who enter *a*fter agent θ), while $b(\theta)$ is the mass of entrants with a type lower than θ (agents who enter *b*efore agent θ). We assume $\beta + \alpha > 0$, such that entering earlier is better than later: we call this property rat-race. We also assume $\beta - \alpha > 0$, such that the average entrant imposes a negative externality on others: we call this property crowding. The two assumptions together imply that $\beta > 0$ while α could be positive or negative. As we specify below, players do not know their type, but can gather information about it through a costly learning process.

While throughout the paper we work with the reduced form payoff (1), to clarify the economic interpretation of the parameters α, β it is useful to build a fully specified economic model microfounding the reduced form (1). In the next part we present such a model in the context of capital arbitrage: this is our leading application. There are many other potential micro foundations of this reduced-form model, such as academic publication tournament, production with externalities and behavioral utility functions that reward early adoption of a trend; several of these are described in Section 4.

2.1.1 Leading application: Capital arbitrage

There are two islands A and B indexed by $i \in \{A, B\}$. There are two types of agents: a worker on each island and a continuum of traders with mass M uniformly distributed over types denoted by $\theta \in [0, 1]$. There are three types of goods: capital, a specialized consumption good produced by capital, and a numeraire good. The numeraire good is used as a method of exchange and all agents are endowed with sufficient numeraire goods to make transactions possible. Time is continuous and denoted by $t \in [0, 1]$. At $t = 0$ capital is inefficiently distributed, the worker on island A is endowed with $k_{A,0}$ capital, while the worker on island B has none ($k_{B,0} = 0$). We think of island B as an emerging idea/industry/country representing a profitable investment opportunity found sequentially by traders.

At time t, trader of type $\theta = t$ has the opportunity to buy capital on island A, transport it to island B and sell it there. The transport is successful with probability $1 - \nu$, with probability $\nu \geq 0$, the trader is hit by an idiosyncratic ("liquidity") shock and has to go back to island A and sell the capital there at $t = 1$. Furthermore, with ex ante probability $\eta \geq 0$, even if the capital transfer is successful, upon arriving on island B, there is an aggregate shock ("crisis") and all traders have to sell their capital in a fire sale. Thus the stock of capital owned by the worker on island i evolves over time as $k_{i,t}$. For simplicity, we assume workers are myopic and buy and sell capital at a price equal to its marginal product at the given capital level at the time.[2] Production only happens at $t = 1$: on island i, k_i capital produces c_i consumption good according to the production function

$$c_i = \gamma \cdot k_i - \delta_i \cdot \frac{k_i^2}{2} \tag{2}$$

thus on island i the marginal product of capital, given the capital level at time t, becomes:

$$\frac{dc_i}{dk_{i,t}} = \gamma - \delta_i \cdot k_{i,t}. \tag{3}$$

On island A, $\delta_i = \delta$ always. On island B, $\delta_i = \delta$ if there is no crisis and $\delta_i = \delta_c > \delta$ if there is a crisis (aggregate shock). All consumption happens at $t = 1$ after production has been completed. Traders only consume the n_j numeraire goods with utility $U_{trader} = n_j$. Workers on both islands consume n_i numeraire goods and c_i consumption goods with utility $U_i = n_i + c_i$.

Denote by $b(t)$ the mass of traders who chose to enter (i.e. engage in capital transport) before time t and $a(t)$ the mass of traders who enter after time t. Thus $k_{A,t} = k_{A,0} - b(t)$ and $k_{B,t} = b(t)$: there is more capital on island B if already $b(t)$ traders have decided to transport capital there from island A. Note that with probability ν the trader is reverted to island A and sells at the average marginal product of capital: $\nu \cdot (a(t) + b(t))$ capital is sold by traders at $t = 1$ in random order. [3] In case of a crisis, $(1 - \nu) \cdot (a(t) + b(t))$ capital is sold on island B in a fire sale, in random order. Overall, the

[2]This is a shortcut to capture that workers know even less than traders and thus cannot capture any of the surplus generated by trade.

[3]Note that this way the expected price is higher than if it was sold at a market clearing price in e.g. an auction. Thus this assumption simplifies the analysis by not allowing workers to capture any of the surplus.

profit of a trader that chooses to transport capital at time t is:

$$(1-\eta)\cdot(1-\nu)\cdot\underbrace{(\gamma-\delta\cdot b(t))}_{\text{sell price (no shock)}}+\eta\cdot(1-\nu)\cdot\underbrace{\left(\gamma-\delta_c\cdot(1-\nu)\cdot\frac{a(t)+b(t)}{2}\right)}_{\text{sell price (crisis)}}+$$

$$\nu\cdot\underbrace{\left[\gamma-\delta\cdot\left(k_{A,0}-[a(t)+b(t)]+\nu\cdot\frac{a(t)+b(t)}{2}\right)\right]}_{\text{sell price (with idiosyncratic shock)}}-\underbrace{\left[\gamma-\delta\cdot[k_{A,0}-b(t)]\right]}_{\text{buy price}} \tag{4}$$

Note that unless there is an idiosyncratic shock, traders buy capital at its marginal return on island A and sell capital at its marginal return on island B. Whenever this activity is profitable, it also decreases the difference between the marginal return on capital across the two islands, that is, it increases market efficiency. Idiosyncratic shock complicates this picutre only to the extent that it introduces some redistribution among traders; an element which washes out by aggregation. Therefore, we can interpret the aggregate profit of traders as a measure of market efficiency.

Choosing $k_{A,0}=\frac{1}{\delta\cdot(1-\nu)}$, the expected payoff of trader θ from transporting capital (given that trader θ can enter at time t) simplifies to (1) if

$$\alpha=\frac{1}{2}\cdot\left(\nu\cdot(2-\nu)-\eta\cdot(1-\nu)^2\right)\cdot\delta-\frac{1}{2}\cdot\eta\cdot(1-\nu)^2\cdot(\delta_c-\delta) \tag{5}$$

$$\beta=\frac{1}{2}\cdot\left((2-\nu)^2-\eta\cdot\left(1-\nu^2\right)\right)\cdot\delta+\frac{1}{2}\cdot\eta\cdot(1-\nu)^2\cdot(\delta_c-\delta) \tag{6}$$

Resulting in crowding and rat-race parameters of:

$$\beta-\alpha=(1-\nu)^2\cdot\eta\cdot(\delta_c-\delta)+(1-\nu)\cdot(2-\eta\cdot\nu-\nu)\cdot\delta \tag{7}$$

$$\alpha+\beta=(2-\eta(1-\nu)-\nu)\cdot\delta \tag{8}$$

To understand the effect of different assumptions, we look at several extreme cases. First assume there are no idiosyncratic, nor aggregate shocks ($\eta=\nu=0$). Then $\beta=\delta$, $\alpha=0$, thus $\alpha+\beta=$

$\beta - \alpha = \delta$. Thus the crowding ($\beta - \alpha > 0$) comes from the fact that there is a downward sloping demand ($\delta > 0$): the more capital is transferred, the less the average payoff to each trader. The rate race ($\alpha + \beta > 0$) also comes from the fact that there is downward sloping demand for capital: early traders get a much higher profit

In the second case, there are only idiosyncratic shocks ($\eta = 0$, $\nu > 0$). We get $\alpha > 0$, thus early entrants benefit from late entrants since if they have to liquidate their position, they can do so at a higher price. β decreases a bit because not all early entrants get to island B, thus the externality exerted on late entrants is weaker. Along this same logic, both rat race and crowding become a bit weaker.

Third, let us assume there are only aggregate shocks that trigger fire sales ($\eta > 0$, $\nu = 0$). We get $\alpha < 0$, thus early entrants suffer from late entrants because if there is a fire sale these will further depress prices. If fire sales are severe enough (δ_c large), then the presence of aggregate shocks increases β and lowers α since the any entry, irrespective of being earlier or later, harms all others in a fire sale. Along this logic, the crowding effect ($\beta - \alpha$) increase in the severeness of a fire sale. On the other hand, the rat race ($\alpha + \beta$) becomes somewhat weaker with aggregate shocks since in a fire sale the the ordering of traders is assumed to irrelevant.

Thus to be able to categorize markets, we have to understand how sensitive market prices are to the amount of capital present (δ) and what is the probability (and the effect) of idiosyncratic and aggregate shocks.

2.2 Learning

Before entry, traders can engage in costly learning about their type. Traders can choose probability $m(\theta)$ of getting signal $s = 1$ (enter) versus signal $s = 0$ (do not enter) subject to a learning cost. Thus $m(\theta)$ is the only choice variable (learning and entry strategy combined) which is the conditional probability of entry. Inspired by the rational inattention literature Yang (2011), the overall amount of learning is assumed to be:

$$L = \frac{1}{2}\int_0^1 (m(\theta) - p)^2 d\theta \tag{9}$$

and the marginal cost of improvement in learning is μ. Thus the learning cost function punishes the trader if it deviates from the average entry under some circumstances, i.e. if it makes the entry decision strongly dependent on its type θ. Learning about the traders' position θ can be interpreted very generally: traders use all information, including the price. We assume here that the price is in general not fully revealing and research or other data can help the trader refine its information.[4] While the above cost function is heuristic and chosen for analytical tractability, one can also use a cost function based on entropy that is the micro-founded using information theory. In Appendix A we show that the results are generally similar to the ones obtained with quadratic cost.

Note that if all traders choose the same informational structure, the mass of lower types entering ("before" trader θ) becomes:

$$b(\theta) = M \cdot \int_0^\theta m(\tilde{\theta}) d\tilde{\theta} \tag{10}$$

mass of higher types entering ("after" trader θ):

$$a(\theta) = M \cdot \int_\theta^1 m(\tilde{\theta}) d\tilde{\theta} \tag{11}$$

and the unconditional probability of entry for any trader is:

$$p = \int_0^1 m(\tilde{\theta}) d\tilde{\theta} = \frac{b(\theta) + a(\theta)}{M}. \tag{12}$$

thus $M \cdot p = b(\theta) + a(\theta)$ is the total mass of traders entering.

We assume that traders have to decide on the amount of information acquisition ex ante without any knowledge about the action of others. We interpret this as the cost of building an information

[4]Note that learning about θ and the amount of traders with lower θ that have chosen to enter, is the same, since each trader correctly infers how many of each θ type traders enter in equilibrium.

gathering and evaluation "machine" which includes the costs of gathering the right data, building the right contacts, research group, devising the best institutional practices, etc.

The expected profit of a trader (which it then consumes in the numeraire good), before taking into account the cost of information acquisition, is:

$$\Pi \equiv \int_0^1 m(\theta) \cdot \Delta u(\theta) \, \mathrm{d}\theta \tag{13}$$

Recall from section 2.1.1 that in our leading application, traders gain in the aggregate if and only if they reduce the difference in marginal returns of capital across locations, i.e. equate the price of capital in the two regions. Therefore, in this economy, total profit $M \cdot \Pi$ can be also interpreted as a measure of market efficiency.

The total expected payoff (value) per unit of trader is the profit from entering net of the ex ante learning cost:

$$V \equiv \Pi - \mu \cdot L \tag{14}$$

which is what traders maximize. Note that in a model with an information capacity constraint, μ can be interpreted as the shadow cost of learning: in this case L must equal the information capacity of the trader and μ is the Lagrange multiplier of the constraint. In equilibrium all surplus ends up with the trader: workers capture none, since they are myopic and sell/buy capital at a price equal to its marginal product. Thus the overall welfare in the whole economy can be computed as $W \equiv M \cdot V$.

2.3 Categorizing markets

We discuss what parameters different types of arbitrage trades correspond to. More specifically, what are the probable values of μ, $\beta - \alpha$ and α. We summarize our heuristic categorization in Table 1.

There is a simple interpretation for μ: using the language of Stein (2009) high μ represents unanchored strategies, while low μ represents anchored strategies. With low μ it is easy for the trader to determine how many traders have entered before, e.g. because the price's relation to the fundamentals

Table 1: **Categorizing trades**

market	μ	α	$\beta - \alpha$
twin stocks	low	high	medium
merger arbitrage	low	low	low
carry trade	high	high	high
on-the-run-off-the run bonds	low	low	medium
momentum	high	high	high
venture capital	high	low	low
IPO in novel industry	high	high	low

In this table we heuristically categorize markets according to the main parameters.

reveals this. Examples for a trade like this would be that of twin stocks or on-the-run-off-the-run bonds: it is clear from the price difference whether a trader is early (large price gap) or late (small price gap). Another example is merger arbitrage, where the price offered by the bidder is known. On the other hand, with high μ, it is very hard for the trader to determine whether to enter, e.g. because there is no clear price signal whether the trade is still profitable. Examples for such trades include: carry trade, momentum, January effect.

Trades with high α are ones where later entrants drive up the price at which earlier entrants can exit. Such trades include: twin stocks, the carry trade, momentum, January effect. Low α trades are ones that converge even without other traders entering, such that later entrants do not affect the profit of earlier entrants. Such trades include merger arbitrage, on-the-run-off-the-run bonds or venture capital.

One can also characterize trades by the amount of crowding $(\beta - \alpha)$ in the trade. That is how much does overall entry decrease the profitability of the trade. $(\beta - \alpha)$ is large, when there is a significant chance of an aggregate liquidity event resulting in a large fire sale. If these events are likely to happen in bad aggregate states that should increase $(\beta - \alpha)$ further. Such trades would include carry trades and potentially momentum strategies.

3 Model Solution

In this section we present our main results. First, we analyze the optimal strategies. Then, we turn to our main questions of how welfare changes as the mass of sophisticated traders increase through its effect on choice of learning and resulting profit. All proofs are relegated to Appendix B.

3.1 Optimal strategies

The problem of any trader is to choose its conditional entry $m(\theta)$ to maximize its value function V, which can be written as the following:

$$\max_{m(\theta)} \int_0^1 \Big(m(\theta) \cdot \Delta u(\theta) - \frac{\mu}{2} \cdot [m(\theta) - p]^2 \Big) \, d\theta \tag{15}$$

using the variation method one can derive an ordinary differential equation (ODE) for $m(\theta)$ with a boundary condition in the form of an integral equation.

Lemma 1. *The competitive symmetric equilibrium strategy $m(\theta)$ has two cases depending on the cutoff $\bar{M}_1 \equiv \frac{1}{\beta}$:*

1. *if $M \leq \bar{M}_1$: $m(\theta) = 1$ for all θ*
2. *if $M > \bar{M}_1$: there exists a non-empty subset $\Theta \subset [0, 1]$, s.t. $m(\theta) \in (0, 1)$ if and only if $\theta \in \Theta$ and for all $\theta \in \Theta$, $m(\theta)$ satisfies the ODE:*

$$M \cdot (\alpha + \beta) \cdot m(\theta) = -\mu \cdot m'(\theta) \tag{16}$$

with boundary condition (for any $\theta \in C(\Theta_s)$ where $C(.)$ denotes the closure of the set)

$$M \cdot (\alpha \cdot a(\theta) - \beta \cdot b(\theta)) + 1 = \mu \cdot (m(\theta) - p) \tag{17}$$

Equation 17 is basically the first order condition of the functional optimization stated in Eq. 15. To be able to solve this integral equation, we differentiate it to get a differential equation (Eq. 16) for $m(\theta)$ which has the original integral equation as its boundary condition.

Now we turn to the solution that a social planner would choose. We assume that the social planner can choose the amount of learning and entry for all agents. This gives us a benchmark against which we can evaluate learning and entry decisions in the competitive equilibrium. The main difference between the competitive solution and the social planer's one is that the social planner takes into account the externalities that traders exert on each other.

Lemma 2. *The social planner's optimal symmetric entry strategy $m_s(\theta)$ has two cases depending on a cutoff $\bar{M}_{1s} \equiv \frac{1}{\beta-\alpha} > 0$:*

1. *if $M \leq \bar{M}_{1s}$: $m_s(\theta) = 1$ for all θ*
2. *if $M > \bar{M}_{1s}$ $m_s(\theta)$ satisfies the ODE:*

$$0 = -\mu \cdot m_s'(\theta) \tag{18}$$

with boundary condition

$$M \cdot (\alpha - \beta) \cdot p_s + 1 = \mu \left(m_s(\theta) - p_s\right) \tag{19}$$

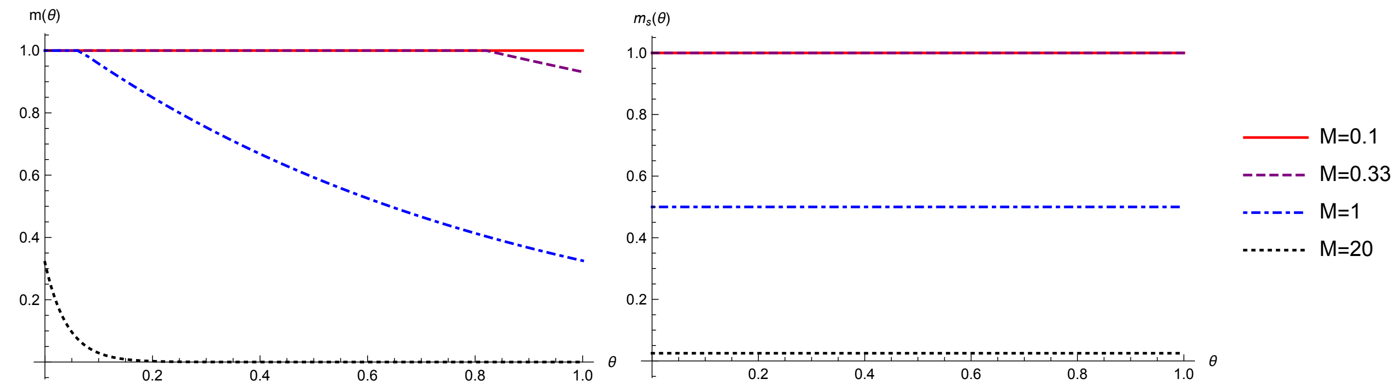
yielding to the solution

$$m_s(\theta) = \frac{1}{M} \cdot \frac{1}{\beta - \alpha} \tag{20}$$

Note that traders want to differentiate between states, but the planner does not. The planner chooses a flat entry function. Comparing ODEs 16 and 53 shows that the reason for this over learning is the rat race ($\alpha + \beta > 0$): every trader wants to know whether he is ahead of the other traders even if this is wasteful from the social planner's point of view.

Figure 1 shows the competitive and social planner's optimal entry function for different levels of M. Note that even when the solution is at least partially interior ($m(\theta) \in (0,1)$ for some θ), it might be a corner solution for some θ either at the value 0, 1 or both.

Figure 1: **Competitive and social entry functions for different M**



Entry functions for the competitive entry (m, left panel) and the social planners entry function (m_s, right panel). Parameters: $\beta = 4$, $\alpha = 2$, $\mu = 5$.

3.1.1 Anchored and unanchored strategies

To better understand the optimal strategies, we first analyze the the extreme cases of $\mu = 0$ (anchored trade) and $\mu = \infty$ (unanchored trade).

Lemma 3. *For fully anchored trades ($\mu = 0$), both the competitive and social planner's entry functions $m(\theta)$ are step functions, resulting in the first p traders entering. The competitive and social planner's optimum are, respectively:*

$$p|_{\mu=0} = \frac{1}{M \cdot \beta} \tag{21}$$

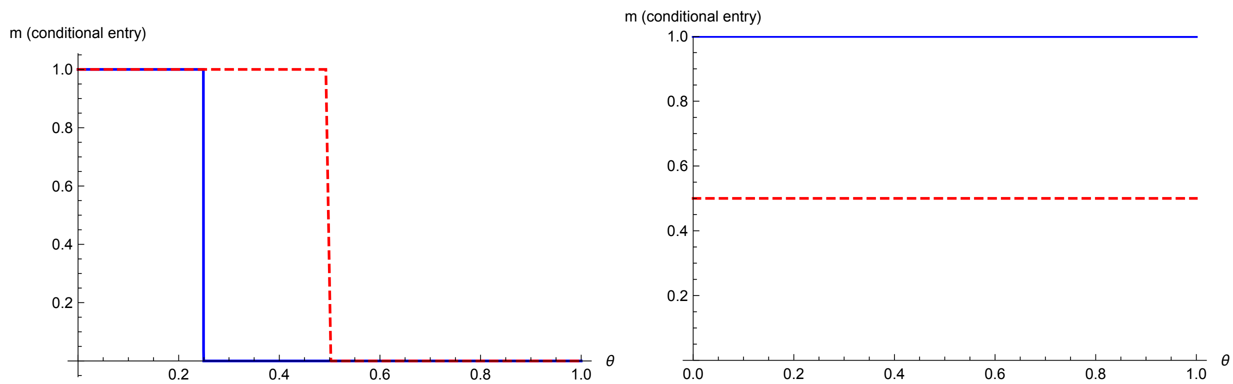
$$p_s|_{\mu=0} = \frac{1}{M \cdot (\beta - \alpha)} \tag{22}$$

For trades without any anchor ($\mu \to \infty$), both the competitive and social planner's entry functions $m(\theta)$ are flat. All traders enter with the same unconditional probability:

$$p|_{\mu\to\infty} = 2 \cdot \frac{1}{M} \cdot \frac{1}{\beta - \alpha} \tag{23}$$

$$p_s|_{\mu \to \infty} = \frac{1}{M} \cdot \frac{1}{\beta - \alpha}. \tag{24}$$

Figure 2: **Entry under perfect anchor and no anchor**



The first panel illustrates the entry decision under a perfect anchor ($\mu = 0$), while the second the entry decision under no anchor ($\mu = \infty$). The competitive choice is the solid (blue) line, the social planner's solution the dashed (red) line. Parameters: $\beta = 4$, $\alpha = 2$, $M = 1$.

Under perfect anchor there is competitive under-entry if $\alpha > 0$, since traders with higher θ do not take into account the positive effect of their entry on entrants with lower θ. If $\alpha = 0$, the competitive solution under perfect anchor is the same as the social planner's solution since it allows players to enter "consecutively". The last entrant takes into account the effect of all others but does not impose any externality on anyone else (since $\alpha = 0$), since there are no more entrants afterwards (thus β does not matter).

Under no anchor there is competitive over-entry similar to the "tragedy of commons" whenever $\beta > \alpha$. Thus one can think of this game as a generalization of the tragedy of commons game. In fact, the game becomes a simple "tragedy of commons" game if one sets $\alpha = -\beta$: the relative position of the trader does not matter and there is competitive over-entry.

Using the above analysis, on can draw implications about specific markets. Anchored trades (low μ) with high α do not have enough agents entering. Thus the model can give a potential explanation of why there is insufficient entry into trades like twin stocks and why mispricing persists. On the other hand, it also shows why unanchored trades (high μ), such as momentum, might see too much entry: in the extreme case of $\mu \to \infty$, even driving the profit of traders to zero, like in a tragedy of commons game.

3.2 Profit, learning and welfare as the proportion of traders increase

In this section we analyze how an increase in the mass of traders affects profit, learning and welfare in our competitive economy and in the planner's solution. We proceed in steps. In sections 3.2.1 and 3.2.2 we present analytical results for the cases when the mass is smaller than a given threshold and when the mass is larger than a given, higher, threshold, respectively. In section 3.2.3 we amend our results with numerical analysis to see the full picture.

3.2.1 Small mass of sophisticated investors

By Proposition 1 when $M < \bar{M}_1 \equiv \frac{1}{\beta}$, the agent and the planner in the decentralized equilibrium chooses to enter for sure without learning. Hence, all equilibrium objects are the same for the planner and in the decentralized solution. In particular, average entry of an agent is $p = 1$ thus expected total entry is M. Total revenue and welfare are

$$M \cdot \Pi = W = M \cdot \Pi_s = W_s = M - \frac{M^2 (\beta - \alpha)}{2}. \tag{25}$$

Note that neither the threshold $\bar{M}_1$, nor the total entry, profit or welfare depend on μ. The total mass of sophisticated investors is small in this range, hence they do not try to beat each other by learning about their relative type. Instead, all decide to enter without putting resources in learning. As their mass is marginally increasing, in terms of our microfoundation, they are able to allocate more capital to the new market, which increases the efficiency of capital allocation.

3.2.2 Large mass of sophisticated investors

In the following proposition, we show that, depending on μ, there is a threshold $\bar{M}_2(\mu)$, that optimal strategies are interior for any $M > \bar{M}_2(\mu)$. We also spell out these strategies.

Lemma 4. *When $\mu > 1 + \frac{\alpha}{\beta}$, there is a unique positive solution $\bar{M}_2(\mu)$ of the equation*

$$\frac{e^{\frac{M}{\mu}(\alpha+\beta)}\left(\frac{M}{\mu}\right)^2(\alpha+\beta)}{\frac{M}{\mu}\alpha + e^{\frac{M}{\mu}(\alpha+\beta)}(\frac{M}{\mu}\beta - 1) + 1} = M. \tag{26}$$

Whenever $M \geq \bar{M}_2(\mu)$ trader's optimal strategy is interior, $m(\theta) \in (0,1)$ and given by

$$m(\theta) = \frac{1}{M} \cdot \left(\frac{M}{\mu}\right)^2 \cdot \frac{(\alpha+\beta)\, e^{-\frac{M}{\mu}(\theta-1)(\alpha+\beta)}}{\left(\frac{M}{\mu} \cdot \beta - 1\right) e^{\frac{M}{\mu}(\alpha+\beta)} + \frac{M}{\mu} \cdot \alpha + 1}. \tag{27}$$

Furthermore $\bar{M}_2(\mu)$ is decreasing in μ, converges to $\frac{2}{\beta-\alpha}$ as $\mu \to \infty$ and converges to ∞ as μ goes to $1 + \frac{\alpha}{\beta}$ from above. When $\mu < 1 + \frac{\alpha}{\beta}$, there is no interior solution for any M.

The competitive entry function is exponential for $\mu > 1 + \frac{\alpha}{\beta}$ and $M \geq \bar{M}_2(\mu)$ because of the rat race $\alpha + \beta > 0$. This can be seen from Equation 16: If all other traders enter with a constant probability, the best response is to condition entry on θ linearly. However, if other traders also enter and learn according to a linear function, the best response is quadratic since traders seek to learn more (higher m') if other traders enter more (m higher), which is exactly what happens when θ is low. Etc. Thus for low θ there is more entry and locally more discrimination between states θ in equilibrium, yielding an exponentially decreasing entry function.

The next proposition describes total entry for $\mu > 1 + \frac{\alpha}{\beta}$ and $M \geq \bar{M}_2(\mu)$.

Lemma 5. *For $\mu > 1 + \frac{\alpha}{\beta}$ and $M \geq \bar{M}_2(\mu)$ total entry is*

$$M \cdot p = \frac{\frac{M}{\mu} \cdot (\alpha+\beta)\left(e^{\frac{M}{\mu}\cdot(\alpha+\beta)} - 1\right)}{(\alpha+\beta)\left((\frac{M}{\mu} \cdot \beta - 1)e^{\frac{M}{\mu}\cdot(\alpha+\beta)} + \frac{M}{\mu} \cdot \alpha + 1\right)} \tag{28}$$

which is monotonically decreasing in M. In the limit, $M \cdot p|_{M\to\infty} = \frac{1}{\beta}$.

Also, note that the total entry $M \cdot p$ only depends on $\frac{M}{\mu}$. Thus increasing the mass of traders M is the same as decreasing the learning cost μ, at least for interior solutions. The intuition is that with the quadratic learning cost, if we distribute the amount of learning among more agents, the incurred learning cost decreases, just as if we had decreased the learning cost itself but held the mass of traders M constant. We use this insight in the proofs.

From Eq. 20, it is easy to see that total entry under the planner is $M \cdot p_s = \frac{1}{\beta-\alpha}$. In our leading application of capital arbitrage, this is the entry which equalizes marginal return on capital across locations, that is, which corresponds to efficient markets. In the following proposition, we compare total entry under the planner and in our competitive solution.

Proposition 1. *For any given $\mu > 1 + \frac{\alpha}{\beta}$, let $M^*(\mu)$ be the unique positive solution of*

$$(M \cdot \alpha - \mu)e^{\frac{M\cdot(\alpha+\beta)}{\mu}} + M \cdot \beta + \mu = 0.$$

and μ^* the unique positive solution of

$$\frac{e^{\frac{M^*(\mu)}{\mu}(\alpha+\beta)}\left(\frac{M^*(\mu)}{\mu}\right)^2(\alpha+\beta)}{\frac{M^*(\mu)}{\mu}\alpha + e^{\frac{M^*(\mu)}{\mu}(\alpha+\beta)}(\frac{M^*(\mu)}{\mu}\beta - 1) + 1} = M^*(\mu).$$

1. *If $\alpha < 0$, there is over-entry, $p > p_s$ for all M.*

2. *If $\alpha > 0$,*

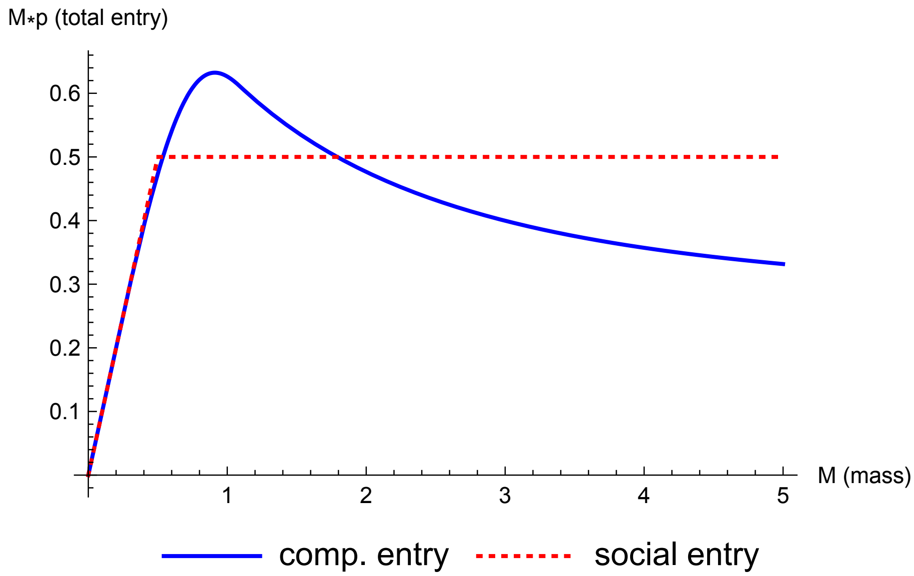
   (a) *whenever $\mu > \mu^*$, $M^*(\mu) > \bar{M}_2(\mu)$.*

   (b) *whenever $\mu^* > \mu$, there is under-entry, $p < p_s$, for any $M > \bar{M}_2(\mu)$.*

   (c) *whenever $\mu > \mu^*$ and $\alpha > 0$, there is over-entry, $p > p_s$ if $M^*(\mu) > M > \bar{M}_2(\mu)$, but there is underentry $p < p_s$ when $M > M^*(\mu)$.*

3. *In the limit, $\frac{M\cdot p|_{M\to\infty}}{M\cdot p^*|_{M\to\infty}} = 1 - \frac{\alpha}{\beta}$, hence, there is over-entry, iff $\alpha < 0$.*

Figure 3: **Total entry in number of traders** M



Total entry in the mass of entrants allowed to invest. The blue solid line is the competitive welfare. The red dotted line is the social welfare. Baseline parameters: $\beta = 4$, $\alpha = 2$, $\mu = 5$.

The proposition states that if μ is sufficiently large, the over-entry for small mass of traders turns to under-entry as the mass of traders grow, see Figure 3. The intuition is that as the mass of traders grows, it has the same effect as improving the anchor, i.e. decreasing μ. Decreasing μ means we move closer to the solution with perfect anchor in which there is under-entry whenever $\alpha > 0$.

Now we turn to welfare. Under the planner, the total value of entering (profit net of learning costs) coincides with welfare as the planner chooses not to learn:

$$M \cdot \Pi_s = W_s = \frac{1}{2(\beta - \alpha)} \tag{29}$$

Before we turn to the analytical expression for welfare under the competitive solution, the following Lemma presents the relationship between learning, profit and welfare in that case.

Lemma 6.

$$V = \frac{\Pi}{2} = \frac{\mu L}{2} \tag{30}$$

In the competitive equilibrium, the amount spent on learning is the same as the net value of entering because of the quadratic-linear optimization problem: the payoff integral is a linear functional of $m(\theta)$. To put it simply, we maximize $u = c \cdot x$ subject to the learning cost $\frac{\mu}{2}x^2$ by choosing the action x taking c as a given constant. Thus the FOC becomes: $c - \mu x = 0$, i.e. $x = \frac{c}{\mu}$. Thus payoff is $\Pi = \frac{c^2}{\mu}$, total learning $\mu L = \frac{c^2}{2\mu}$, thus $V = \Pi - \mu L = \frac{c^2}{2\mu}$. Note that value of entry net of learning costs is zero $V = 0$ if $\mu \to \infty$. This observation implies that analyzing how the profit changes as we increase the mass of traders also tells us how welfare changes. This is what we show in the next proposition.

Proposition 2. *For $\mu > 1 + \frac{\alpha}{\beta}$ and $M \geq \bar{M}_2(\mu)$:*

1. *if $\alpha < 0$, total profit, $M \cdot \Pi$, and welfare, W are monotonically increasing in M.*
2. *if $\alpha < 0$*
   - (a) *whenever $\mu < \mu^*$, total profit, $M \cdot \Pi$, and welfare, W are monotonically decreasing in M,*
   - (b) *whenever $\mu > \mu^*$ then total profit, $M \cdot \Pi$, and welfare, W are hump shaped, with a maximum at $M^*(\mu)$. At the maximum, the expected profit per trader is:*

$$M \cdot \Pi|_{M=M^*(\mu)} = M \cdot \Pi_s|_{M>\bar{M}_{1s}} = \frac{1}{2(\beta - \alpha)} \tag{31}$$

   which is equal to the total profit in the planner's case $M\Pi_s$. Welfare, W, is

$$W|_{\mu=\mu^*} = \frac{1}{4(\beta - \alpha)} \tag{32}$$

   and total entry is the same as under the planner at this point

$$M \cdot p|_{M=M^*(\mu)} = M \cdot p_s|_{M>\bar{M}_{1s}} = \frac{1}{\beta - \alpha}.$$

3. *In the limit*

$$W|_{M\to\infty} = \frac{\alpha + \beta}{4\beta^2} \tag{33}$$

and the total amount spent on learning is equal to $W|_{M\to\infty}$.

Thus if $\alpha > 0$ and μ is relatively large $(\mu > \mu^*)$, then in the region $M \geq \bar{M}_2(\mu)$, welfare has an interior maximum in M^* at exactly the same point as the payoff Π since $W = \frac{\Pi}{2}$. In this point, both total profit and total entry are the same under the planner and in the competitive equilibrium. This means that there is an optimal level of the mass of traders for which learning, even if socially wasteful on its own, is optimal because it balances the problem of over-entry and under-entry. That is, in our leading application, this is the point where markets are efficient. The market reaches the same total profit as the planner could, but welfare is lower, because the market reaches this outcome with wasteful learning.

From the last part of the proposition, note that regardless of μ, in the limit welfare goes to the same point, $\frac{\alpha+\beta}{4\beta^2}$. The intuition is that as M grows, μ becomes irrelevant. Nevertheless it is still true that half of the profit is spent on learning.

3.2.3 Overall effect

Now we are ready to turn to our main question: what happens to the welfare in a market as more and more sophisticated traders enter? We analyze this question in our model by increasing M, starting from 0. We show that the presence of some sophisticated trader unambiguously increases welfare in the competitive equilibrium. This remains the case for larger M in case of the social planner's optimum. However, in the competitive equilibrium, raising M can be very harmful to welfare. The reason is that as the amount of sophisticated traders in the market grows, these start worrying about crowding and their relative type θ and start learning about it. A rat race ensues with huge amounts wasted in learning costs and reduced welfare. Thus large amount of sophisticated traders leads to a drop in welfare not simply because of crowding but because of wasteful learning. We analyze how a permanent increase in the mass of (sophisticated) traders changes welfare in our economy.

We also add that, even after the initial increase in welfare, the effect of more sophisticated investors is non-monotonic. Typically, the minimal welfare is at an interim point, where the mass of traders is neither too high nor too low.

We also check how this picture changes with the nature of the arbitrage trade in question. In particular, is the detrimental effect of more sophisticated traders larger for unanchored trades (larger μ), in trades with high crash-risk (larger $\beta - \alpha$) or in markets where late entrants are important for profitability (larger α)?

As a starting point, we summarize the extreme cases of perfect anchor ($\mu = 0$) and no anchor in the next proposition.

Proposition 3. *With a perfect anchor, $\mu = 0$,*

$$W = \left\{ \begin{array}{ccc} M - \frac{M^2(\beta-\alpha)}{2} & if & M \leq \frac{1}{\beta} \\ \frac{(\alpha+\beta)}{2\beta^2} & & M > \frac{1}{\beta} \end{array} \right\}.$$

without anchor,

$$W = \left\{ \begin{array}{ccc} M - \frac{M^2(\beta-\alpha)}{2} & if & M \leq \frac{2}{\beta-\alpha} \\ 0 & if & \frac{2}{\beta-\alpha} < M. \end{array} \right\}.$$

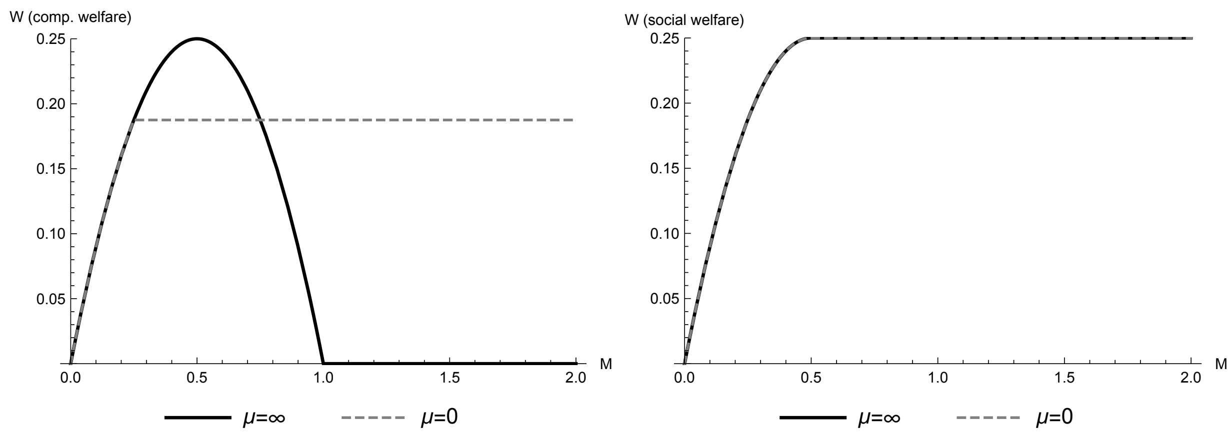
The next proposition summarizes our previous results for the planner for any μ.

Proposition 4. *Under the planner, for any μ*

$$W_s = \left\{ \begin{array}{ccc} M - \frac{M^2(\beta-\alpha)}{2} & if & M \leq \frac{1}{\beta-\alpha} \\ \frac{1}{2(\beta-\alpha)} & & otherwise \end{array} \right\}.$$

Figure 4 illustrates the two propositions showing the change in welfare as the function of the mass M of traders. With full information, welfare monotonically increases up to the point $\bar{M}_1$ and is flat afterwards at the level $W = \frac{(\alpha+\beta)}{2\beta^2}$. Qualitatively, it is a similar picture to the planner's outcome, except that, for $\alpha > 0$ ($\alpha < 0$), we have permanent over-entry (under-entry) lowering welfare compared to the social level. Naturally, there is no over-learning in this case. Without anchor, welfare follows the same increasing path up to $\bar{M}_1$, but continues to grow as M increases and peaks at $M^{**} \equiv \frac{1}{\beta-\alpha}$ where $W = \frac{1}{2(\beta-\alpha)}$. As the mass of traders increases further, all profits are gradually competed away and $W = 0$ for any $M > \frac{2}{\beta-a}$. It is immediately clear from the Figure 4 that perfect anchor is not always preferred to no anchor.

Figure 4: **Competitive and social welfare with extreme μ**

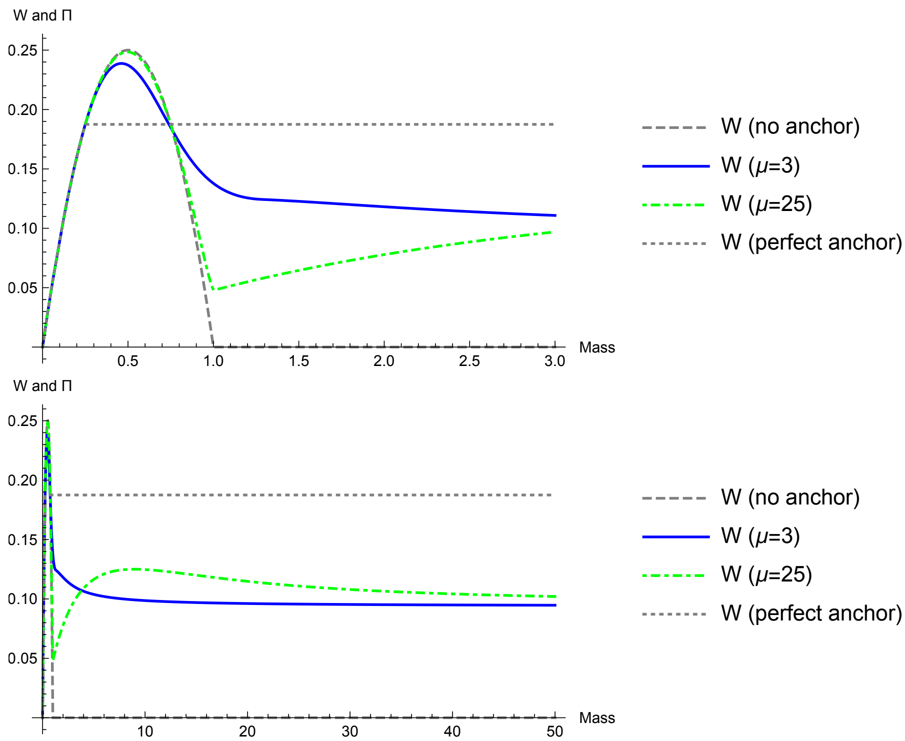


Welfare for competitive (left) and social (right) entry as a function of the mass M of traders. The no anchor ($\mu = \infty$) case is represented by the solid line, the perfect anchor ($\mu = 0$) case by the dashed line. Parameters: $\beta = 4$, $\alpha = 2$.

Figure 5 shows a similar plot, but for interim levels of anchoring. (We also show the perfect anchor case and the no anchor case for comparison as light curves.) We show two typical cases. The solid curve represents a market where anchoring is strong ($\mu < \mu^*$ in terms of Proposition 2), while the heavy dashed curve represents a market where anchoring is weak ($\mu > \mu^*$). We also show the critical points $\bar{M}_1, M^{**}, \bar{M}_2(\mu), M^*$ as we introduced in the previous section. First note that for $M < \bar{M}_1$ and for $M > \bar{M}_2$ both curves behave as stated in sections 3.2.1 and in Proposition 2. That is, both increases monotonically for $M < \bar{M}_1$, and for $M > \bar{M}_2$ welfare is decreasing when anchoring is strong and hump shaped when it is weak.

While, we do not have analytical results for $\bar{M}_2 > M > \bar{M}_1$, it is clear from the figure that the same forces shape this part as the ones driving the limit cases in Proposition 3. Indeed, both curves are

Figure 5: **Welfare in number of traders M at different levels of μ**



Welfare in the mass of entrants allowed to invest. The blue solid line is the competitive welfare. The gray dotted line is the welfare with $\mu = 0$, while the gray dashed line is that with $\mu \to \infty$. The second graph plots M until higher levels. Baseline parameters: $\beta = 4$, $\alpha = 2$, $\mu = 3$ and 25.

between the two limit cases and the curve where anchoring is weaker is closer to the no-learning case. Also, similarly, to the no learning case, we see a monotonic increase for small M up to a point close to M^*, turning to a sharp decrease decrease. The main qualitative difference between the two cases is that when anchoring is strong, welfare remains monotonically decreasing, while when anchoring is weak, there is a second local maximum at M^*. As we discussed in the previous section, the main reason for that that for weak anchoring, the competitive outcome might reproduce the optimal total entry at M^*, while there is no such point when anchoring is strong.

Note that our curves for different anchoring are getting closer as the mass of traders increase. Indeed, as Proposition 2 states, for any finite μ, welfare converges to the same point $\frac{\alpha+\beta}{4\beta^2}$. Thus, there is a the discontinuity for $M \to \infty$ between very large μ and no anchor. Without learning, welfare is zero for any $\frac{2}{\beta-\alpha} < M$. For finite, but large μ, as the mass of traders increase, the level of μ eventually becomes irrelevant.

With the help of the next proposition, we gather intuition on how the parameters β and α influences the affect of more sophisticated traders on welfare.

Proposition 5. *In the competitive equilibrium, for very large μ,*

$$\lim_{\mu\to\infty} \max_M W = \frac{1}{2(\beta-\alpha)} \tag{34}$$

$$\lim_{\mu\to\infty} W|_{\bar{M}_2(\mu)} = 0 \tag{35}$$

$$\lim_{\mu\to\infty} \lim_{M\to\infty} W = \frac{\alpha+\beta}{4\beta^2}. \tag{36}$$

Therefore,

$$\begin{aligned} W^s|_{M>\bar{M}_{1s}} - \lim_{\mu\to\infty} W|_{\bar{M}_2(\mu)} &= \max_M W - \lim_{\mu\to\infty} W|_{\bar{M}_2(\mu)} \to \frac{1}{2(\beta-\alpha)} \\ \frac{W^s|_{M>\bar{M}_{1s}}}{\lim_{\mu\to\infty}\lim_{M\to\infty} W} &= \frac{\max_M W}{\lim_{\mu\to\infty}\lim_{M\to\infty} W} \to 2\left(1+\frac{\alpha^2}{\beta^2-\alpha^2}\right) \end{aligned}$$

The proposition characterizes the welfare loss at the point when welfare is the lowest $W|_{\bar{M}_2(\mu)}$, and the welfare loss at the point where mass of traders increase without limit, $\lim_{M\to\infty} W$ for unanchored strategies. We show that results do not depend on whether the benchmark is the social planner's solution or the welfare in the competitive market with the optimal number of entrants.

In both cases, we find that the maximum welfare loss is larger, in markets with less crash risk (small $\beta-\alpha$). The welfare loss in the large mass limit is also decreasing in crash risk, but it is larger in markets where late entrants are more important for the profitability of the trade (large α). The first part of the Proposition gives the intuition for the effect of crash risk. Larger $(\beta-\alpha)$ decreases the loss, because it decreases the benefit of sophisticated traders at the first place. That is, $\max_M W$ is decreasing in $(\beta-\alpha)$ for unanchored strategies. Larger α increases the loss at very large mass, because the benefit of late entrants is responsible for the externality resulting underentry for very large mass.

In the micro-foundation we saw that the deep parameters of the model influence α and β in a non-trivial way. In Proposition 6, we show that if only idiosyncratic or only aggregate shocks are present then a higher probability of these shocks implies a larger welfare drop for the maximum to that at $M \to \infty$. The intuition is that the welfare drop is at least partially driven by over- and under-entry due to α when M is high (thus effectively μ is low). Since both shock increase α in absolute value, they both mean welfare deteriorates more due to the presence of more sophisticated agents.

Proposition 6. *In the microfounded model, the relative welfare loss* $\frac{\max_M W}{\lim_{\mu\to\infty}\lim_{M\to\infty} W}$ *due to entry of many sophisticated agents:*

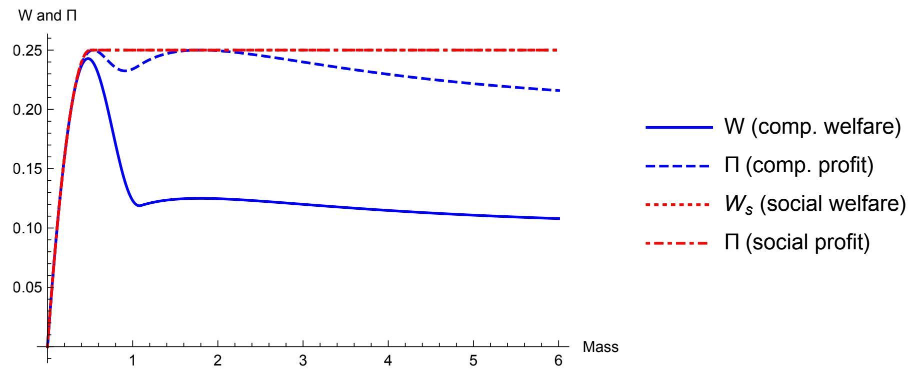
1. *increases in the probability of the idiosyncratic shock ν (if there is no aggregate shock $\eta = 0$)*
2. *increases in the probability of the aggregate shock η (if there is no idiosyncratic shock $\nu = 0$)*
3. *increases in the severity of the fire sale in the aggregate shock δ_c (if $\nu = 0$)*

Next we turn to the decomposition of welfare in the comparative market to the profit part and the learning part. An interesting insight is shown on Figure 6: while the profit of traders might be (weakly) increasing as M increases, the gain in profit is more than offset by the loss from learning. Thus while learning is good in limiting crowding, it does use large amounts of resources. In the context of our leading application, it implies that more sophisticated traders might make markets more efficient and decrease welfare at the same time. It is so, because the increased market efficiency is achieved at the cost of over-learning.

3.3 Exclusive markets

If there are other ways to limit entry behind learning, that might be welfare improving. We limit the mass of traders to $p_P \cdot M$ before they learn their type. Entry is limited indiscriminately, thus

Figure 6: **Welfare and profit in the number of traders** M



Welfare in the mass of entrants allowed to invest. The blue solid line is the competitive welfare. The red dotted line is the social welfare. Baseline parameters: $\beta = 4$, $\alpha = 2$, $\mu = 5$.

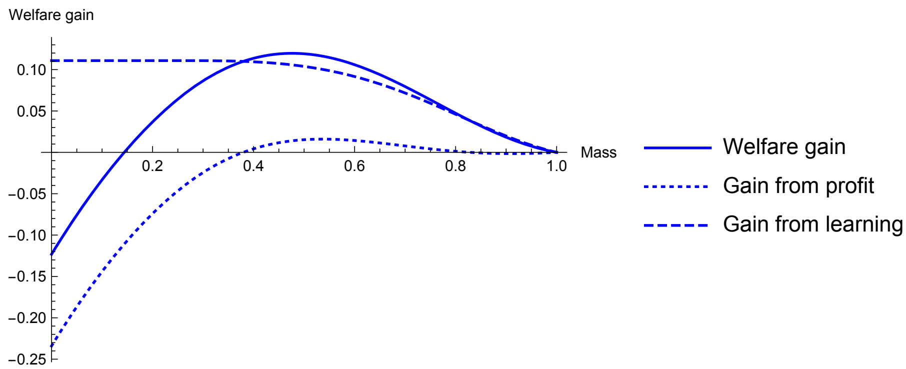
the traders who are allowed to enter still have a uniform distribution over θ. Proposition 7 gives a sufficient condition for limited entry to be welfare improving.

Proposition 7. *If $\alpha \in [-\frac{\beta}{\sqrt{2}}, \frac{\beta}{\sqrt{2}}])$, charging an entry fee of $\frac{(\beta+\alpha)}{2\beta^2}$ is welfare improving for any μ for which the entry decision is originally an interior solution. For low μ, the welfare improvement is due to the policy discouraging learning, not from limiting crowding.*

The intuition is that limiting entry also limits the incentive to learn since if there are sufficiently small numbers of traders who can enter, traders will low enough θ choose to enter all the time since they are sure entry yields positive payoff. Mathematically, m becomes a corner solution instead of an interior one. While the Proposition is only a sufficient condition, Figure 7 shows the decomposition of welfare gain from the policy at different levels of M allowed to enter. The figure clearly shows that the welfare gain happens despite reduced profits from restricted entry and is thus solely due to welfare gains from limited learning.

In terms of our lead application from Section 2.1.1, limiting entry means that the marginal product of capital on the two islands ((3)) does not converge as much as if entry would be unrestricted. An example for such an arrangements of limited entry are dark pools which are modern equivalents of upstairs market, designed for secret executions of large, uninformed trades. These are often attacked

Figure 7: **Dissecting welfare gain due to exclusivity.**



Welfare gain in the mass of entrants allowed to invest, compared to the benchmark when all are allowed to invest. Baseline parameters: $\beta = 4$, $\alpha = 2$, $\mu = 5$.

because of their weak anchor and exclusivity but as Proposition 7 shows, this might be an optimal design to avoid a harmful rat-race (i.e. front-running).

Note that in a more general model where workers (i.e. consumers or workers) benefit from entry of traders, the welfare analysis changes, see Section 4.4.

4 Additional microfoundations and benefits of a weak anchor

4.1 Consumer/producer surplus

It is reasonable to assume that traders are not the only agents in the economy who can benefit from the entry of traders. Here we show a few examples where passive traders (consumers or producers) benefit from entry and thus the welfare consequences of over-entry change. First we present a model of competition for scarce resources and then a model of publication tournaments. In Section 4.4 we then present the welfare implication of the presence of these agents.

Assume there are some passive investors with a payoff (value) of

$$V^P \equiv \alpha_w \cdot \int_0^1 m(\theta) \cdot a(\theta)\, \mathrm{d}\theta = \alpha_w \cdot \frac{(M \cdot p)^2}{2} \tag{37}$$

with some parameter $\alpha_w \geq 0$. Thus the overall welfare in the whole economy can be computed as:

$$W \equiv M \cdot V + V^P \tag{38}$$

One can interpret all our results up to now as a special case setting $\alpha_w = 0$. Thus the relevant payoff of entry from the social planner's perspective becomes:

$$\Delta u = (\alpha + \alpha_w) \cdot a(\theta) - \beta \cdot b(\theta) + 1 \tag{39}$$

Note that the traders' competitive payoff function is unchanged, i.e. it does not include α_w.

4.2 Product market with scarce resources

The producers produce a single good traded in a competitive market. Heterogenous producers that compete for scarce resources thus producing the good is subject to local spill-overs. In this example the producers are the active traders making the entry decision while the consumers are passive.

The goods are purchased by a representative consumer with quadratic utility:

$$U(q) = q - \frac{\alpha_w}{2} q^2 \tag{40}$$

such that $MU = p$ yields linear demand for the good.

$$\text{price}(q) = 1 - \alpha_w \cdot q \tag{41}$$

A unit mass of firms indexed by θ decide to move to a specific area (e.g. the new Silicon Valley) where other firms might also move. Producers are heterogenous θ and lower θ producers have lower costs both because they have better technology and can also secure a better (heterogenous) input. The cost of building the plant is subject to weakly increasing marginal cost (and price of the building). Lower θ consumers manage to purchase the land earlier, profit from moving.

$$\text{cost}(\theta) = (\alpha + \beta)b(\theta) - (\alpha + \alpha_w)(a(\theta) + b(\theta)) \tag{42}$$

The $(\alpha + \beta)$ term multiplying $b(\theta)$ of the cost function comes from the price of the production input (e.g. land) which increases as more (and worse) type producers enter. Better types can choose a better quality input before the others at a fixed price (given by the value of external use). The $\alpha + \alpha_w$ term multiplying $(a(\theta) + b(\theta))$ captures network externalities that some resources might become cheaper if many producers use it because of economies of scale.

The payoff of firm θ conditional on moving is:

$$\Delta u = \text{price}(q) - \text{cost}(\theta) = 1 - \alpha_w \cdot q - (\alpha + \beta)b(\theta) + (\alpha + \alpha_w)(a(\theta) + b(\theta)) \tag{43}$$

Where in equilibrium the amount of goods produced depends on how many firm decide to produce: $q = a(\theta) + b(\theta)$ yielding the payoff function of Equation 1 in the reduced form.

The payoff of the passive trader in this example is the consumer surplus which can be computed as:

$$\text{CS} = \int_0^q (1 - \alpha_w \cdot \hat{q} - \text{price}(q))d\hat{q} = \int_0^q \alpha_w \cdot (q - \hat{q})d\hat{q} = \alpha_w \cdot \frac{q^2}{2} \tag{44}$$

A different way to compute the consumer surplus is to add a term of $\alpha_w \cdot a(\theta)$ (alternatively $\alpha_w \cdot b(\theta)$) to Δu

$$\text{CS} = \int_0^1 m(\theta) \cdot [\alpha_w \cdot a(\theta)]\, d\theta = -\frac{\alpha_w}{2}a(\theta)^2\Big|_0^1 = \frac{\alpha_w}{2}a(1)^2 = \alpha_w \cdot \frac{q^2}{2} \tag{45}$$

This means that the objective function for the social planner maximizing welfare is the same as (1) substituting α with $\alpha + \alpha_w$ yielding (39).

4.3 Academic publications

The simplest interpretation is an academic tournament: e.g. the strategic choice of field of an aspiring academic. The academic wants to choose a topic that has not yet been done and that will have many followers who cite him. Both reading through and understanding the previous literature is time consuming and also trying to figure out whether others will find the same topic interesting and cite you. The academic's payoff is the probability of publishing and being cited. θ can be interpreted as time. If lots of researchers finish their paper before (or other old but similar papers are discovered), the lower your chance of publication: this is captured by β. If lots of researchers write a paper on the same topic afterwards that increases citation and the academic's chance of publication, captured by α. Thus the payoff to the academic (producer of knowledge) is:

$$\Delta u = \alpha \cdot a(\theta) - \beta \cdot b(\theta) + 1 \tag{46}$$

where we assumed that the fixed payoff already incorporates the fixed cost of producing knowledge. This yields Equation 1 in the reduced form model. Adding a consumer of knowledge with quadratic utility:

$$U(q) = q - \frac{\alpha_w}{2} q^2 \tag{47}$$

yields a similar additional term in welfare as in Section 4.2.

4.4 Welfare and opaqueness

Thus crowding is not necessarily welfare decreasing if there is a consumer or producer in the economy who benefits from higher consumption or production. While there is over-entry from the traders' point of view, from a welfare point of view this might be beneficial. Since $\alpha + \beta > 0$, opaqueness is never preferred if there is no passive agent ($\alpha_w = 0$).

Proposition 8. *Welfare is higher in case of no anchor than in case of perfect anchor if*

$$\alpha_w > \frac{(\beta - \alpha)^2(\alpha + \beta)}{\alpha(2\beta - \alpha)} \tag{48}$$

If the denominator is negative, it is true for all $\alpha_w \geq 0$*.*

Note that while in Proposition 7 we assume α_w is zero it is also true for sufficiently small α_w since the additional term in the welfare V_P is of the form $\alpha_w \cdot \text{const}$.

5 Conclusions

We present a tractable model of joint learning and entry in a crowded market. Costly over-learning in general alleviates the effect of crowding. Furthermore, a policy intervention to drastically restrict ex ante entry can also be welfare improving: not simply because it alleviates crowding but because it reduces the incentive to over-learn, which is costly in terms of welfare.

References

Abreu, Dilip and Markus K. Brunnermeier, "Bubbles and Crashes," *Econometrica*, 2003, *71* (1), 173–204.

Acemoglu, Daron, Kostas Bimpikis, and Asuman Ozdaglar, "Experimentation, Patents, and Innovation," *American Economic Journal: Microeconomics*, 2011, *3*, 37–77.

Bolton, Patrick and Joseph Farrell, "Decentralization, Duplication, and Delay," *Journal of Political Economy, Vol. 98, No. 4 (Aug., 1990)*, 1990, *98* (4), 803–826.

Fudenberg, Drew and Jean Tirole, "Preemption and Rent Equalization in the Adoption of New Technology," *Review of Economic Studies*, 1985, *52* (3), 383–401.

_ and _ , "A Theory of Exit in Duopoly," *Econometrica*, 1986, *54* (4), 943–960.

Glode, Vincent, Richard C. Green, and Richard Lowery, "Financial Expertise as an Arms Race," *Journal of Finance*, 2012, *67* (5).

Hellwig, Christian and Laura Veldkamp, "Knowing what others know: Coordination motives in information acquisition," *Review of Economic Studies*, 2009, *76*, 223–251.

Hopenhayn, Hugo A. and Francesco Squintani, "Preemption Games with Private Information," *Review of Economic Studies*, 2011, *78*, 667–692.

Khandani, Amir E and Andrew W Lo, "What happened to the quants in August 2007? Evidence from factors and transactions data," *Journal of Financial Markets*, 2011, *14* (1), 1–46.

Krueger, Anne O., "The Political Economy of the Rent-Seeking Society," *American Economic Review*, 1974, *64* (3), pp. 291–303.

MacKenzie, Donald, "Long-Term Capital Management and the sociology of arbitrage," *Economy and Society*, 2003, *32* (3), 349–380.

Moinas, Sophie and Sebastien Pouget, "The Bubble Game: An Experimental Study of Speculation," 2011. University of Toulouse working paper.

Myatt, David P. and Chris Wallace, "Endogenous Information Acquisition in Coordination Games," *Review of Economic Studies*, 2012, *79*, 340–374.

Rob, Rafael, "Learning and Capacity Expansion under Demand Uncertainty," *Review of Economic Studies*, 1991, *58* (4), 655–675.

Stein, Jeremy C., "Presidential Address: Sophisticated Investors and Market Efficiency," *Journal of Finance*, 08 2009, *64* (4), 1517–1548.

Sun, Zheng, Ashley Wang, and Lu Zheng, "The Road Less Traveled: Strategy Distinctiveness and Hedge Fund Performance," *Review of Financial Studies*, 2012, *25* (1), 96–143.

Tullock, Gordon, "The welfare costs of tariffs, monopolies, and theft," *Economic Inquiry*, 1967, *5* (3), 224–232.

Woodford, Michael, "Inattention as a Source of Randomized Discrete Adjustment," 2008. New York University working paper.

Yang, Ming, "Coordination with Rational Inattention," 2011. Duke University working paper.

Zeira, Joseph, "Informational Cycles," *Review of Economic Studies*, 1994, *61*, 31–44.

_ , "Informational Overshooting, Booms and Crashes," *Journal of Monetary Economics*, 1999, *43*, 237–257.

A Using entropy

As a robustness check we specify the learning cost to be based on entropy. The advantage of this cost specification is that it can be derived based on information theory and is not simply assumed like the quadratic cost specification in the main text. We use the framework of Yang (2011) and Woodford (2008) who show that the optimal signal is $s \in \{0, 1\}$ where $s = 1$ with probability $m(\theta)$. The optimal entry decision conditional on the signal is: enter if $s = 1$, stay out if $s = 0$. Thus traders enter with conditional probability $m(\theta)$.

Yang (2011) shows that the learning cost is specified as:

$$L = \int_0^1 (m(\theta) \log[m(\theta)] + (1 - m(\theta)) \log[1 - m(\theta)]) \, d\theta - p \log [p] - (1 - p_1) \log [1 - p]. \tag{49}$$

Following Yang (2011), the first order condition for entry is:

$$\Delta u(\theta) = \mu \cdot \left[\log \left(\frac{m(\theta)}{1 - m(\theta)} \right) - \log \left(\frac{p}{1 - p} \right) \right] \tag{50}$$

We now show that the entry functions are similar to those we got in case of the quadratic case. One can solve for the resulting ODE but cannot in general solve for it in closed form and the boundary condition is only given by an implicit equation.

Lemma 7. *The competitive equilibrium strategy $m(\theta)$ in the symmetric equilibrium has to solve the differential equation*

$$(M \cdot \alpha + M \cdot \beta) \cdot m(\theta) = -\frac{\mu \cdot m'(\theta)}{m(\theta) \cdot (1 - m(\theta))} \tag{51}$$

with the boundary condition

$$M \cdot \alpha \cdot p + 1 = \mu \cdot \left[\log \left(\frac{m(0)}{1 - m(0)} \right) - \log \left(\frac{p}{1 - p} \right) \right] \tag{52}$$

Lemma 8. *The socially optimal strategy $m_s(\theta)$ in the symmetric equilibrium that solves the differential equation*

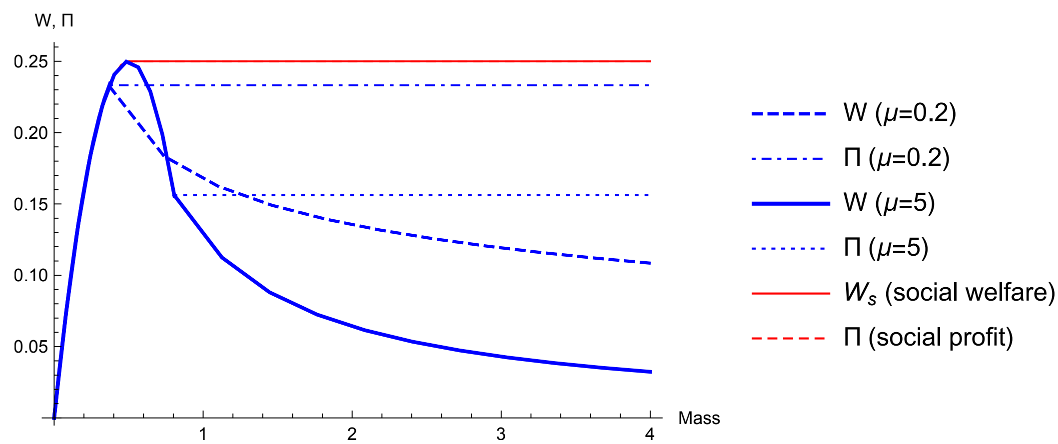
$$0 = -\frac{\mu \cdot m_s'(\theta)}{m_s(\theta) \cdot (1 - m_s(\theta))} \tag{53}$$

subject to the boundary condition

$$(\alpha - \beta) \cdot p_{1s} + 1 = \mu \cdot \left[\log\left(\frac{m_s(0)}{1 - m_s(0)} \right) - \log\left(\frac{p_{1s}}{1 - p_{1s}} \right) \right]. \tag{54}$$

Comparing Lemmas 7 and 8 it is obvious that there is excessive learning in the competitive equilibrium since the social planner would prefer no learning at all.

Figure 8: **Welfare in the number of traders M, using entropy**

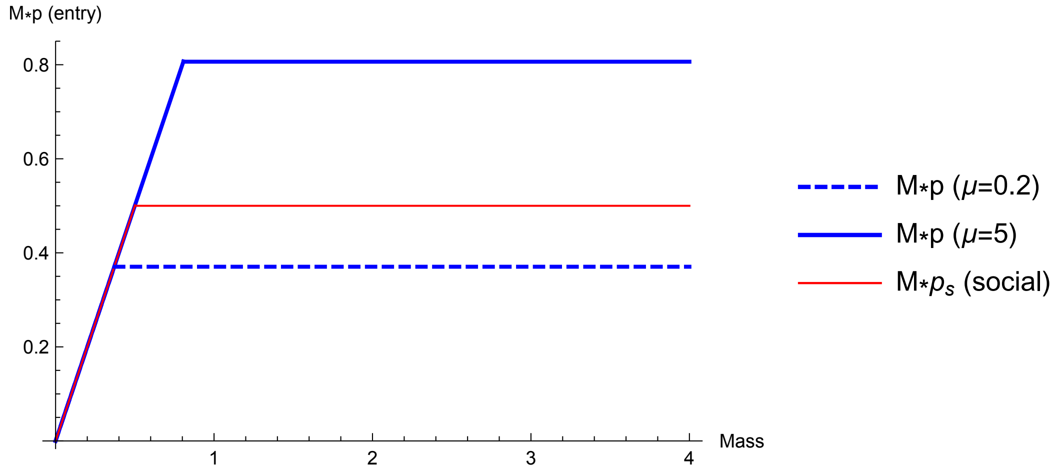


Welfare in the mass of entrants allowed to invest. The thick blue solid line is the competitive welfare. Baseline parameters: $\beta = 4$, $\alpha = 2$, $\mu = 0.2$ and 5.

Since these ODE's do not have a closed form solution, all further analysis is fully numeric. Figure 8 shows that as more sophisticated agents enter, welfare eventually deteriorates even though the profit of the market (and thus efficiency) is close to unchanged. This result is consistent with our main result obtained using the quadratic cost function.

The major difference between the two cost functions can be seen in Figure 9: overall entry does not diminish as M increases. The reason is that while with the quadratic cost function if one splits learning among many agents, it becomes ver cheap in the limit. This intuition does not hold for

Figure 9: **Total entry in the number of traders M, using entropy**



Total entry in the mass of entrants allowed to invest. The thick blue solid line is the competitive welfare. Baseline parameters: $\beta = 4$, $\alpha = 2$, $\mu = 0.2$ and 5.

entropy; it is the relative probabilities of entry that count and splitting learning does not decrease the costs.

B Proofs

Proof of Lemma 1

Proof. First we show that if only M traders are present in the market and all enter (indiscriminately with $m(\theta) = 1$) then no trader has an incentive to learn if M is low enough. This is because even the last trader with $\theta = 1$ has an incentive to enter if:

$$-M \cdot \beta + 1 \geq 0 \tag{55}$$

thus the maximum mass of traders that can enter without anyone having an incentive to acquire information is

$$\tilde{M} = \frac{1}{\beta} \tag{56}$$

Note that we need $M \cdot \beta > 1$ to ensure $p_P > 0$.

Using the variation method (following the proof of Yang (2011)), the optimal solution of the problem has to satisfy the following first order condition:

$$\Delta u = \alpha \cdot \int_{\theta}^{1} \tilde{m}(\tilde{\theta}) d\tilde{\theta} - \beta \cdot \int_{0}^{\theta} \tilde{m}(\tilde{\theta}) d\tilde{\theta} + 1 = \mu \cdot (m(\theta) - p) \tag{57}$$

where $\tilde{m}$ is the choice of other agents. We focus on symmetric equilibria, thus we assume all others chose the same $\tilde{m}$.

Denote the strategy function of all other players as $\tilde{m}(\theta)$. Substituting our payoff functions into Equation 57:

$$M \cdot \alpha \cdot \int_{\theta}^{1} \tilde{m}(\tilde{\theta}) d\tilde{\theta} - M \cdot \beta \cdot \int_{0}^{\theta} \tilde{m}(\tilde{\theta}) d\tilde{\theta} + 1 = \mu \cdot (m(\theta) - p)\,. \tag{58}$$

Differentiating this we arrive at the differential equation:

$$(M \cdot \alpha + M \cdot \beta) \cdot \tilde{m}(\theta) = -\mu \cdot m'(\theta). \tag{59}$$

Imposing symmetry $\tilde{m}(\theta) = m(\theta)$ results in Equation 16. The boundary condition is given by the original integral-differential Equation 58 evaluated at any θ: in Equation 17 we set $\theta = 0$.

The differential equations pins down the solution for interior μ up to a constant C:

$$m(\theta) = Ce^{-\frac{\theta(M\cdot\alpha+M\cdot\beta)}{\mu}} \tag{60}$$

where the constant C is pinned down by the boundary condition at $\theta = 0$ yielding

$$\begin{aligned} M\cdot\alpha p + 1 &= \mu\cdot(m(0)-p) \\ M\cdot\alpha p + 1 &= \mu\cdot(C-p) \\ C &= \frac{1}{\mu}(1+M\cdot\alpha p+\mu\cdot p) \end{aligned}$$

From integrating $m(\theta)$ solution:

$$p = \frac{C\mu\left(1-e^{-\frac{M\cdot\alpha+M\cdot\beta}{\mu}}\right)}{M\cdot\alpha+M\cdot\beta} \tag{61}$$

Using two above equations to express C, implying

$$C = -\frac{(-1(M\cdot\alpha+M\cdot\beta))\,e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}}{\mu\left((M\cdot\beta-\mu)e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}+M\cdot\alpha+\mu\right)} \tag{62}$$

substituting yields the expression in the Proposition. □

Proof of Lemma 2

Proof. The social planner's chooses the symmetric function $m_s(\theta)$ to maximize

$$\int_0^1 m_s(\theta)\cdot\Delta u(\theta, m_s)d\theta - \mu\cdot I(m_s) \tag{63}$$

where it takes into account that Δu depends not only on θ but on the information choice function of all other traders m.

We use a perturbation method similar to the proof in Yang (2011). In the first order perturbation we set $m_s(\theta)+\nu\cdot\epsilon(\theta)$ as $m_s(\theta)$, take derivative wrt ν and then set $\nu = 0$ in order to arrive at the following equation that has to hold for any function $\epsilon(\theta)$:

$$\int_0^1 \epsilon(\theta)\cdot\left(M\cdot\alpha\cdot\int_\theta^1 m_s(\tilde\theta)d\tilde\theta - M\cdot\beta\cdot\int_0^\theta m_s(\tilde\theta)d\tilde\theta - \mu\cdot\left(m_s(\theta)-\int_0^1 m_s(\tilde\theta)d\tilde\theta\right)\right)d\theta + \tag{64}$$

$$+\int_0^1 m_s(\theta)\cdot\left(M\cdot\alpha\cdot\int_\theta^1 \epsilon(\tilde{\theta})d\tilde{\theta} - M\cdot\beta\cdot\int_0^\theta \epsilon(\tilde{\theta})d\tilde{\theta}\right)d\theta = 0 \tag{65}$$

We choose $\epsilon(\theta) = \delta_{\hat{\theta}}(\theta)$ where $\delta_{\hat{\theta}}$ is the Dirac-Delta function. Thus $\int_\theta^1 \epsilon(\tilde{\theta})d\tilde{\theta} = \mathbf{1}_{\theta<\hat{\theta}}$ where $\mathbf{1}$ is the heaviside function. Substituting $\hat{\theta} = \theta$, the equation becomes:

$$M\cdot\alpha\cdot\int_\theta^1 m_s(\tilde{\theta})d\tilde{\theta} - M\cdot\beta\cdot\int_0^\theta m_s(\tilde{\theta})d\tilde{\theta} + 1 - \theta - \mu\cdot\left(m_s(\theta) - \int_0^1 m_s(\tilde{\theta})d\tilde{\theta}\right) + \tag{66}$$

$$+ M\cdot\alpha\cdot\int_0^\theta m_s(\tilde{\theta})d\tilde{\theta} - M\cdot\beta\cdot\int_\theta^1 m_s(\tilde{\theta})d\tilde{\theta} = 0 \tag{67}$$

which simplifies to:

$$(M\cdot\alpha - M\cdot\beta)\cdot p_s + 1 - \mu\cdot(m(\theta) - p_s) = 0 \tag{68}$$

The derivative of Equation 68 w.r.t. θ delivers Equation 53, while setting $\theta = 0$ in Equation 68 gives the boundary condition Equation 54. Differentiating:

$$0 = \mu m'(\theta) \tag{69}$$

giving

$$m(\theta) = C \tag{70}$$

with boundary condition

$$\frac{(M\cdot\alpha - M\cdot\beta + \mu)\,p + 1}{\mu} = m(0) = C \tag{71}$$

implying

$$m_s(\theta) = \frac{(M\cdot\alpha - M\cdot\beta + \mu)\,p_s + 1}{\mu} \tag{72}$$

$$p_s = \int_0^1 m(\theta)\,d\theta = \frac{(M\cdot\alpha - M\cdot\beta + \mu)\,p_s + 1}{\mu} \tag{73}$$

Solving for p_s

$$p_s = \frac{1}{2M\cdot\alpha - 2M\cdot\beta} \tag{74}$$

□

Proof of Lemma 3

Proof. Under complete information, in the competitive optimum the last one to enter $\bar{\theta}$ is indifferent between entering and not:

$$-M \cdot \beta \cdot \bar{\theta} + 1 = 0 \tag{75}$$

yielding Eq. 21. For the social planner's optimum one has to find $\bar{\theta}$ that everyone with $\theta < \bar{\theta}$ enters, the others stay out, maximizing:

$$\int_0^{\bar{\theta}} \big((M \cdot \alpha + M \cdot M \cdot \alpha_w) \cdot (\bar{\theta} - \theta) - M \cdot \beta \cdot \theta + 1\big)\, d\theta = \frac{(M \cdot \alpha + M \cdot M \cdot \alpha_w) - M \cdot \beta}{2} \cdot \bar{\theta}^2 + 1 \cdot \bar{\theta} \tag{76}$$

yielding the interior optimum in Eq. 22 if $0 < M \cdot \beta - (M \cdot \alpha + M \cdot M \cdot \alpha_w) < 1$. Note that $M \cdot \beta - (M \cdot \alpha + M \cdot M \cdot \alpha_w) > 0$ holds since $M \cdot \beta - (M \cdot \alpha + M \cdot M \cdot \alpha_w) > 0$. If on the other hand, $1 \leq M \cdot \beta - (M \cdot \alpha + M \cdot M \cdot \alpha_w)$, no one enters: $m(\theta) = 0$ is optimal.

Under no anchor, in the competitive equilibrium every trader enters with probability p and they are all indifferent given they do not know their θ and use a uniform prior. Expected payoff to entering:

$$\int_0^1 (M \cdot \alpha \cdot (1 - \theta) \cdot p - M \cdot \beta \cdot \theta \cdot p + 1)\, d\theta = 0 \tag{77}$$

yielding the unconditional entry probability in Eq. 23. In the social planner's optimum every trader enters with probability p and they maximize social planner's welfare

$$\int_0^1 p \cdot ((M \cdot \alpha + M \cdot \alpha_w) \cdot (1 - \theta) \cdot p - M \cdot \beta \cdot \theta \cdot p + 1)\, d\theta \tag{78}$$

taking derivative w.r.t. p and setting to zero, this implies the entry probability in Eq. 24. Note that there are infinite other solutions since the social planner's planner does not care about who exactly enters. □

Proof of Lemma 4

Proof. It is straightforward to see from Equation 16 that $m(\theta)$ is monotonically decresing. Hence, for us to ensure an interior solution, we must constrain only $m(0)$ and $m(1)$. Substituting $\theta = 0$ into Equation 27:

$$m(0) = \frac{e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}\left(M\cdot\alpha+M\cdot\beta\right)}{\mu\left(M\cdot\alpha+e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}\left(M\cdot\beta-\mu\right)+\mu\right)} < 1$$

Which after a few algebraic manipulations takes the form of the following inequality:

$$\mu\left(M\cdot\beta-\mu+e^{-\frac{M\cdot\alpha+M\cdot\beta}{\mu}}\left(M\cdot\alpha+\mu\right)\right) > \frac{\left(M\cdot\alpha+M\cdot\beta\right)\left(M\cdot\alpha+M\cdot\beta\right)}{M\cdot\alpha+M\cdot\beta}$$

On the other hand, substituting $\theta = 1$ into Equation 27:

$$m(1) = \frac{M\cdot\alpha+M\cdot\beta}{\mu\left(M\cdot\alpha+e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}\left(M\cdot\beta-\mu\right)+\mu\right)} > 0$$

Which after a few algebraic manipulations takes the form of the following inequality:

$$\left(M\cdot\alpha+M\cdot\beta\right)\left(M\cdot\alpha+M\cdot\beta\right) > 0$$

Combining both restrictions ensures the condition of the proposition.

The equation $m\left(1\right) = 0$ gives (**??**) with the subsitution x. Simple differentiation shows that $\frac{x^2(\alpha+\beta)}{\alpha x+e^{x(\alpha+\beta)}(x\beta-1)+1}$ is monotonically decreasing iff

$$-2-x\alpha+e^{x(\alpha+\beta)}\left(2+x^2\beta\left(\alpha+\beta\right)-x\left(\alpha+2\beta\right)\right)$$

is positive. Bounding $e^{x(\alpha+\beta)}$ from below by its third-order Taylor approximation and simplifiying shows that this expression is larger than

$$\frac{1}{6}x^3(\alpha+\beta)^2(-\alpha+2\beta+x^2\left(\alpha^2\beta+2\alpha\beta^2+\beta^3\right)+x\left(\beta^2-\alpha^2\right))$$

which is positive for our parameter restrictions. Given that

$$\lim_{x\to 0}\frac{x^2\left(\alpha+\beta\right)}{\alpha x+e^{x(\alpha+\beta)}(x\beta-1)+1} = \frac{2\left(\alpha+\beta\right)}{\beta^2-\alpha^2} > 0$$

and it goes to zero as $x \to \infty$, there is exactly one solution.

The equation $m(0) = 1$ gives (**??**) and the definition of x_2.

$$\lim_{x\to 0} \frac{e^{x(\alpha+\beta)} x^2 (\alpha+\beta)}{\alpha x + e^{x(\alpha+\beta)}(x\beta - 1) + 1} - \mu x = \frac{2(\alpha+\beta)}{\beta^2 - \alpha^2} > 0$$

and

$$\lim_{x\to\infty} \frac{e^{x(\alpha+\beta)} x (\alpha+\beta)}{\alpha x + e^{x(\alpha+\beta)}(x\beta - 1) + 1} = 1 + \frac{\alpha}{\beta}$$

implying that the condition ensures that $\frac{e^{x(\alpha+\beta)} x^2 (\alpha+\beta)}{\alpha x + e^{x(\alpha+\beta)}(x\beta-1)+1} - \mu x$ starts above 0 and goes below this value as x increases, so it has at least one solution.

Observe that the sign of the derivative of $\frac{e^{x(\alpha+\beta)}(\alpha+\beta)}{\alpha x + e^{x(\alpha+\beta)}(x\beta-1)+1}$ in x determines the sign of $\frac{\partial x_2}{\partial \mu}$ by the implicit function theorem. By differentiating $\frac{e^{x(\alpha+\beta)}(\alpha+\beta)}{\alpha x + e^{x(\alpha+\beta)}(x\beta-1)+1}$ in x reveals that this sign determined by $(e^{x(\alpha+\beta)}-1)\beta - x\alpha(\alpha+\beta)$ which, using the second order Taylor approximation is bounded from below by $\frac{x^2}{2}\left(\beta^2 - \alpha^2\right) > 0$.

Eq. **??** reduces to $\frac{e^{x(\alpha+\beta)} x(\alpha+\beta)}{\alpha x + e^{x(\alpha+\beta)}(x\beta-1)+1} = \mu$ where the right hand side is monotonically decreasing in x from infinity to $1 + \frac{\alpha}{\beta}$, this equation has exactly one solution iff $\mu > 1 + \frac{\alpha(1)}{\beta}$ and zero otherwise. All the other results for this case follows immediately. ☐

Proof of Lemma 5

Proof. TBA ☐

Proof of Proposition 1

Proof. In the limit as $\mu \to \infty$ the privately optimal level of average entry p is given by Equation 23, while the social planner's average entry p_s is given by Equation **??**. Thus $p > p_s$.

In the limit as $\mu \to 0$, the private entry becomes $\frac{1}{M\cdot\beta}$ (even though it is not an interior solution any more) while the social entry stays unchanged and thus $p < p_s$ follows if

$$-2(M\cdot\alpha + M\cdot\alpha_w) < 0. \tag{79}$$

Thus this is the condition for having a interior $\bar{\mu}$ for at which $p = p_s$.

Thus there exists a $\bar{\mu}$ at which the average entries (Equation 28 and ??) are the same and this is pinned down by the implicit equation:

$$-\frac{(-M\cdot\alpha-M\cdot\beta)\left(e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}-1\right)}{(M\cdot\alpha+M\cdot\beta)\left((M\cdot\beta-\mu)e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}+M\cdot\alpha+\mu\right)}+\frac{2}{2M\cdot\alpha-2M\cdot\beta}=0 \tag{80}$$

[TBA: SHOW $\bar{\mu}$ IS INTERIOR SOLUTION] □

Proof of Lemma 6

Proof. Proof by straightforward algebra. □

Proof of Proposition 2

Proof.

$$\frac{\partial\Pi}{\partial\mu}=-\frac{e^{\frac{3(M\cdot\alpha+M\cdot\beta)}{2\mu}}\left((M\cdot\alpha+M\cdot\beta)^2+2\mu^2-2\mu^2Cosh\left(\frac{M\cdot\alpha+M\cdot\beta}{\mu}\right)\right)}{(M\cdot\alpha+M\cdot\beta)^2\mu^2\left(M\cdot\alpha+e^{\frac{(M\cdot\alpha+M\cdot\beta)}{\mu}}(M\cdot\beta-\mu)+\mu\right)^3}$$
$$\times\left((M\cdot\alpha+M\cdot\beta)\,Cosh\left(\frac{M\cdot\alpha+M\cdot\beta}{2\mu}\right)+(M\cdot\alpha-M\cdot\beta-2\mu)\,Sinh\left(\frac{M\cdot\alpha+M\cdot\beta}{2\mu}\right)\right)(M\cdot\alpha+M\cdot\beta-M\cdot)^2 \tag{81}$$

Which can be rewritten as:

$$\frac{\partial\Pi}{\partial\mu}=-\frac{e^{\frac{3(M\cdot\alpha+M\cdot\beta)}{2\mu}}\left((M\cdot\alpha+M\cdot\beta)^2+2\mu^2-2\mu^2Cosh\left(\frac{M\cdot\alpha+M\cdot\beta}{\mu}\right)\right)}{\mu^2\left(M\cdot\alpha+e^{\frac{(M\cdot\alpha+M\cdot\beta)}{\mu}}(M\cdot\beta-\mu)+\mu\right)^3}$$
$$\times\frac{(M\cdot\alpha+M\cdot\beta)^2}{(M\cdot\alpha+M\cdot\beta)^2}\left(e^{-\frac{(M\cdot\alpha+M\cdot\beta)}{2\mu}}\left(M\cdot\beta+e^{\frac{(M\cdot\alpha+M\cdot\beta)}{\mu}}(M\cdot\alpha-\mu)+\mu\right)\right) \tag{82}$$

It is clear that the second part of the equation is always positive. Dividing both the numerator and denominator of the first part of the equation by μ^2, and denoting $y=\frac{M\cdot\alpha+M\cdot\beta}{\mu}$:

$$-\frac{e^{\frac{3}{2}y}\left(y^2+2-2Cosh\left(y\right)\right)}{\left(M\cdot\alpha+e^y\left(M\cdot\beta-\mu\right)+\mu\right)^3} \tag{83}$$

From the constant C derived earlier, we know that the denominator is always positive, while the part inside the brackets of the numerator can be rewritten as:

$$y^2+2-e^y-e^{-y} \tag{84}$$

Which is 0 for $y = 0$, and straightforward to see that it is monotonically decreasing, thus always negative. This implies that the sign of the derivative depends on the third part of the above equation, in particular:

$$sign\left(\frac{\partial\Pi}{\partial\mu}\right) = sign\left(M\cdot\beta + \mu\left(1 - e^{\frac{M\cdot\alpha + M\cdot\beta}{\mu}}\right) + M\cdot\alpha e^{\frac{M\cdot\alpha + M\cdot\beta}{\mu}}\right) \tag{85}$$

$\mu\left(1 - e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}\right) \in [-\infty, -(M\cdot\alpha + M\cdot\beta)]$, and $M\cdot\alpha e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}} \in [M\cdot\alpha\infty, M\cdot\alpha]$, thus if $\alpha < 0$, both parts are always negative, with the highest combined value being $-M\cdot\beta$ reached as $\mu \to \infty$. Consequently the sign of this part of the function is always negative, leading to a monotonically decreasing Π function.

On the other hand if $0 < \alpha < M\beta$, the rates of change of the two above specified parts matter, and since $M\cdot\alpha e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}$ initially falls faster than $\mu\left(1 - e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}\right)$ increases, the function eventually becomes negative, and afterwards converges to 0 as $\mu \to \infty$. Consequently Π initially increases and after a certain threshold monotonically decreases to 0 as $\mu \to \infty$.

For $0 < \alpha < M\beta$, from previous analysis we know that Π reaches its maximum value when $M\cdot\beta + e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}(M\cdot\beta - \mu) + \mu = 0$. Since $M\cdot\beta + e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}(M\cdot\beta - \mu) + \mu$ is the difference between $\left(M\cdot\alpha + M\cdot\beta + e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}(M\cdot\alpha + M\cdot\beta - 2\mu) + 2\mu\right)$ and $\left(M\cdot\alpha + e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}(M\cdot\beta - \mu) + \mu\right)$, these terms cancel out reducing the Π function to:

$$\Pi = \frac{\left(e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}} - 1\right)(M\cdot\alpha + M\cdot\beta)^2}{2(M\cdot\alpha + M\cdot\beta)^2\left(M\cdot\alpha + e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}(M\cdot\beta - \mu) + \mu\right)} \tag{86}$$

From $M\cdot\beta + e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}(M\cdot\beta - \mu) + \mu = 0$, we can see that at this point $e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}} = \frac{M\cdot\beta+\mu}{\mu-M\cdot\alpha}$, replacing this into the above equation:

$$\Pi = \frac{\left(\frac{M\cdot\beta+\mu}{\mu-M\cdot\alpha} - 1\right)(M\cdot\alpha + M\cdot\beta)^2}{2(M\cdot\alpha + M\cdot\beta)^2\left(M\cdot\alpha + \frac{M\cdot\beta+\mu}{\mu-M\cdot\alpha}(M\cdot\beta - \mu) + \mu\right)} \tag{87}$$

Which simplifies to $\frac{(M\cdot\alpha+M\cdot\beta)^2}{2(M\cdot\beta-M\cdot\alpha)(M\cdot\alpha+M\cdot\beta)^2}$, thus this value represents the maximum value of Π when $0 < \alpha < M\cdot\beta$, while for the $\alpha < 0$ case, since Π monotonically decreases in μ from its initial value, its maximum value is exactly $\frac{(M\cdot\alpha+M\cdot\beta)^2}{2M\cdot\beta^2(M\cdot\alpha+M\cdot\beta)}$.

We continue by computing:

$$\lim_{\mu\to\infty}\Pi = \frac{(M\cdot\alpha+M\cdot\beta)^2}{2\,(M\cdot\alpha+M\cdot\beta)^2}\lim_{\mu\to\infty}\frac{\left(e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}-1\right)\left(M\cdot\alpha+M\cdot\beta+e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}\,(M\cdot\alpha+M\cdot\beta-2\mu)+2\mu\right)}{\left(M\cdot\alpha+e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}\,(M\cdot\beta-\mu)+\mu\right)^2} \tag{88}$$

The limit is of the 0/0 form, allowing us to apply L'Hopital's rule. Applying L' Hopital' s rule three times and afterwards simplifying the fraction, the limit of the numerator converges to $5\,(M\cdot\alpha+M\cdot\beta)^3$, while the limit of the denominator increases to ∞, thus the whole expression decreases to 0 as $\mu\to\infty$.

The payoff function is:

$$\Pi = \frac{\left(e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}-1\right)\left(M\cdot\alpha+M\cdot\beta+e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}\,(M\cdot\alpha+M\cdot\beta-2\mu)+2\mu\right)}{2\left(M\cdot\alpha+e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}\,(M\cdot\beta-\mu)+\mu\right)^2} \tag{89}$$

$$\Pi_s = \frac{1}{2\,(M\cdot\beta-M\cdot\alpha)} \tag{90}$$

The maximizing value of Π reduces from $\frac{(M\cdot\alpha+M\cdot\beta)^2}{2(M\cdot\beta-M\cdot\alpha)(M\cdot\alpha+M\cdot\beta)^2}$ to $\frac{1}{2(M\cdot\beta-M\cdot\alpha)}$, which is exactly the value of Π_s. For any other value of μ, Π is lower than Π_s. On the other hand, if $M\cdot\alpha<0$, Π monotonically decreases from $\frac{(M\cdot\alpha+M\cdot\beta)}{2M\cdot\beta^2}$ (which is the reduced form of $\frac{(M\cdot\alpha+M\cdot\beta)^2}{2M\cdot\beta^2(M\cdot\alpha+M\cdot\beta)}$) to 0 as $\mu\to\infty$, implying that if $M\cdot\alpha<0$, Π is lower than Π_s for any value of μ.

If $0<\alpha<\beta$, there exists a μ^* that equates Π to Π_s if and only if:

$$\Pi = \frac{e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}-1}{2\left(M\cdot\alpha+e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}\,(M\cdot\beta-\mu)+\mu\right)} = \frac{1}{2\,(M\cdot\beta-M\cdot\alpha)} = \Pi_s \tag{91}$$

where we used the expression of Π from Eq. 86. Which follows directly using Eq. **??**. Similarly one can show:

$$p = \frac{e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}-1}{M\cdot\alpha+e^{\frac{M\cdot\alpha+M\cdot\beta}{\mu}}\,(M\cdot\beta-\mu)+\mu} = \frac{1}{M\cdot\beta-M\cdot\alpha} = p_s \tag{92}$$

At this point it is straightforward to see that the two equalities are almost identical, with the first one being divided by 2 on both sides, implying that the same μ satisfies both equalities. □

Proof of Proposition 3

TBA. □

Proof of Proposition 4

TBA. □

Proof of Proposition 5

Proof. Thus irrespective of μ, if there are abundant sophisticated traders $\frac{M}{\mu} \to \infty$, welfare converges to the same amount, which is exactly halfway between the welfare if $\mu = 0$ ($\frac{\alpha+\beta}{2\beta^2}$) and that when $\mu = \infty$ ($W = 0$). The reason why the exact value of μ does not matter is that if M is large enough, μ becomes irrelevant as the same amount of aggregate learning gets divided between more and more agents. Thus what matters is that the quadratic learning costs implies that half of the profits are used for learning.

Without learning ($\mu \to \infty$), entry is a constant in θ. Whenever the profit is positive, traders increase their entry to the maximum possible: $m(\theta) = 1$, afterwards they choose m s.t. profits (and thus welfare) is exactly zero. Welfare reaches its maximum at: $M^* = \frac{1}{2} \cdot \frac{2}{\beta-\alpha}$, while for any $M > M^{**} = \frac{2}{\beta-\alpha}$, profits are zero. The maximal welfare (and thus profit) is: $W^* = \frac{1}{8} \cdot \frac{(2)^2}{\beta-\alpha}$. One can show that the maximum interior profit is strictly larger than this for all parameters. □

Proof of Proposition 6

Proof. Using the the parameters of the model, the only part changing in the welfare ratio is:

$$\frac{\alpha^2}{\beta^2-\alpha^2} = \frac{\left(\delta(\nu-2)\nu + \delta_c\eta(\nu-1)^2\right)^2}{4\delta(\nu-1)(\eta(\nu-1)-\nu+2)(\delta(\eta+\nu-2)+\delta_c\eta(\nu-1))} \tag{93}$$

In the special case of $\nu = 0$ this simplifies to

$$\left.\frac{\alpha^2}{\beta^2-\alpha^2}\right|_{\nu=0} = \frac{\delta_c^2\eta^2}{4\delta(2-\eta)(\eta(\delta_c-\delta)+2\delta)} \tag{94}$$

which can be shown to be strictly increasing in both η and δ_c.

In the special case of $\eta = 0$ this simplifies to

$$\left.\frac{\alpha^2}{\beta^2-\alpha^2}\right|_{\eta=0} = \frac{\nu^2}{4-4\nu} \tag{95}$$

which can be shown to be strictly increasing in ν. □

Proof of Proposition 7

Proof. The welfare of the trader equals the payoff in this case since there is no anchor cost. To ensure that only p_P traders enter, the planner has to charge an entry fee equal to the profit to make non-entrants indifferent. Thus the entry fee that has to be charged is equal to the ex-ante value V of the traders to entering:

$$\text{fee} = \Pi = \int_0^1 p_P \left(M \cdot \alpha \cdot (1-\theta) \cdot p_P - M \cdot \beta \cdot \theta \cdot p_P + 1 \cdot \theta\right) d\theta = \frac{(\beta+\alpha)}{2M \cdot \beta^2} \tag{96}$$

This is an equilibrium since if more traders enter, the overall value V will decrease due to crowding and the cost of starting to learn thus choosing to enter entails a negative payoff. A similar argument holds if too few traders choose to enter.

The policy is welfare improving for any μ if it yields higher welfare than the maximum welfare in the competitive equilibrium, substituting $x = \frac{\alpha}{\beta}$ (where by assumption $x \in [-1, 1]$), the condition simplifies to:

$$2(1-x)(1+x)^2 - x > \frac{1}{1} \tag{97}$$

The condition is true for $x \in [-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]$, i.e. $\alpha \in [-\frac{\beta}{\sqrt{2}}, \frac{\beta}{\sqrt{2}}]$. Note that another solution, $\alpha < -\beta$ contradicts the assumption of a rat race ($\alpha + \beta > 0$) and thus can be ruled out. [TBA: COMPARISON TO NO INFO] [TBA: WELFARE GAIN FROM LEARNING] □

Proof of Proposition 8

Proof. Welfare under no anchor is better than under perfect anchor if and only if:

$$\frac{4\alpha_w}{2(\alpha-\beta)^2} > \frac{M^2(\alpha+\beta+\alpha_w)}{2(\beta M)^2} \tag{98}$$

if $(2)(\beta M) > M(\beta - \alpha)$, one can rearrange this as a positive lower bound on α_w given in the proposition. If $(2)(\beta M) < M(\beta - \alpha)$, then the above is true for any non-negative α_w. □

Proof of Lemma 7

Proof. Denote the strategy function of all other players as $\tilde{m}(\theta)$. Substituting our payoff functions into Equation 50:

$$M \cdot \alpha \cdot \int_\theta^1 \tilde{m}(\tilde{\theta}) d\tilde{\theta} - M \cdot \beta \cdot \int_0^\theta \tilde{m}(\tilde{\theta}) d\tilde{\theta} + 1 = \mu \cdot \left[\log\left(\frac{m(\theta)}{1-m(\theta)}\right) - \log\left(\frac{p}{1-p}\right)\right]. \tag{99}$$

Differentiating this we arrive at the differential equation:

$$(M \cdot \alpha + M \cdot \beta) \cdot \tilde{m}(\theta) = -\frac{\mu \cdot m'(\theta)}{m(\theta) \cdot (1 - m(\theta))}. \tag{100}$$

Imposing symmetry $\tilde{m}(\theta) = m(\theta)$ results in Equation 51. The boundary condition is given by the original integral-differential Equation 99 evaluated at any θ: in Equation 17 we set $\theta = 0$. $\square$

Proof of Lemma 8

Proof. The social planner's planner chooses the symmetric choice function $m_s(\theta)$ to maximize

$$\int_0^1 m_s(\theta) \cdot \Delta u(\theta, m_s) d\theta - \mu \cdot I(m_s) \tag{101}$$

where it takes into account that Δu depends not only on θ but on the information choice function of all other traders m.

We use a perturbation method similar to the proof in Yang (2011). In the first order perturbation we set $m_s(\theta) + \nu \cdot \epsilon(\theta)$ as $m_s(\theta)$, take derivative wrt ν and then set $\nu = 0$ in order to arrive at the following equation that has to hold for any function $\epsilon(\theta)$:

$$\int_0^1 \epsilon(\theta) \cdot \left(M \cdot \alpha \cdot \int_\theta^1 m_s(\tilde{\theta}) d\tilde{\theta} - M \cdot \beta \cdot \int_0^\theta m_s(\tilde{\theta}) d\tilde{\theta} - \mu \cdot \left[\log\left(\frac{m_s(\theta)}{1 - m_s(\theta)} \right) - \log\left(\frac{\int_0^1 m_s(\tilde{\theta}) d\tilde{\theta}}{1 - \int_0^1 m_s(\tilde{\theta}) d\tilde{\theta}} \right) \right] \right) d\theta + \tag{102}$$

$$+ \int_0^1 m_s(\theta) \cdot \left(M \cdot \alpha \cdot \int_\theta^1 \epsilon(\tilde{\theta}) d\tilde{\theta} - M \cdot \beta \cdot \int_0^\theta \epsilon(\tilde{\theta}) d\tilde{\theta} \right) d\theta = 0 \tag{103}$$

We choose $\epsilon(\theta) = \delta_{\hat{\theta}}(\theta)$ where $\delta_{\hat{\theta}}$ is the Dirac-Delta function. Thus $\int_\theta^1 \epsilon(\tilde{\theta}) d\tilde{\theta} = \mathbf{1}_{\theta < \hat{\theta}}$ where $\mathbf{1}$ is the heaviside function. Substituting $\hat{\theta} = \theta$, the equation becomes:

$$M \cdot \alpha \cdot \int_\theta^1 m_s(\tilde{\theta}) d\tilde{\theta} - M \cdot \beta \cdot \int_0^\theta m_s(\tilde{\theta}) d\tilde{\theta} + 1 - \mu \cdot \left[\log\left(\frac{m_s(\theta)}{1 - m_s(\theta)} \right) - \log\left(\frac{\int_0^1 m_s(\tilde{\theta}) d\tilde{\theta}}{1 - \int_0^1 m_s(\tilde{\theta}) d\tilde{\theta}} \right) \right] + \tag{104}$$

$$+ M \cdot \alpha \cdot \int_0^\theta m_s(\tilde{\theta}) d\tilde{\theta} - M \cdot \beta \cdot \int_\theta^1 m_s(\tilde{\theta}) d\tilde{\theta} = 0 \tag{105}$$

which simplifies to:

$$(M \cdot \alpha - M \cdot \beta) \cdot p_s + 1 - \mu \cdot \left[\log\left(\frac{m(\theta)}{1 - m(\theta)}\right) - \log\left(\frac{p_s}{1 - p_s}\right) \right] = 0 \tag{106}$$

The derivative of Equation 106 w.r.t. θ delivers Equation 53, while setting $\theta = 0$ in Equation 106 gives the boundary condition Equation 54. $\square$