Learning in Crowded Markets *

Péter Kondor[†] Adam Zawadowski[‡]

Central European University Boston University

May 27, 2015

Abstract

We develop a model of capital reallocation to analyze whether the presence of more sophisticated traders improve capital allocation and welfare. Trades can become crowded due to externalities but traders can devote resources to learn about the number of earlier entrants. In general, more traders having the choice to enter neither improves the efficiency of capital allocation nor does it aggravate crowding. In fact, whether there is eventually too little or too much capital allocated to the new sector is determined solely by the technology in that sector, the cost of learning, the depth of the market, and the severity of the potential shocks, but not the mass of sophisticated traders present. However, the presence of more traders decreases welfare, as they waste more aggregate resources in learning about each others' position.

*For comments and suggestions, we are grateful to Patrick Bolton, Thomas Chemmanur and seminar participants at Boston University, Central European University, ESSET 2014 (Gerzensee), Brigham Young University, GLMM 2015 (Boston). 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 an increasing 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 easily learn from prices 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]

We analyze this question through a capital reallocation problem with learning and externalities. Traders' main problem is to decide whether to invest in a new sector with scarce capital. Because of decreasing returns to scale, trades can become crowded and traders who invest earlier benefit the most. It is costly for traders to learn how many already entered. Hence, each trader jointly decides how much to learn and whether to enter. Each trader's decision might impose positive or negative external effects on others mostly due to potential idiosyncratic and aggregate liquidity shocks. We are interested in how the allocation of capital and welfare changes as the number of traders facing the same problem is increasing.

Our main observation is that in general, more traders having the choice to enter neither improves the efficiency of capital allocation nor does it aggravate crowding. In fact, whether there is eventually too little or too much capital allocated to the new sector is determined solely by the technology in that sector, the cost of learning, the depth of the market, and the severity of the potential shocks, but not the mass of sophisticated traders present. However, more traders decrease welfare, as in total, they invest more resources in learning about each others' position without any added social value.

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

The sole exception is when there are only few traders present. In this case, they all enter without learning and marginally increasing their mass unambiguously increases both efficiency and welfare.

In our model, the order in which each trader learns the existence of the new trading opportunity is it unknown type. The characteristics of the new sector, and the potential shocks determine the relative payoff to early and late entrants together with the externalities they impose on each other. If the marginal product of capital is more strongly decreasing in the new sector, early entrants gain more relative to late entrants. Also, a market with more traders provides an easier exit opportunities for those hit by an idiosyncratic liquidity shock. On the other hand, in a market with more traders an aggregate liquidity shock creates a more severe price crash making exit more costly for all. We show that the first entrant prefers others to enter only when the idiosyncratic shock is sufficiently large. However, regardless of the parameters, the average type (and all those who arrive later than the average type) would always prefer a smaller crowd, that is, there is a negative crowding externality.

The trader's choice is a function which maps it potential type into a probability of entry. Following the rational inattention approach of Sims (1998), Yang (2015) and Woodford (2008), we model the learning cost of this choice by assuming that reducing the entropy of this function has a constant marginal cost. For example, the choice of entering with a constant probability regardless of being an early or a late trader is free of learning cost as this function has the maximum entropy. However, entering only when very few entered before is very expensive, because this requires a large reduction in the entropy of the entry function. Our interpretation is that an unanchored strategy as momentum, or carry trade is equivalent to high marginal cost of reducing entropy as prices give little information on how many traders are already following that strategy. Apart from the parsimony of jointly capturing choice and learning, the main advantage of this cost specification is that it has an axiomatic foundation based on information theory.

Our main focus is the effect of increasing the total mass of traders who eventually learn about the new investment opportunity. We compare all the outcomes with the choice of a planner who internalizes the externalities, but is also subject to costly learning. It is not surprising that when the total mass of traders is small, all traders enter, therefore they do not engage in costly learning. At that stage, more traders lead to more entrants, more efficient allocation of capital and higher welfare.

However, when the total mass of traders is sufficiently large, traders learn and differentiate their entry probability by their type. There are two main effects which both get stronger as the number of traders increases. First, the relative benefit of entering early compared to late is increasing because there are more traders who may potentially enter before. Thus traders choose entry strategies that are more contingent on their types. We refer to this as the “rat race effect”. Second, the negative effect from crowding for the average entrant is increasing. This pushes average entry of a given trader down. We refer to this as the “crowding effect”. While the rat race has little effect on average entry, the crowding effect decreases average entry. It turns out, as we increase the mass of traders, the average entry decreases proportionally leading to constant aggregate entry as a function of the mass of traders. That is, a larger mass of sophisticated traders does not improve the efficiency of capital allocation.

As more traders with the choice of entry does not affect aggregate entry, whether there is over- or under-entry compared to the planners’ solution is also independent from the mass of traders. Instead, we have permanent over- or under-entry, depending on the parameters. Less idiosyncratic and more aggregate shocks, financial markets with less depth, a stronger decreasing returns to scale techonolgy and less anchored strategies all make over-entry more likely and/or more severe.

Importantly, as the number of potential entrants increase, welfare decreases. This is because the rat race effect increases the private incentives to learn, but does not change the social value of learning. In fact, the rat race effect captures the zero-sum part of the trading process: early entrants gain for the expense of late entrants. In contrast, the planner is interested only in aggregate entry (as opposed to the order of entry), because this determines the efficiency of capital allocation. Therefore, even though in the presence of more traders, aggregate entry does not increase, they do invest more in learning whether to enter. This leads to a wasteful use of resources and decreasing welfare.

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. The main reason is that in most other applications than our baseline of capital reallocation, there is a group of traders whose welfare increases when others enter too much. 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.

The rest of the paper is structured as follows. In Section 2 we present our reduced form model and also give a structural microfoundation. In Section 3 we first solve the problem for a perfect price anchor and no anchor at all. Then we present the general results for the entry and learning decision and analyze how this relates to aggregate entry and welfare. In Section 4 we present further microfoundations of the reduced form model and discuss how our main insights are affected by these different settings.

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) traders, 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 trader 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 traders who enter *a*fter trader θ), while $b(\theta)$ is the mass of entrants with a type lower than θ (traders who enter *b*efore trader θ). We show in the microfoundation, that the following two assumptions are natural. First, $\beta + \alpha > 0$, such that entering earlier is better than later: we call this property rat-race. Second, $\beta - \alpha > 0$, 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 microfoudation. 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 Microfoundation: Capital arbitrage

There are two islands A and B indexed by $i \in \{A, B\}$. There are two types of traders: 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 traders 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) where $\delta_c > 0$. 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 revenue of a trader that chooses to transport capital at time t is:

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

[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.

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

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 picture only to the extent that it introduces some redistribution among traders; an element which washes out by aggregation. Therefore, we can interpret the aggregate revenue of traders as a measure of market efficiency.

We interpret our deep parameters as follows. δ captures the extent of decreasing return to scale in each market. Conceptually, this is a technological parameter of the sectors or firms which are subject to the capital reallocation. δ_c characterizes the depth of the financial market where claims on these firms and sectors are traded. When δ_c is large, a sudden selling presure of the participating traders drives the price down significantly. In contrast, η and ν characterizes the traders as opposed to the markets. A large η is interpreted as a large probability for a common liquidity shock for all traders, for example, because of large common exposure to risk factors outside of the model. A large v is a large probability of an idiosyncratic liquidity shock, for example, because traders are professional investors with a volatile investor base.

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

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

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

Resulting in positive crowding and rat-race parameters of:

$$\beta - \alpha = (1-\nu)^2 \cdot (\eta \cdot \delta_c + 2\delta) > 0 \tag{7}$$

$$\alpha + \beta = \left(1 + (1-\nu)^2 (1-\eta)\right) \cdot \delta > 0 \tag{8}$$

for all parameter values.

To interpret α and β, it is useful to first consider the case without idiosyncratic shock $(\nu = 0)$. In this case, $\alpha = -\frac{1}{2}\eta(\delta + \delta_c)$ and $\beta = \delta\left(2 - \frac{1}{2}\eta\right) + \frac{1}{2}\eta\delta_c$ implying $\beta - \alpha = (\eta \cdot \delta_c + 2\delta)$ and $\alpha + \beta = (1 + (1-\eta)) \cdot \delta$. Note that our rat-race parameter, $\alpha + \beta$, is mainly driven by δ measuring the extent of decreasing return to scale. That is, entering early is beneficial because the investment opportunity is more revenueable when not many traders have entered yet. The crowding parameter, $\beta - \alpha$, is also increasing in δ and positive if $\delta > 0$: this shows that in financial market it is natural to assume that crowding comes hand-in-hand with a rat race. It is important to note that the crowding parameter, $\beta - \alpha$, is also increasing in the probability of the aggregate liquidity shock, η, and the illiqidity of the market, δ_c. That is, entrants impose a negative externality on each other, because it is more costly to exit when more want to exit the same time. Finally, note that α is negative whithout idiosyncratic shock, because of the same logic. Without idiosyncratic shock, the effect of more late entrants in a liquidity crisis is the same as the effect of more entrants, late or early.

While the introduction of idiosyncratic shock affects all our reduced form parameters, its main qualitative effect is that it changes the sign of α. Indeed, α is monotonically increasing in ν, reaching $\frac{1}{2} \cdot \delta > 0$ when $v = 1$. The intuition is that for large v early entrants benefit from late entrants since if they have to liquidate their position, they can do so at a higher price. This means that α is likely to be positive in markets where entrants need enough subsequent liquidity to exit at a reasonable price.

2.2 Learning cost based on entropy

Before entry, traders can engage in costly learning about their type. We specify the learning cost to be proportional to the reduction in entropy.[4] Sims (1998) argues that the advantage of such a specification is that it both allows for flexible information acquisition and can be derived based on

[4] The entropy of a discrete variable is defined as $\sum_x P(x) \log \frac{1}{P(x)}$, where the random variable takes on the value x with probability $P(x)$, see MacKay (2003).

information theory. Note that the payoff (1) for a given θ in our model is linear in entry. Woodford (2008) proves the optimal signal structure and entry decision rule for such problems.

Lemma 1. *The optimal signal structure is binary: given their type θ, traders choose to receive signal $s = 1$ with probability $m(\theta)$ and $s = 0$ with probability $1 - m(\theta)$. The optimal entry decision conditional on the signal is: enter if $s = 1$, stay out if $s = 0$.*

Thus, similar to Yang (2015), $m(\theta)$ is the only choice variable (learning and entry strategy combined) which in turn is the conditional probability of entry. The intuition for the binary signal structure is that the only reason traders want to learn about θ is to be able to make a binary decision of whether or not to enter. Given the linearity of the problem, the "cheapest" signal to implement the optimal entry strategy is also binary, it simply tells the trader whether or not to enter.

We now write the cost of learning, defined by the reduction in entropy, in case of a binary information structure. Denote the marginal cost of learning (entropy reduction) as μ and using the definition of entropy,

$$L = \mu \left[\underbrace{\left(p \log \left[\frac{1}{p} \right] + (1-p) \log \left[\frac{1}{1-p} \right] \right)}_{\text{entropy of no info signal } p} - \underbrace{\int_0^1 \left(m(\theta) \log \left[\frac{1}{m(\theta)} \right] + (1-m(\theta)) \log \left[\frac{1}{1-m(\theta)} \right] \right) d\theta}_{\text{entropy of signal } m(\theta)} \right], \tag{9}$$

where p denotes the unconditional probability of entry and is defined by:

$$p = \int_0^1 m(\tilde{\theta}) d\tilde{\theta}. \tag{10}$$

The expression for learning (9) can be understood in the following way. Entropy is always highest for the signal that reveals the least amount of information, in this case this is the uninformative signal which prompts the trader to enter with probability p unconditional on its type θ. Thus learning depends on how much lower the entropy is for the signal $m(\theta)$ than for the flat (uninformative) signal p that gives the same amount of average entry. Intuitively, the steeper $m(\theta)$ becomes in θ (keeping average entry p constant), the more the trader can differentiate its entry decision according to its type and the higher the entropy reduction, thus the higher the learning cost. Note that the minimum of L is zero when $m(\theta) = p$ and L is bounded from above but might generate infinite marginal cost of learning.

If all traders choose the same amount of learning and entry (i.e. the same $m(\theta)$), the mass of lower types entering ("before" trader θ) becomes:

$$b(\theta) = M \cdot \int_0^\theta m(\tilde{\theta}) d\tilde{\theta} \tag{11}$$

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

$$a(\theta) = M \cdot \int_\theta^1 m(\tilde{\theta}) d\tilde{\theta} \tag{12}$$

thus $M \cdot p = b(\theta) + a(\theta)$ is the aggregate entry of traders.

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 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. An alternative would be to think of capacity as limited and μ being the Lagrange multiplier of the capacity constraint. We choose to use a fixed μ instead of a fixed capacity because we think in this context learning capacity can be expanded freely at a fixed marginal cost: e.g. hiring new staff is always possible or allocating more attention to this specific trade at the expense of other trades.

Using the language of Stein (2009), we interpret μ as a characteristic of the trading strategy. 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 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.

2.3 Definition of allocative efficiency and welfare

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

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

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, aggregate revenue $M \cdot R$ can be also interpreted as a measure of allocative efficiency.

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

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

which is what traders maximize. 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 $W \equiv M \cdot V$.

3 Model Solution

In this section we present our main results. First, we establish benchmark results for markets with a perfect anchor and those with no anchor at all. Second, we drive the optimal strategies of traders for general levels of the strength of the anchor. Third, we analyze how aggregate entry and welfare changes as the mass of sophisticated traders increases. All proofs are relegated to Appendix A.

3.1 Fully anchored and unanchored trades

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

Lemma 2. *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:*

$$p|_{\mu=0} = \max\left(\frac{1}{M \cdot \beta}, 1\right) \tag{15}$$

$$p_s|_{\mu=0} = \max\left(\frac{1}{M \cdot (\beta - \alpha)}, 1\right) \tag{16}$$

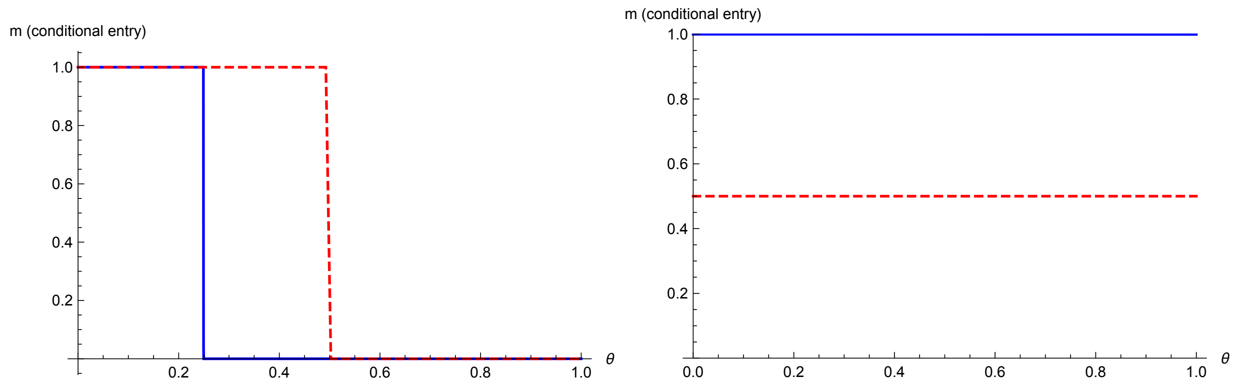
For trades without any anchor ($\mu \to \infty$), both the competitive and social planner's entry functions $m(\theta)$ are flat. All traders enter with the same unconditional probability:

$$p|_{\mu\to\infty} = \max\left(\frac{2}{M \cdot (\beta - \alpha)}, 1\right) \tag{17}$$

$$p_s|_{\mu\to\infty} = \max\left(\frac{1}{M \cdot (\beta - \alpha)}, 1\right). \tag{18}$$

Under a perfect anchor whether there is under- or over-entry (compared to that under the social planner) depends on the sign of α, i.e. the type of liquidity shock in the structural model. We analyze three cases: First, when idiosyncratic liquidity shocks are important, there is competitive under-entry if $\alpha > 0$, since traders with higher θ do not take into account the positive effect of their entry since that, at least partially, accrues to entrants with lower θ. When the order of entry is known, traders do not internalize that if they enter, market liquidity increases for everyone, that is, a market with more traders provides easier exit opportunities for those hit by an idiosyncratic liquidity shock. Second,

Figure 1: **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 line, the social planner's solution the dashed line. Parameters: $\beta = 4$, $\alpha = 2$, $M = 1$.

when aggregate liquidity shocks are more important, $\alpha < 0$ and there is over-entry. It is so, because traders do not take into account that as they enter, in case of an aggregate liquidity shock, they all want to sell reducing the price for everyone. This is a fire-sale externality. Finally, when there is neither idiosyncratic, nor aggregate shocks, $\alpha = 0$. Hence, the competitive solution under perfect anchor is the same as the social planner's solution. 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 under any structural parameter values: traders enter twice as often than they should. The intuition is analogous to the "tragedy of commons". While each trader internalizes that if others enter more often, that reduces it valuation, she does not internalize that when she enters that reduces the benefit of entry for everyone else. Note that as $\beta - \alpha >$ for any of our structural parameters, there is over-entry even if all type of liquidity shocks have zero probability. In this particular case, the externality comes from the decreasing returns to scale technology.

Lemma 2 makes it clear that simply increasing the mass of sophisticated traders in a market does not mean that entry converges to the socially optimal level. Using the above analysis, on can draw implications about specific markets. Anchored trades (low μ) with high α do not have enough traders entering. One example for such a market is twin stocks: early entrants need later entrants to be able to exit the trade with a profit. 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 revenue of traders to zero, like in a tragedy of commons game.

We now turn to welfare as a function of the mass of traders M. In both extreme cases it is the same as the aggregate revenue of traders since traders do not spend anything on learning: either there is no learning (under $\mu \to \infty$) or learning is free ($\mu = 0$). Thus in these extreme cases, welfare equals the measure of the efficiency of capital allocation.

Lemma 3. *With a perfect anchor ($\mu = 0$):*

$$W = \begin{cases} M - \frac{M^2(\beta-\alpha)}{2} & \text{if } M \leq \frac{1}{\beta} \\ \frac{(\alpha+\beta)}{2\beta^2} & \text{if } M > \frac{1}{\beta}, \end{cases}$$

without anchor ($\mu \to \infty$):

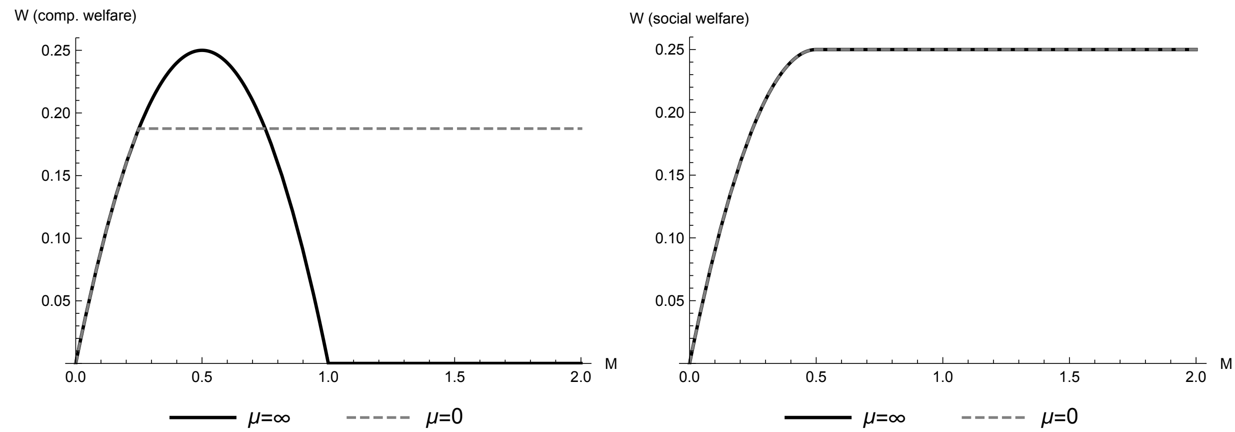
$$W = \begin{cases} M - \frac{M^2(\beta-\alpha)}{2} & \text{if } M \leq \frac{2}{\beta-\alpha} \\ 0 & \text{if } M > \frac{2}{\beta-\alpha}. \end{cases}$$

Under the planner, both with and without anchor:

$$W_s = \begin{cases} M - \frac{M^2(\beta-\alpha)}{2} & \text{if } M \leq \frac{1}{\beta-\alpha} \\ \frac{1}{2(\beta-\alpha)} & \text{if } M > \frac{1}{\beta-\alpha}. \end{cases}$$

Thus increasing the amount of sophisticated traders M in a market has a different effect on welfare depending on the strength of the anchor μ, see Figure 2. With a perfect anchor, welfare monotonically increases up to the point $M = \frac{1}{\beta}$ and is flat afterwards. 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)

Figure 2: **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: $\beta = 4$, $\alpha = 2$.

lowering welfare compared to the social level. Without an anchor, welfare follows the same increasing path as with a perfect anchor but continues to grow to a higher level $M = \frac{1}{\beta-\alpha}$ if $\alpha > 0$. The reason is that without an anchor, if $\alpha > 0$, late traders cannot figure out they are late and do not stay out of the market. As the mass of traders increases further, all revenues are gradually competed away and $W = 0$ for any $M > \frac{2}{\beta-\alpha}$. In this case the model reverts to a simple tragedy of commons. It is immediately clear from the Figure 2 that a perfect anchor is not always preferred to no anchor for a given level of the mass of traders M, we further analyze this question in a more general setup in Section 4.4.

3.2 Optimal strategies

The private 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:

$$\max_{m(\theta)} \int_0^1 \left(m(\theta) \cdot \Delta u(\theta) - \mu \cdot L(m)\right) d\theta. \tag{19}$$

We contrast the private solution with that of a social planner who can choose the amount of learning and entry for all traders. 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: it takes into account that Δu depends not only on θ but on the choice function of all other traders m. The social planner chooses the symmetric function $m_s(\theta)$ to maximize

$$\max_{m_s(\theta)} \int_0^1 \left(m_s(\theta) \cdot \Delta u(\theta, m_s) d\theta - \mu \cdot L(m_s)\right) d\theta \tag{20}$$

We derive the first order condition (FOC) of these problems using the variation method, i.e. we look for the function $m(\theta)$ such that if we take a very small variation around the function, the value function of the traders does not change.

Lemma 4. *Denote the strategy function of all other players as $\tilde{m}(\theta)$. The first-order condition of the private problem is:*

$$M \cdot \alpha \cdot \int_\theta^1 \tilde{m}(\tilde{\theta}) d\tilde{\theta} - M \cdot \beta \cdot \int_0^\theta \tilde{m}(\tilde{\theta}) d\tilde{\theta} + 1 = \mu \cdot \left[\log\left(\frac{m(\theta)}{1 - m(\theta)}\right) - \log\left(\frac{p}{1-p}\right)\right]. \tag{21}$$

The first-order condition of the social problem (assuming the same entry function m_s for all traders) is:

$$M \cdot (\alpha - \beta) \cdot p_s + 1 = \mu \cdot \left[\log\left(\frac{m_s(\theta)}{1 - m_s(\theta)}\right) - \log\left(\frac{p}{1-p}\right)\right] \tag{22}$$

Using the FOC one can derive an ordinary differential equation for $m(\theta)$ where the original FOC at $\theta = 0$ (an integral equation) is the boundary condition. The solution of this ordinary differential equation can be expressed implicitly up to a constant (boundary value) $m(0)$.

Proposition 1. *If $M \leq \bar{M}$, $m(\theta) = 1$. If $M > \bar{M}$, $m(\theta)$ is given by the following implicit functional equation*

$$\frac{1}{m(\theta)} + \log\left(\frac{1 - m(\theta)}{m(\theta)}\right) - \frac{M(\alpha + \beta)}{\mu} \cdot \theta = \frac{1}{m(0)} + \log\left(\frac{1 - m(0)}{m(0)}\right), \tag{23}$$

where $m(0)$ is pinned down by the boundary condition ((21) evaluated at $\theta = 0$):

$$M \cdot \alpha \cdot p + 1 = \mu \cdot \left[\log \left(\frac{m(0)}{1 - m(0)} \right) - \log \left(\frac{p}{1-p} \right) \right]. \tag{24}$$

$\bar{M}$ is pinned down by the following implicit equation:

$$\frac{\bar{M} \cdot (\alpha + \beta)}{\mu} = e^{-\frac{1-\beta \cdot \bar{M}}{\mu}} - e^{-\frac{1+\alpha \cdot \bar{M}}{\mu}}. \tag{25}$$

Proposition 2. *If $M \leq \bar{M}_s = \frac{1}{\beta - \alpha}$ then $m_s(\theta) = 1$. If $M > \bar{M}_s$ then:*

$$m_s(\theta) = \frac{1}{M \cdot (\beta - \alpha)}. \tag{26}$$

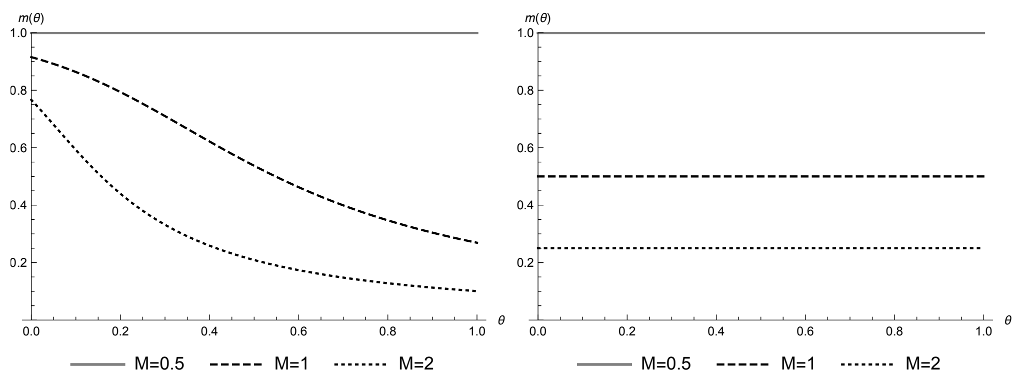
Note that traders want to differentiate between states, but the planner does not. The planner chooses a flat entry function. The reason for this over-learning is the rat race ($\alpha + \beta > 0$) that accompanies the crowding: every trader wants to know whether he is ahead of the other traders even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

To understand the effect of M on incentives, see Figure 3 which shows the competitive and social planner's optimal entry function for different levels of M. For small $M = 0.5$, traders enter for sure and there is no need for learning since revenues in the market are high. To gain intuition on how $m(\theta)$ changes as M increases from 1 to 2, consider the effect of larger M on the benefit of entry for a given trader $\Delta u(\theta)$, keeping the others' strategy constant. First, note that we can measure the relative incentive for entering earlier by differentiating $\Delta u(\theta)$ in θ, giving $M \cdot (\alpha + \beta) \tilde{m}(\theta)$. Therefore, keeping other trader's strategy fixed, the incentive to learn more and follow a more differentiated strategy is increasing in M. Loosely speaking, this results in a steeper $m(\theta)$. Given that this effect is scaled by the rat race parameter, $(\alpha + \beta)$, we refer to this as the rat race effect. Second, note that the benefit of entry for the average trader is $\Delta u\left(\frac{1}{2}\right) = M \cdot \alpha \cdot \int_{\frac{1}{2}}^{1} \tilde{m}(\tilde{\theta}) d\tilde{\theta} - M \cdot \beta \cdot \int_{0}^{\frac{1}{2}} \tilde{m}(\tilde{\theta}) d\tilde{\theta}$, hence, keeping others' strategy constant

$$\frac{\partial \Delta u\left(\frac{1}{2}\right)}{\partial M} = \alpha \cdot \int_{\frac{1}{2}}^{1} \tilde{m}(\tilde{\theta}) d\tilde{\theta} - \beta \cdot \int_{0}^{\frac{1}{2}} \tilde{m}(\tilde{\theta}) d\tilde{\theta} < (\alpha - \beta) \frac{p}{2} < 0$$

where the first inequality comes from the fact that $\tilde{m}(\theta)$ is decreasing in equilibrium. This suggests that for the average trader, increasing M is deceasing the incentive to enter. Given that this effect is scaled by the crowding parameter $(\alpha - \beta)$, we refer to this as the crowding effect. While in equilibrium the strategy of other traders, $\tilde{m}(\tilde{\theta})$, also changes, implying further adjustments, as Figure 3 demonstrate, the total effect is still driven by this intuition.

Figure 3: **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) for different levels of M. Parameters: $\beta = 4$, $\alpha = 2$, $\mu = 1$.

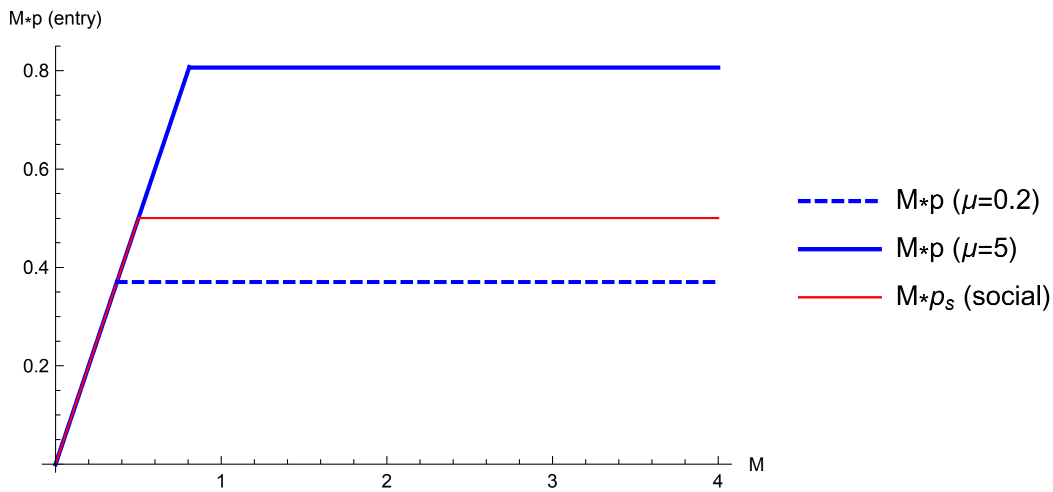
3.3 Allocative efficiency: over- and under-entry

Since the allocative efficiency of capital in the markets depends on the overall entry of all traders, in this subsection we analyze how aggregate entry $M \cdot p$ changes as the mass of traders M grows.

Proposition 3. *If $M \leq \bar{M}$, the competitive aggregate entry is $M \cdot p = M$. If $M > \bar{M}$, aggregate entry is constant in M at $M \cdot p = \bar{M}$. The aggregate entry in the social planner's solution is $M \cdot p_s = \min\left(M, \frac{1}{\beta - \alpha}\right)$.*

Just as in the planner's solution, traders in the competitive equilibrium also enter with probability one for small M and aggregate entry $M \cdot p$ is constant when M is large. However, that constant level, $\bar{M}$, is different from the social optimum $\bar{M}_s$. That is, whenever $M > \bar{M}$ increasing the number of sophisticated traders neither improves the efficiency of capital allocation, nor does it lead to additional crowding. Figure 4 illustrates this part by showing the amount of total entry $M \cdot p$ as a function of the mass of traders M.

Figure 4: **aggregate entry in the number of traders M**



Aggregate entry in the mass M of traders allowed to invest. The thin solid line is the socially optimal level of aggregate entry. The thick solid line refers to the case $\mu = 5$ when there is over-entry in the competitive equilibrium. The thick dashed line refers to $\mu = 0.2$ when there is under-entry in the competitive equilibrium. Parameters: $\beta = 4$, $\alpha = 2$.

For the intuition, recall that changing M changes the optimal strategy $m(\theta)$ for every trader through the rat race effect and the crowding effect. As the rat race effect primarily affects the slope of the entry function, as opposed to its level, it has little influence on $M \cdot p$. In contrast, due to the crowding effect the average entry, p, decreases. It turns out that in equilibrium, the decrease in p is exactly proportional to the increase in M, keeping $M \cdot p$ constant. While the exact proportionality is due to the technical properties of entropy, the choice of entropy is not ad-hoc but has a sound axiomatic foundation.

Proposition 3 shows that whether there is under- or over-entry for $M > \bar{M}$ is independent of the mass of traders considering to enter. Instead, as we state in Proposition 4, it is determined by all the other characteristics of the capital reallocation problem. More frequent aggregate liquidity shocks (larger η) and less market depth (higher δ_c) make markets more crowded since they increase fire sales externalities. Less anchored strategies lead to more crowding, because the game is closer to a tragedy of commons problem as explained in Section 3.1. A faster decrease in marginal product of capital (higher δ) in the technology also makes the market more crowded. On the other hand, more frequent idiosyncratic liquidity shocks (larger ν) makes the market less crowded since it leads to under-entry due to late entrants not internalizing the positive effect they have on earlier entrants.

Proposition 4. *For $M > \max[\bar{M}, \bar{M}_s]$ the relative amount of competitive aggregate entry to social aggregate entry $\frac{\bar{M}}{\bar{M}_s}$ is*

i) increasing in μ

ii) increasing in δ

iii) increasing in δ_c

iv) increasing in η

v) decreasing in ν

3.4 Decoupling of welfare and allocative efficiency

Now we turn to the question of welfare as more and more sophisticated traders enter. We show that the presence of some sophisticated traders ($M < \bar{M}$) unambiguously increases welfare in the competitive equilibrium. Note that in the small M case welfare does not depend on μ since no resources are spent on learning. 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. Thus allocative efficiency and welfare go hand-in-hand. The above insight that larger M means (at least weakly) higher welfare and a more efficient capital allocation remains true in case of the social planner's optimum since no learning is chosen in that case.

In the competitive equilibrium, raising M above $\bar{M}$ leads to decoupling of welfare and allocative efficiency. While allocative efficiency stays constant (even though at a suboptimal level), welfare drops. 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 because of crowding (the total amount of traders entering is constant) but because of wasteful learning.

Proposition 5. *If $M > \bar{M}$, the efficiency of capital allocation (aggregate revenue of traders) stays constant as we increase M. However welfare becomes decoupled from allocative efficiency, welfare converges to zero from above as $M \to \infty$:*

$$W(\bar{M}) > W(M \to \infty) = 0 \tag{27}$$

In general, one can write:

$$W(M) = M \cdot \int_0^1 \log\left(\frac{1-p}{1-m(\theta)}\right) d\theta \tag{28}$$

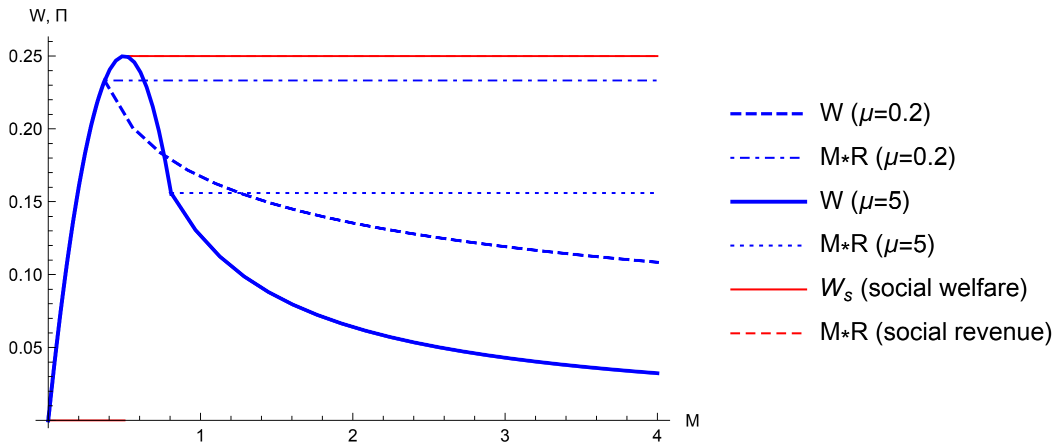
The welfare in the social planner's optimum for $M > \bar{M}_s$ is constant:

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

Thus while learning is useful in limiting crowding (compare the above welfare to $W = 0$ for $M > \frac{2}{\beta - \alpha}$ if $\mu \to \infty$), it does use large amounts of resources. In the context of our structural model, it implies that increasing the mass of traders e.g. from $M < \bar{M}$ to $M > \bar{M}$ might make markets more efficient from an allocative point of view and decrease welfare at the same time. This is due to the fact that increased market efficiency is achieved at the cost of over-learning.

Figure 5 shows that as more sophisticated traders enter, welfare eventually deteriorates even though the revenue of the market (and thus efficiency) is unchanged. Thus for $M \to \infty$, while crowding is limited by learning, there is no welfare gain compared to $\mu \to \infty$ because all resources that are saved in limiting crowding are used in learning.

Figure 5: **Welfare in the number of traders** M



Total revenue and welfare in the mass M of traders allowed to invest. The thin solid line is the revenue and welfare in the social optimum. The thick solid line is welfare and the thin dashed line revenue in the case $\mu = 5$. The thick dashed line is welfare and the thin dot-dashed line is total revenue in the case $\mu = 0.2$. Parameters: $\beta = 4$, $\alpha = 2$.

If there are other ways to limit entry without learning, that might be welfare improving. In fact if we limit the mass of traders before they learn their type (indiscriminately) one can improve welfare. The intuition is that limiting entry decreases the effective M and thus limits the incentive to learn, see the discussion following Proposition 2. 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 More microfoundations and benefits of a weak anchor

4.1 Consumer/producer surplus

It is reasonable to assume that traders are not the only traders 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 traders.

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

$$V^P \equiv \alpha_w \cdot \int_0^1 m(\theta) \cdot a(\theta) \, \mathrm{d}\theta = \alpha_w \cdot \frac{(M \cdot p)^2}{2} \tag{30}$$

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

$$W \equiv M \cdot V + V^P \tag{31}$$

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:

$$\Delta u = (\alpha + \alpha_w) \cdot a(\theta) - \beta \cdot b(\theta) + 1 \tag{32}$$

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:

$$U(q) = q - \frac{\alpha_w}{2} q^2 \tag{33}$$

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

$$\text{price}(q) = 1 - \alpha_w \cdot q \tag{34}$$

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, revenue from moving.

$$\text{cost}(\theta) = (\alpha + \beta) b(\theta) - (\alpha + \alpha_w)(a(\theta) + b(\theta)) \tag{35}$$

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:

$$\Delta u = \text{price}(q) - \text{cost}(\theta) = 1 - \alpha_w \cdot q - (\alpha + \beta)b(\theta) + (\alpha + \alpha_w)(a(\theta) + b(\theta)) \tag{36}$$

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:

$$\text{CS} = \int_0^q (1 - \alpha_w \cdot \hat{q} - \text{price}(q))d\hat{q} = \int_0^q \alpha_w \cdot (q - \hat{q})d\hat{q} = \alpha_w \cdot \frac{q^2}{2} \tag{37}$$

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

$$\text{CS} = \int_0^1 m(\theta) \cdot [\alpha_w \cdot a(\theta)]\, d\theta = -\frac{\alpha_w}{2} a(\theta)^2 \Big|_0^1 = \frac{\alpha_w}{2} a(1)^2 = \alpha_w \cdot \frac{q^2}{2} \tag{38}$$

This means that the objective function for the social planner maximizing welfare is the same as (1) substituting α with $\alpha + \alpha_w$ yielding (32).

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:

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

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:

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

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 trader ($\alpha_w = 0$).

Proposition 6. *Welfare is higher in case of no anchor than in case of perfect anchor if*

$$\alpha_w > \frac{(\beta - \alpha)^2 (\alpha + \beta)}{\alpha (2\beta - \alpha)} \tag{41}$$

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

5 Conclusions

We develop a model of capital reallocation to analyze whether the presence of more sophisticated traders improve capital allocation and welfare. Trades can become crowded due to externalities but traders can devote resources to learn about the number of earlier entrants. In general, more traders having the choice to enter neither improves the efficiency of capital allocation nor does it aggravate crowding. In fact, whether there is eventually too little or too much capital allocated to the new sector is determined solely by the technology in that sector, the cost of learning, the depth of the market, and the severity of the potential shocks, but not the mass of sophisticated traders present. However, the presence of more traders decreases welfare, as they waste more aggregate resources in learning about each others' position.

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.

MacKay, David J.C., *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, 2003.

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.

Sims, Christopher A., "Stickiness," *Carnegie-Rochester Conference Series on Public Policy*, December 1998, *49* (1), 317–356.

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," *Journal of Economic Theory*, 2015, *forthcoming.* 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 Proofs

Proof of Lemma 1

Proof. The proof follows that of Lemma 1 in Woodford (2008). □

Proof of Lemma 2

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

$$-M\cdot\beta\cdot\bar{\theta}+1=0 \tag{42}$$

yielding Eq. 15. For the social planner's optimum one has to find $\bar{\theta}$ that everyone with $\theta<\bar{\theta}$ enters, the others stay out, maximizing:

$$\int_0^{\bar{\theta}}\big((M\cdot\alpha)\cdot(\bar{\theta}-\theta)-M\cdot\beta\cdot\theta+1\big)\,d\theta=\frac{M\cdot\alpha-M\cdot\beta}{2}\cdot\bar{\theta}^2+1\cdot\bar{\theta} \tag{43}$$

yielding the interior optimum in Eq. 16 if $M\cdot(\beta-\alpha)>1$. If on the other hand, $M\cdot(\beta-\alpha)<1$, everyone enters: $m(\theta)=1$ 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:

$$\int_0^1(M\cdot\alpha\cdot(1-\theta)\cdot p-M\cdot\beta\cdot\theta\cdot p+1)\,d\theta=0 \tag{44}$$

yielding the unconditional entry probability in Eq. 17. If M is low and the implied entry is >1, then the revenue is not driven to zero and everyone enters for sure implying $p=1$. In the social planner's optimum every trader enters with probability p and they maximize social planner's welfare

$$\int_0^1 p\cdot(M\cdot\alpha\cdot(1-\theta)\cdot p-M\cdot\beta\cdot\theta\cdot p+1)\,d\theta \tag{45}$$

taking derivative w.r.t. p and setting to zero, this implies the entry probability in Eq. 18. As before, of the implied entry probability is >1, then everyone enters for sure $m(\theta)=1$ implying $p_s=1$. Note that there are infinite other solutions since the social planner does not care about who exactly enters. □

Proof of Lemma 3

Proof. Note that in both benchmark cases $\mu \cdot L = 0$ thus $W = M \cdot R$. The values for R follow from straightforward algebra using the entry functions from Lemma 2. □

Proof of Lemma 4

Proof. For the private FOC we use a perturbation method similar to the proof in Yang (2015). In the first order perturbation we set $m(\theta) + \nu \cdot \epsilon(\theta)$ as $m(\theta)$, while we keep the entry decision of the others $\tilde{m}$ fixed:

$$\int_0^1 \left((m(\theta) + \nu \cdot \epsilon(\theta)) \cdot \Delta u(\tilde{m}, \theta) - \mu \cdot L(m(\theta) + \nu \cdot \epsilon(\theta)) \right) d\theta. \tag{46}$$

We then take derivative wrt ν and then set $\nu = 0$ yielding the FOC:

$$\int_0^1 \epsilon(\theta) \cdot \left(\Delta u(\tilde{m}, \theta) - \mu \cdot \left[\log \left(\frac{m(\theta)}{1 - m(\theta)} \right) - \log \left(\frac{\int_0^1 m(\tilde{\theta}) d\tilde{\theta}}{1 - \int_0^1 m(\tilde{\theta}) d\tilde{\theta}} \right) \right] \right) d\theta = 0. \tag{47}$$

Since the original equation is an optimum, the above equality has to hold for any $\epsilon(\theta)$: thus the part multiplying $\epsilon(\theta)$ has to be zero for all θ. Setting $\tilde{m} = m$ we arrive at the symmetric solution we get (21).

For the social FOC we also use a perturbation method similar to the proof in Yang (2015). 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 $\epsilon(\theta)$:

$$\int_0^1 \epsilon(\theta) \cdot \left(M \cdot \alpha \cdot \int_\theta^1 m_s(\tilde{\theta}) d\tilde{\theta} - M \cdot \beta \cdot \int_0^\theta m_s(\tilde{\theta}) d\tilde{\theta} - \mu \cdot \left[\log \left(\frac{m_s(\theta)}{1 - m_s(\theta)} \right) - \log \left(\frac{\int_0^1 m_s(\tilde{\theta}) d\tilde{\theta}}{1 - \int_0^1 m_s(\tilde{\theta}) d\tilde{\theta}} \right) \right] \right) d\theta + \tag{48}$$

$$+ \int_0^1 m_s(\theta) \cdot \left(M \cdot \alpha \cdot \int_\theta^1 \epsilon(\tilde{\theta}) d\tilde{\theta} - M \cdot \beta \cdot \int_0^\theta \epsilon(\tilde{\theta}) d\tilde{\theta} \right) d\theta = 0 \tag{49}$$

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:

$$M \cdot \alpha \cdot \int_\theta^1 m_s(\tilde{\theta}) d\tilde{\theta} - M \cdot \beta \cdot \int_0^\theta m_s(\tilde{\theta}) d\tilde{\theta} + 1 - \mu \cdot \left[\log \left(\frac{m_s(\theta)}{1 - m_s(\theta)} \right) - \log \left(\frac{\int_0^1 m_s(\tilde{\theta}) d\tilde{\theta}}{1 - \int_0^1 m_s(\tilde{\theta}) d\tilde{\theta}} \right) \right] + \tag{50}$$

$$+ M \cdot \alpha \cdot \int_0^\theta m_s(\tilde{\theta}) d\tilde{\theta} - M \cdot \beta \cdot \int_\theta^1 m_s(\tilde{\theta}) d\tilde{\theta} = 0 \tag{51}$$

which simplifies to (22). □

Proof of Proposition 1

Proof. Differentiating the FOC (21) we arrive at the following differential equation:

$$(M \cdot \alpha + M \cdot \beta) \cdot \tilde{m}(\theta) = -\frac{\mu \cdot m'(\theta)}{m(\theta) \cdot (1 - m(\theta))}. \tag{52}$$

thus the competitive equilibrium strategy $m(\theta)$ in the symmetric equilibrium ($m = \tilde{m}$) has to solve the above differential equation with the original FOC (e.g. evaluated at $\theta = 0$) as a boundary condition which is (24). If there is an interior solution (s.t. $m(\theta) \neq 1$), it can be written in the form

$$\frac{\frac{1}{m(\theta)} + \log\left(\frac{1-m(\theta)}{m(\theta)}\right)}{M(\alpha+\beta)} = C + \frac{1}{\mu} \tag{53}$$

for an appropriate constant C. Setting $\theta = 0$ and $\theta = 1$ above and eliminating C yields (23).

To calculate the level of $\bar{M}$ we use the observation (independently proven in Proposition 3) that $M \cdot p$ is constant, including in the limit as $M \to \infty$. At $\bar{M}$ still all traders enter with probability 1, thus $p = 1$ and $\bar{M}$ can be expressed as:

$$\bar{M} = \lim_{\mu \to \infty} (M \cdot p) \tag{54}$$

Thus we focus on expressing $M \cdot p$ in the limit for large M. As a first step note that as $M \to \infty$, given that $M \cdot p$ is constant, $m(\theta) \to 0$ for every θ. Thus the implicit equation (23) for $m(\theta)$ can be approximated by

$$\frac{1}{m} - M(\alpha+\beta)\left(C + \frac{\theta}{\mu}\right) = 0 \tag{55}$$

since for $m \approx 0$: $\frac{1}{m} \gg \log\left(\frac{1}{m}\right)$. A closed form solution can be obtained in this limit case:

$$m(\theta) = \frac{\mu}{M(\alpha+\beta)(C\mu+\theta)} \tag{56}$$

for a specific C. By the definition of the average entry p this implies

$$M \cdot p = M \cdot \int_0^1 m(\theta) d\theta = \frac{\mu}{\alpha+\beta} \cdot \log\left(\frac{1}{C\mu} + 1\right). \tag{57}$$

Substituting this into the boundary condition (24) yields:

$$\alpha \frac{\mu}{\alpha+\beta} \log\left(\frac{1}{C\mu}+1\right)+1=\mu\left[\log\left(\frac{1}{CM(\alpha+\beta)-1}\right)-\log\left(\frac{M(\alpha+\beta)-\mu\log\left(\frac{1}{C\mu}+1\right)}{\mu\log\left(\frac{1}{C\mu}+1\right)}\right)\right] \tag{58}$$

Since $M \cdot p$ is a constant for any $M > \bar{M}$, C also has to converge to a finite constant as $M \to \infty$. Using this insight, one can take the limit of the above equation as $M \to \infty$:

$$\mu(-\alpha-\beta)\log\left(\frac{1}{C}\right)+(\alpha+\beta)\left(\mu\log\left(\mu\log\left(\frac{1}{C\mu}+1\right)\right)+1\right)+\alpha\mu\log\left(\frac{1}{C\mu}+1\right)=0 \tag{59}$$

Using the relation between C and $M \cdot p$ in (57), one can eliminate C:

$$(\alpha+\beta)\left(\mu\log(M\cdot p\cdot(\alpha+\beta))-\mu\log\left(\mu\left(e^{\frac{M\cdot p\cdot(\alpha+\beta)}{\mu}}-1\right)\right)+\alpha M\cdot p+1\right)=0 \tag{60}$$

using (54) and rearranging yields equation (25) in the proposition. ☐

Proof of Proposition 2

Proof. The derivative of FOC (22) w.r.t. θ delivers the differential equation

$$0=-\frac{\mu\cdot m_s'(\theta)}{m_s(\theta)\cdot(1-m_s(\theta))} \tag{61}$$

subject to the boundary condition (setting $\theta=0$ in (22))

$$M\cdot(\alpha-\beta)\cdot p_s+1=\mu\cdot\left[\log\left(\frac{m_s(0)}{1-m_s(0)}\right)-\log\left(\frac{p_s}{1-p_s}\right)\right]. \tag{62}$$

This trivially yields

$$m_s(\theta)=C \tag{63}$$

for some constant C, implying $p_s=C$. The boundary condition (62) simplifies to

$$M\cdot(\alpha-\beta)p_s+1=0 \tag{64}$$

implying (26). If the implied entry probability is > 1, then we have the corner solution that all enter with $m(\theta) = 1$. $\square$

Proof of Proposition 3

Proof. To show that $M \cdot p$ is constant in M once the solution m is interior, first write the system of 3 equations determining p. First, the difference of FOC (21) at $\theta = 0$ and $\theta = 1$.

$$p = \frac{\mu\left(\log\left(\frac{m(0)}{1-m(0)}\right) - \log\left(\frac{m(1)}{1-m(1)}\right)\right)}{M(\alpha+\beta)} \tag{65}$$

Second, the boundary condition (21) at $\theta = 0$

$$\alpha M p + 1 = \mu\left(\log\left(\frac{m(0)}{1-m(0)}\right) - \log\left(\frac{p}{1-p}\right)\right). \tag{66}$$

Third, the implicit equation for $m(\theta)$ evaluated at $\theta = 1$.

$$\log\left(\frac{m(0)}{1-m(0)}\right) - \log\left(\frac{m(1)}{1-m(1)}\right) = \frac{M(\alpha+\beta)}{\mu} + \frac{1}{m(0)} - \frac{1}{m(1)} \tag{67}$$

Substituting

$$x_0 = \log\left(\frac{m(0)}{1-m(0)}\right) \tag{68}$$

and

$$x_1 = \log\left(\frac{m(1)}{1-m(1)}\right) \tag{69}$$

the system of three equations can be written as:

$$p = \frac{\mu(x_0 - x_1)}{M(\alpha+\beta)} \tag{70}$$

$$\alpha M p + 1 = \mu\left(x_0 - \log\left(\frac{p}{1-p}\right)\right) \tag{71}$$

$$x_0 - x_1 = \frac{M(\alpha+\beta)}{\mu} + e^{-x_0} - e^{-x_1} \tag{72}$$

Substituting out p from (70), (71), (72) we arrive at a system of two equations:

$$F = \mu\left(x_0 - \log\left(\frac{\mu(x_0 - x_1)}{M(\alpha+\beta) + \mu(x_1 - x_0)}\right)\right) - \left(\frac{\alpha\mu(x_0 - x_1)}{\alpha+\beta} + 1\right) = 0 \tag{73}$$

$$G = \frac{M(\alpha+\beta)}{\mu} - (x_0 - x_1) + e^{-x_0} - e^{-x_1} = 0 \tag{74}$$

To prove $M \cdot p$ is constant, it is sufficient to prove $\frac{\partial(M \cdot p)}{\partial M} = 0$ which from (70) is equivalent to

$$\frac{\partial x_0}{\partial M} = \frac{\partial x_1}{\partial M} \tag{75}$$

We apply Cramer's rule both for x_0 and x_1 to the system of equations (73) and (74):

$$\frac{\partial x_0}{\partial M} = \frac{\begin{vmatrix} \frac{\partial F}{\partial x_0} & -\frac{\partial F}{\partial M} \\ \frac{\partial G}{\partial x_0} & -\frac{\partial G}{\partial M} \end{vmatrix}}{\begin{vmatrix} \frac{\partial F}{\partial x_0} & \frac{\partial F}{\partial x_1} \\ \frac{\partial G}{\partial x_0} & \frac{\partial G}{\partial x_1} \end{vmatrix}} \tag{76}$$

$$\frac{\partial x_1}{\partial M} = \frac{\begin{vmatrix} -\frac{\partial F}{\partial M} & \frac{\partial F}{\partial x_1} \\ -\frac{\partial G}{\partial M} & \frac{\partial G}{\partial x_1} \end{vmatrix}}{\begin{vmatrix} \frac{\partial F}{\partial x_0} & \frac{\partial F}{\partial x_1} \\ \frac{\partial G}{\partial x_0} & \frac{\partial G}{\partial x_1} \end{vmatrix}} \tag{77}$$

We check numerically that the denominator (which is the same for both derivatives) is non-zero thus the two equations are indeed independent. It is thus sufficient to show that the numerators of the Cramer rule for the two derivatives are equal, yielding the sufficient condition

$$\frac{(\alpha+\beta)e^{-x_0-x_1}\left(e^{x_0+x_1}(M(\alpha+\beta) + \mu(x_1 - x_0)) - \mu e^{x_0} + \mu e^{x_1}\right)}{M(\alpha+\beta) + \mu(x_1 - x_0)} = 0. \tag{78}$$

It follows from (72) that the denominator is non-zero if $x_0 \neq x_1$. Thus it is sufficient to prove that

$$\frac{M(\alpha+\beta)}{\mu} + (x_1 - x_0) + \frac{1}{e^{x_0}} - \frac{1}{e^{x_1}} = 0, \tag{79}$$

which is exactly the function $G = 0$ defined in (74). Thus the identity holds and we have proved that $M \cdot p$ is constant in M for interior solutions.

$\square$

Proof of Proposition 4

Proof. [TBA: WE CURRENTLY ONLY HAVE NUMERIC PROOFS]

$\square$

Proof of Proposition 5

Proof. By Proposition 1 when $M < \bar{M}$ all traders enter with probability 1. Hence, all equilibrium objects are the same for the planner and in the decentralized solution. In particular, average entry of an trader is $p = 1$ thus expected aggregate entry is M. Total revenue and welfare are

$$M \cdot R = W = M \cdot R_s = W_s = M \cdot \int_0^1 (M \cdot \alpha \cdot (1-\theta) - M \cdot \beta \cdot \theta + 1)\, d\theta = M - \frac{M^2(\beta - \alpha)}{2}. \tag{80}$$

To arrive at the formula for $W(M)$ one can rearrange the aggregate learning from (9) to get:

$$M \cdot L = M \int_0^1 m(\theta) \cdot \mu \cdot \left(\log\left(\frac{m(\theta)}{1-m(\theta)}\right) - \log\left(\frac{p}{1-p}\right)\right) d\theta - M \int_0^1 \mu \log\left(\frac{1-p}{1-m(\theta)}\right) d\theta \tag{81}$$

where the interior part of the first integral multiplying $m(\theta)$ can be replaced using the FOC (21) to yield:

$$M \cdot L = M \int_0^1 m(\theta) \cdot \left[M \cdot \alpha \cdot \int_\theta^1 \tilde{m}(\tilde{\theta}) d\tilde{\theta} - M \cdot \beta \cdot \int_0^\theta \tilde{m}(\tilde{\theta}) d\tilde{\theta} + 1 \right] d\theta - M \int_0^1 \mu \log\left(\frac{1-p}{1-m(\theta)}\right) d\theta \tag{82}$$

thus the first integral is exactly the definition of aggregate revenue. Since $M \cdot p$ is constant if $M \geq \bar{M}$ (Propostion 3), so is aggregate revenue $M \cdot R$. Rearranging yields expression (28) for W in the proposition.

We now show that welfare converges to zero for large M. For large M, $m \approx 0$ and $p \approx 0$ thus in the $M \to \infty$ limit (28) converges to zero. This convergence happens from above, since the payoff per trader $\frac{W}{M}$ cannot be negative, otherwise traders would choose not to enter.

In the social planner's interior optimum every trader enters with probability $p = \frac{1}{M \cdot (\beta - \alpha)}$ and thus welfare becomes

$$W_s = M \cdot \int_0^1 p \cdot (M \cdot \alpha \cdot (1-\theta) \cdot p - M \cdot \beta \cdot \theta \cdot p + 1)\, d\theta = \frac{1}{2 \cdot (\beta - \alpha)}. \tag{83}$$

□

Proof of Proposition 6

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

$$\frac{4\alpha_w}{2(\alpha-\beta)^2} > \frac{M^2\left(\alpha+\beta+\alpha_w\right)}{2(\beta M)^2} \tag{84}$$

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. □