Learning in Crowded Markets *

Péter Kondor†

London School of Economics

Adam Zawadowski‡

Central European University

May 21, 2019

Abstract

We present a novel entry-game with endogenous information acquisition to study the welfare effects of opacity and competition. Potential entrants to an opaque market are uncertain about their competitive advantage relative to other investors, i.e. their type. They construct optimal costly signals to learn about their types, where the marginal cost of learning captures the opacity of the market. In general, the individually optimal entry and learning decisions are socially suboptimal. Players over-invest in learning and more opaque markets are associated with more crowding. Nevertheless, more opaque markets might still lead to higher welfare by implying a better trade-off between the degree of crowding and the total cost of learning. Similarly, decreasing the share of smart investors in the market might also improve welfare. However, fierce competition is always detrimental to welfare as it leads to more wasteful learning without changing the level of crowding.

*For comments and suggestions, we are grateful to Patrick Bolton, Thomas Chemmanur, Johannes Hörner, Victoria Vanasco, Marianne Andries, Ming Yang, Martin Oehmke, Gábor Lippner and anonymous referees. We also thank seminar participants at Boston University, Central European University, ESSET 2014 (Gerzensee), Brigham Young University, GLMM 2015 (Boston), Paul Woolley Conference 2015 (LSE), University of Vienna, 6th Annual Financial Market Liquidity Conference (Budapest), American Finance Association 2016 Annual Meeting (San Fransisco), Adam Smith Workshop 2016 (Oxford), University of Maryland, University of Naples, University of Copenhagen, 4nation cup workshop (Copenhagen), Federal Reserve Bank of New York, Toulouse School of Economics, University of Zurich. We thank Pellumb Reshidi, Dániel Szabó and Brandon (Yueyang) Han for their excellent research assistance. The support of the European Research Council Starting Grant #336585 is acknowledged.

†e-mail: p.kondor@lse.ac.uk, http://personal.lse.ac.uk/kondor

‡e-mail: zawadowskia@ceu.edu, http://www.personal.ceu.hu/staff/Adam_Zawadowski/

1 Introduction

In our rapidly changing global economy, entrepreneurs and financial investors have to be nimble and well prepared to survive. Fleeting investment opportunities arise routinely in global financial markets. Technological developments create new market segments by the day and make existing ones obsolete. It is insufficient to simply recognize new opportunities: The key to success for the potential entrant is to tell whether she was sufficiently early in noticing the opportunity and whether her existing skill sets make her able to compete efficiently. In light of this, it is not surprising that market participants, ranging from venture capitalists, global banks to leading technological companies, invest vast sums into analyzing markets, developing know-how and technology to help them decide whether to undertake investment opportunities. Also, new opportunities are heterogeneous. Some opportunities are in the public eye, leading to a large mass of would-be entrants standing on the sideline considering whether to jump into the fray. Some opportunities are opaque making it costly or even impossible to predict whether the existing skill sets will lead to success or failure. Which kinds of new market segments or trading opportunities are subject to overcrowding? Are the vast resources invested in learning about opaque opportunities socially useful? Should regulators push for transparency? In general, what are the welfare implications of learning about such new investment opportunities in the face of competition and opaqueness?

In this paper, we present a unique, parsimonious framework to answer these questions. We analyse a novel entry-game with endogenous information acquisition to study the welfare effects of opacity and competition. Potential entrants (players hereon) to an opaque market are uncertain about their competitive advantage relative to others, i.e. their type. They construct optimal signals to learn about their types subject to an entropy cost. The opacity of the market is captured by the marginal cost of learning, while the extent of competition is modelled by the mass of players standing on the sidelines. In general, the individually optimal entry and learning decisions are socially suboptimal. Players always over-invest in learning and more opaque markets tend to be more crowded. Nevertheless, transparency is not always optimal as more opaqueness might still lead to higher welfare by discouraging costly learning without excessively increasing crowding. Opaqueness is more likely to be beneficial if competition is not too fierce, while transparency is preferred if competition is excessive. Fiercer competition leads to more wasteful learning leading to deteriorating welfare without affecting crowding.

In our game, players' type is distributed over the unit interval. Each player's pay-off upon entry is a linear function of the mass of entrants with better types, and the mass of entrants with worse types. We focus on the case for which an entrant's revenue is decreased by an additional better entrant and increased by a worse entrant such that players engage in a "rat race". We also assume that the effect of an additional random entrant (if she is better or worse with equal probability) is also negative, leading to potential "crowding" in the market. Each player constructs an optimal signal structure to learn about her type. Building on Sims (1998, 2003), the cost of any given signal structure is proportional to the implied reduction in entropy. Given the equilibrium strategies, a player's posterior on her type implies a posterior on the mass of better and worse entrants. Hence, each player learns about her competitive advantage on the given market relative to other entrants.

Our formalization results in a parsimonious structure. In equilibrium, the optimal information acquisition and entry strategies are reduced to a single function, mapping player's possible types into a probability of entry. For example, if the player were to decide to enter with a given probability independently of her type, we represent this choice with a constant function. This strategy does not require learning. In contrast, if the player were to decide to enter if and only if her type is better than a given threshold, this can be represented by a step function. However, for this, her signal has to be sufficiently precise to know with certainty whether her type is above this threshold. This strategy turns out to be very costly under our specification. In equilibrium, players typically choose an interior strategy represented by a smooth, monotonic function, implying higher entry probability for better types.

Our main result is that whether a regulator should aim for more transparency depends on the level of competition. Typically, full transparency is preferred only if there is a large mass of players aware of the opportunity and ready to enter, i.e. if competition is fierce. Otherwise either an interior level of opacity or full opacity maximizes welfare. The intuition relies on two effects. First, full transparency leads to insufficient entry because worse players do not internalize their beneficial effect on better entrants. On the other hand, full opacity leads to crowding which is more severe if competition is more fierce because players unaware of their types exert a negative externality on other entrants. Therefore, less than full transparency might push entry closer to its efficient level especially when competition is limited. Second, there is costly over-learning in our model because of the rat-race between the players. This leads to a more subtle benefit of increased opacity. Less than full transparency might help to reduce the overall cost of learning. While decreasing transparency, by definition, increases the marginal cost of learning, it might reduce the amount of learning sufficiently that the overall cost decreases. The benefit of this reduction in learning expenditure can more than offset the welfare loss of increased crowding due to more opacity. Thus increasing opacity is more likely to improve welfare if competition is limited because the welfare loss from crowding is less of a concern.

An additional result is that fiercer competition decreases welfare. First we show that unless the mass of players is so small that the entry decision is trivial, increasing competition does not change aggregate entry, which stabilizes at an inefficiently low or high level. To understand this result, note that players adjust their entry decisions along two main dimensions as competition increases. First, the marginal benefit of knowing your type more precisely before entering is increasing in competition because there are more players with a better type, increasing the "rat race" among players. Second, with more players, "crowding" becomes a bigger concern. We show that these two effects have exactly offsetting effects on aggregate entry. Nevertheless, as competition increases, welfare decreases. The key insight is that the "rat race" effect increases with competition, thus players choose to learn more. While ceteris paribus more learning can alleviate inefficient over-entry (as stated above), more learning due to increased competition does not change the amount of entry. Thus the higher learning cost implied by more competition is socially wasteful.

We analyze two extensions. First, we show that when better types find socially more valuable deals in the new market, then more competition often leads to an allocation that is "too efficient" compared to the planner's solution. The reason is similar to that in the baseline model, it is due to over-learning that decreases welfare. Second, we also extend our model to the case in which there is heterogeneity across players: some are more sophisticated and thus can learn at a lower cost than others. Keeping the mass of all players fixed but increasing the share of sophisticated players might also decrease welfare. Initially, increasing the fraction of sophisticated players increases welfare since it raises the average sophistication of players and this can alleviate over-entry. However, further increasing the fraction of sophisticated players beyond a certain threshold, less sophisticated players are afraid of being ripped off and exit the market. Once less sophisticated players exit, sophisticated players engage in a vicious "rat race" of learning which leads to decreasing welfare, similar to the baseline model. Thus like in opacity, in many cases there is an intermediate mix of sophisticated and unsophisticated players that maximizes welfare. Identifying the most sophisticated players as high-frequency traders connects this result to the policy debate on the social benefit of ultra-high frequency trading.[1]

Our main contribution is to study the welfare effects of optimal learning in an entry game with uncertain competitive advantage. Our paper is connected to various branches of literature. First, there is a growing literature on the welfare effects of endogenous information acquisition, e.g. Myatt and Wallace (2012) and Colombo, Femminis, and Pavan (2014). While this literature focuses on a common-value learning, we analyze an environment when players learn about their relative advantage compared to the other entrants.

Second, from a methodological view point we rely on the rational inattention approach pioneered by Sims (1998, 2003). We follow the branch of the literature which allows for fully flexible informa-

[1]See, for example, Securities and Commission (2010).

tion acquisition as Matějka and McKay (2015), but restrict ourselves to binary actions similarly to Woodford (2008), Yang (2015,b).[2]

Third, there is a literature analyzing entry/exit in financial markets in the presence of externalities induced by other investors. Stein (2009) introduces a simple model of crowded markets. More generally, there is a classic literature on socially inefficient entry, e.g. Tullock (1967), Krueger (1974), and Loury (1979). We contribute to this literature by introducing a flexible, but costly learning technology and studying its welfare effects. Relatedly, Abreu and Brunnermeier (2003) and Moinas and Pouget (2013) show that the inability to learn about one's relative position versus that of other investors' is a key ingredient in sustaining excessive investment in bubbles. This highlights our contribution in adding learning to a model of crowded markets with potential over-entry.

Finally, there are numerous papers showing excessive investment in learning or effort. There is a literature on the social value of private learning: e.g. in Hirshleifer (1971), private information can be detrimental as it changes ex ante incentives for insurance, in Glode, Green, and Lowery (2012), learning affects ex-post trading opportunities. These papers study welfare effects in markets with asymmetric private information, while in our framework information is imperfect but symmetric.

The rest of the paper is structured as follows. In Section 2 we present our model. In Section 3 we analyze the optimal choice of entry and learning and the effects of opaqueness and competition on crowding and welfare. In Section 4 we consider extensions of the payoff function and also allow for heterogeneity in player sophistication. Section 5 concludes. All proofs are relegated to Appendix A. Further analysis can be found in the online appendices: In Appendix B we analyze median entrants. In Appendix C we give a structural microfoundation for the reduced form model and analyze its economic implications. We analyze Gaussian signals instead of fully flexible learning in Appendix D.

[2]The other successful approach is to allow for continuous actions, but restrict the signals to be Gaussian. See Maćkowiak and Wiederholt (2009), Hellwig and Veldkamp (2009) and Kacperczyk, Nieuwerburgh, and Veldkamp (2016) for intriguing models using this approach.

2 A model of learning and investing in crowded markets

In this part we describe our setup. We first present the payoff function, then introduce the flexible learning technology and define the real outcomes. Finally, we discuss potential interpretations and microfoundations of our reduced form model.

2.1 Payoffs

Consider an entry game with a continuum of players of mass M, each with a type θ uniformly distributed over $[0, 1]$. M measures the level of competition between the players. Each player can decide to take an action: whether to enter the market or not. θ characterizes the player's ability to identify better investment opportunities in this new market than others. Lower θ implies a better type. The utility gain (or loss, if negative) from entry is given by

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

where α and β are constant parameters. $b(\theta)$ denotes the equilibrium mass of entrants with a type better than θ. $a(\theta)$ denotes the equilibrium mass of entrants whose type is worse than θ. We show in the microfoundation in Appendix C that it is natural to assume that that $\beta > |\alpha|$. First, this implies that, $\beta + \alpha > 0$, which is without loss of generality since it is simply consistent with the interpretation that a lower θ represents a better type. Second, it follows that $\beta - \alpha > 0$, such that the median entrant imposes a negative externality on others, that is, the market is prone to getting crowded from a social point of view. It also follows that $\beta > 0$ while α could be positive or negative, though we focus most of our analysis on the more interesting case of $\alpha > 0$. When $\kappa > 0$, better players have an absolute advantage, that is, better types derive more utility from entering regardless of the entry decision of others. We discuss this case in Section 4.1, otherwise we analyze the simpler case of $\kappa = 0$.[3]

As we specify below, players do not know their type, but can gather information about it through a costly learning process.

2.2 Learning cost based on entropy

Before entry, players can engage in costly learning about their type. Observe that if $H(\cdot)$ is any intuitive measure of uncertainty then $H(\theta) - H(\theta|s)$, the reduction of uncertainty after observing signal s, is a measure of learning induced by signal s. Following Sims (1998), we measure uncertainty by specifying $H(\cdot)$ as the Shannon-entropy of a random variable.[4] Therefore, we specify the cost of learning a signal s as being proportional to the induced reduction in entropy of $\theta : H(\theta) - H(\theta|s)$. This quantity is often called the mutual information in θ and s. As Sims (1998) argues, the advantage of such a specification is that it both allows for flexible information acquisition and can be derived based on information theory. Note that the payoff (1) for a given θ in our model is linear in entry. Woodford (2008) derives the optimal signal structure and entry decision rule for such problems which we restate in the lemma below.

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

Thus, similar to Yang (2015) the conditional probability of entry $m(\theta)$, or equivalently, the conditional probability of getting a signal 1, is the only choice variable. The intuition for the binary signal

[3]In the main text, we work with the reduced form payoff (1). In Appendix C, we embed the reduced form game into an explicit model of capital reallocation. Also, in an earlier version Kondor and Zawadowski (2016), we provide microfoundations in various other contexts, including production with local spill-overs, consumption with externalities, and academic publications. The critical feature of all microfoundations is that each player's pay-off is lower if better types also enter, while worse entrants can either help or hurt. These applications provide further insights on the interpretation of parameters α,β and κ. We summarize these applications in section 2.4.

[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).

structure is that the only reason players 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 player 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 amount of learning L using the mutual information in type θ and signal s (defined in Lemma 1) as

$$L(m) \equiv H(\theta) - H(\theta|s) = H(s) - H(s|\theta) = \tag{2}$$
$$\left(p \log\left[\frac{1}{p}\right] + (1-p)\log\left[\frac{1}{1-p}\right]\right) - \int_0^1 \left(m(\theta)\log\left[\frac{1}{m(\theta)}\right] + (1-m(\theta))\log\left[\frac{1}{1-m(\theta)}\right]\right) \mathrm{d}\theta$$

where the first equation is a property of Shannon-entropy. p denotes the unconditional probability of entry and is defined by:

$$p = \int_0^1 m(\tilde{\theta})\mathrm{d}\tilde{\theta} \tag{3}$$

The expression for learning (2) can be understood in the following way. There is no learning if the signal is uninformative about the state, that is, if it prompts the player to enter with probability p unconditional on its type θ. Indeed, it is easy to check that when $m(\theta)$ is constant at p then $L(m) = 0$. Thus, learning depends on how much information the signal contains about the state. Intuitively, the steeper $m(\theta)$ becomes in θ (keeping average entry p constant), the more the player is differentiating its entry decision according to its type and the higher the entropy reduction, thus the higher the learning cost. The highest cost is achieved when $m(\theta)$ is a step function. Note that L is bounded from above but might generate infinite marginal cost of learning.

Our measure of the cost of learning induced by a signal defined in Lemma 1 is $\mu \cdot L(m)$ where μ is an exogenous marginal cost parameter. μ is our measure of the opacity of the market. We assume that players have to decide about 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 and optimally evaluating the right data.

Conveniently, standard results in information theory imply that the entropy of a random variable is proportional to the average number of bits needed to optimally convey its realization. Hence, the parameter μ can be interpreted as the cost of building a marginally larger information gathering and evaluating machine or writing a longer "code".[5]

2.3 Competitive and planner's solution

We define the competitive solution of the game as a Nash equilibrium: strategy profiles $m_i(\theta) : [0,1] \to [0,1]$ for all $i \in [0,1]$, such that player i's strategy is a best response to all other players' strategy. We restrict our attention to looking for a symmetric Nash equilibrium in which all players choose the same $m(\theta)$ function, thus the i subscript is suppressed in what follows. Remember that the payoff of player θ depends on the mass of players with higher and lower θ entering. In case of symmetric strategies, the mass of players with types lower (i.e. better) than θ who choose to enter is

$$b(\theta) = M \cdot \int_0^{\theta} m(\tilde{\theta}) \mathrm{d}\tilde{\theta}, \tag{4}$$

while the mass of types higher (i.e. worse) than θ who choose to enter as

$$a(\theta) = M \cdot \int_{\theta}^{1} m(\tilde{\theta}) \mathrm{d}\tilde{\theta}, \tag{5}$$

and $M \cdot p = b(\theta) + a(\theta)$ is the *aggregate entry* of players.

Players aim to maximize their expected payoff from entering, net of learning costs:

$$V = \int_0^1 m(\theta) \cdot \Delta u(\theta) \; \mathrm{d}\theta - \mu \cdot L(m), \tag{6}$$

[5] An alternative would be to think of capacity as limited and μ being the Lagrange multiplier of the capacity constraint. Instead, our modelling choice captures the idea that in most relevant contexts learning capacity can be expanded, even if for a cost. That is, the player can decide to use more complex code to evaluate data, hire new staff or spend more time with the analysis before entry.

where Δu is the utility gain of entrants defined by (1). We define *welfare*[6] as the total utility gain from entry $W \equiv M \cdot V$, and *aggregate revenue* of the players as their expected payoff before taking into account learning costs $M \cdot R \equiv M \cdot \int_0^1 m(\theta) \cdot \Delta u(\theta)\, \mathrm{d}\theta$. In the competitive solution, each player chooses $m(\theta)$ and takes Δu as given. The social planner can choose the strategy $m(\theta)$ of all players and takes into account that Δu also depends on $m(\theta)$.

2.4 Applications

There are various applications which imply the reduced form payoff (1). The common theme in these applications is that each player's pay-off is lower if better types also enter ($\beta > 0$), while worse entrants can typically increase the payoff ($\alpha > 0$). Applications differ in the interpretation of a better type, and in what the source of the externalities are.

In Online Appendix C, we consider the problem of a financial investor who develops a novel trading strategy. Her problem is that she does not know whether she is among the first investors with this trading idea or the strategy is already "crowded" as described by Stein (2009). In the latter case, not only her realized return is expected to be smaller, but in the case of an aggregate liquidity shock, her losses induced by fire-sales are larger. On the other hand, a larger mass of late players can help her if she is subject to an idiosyncratic liquidity shock. This is so, because they can provide her liquidity, i.e., better terms for exit. We show that in this application, a better type is the player who finds the new opportunity early. The size of β is related to the price impact of early entrants, while α is related to the benefit of liquidity provided by late entrants. $\beta - \alpha$ is higher when fire-sales are more frequent and more severe. In this setting opacity can be interpreted as the amount of information that is available about other traders' trading strategy and the informativeness of price. The regulator

[6] In principle, there might exist applications were V, the per capita utility is a better measure of welfare than total utility W. We have opted to focus on W because of two reasons. First, in the applications we have considered and summarized in section 2.4, W tends to be a more relevant welfare measure. In these applications, Δu tends to measure the productivity gain on a reallocated unit of capital a potential entrant can transfer towards a new opportunity. Potentially, the economy benefits more when more capital is reallocated. This potential benefit is measured by total utility and not per capita utility. Second, our results showing that welfare might be decreasing in M are stronger if one consider the measure W. In fact, whenever W is decreasing in M, V must be also decreasing.

can potentially make markets more transparent by e.g. collecting and disclosing the amount of capital devoted to different trading strategies.

In the working paper version Kondor and Zawadowski (2016), we provide further microfoundations. For instance, our payoff function can be derived from the problem of potential entrants into a market with scarce inputs and local spill-overs. Consider a technology start-up developing a new service with network externalities. A late entrant might struggle to attract the best specialized engineers and other scarce resources paying a premium. On the other hand, more late entrants can increase the value of the product for all firms through the network externality. Large levels of opacity could result from secretive product development to guard intellectual property. A planner could increase transparency by mandating the reporting and publishing of product development plans and the amounts invested in certain activities.

Finally, we also consider a tournament model in academic publications. When a researcher chooses a subfield to work on, she has to make sure that she can write better papers than others. This might not be clear without significant investment in understanding the connected literature and methodology. At the same time, she would like to enter to a field where many others enter, otherwise her impact might remain very low, even if her quality of work is high. In this setting opacity measures how hard it is to figure out what other researchers are working on or planning to work on. A planner could increase transparency by publicizing research plans and encouraging dissemination and feedback on early stage research.

3 Model Solution

In this section we analyze the model. We start by the characterization of the no information ($\mu = \infty$) and full information ($\mu = 0$) benchmark. We then formulate and solve the players' and the planner's problem for general levels of opacity μ. Finally, we derive the welfare effects of more competition (higher M) and less transparency (higher μ). This section contains the main insights of the model.

We argue that opaque markets tend to be more crowded, opaqueness is likely to be beneficial if competition is not too fierce, and that fiercer competition leads to more wasteful learning leading to deteriorating welfare without affecting crowding. We also explain the main mechanisms behind these insights.

In this section, we allow for any β and α in (1) satisfying our parameter restrictions but restrict κ to 0. We analyze the $\kappa > 0$ case in an extension in Section 4.1 to show that results are qualitatively similar in that case.

3.1 Full and no information benchmark

To better understand the optimal strategies and aggregate entry, we first look at the extreme cases of full information and no information. These extreme cases highlight that the nature of externality changes depending on the amount of information. While without information both α and β represent externality, with full information only β does. We show that this implies under-entry with full information whenever $\alpha > 0$, and over-entry with no information regardless of the sign of α, both compared to the planner's solution.

The next Lemma characterizes optimal strategies and entry in the competitive solution and under the planner in the full information benchmark, that is, when the marginal cost of reducing entropy, μ, is zero. We also refer to this case as full transparency.

Lemma 2. ***Full information benchmark.*** *Under full information ($\mu = 0$), there is too little entry in the competitive equilibrium if $\alpha > 0$, and excessive entry if $\alpha < 0$. In the symmetric competitive equilibrium players' equilibrium strategy is a unit step function $m(\theta)|_{\mu=0} = \mathbb{1}\left\{\theta \leq \frac{1}{M\beta}\right\}$, resulting in aggregate entry*

$$M \cdot p|_{\mu=0} = \min\left(\frac{1}{\beta}, M\right). \tag{7}$$

In the social planner's optimum, one of the many symmetric optimal strategies is a unit step function $m_s(\theta)|_{\mu=0} = \mathbb{1}\left\{\theta \leq \frac{1}{M\cdot(\beta-\alpha)}\right\}$*, and all of the planner's optimal strategies imply aggregate entry*

$$M \cdot p_s = \min\left(\frac{1}{\beta-\alpha}, M\right). \tag{8}$$

Under full information, in the competitive equilibrium, each player enters if and only if her pay-off is non-negative. The worst type entrant (with $\theta = \frac{1}{M\beta}$) earns zero revenue. This results in strategies $m(\theta)$ that are unit step functions. Comparing (7) and (8) shows that in case of $\alpha > 0$ the competitive and social solutions lead to the same entry if and only if $M < \frac{1}{\beta}$. In that case, the mass of players are so small that under both solutions all players enter. Apart from this trivial case, whether the competitive solution implies under- or over-entry – compared to the social planner's choice – depends on the sign of α. There is excessive entry in the competitive equilibrium if $\alpha > 0$, since players with higher θ do not take into account the positive effect of their entry that accrues to entrants with lower θ. This highlights the fact that under full information, externalities are fully captured by α.

We now turn to the of no information case, which we also refer to as full opacity. In this case each player enters with a constant probability irrespective of its type θ since learning about its type is prohibitively expensive. This results in strategies $m(\theta)$ that are flat functions.

Lemma 3. ***No information benchmark.*** *In the absence of information ($\mu \to \infty$), there is always excessive entry in the competitive equilibrium. The competitive entry function is constant at* $m(\theta)|_{\mu\to\infty} = \min\left(\frac{2}{M(\beta-\alpha)}, 1\right)$*, implying aggregate entry of*

$$M \cdot p|_{\mu\to\infty} = \min\left(\frac{2}{\beta-\alpha}, M\right) \tag{9}$$

The social planner's entry functions is also constant at $m_s(\theta)|_{\mu\to\infty} = \min\left(\frac{1}{M(\beta-\alpha)}, 1\right)$*, implying aggregate entry of (8).*

Similarly, to the full information benchmark, the competitive and social solutions lead to the same entry when the mass of total entrants is small, $M < \frac{2}{\beta-\alpha}$, because all enter with probability 1. This threshold is determined by a zero utility condition for the average entrant. This leads to a larger threshold than in the case of full information where the threshold was determined by a zero utility condition for the worst type (highest θ) entrant. For interior solutions in the competitive equilibrium, players are indifferent between entering and not. The planner restricts entry in order to increase the payoff to players. In fact the interior solution implies excessive entry under any other parameter values: twice as many players enter in the competitive equilibrium than under the planner's choice. The intuition is analogous to the "tragedy of commons": Each player fails to internalize that her own entry reduces the expected benefit of entry for all other entrants. This follows from the assumption of $\beta > |\alpha|$, thus entering harms other entrants on average and the market is prone to "crowding". Since players enter irrespective of their type, what matters is the net average externality on others, which is captured by $\beta - \alpha$.

Why does only α corresponds to an externality under full information, while both α and β do under no information? After all, expression (1) suggest that both low type entrants affect the payoff of better type entrants and vice-versa regardless of the amount of information. In the full information benchmark the worst type entrant is the marginal type. This entrant finds that if no one enters with a worse type, and everyone enters with a better type, her utility is exactly 0. When $\alpha > 0$, the planner would like to force a slightly worse type to enter also, as it would increase each better type's pay-off through the term $a(\theta)$, compensating for this worst type's negative private pay-off. The planner could also take into account the negative effect of this additional entrant on all the worse type entrants through the term $b(\theta)$, but those do not enter anyway. This is why β does not correspond to an externality under full information.

In contrast, under the no information benchmark, each type enters with the same positive probability. That is, the median entrant is the marginal type. Her utility is exactly zero in equilibrium. The median entrant effects the pay-offs of both the better type entrants through the term $a(\theta)$ and the worse type entrants through the term $b(\theta)$. These effects are ignored in the competitive solution, but not under the planner. This is why both α and β corresponds to externalities under no information.

3.2 Optimal strategies in the general problem

In this part, we derive and describe the competitive equilibrium.

The following Lemma restates Proposition 1 of Yang (2015) for our model, which delivers our first order conditions.

Lemma 4. ***Competitive best responses.*** *The unique best response of a player to all other players playing $\tilde{m}(\theta)$ is:*

i) $m(\theta) = 1$ *if and only if* $\int_0^1 e^{-\frac{\Delta u(\theta)}{\mu}} \mathrm{d}\theta \leq 1$*;*

ii) $m(\theta) = 0$ *if and only if* $\int_0^1 e^{\frac{\Delta u(\theta)}{\mu}} \mathrm{d}\theta \leq 1$*;*

iii) $m(\theta) \in (0,1)$ *if and only if* $\int_0^1 e^{\frac{\Delta u(\theta)}{\mu}} \mathrm{d}\theta > 1$ *and* $\int_0^1 e^{-\frac{\Delta u(\theta)}{\mu}} \mathrm{d}\theta > 1$*, the optimal* $m(\theta)$ *is pinned down by the first-order condition:*

$$1 - M \cdot \beta \cdot \int_0^\theta \tilde{m}(\tilde{\theta}) d\tilde{\theta} + M \cdot \alpha \cdot \int_\theta^1 \tilde{m}(\tilde{\theta}) d\tilde{\theta} = \mu \cdot \left[\log\left(\frac{m(\theta)}{1 - m(\theta)}\right) - \log\left(\frac{p}{1-p}\right) \right]. \tag{10}$$

To solve for the equilibrium $m(\theta)$, we differentiate the first-order condition (FOC) (10) with respect to θ. That gives an ordinary differential equation for $m(\theta)$ where the original first order condition at $\theta = 0$ is the boundary condition. This ordinary differential equation can be solved up to the boundary value $m(0)$.

Proposition 1. ***Competitive equilibrium.*** *There is a unique threshold $\bar{M} > 0$ pinned down by*

$$\frac{\bar{M} \cdot (\alpha + \beta)}{\mu} = e^{-\frac{1 - \beta \cdot \bar{M}}{\mu}} - e^{-\frac{1 + \alpha \cdot \bar{M}}{\mu}}, \tag{11}$$

such that

a) if competition is weak, i.e. $M \leq \bar{M}$*, all players enter without learning* $(m(\theta) = 1)$ *and this is the unique symmetric competitive equilibrium,*

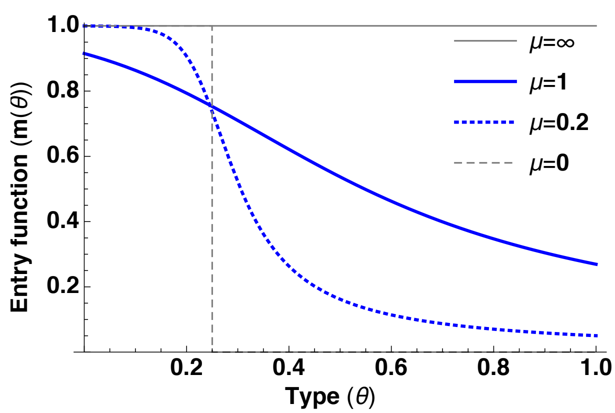
b) if competition is large enough, i.e. $M > \bar{M}$*, there is a competitive symmetric equilibrium in which the optimal entry function is given by:*

$$m(\theta) = \frac{1}{1 + W_0\left(e^{M \cdot \frac{\alpha+\beta}{\mu} \cdot \theta + \frac{1-m(0)}{m(0)} + \log\left(\frac{1-m(0)}{m(0)}\right)}\right)}, \tag{12}$$

where W_0 *denotes the upper branch of the Lambert function*[7] *and* $m(0)$ *is pinned down by the boundary condition:*

$$M \cdot \alpha \cdot p + 1 = \mu \cdot \left[\log\left(\frac{m(0)}{1 - m(0)}\right) - \log\left(\frac{p}{1-p}\right)\right]. \tag{13}$$

Figure 1: **Competitive entry functions for different levels of opacity**



The figure shows the probability of entry depending on the player's type (entry function $m(\theta)$) in competitive equilibrium for different levels of opacity μ and fixed $\beta = 4$, $\alpha = 2$, $M = 1$, $\kappa = 0$. The optimal entry function is a step function for full transparency $\mu = 0$ (thin dashed line), and is flattening as opacity μ increases: $\mu = 0.2$ (thick dashed line) and $\mu = 1$ (thick line). Eventually for full opacity $\mu = \infty$ it is flat (thin line).

The entry strategy $m(\theta)$ for different levels of opacity μ is shown in Figure 1. For $\mu = 0$ we have a step function as described in Lemma 2. As opacity μ increases, the step is first smoothed at moderate levels of opacity ($\mu = 0.2$) reflecting that perfect learning of one's type is too costly and

[7]The upper branch of Lambert function is defined by the following implicit equation: $z = W_0(z) \cdot e^{W_0(z)}$ if $z > 0$.

not thus optimal. $m(\theta)$ then flattens out for higher levels ($\mu = 1$) as learning becomes more costly and eventually becomes completely flat as with full opacity ($\mu = \infty$) there is no learning as shown in Lemma 3.

Note that for any level of opacity μ, the planner's problem is the same as in the benchmark described in Lemma 3 and there is no learning in the planner's solution. Players want to differentiate between states, but the planner does not. Every player wants to know whether she is better than the other players even if this is wasteful from the social planner's point of view. We label this as the "rat-race" effect and it is driven by the assumption $\beta > |\alpha|$.

3.3 The Welfare Effect of Competition and Opacity

This part contains our main results on the welfare effects of competition and opacity. First, we describe the two benchmark cases of full opacity and full transparency as a function of the degree of competition M. Then, we proceed to the welfare effects of competition. Finally, we analyze the welfare effects of opacity.

The welfare in the benchmark cases is presented in the following Lemma and illustrated in the bottom right panel of Figure 2.

Lemma 5. ***Welfare in the no information and full information benchmarks.***

1. *In all cases, if competition M is low enough that all players enter, welfare is given by:*

$$W_{full\ entry} = M \cdot \left(1 - M \cdot \frac{\beta - \alpha}{2}\right). \tag{14}$$

2. *In the competitive equilibrium, if $M \geq \frac{1}{\beta}$, with full information: $W|_{\mu=0} = \frac{\alpha+\beta}{2 \cdot \beta^2}$.*

3. *In the competitive equilibrium, if $M \geq \frac{2}{\beta - \alpha}$, with no information: $W|_{\mu \to \infty} = 0$.*

4. *In the planner's solution, if $M \geq \frac{1}{\beta - \alpha}$, both under full and no information: $W^S = \frac{1}{2 \cdot (\beta - \alpha)}$.*

Note first that welfare under the planner's solution is the same under both benchmarks. This is so, because $\kappa = 0$, hence the planner cares only about aggregate entry as opposed to entry by each type. Therefore, the planner can implement the full information entry in the no information benchmark by imposing identical probability of entry for all types. Welfare is increasing in competition M as long as competition is not excessive, i.e. $M < \frac{1}{\beta-\alpha}$ because below this level there are not enough players to implement the socially efficient level of entry. Welfare is constant above that level as more players do not influence the level of aggregate entry.

Turning to the competitive solution, under no information all players enter with probability 1 as long as the average player's utility is positive. This leads to a hump shape pattern in M as shown in Figure 2. Welfare is increasing as long as the mass of players is smaller than what is required for the efficient level, but decreases afterwards. Once welfare is zero, any additional increase in M proportionally reduces the probability of entry keeping aggregate entry and welfare constant.

Under full information, only those players enter whose payoff is positive, keeping the average payoff strictly positive. Whether there is over or under entry in the interior equilibrium, depends on the sign of α as we discussed above.

Finally, given Lemma 5, it is easy to see that for some levels of competition M, welfare is higher in the no information benchmark than in the full information benchmark whenever $\alpha > 0$. The intuition is that under full information entry is inefficiently low. No information leads to higher entry which can be closer to the efficient level depending on M leading to higher welfare. We summarize this observation in the following Corollary and return to this intuition in Subsection 3.3.2 when discussing the effect of opacity on welfare.

Corollary 1. *When the mass of players are in an intermediate range, $M \in \left(\frac{1}{\beta}, \frac{2}{\beta-\alpha}\right)$, welfare is higher in the no information benchmark than in the full information benchmark.*

3.3.1 Competition

In this part, we analyze the effect of competition M on welfare for general levels of opacity μ. As a preliminary step we show in Propositions 2 and 3 that the aggregate entry of players is insensitive to the degree of competition once M is above the threshold $\bar{M}$ defined in (11). However, $\bar{M}$ is in general different from the socially optimal level of aggregate entry defined in (8).

Proposition 2. ***Crowding in competitive equilibrium.*** *The level of aggregate entry in the competitive equilibrium is* $M \cdot p = \min\left(M, \bar{M}\right)$.

For the intuition, recall that changing M changes the optimal strategy $m(\theta)$ for every player 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 the average probability of entry p. In contrast, due to the crowding effect the average entry p decreases in the mass of potential competitors M. In equilibrium, the decrease in p is exactly proportional to the increase in M, keeping $M \cdot p$ constant.[8] While aggregate entry is insensitive to the level of competition, it does depend on all the other parameters of the model. In the next Proposition we characterize how the level of under- or over-entry is determined by the relative strength of the externalities implied by the parameters α and β for fixed level of opacity μ.

Proposition 3. ***Comparative statics of crowding.*** *If there is sufficient competition,* $M > \max[\bar{M}, \frac{1}{\beta-\alpha}]$, *the amount of aggregate entry in the planner's solution is* $\bar{M}_s = \frac{1}{\beta-\alpha}$. *Crowding is defined by* $\frac{\bar{M}}{\bar{M}_s}$, *the relative amount of competitive aggregate entry to planner's optimum, and is*

1. *decreasing in* α, *and*

2. *increasing in* β *if* $\alpha > 0$, *decreasing in* β *if* $\alpha < 0$ *and does not depend on* β *if* $\alpha = 0$.

[8]Allowing players to flexibly choose their information structure is crucial in generating the result of constant entry as the mass of players increases. With flexible learning, the players can optimally devise their information to exactly counter the increase in the mass of players and thus enter at a constant aggregate rate. When learning is constrained, this is not necessarily the case: we demonstrate this in Online Appendix D in which players can only buy Gaussian signals about their type, subject to the same entropy cost as before.

To better understand the effects of β and α for fixed μ, consider their effect on the difference of social and private incentives in the benchmarks presented in Lemmas 2 and 3. In the full information benchmark, crowding is $\frac{\bar{M}}{M_s}|_{\mu\to 0} = \frac{\beta-\alpha}{\beta}$. As α measures the social benefit of low-type entrants which players fails to internalize, crowding is decreasing in α. In contrast, for fixed α, larger β alleviates the effect of this externality. If players suffer more by better entrants (larger β), aggregate entry decreases both under the planner's and the competitive solutions. However, when $\alpha > 0$, this decrease in entry is relatively smaller in the competitive solution, because $\alpha > 0$ implies under-entry. In contrast, in the no learning benchmark, crowding is $\frac{\bar{M}}{M_s}|_{\mu\to\infty} = 2$, a ratio unaffected by the parameters. Therefore, the effects driving the full information benchmark carry over for intermediate values of μ.

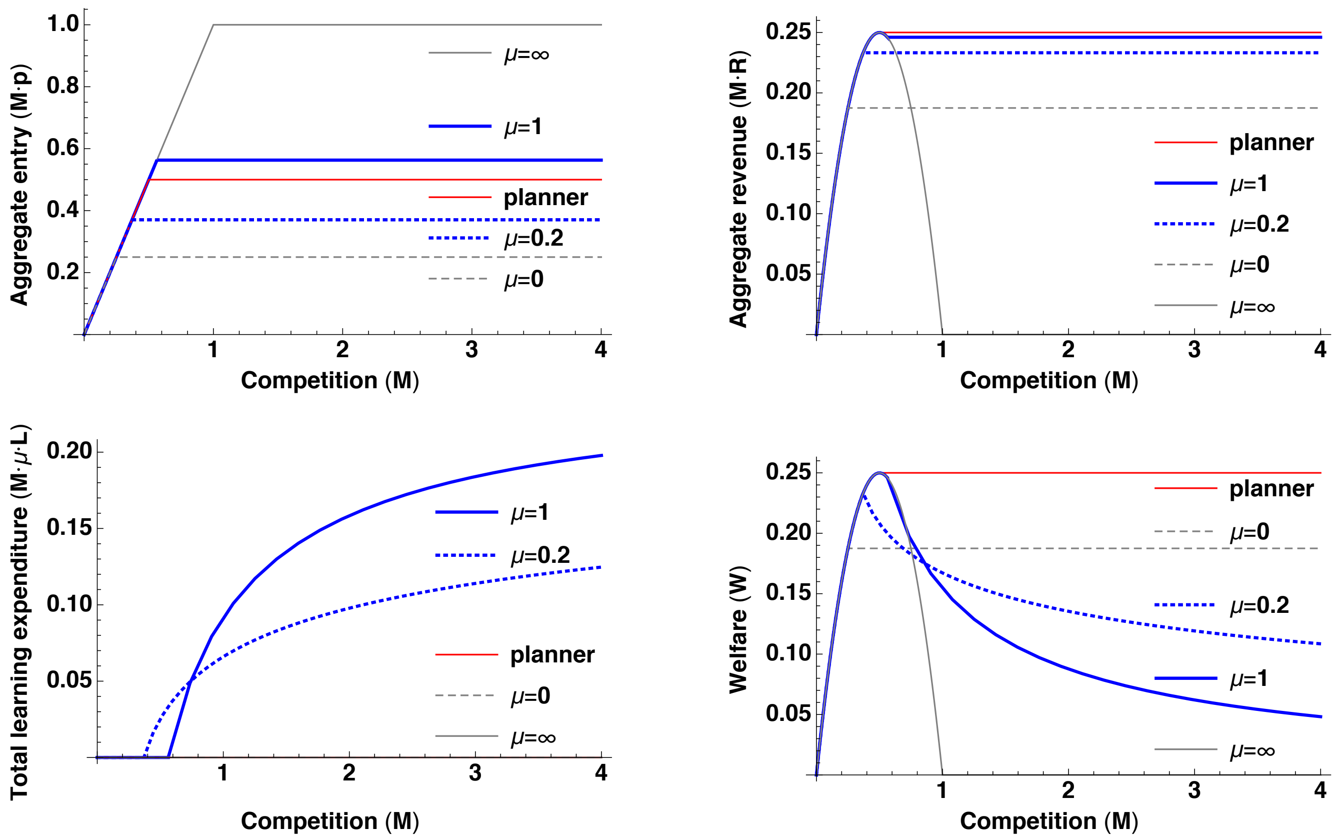
Next, we show that while high levels of competition do not effect aggregate entry and thus the level of crowding, it does have a detrimental effect on welfare because it enhances the rat-race between players. Figure 2 shows the level of welfare, aggregate entry and revenue and total learning expenditure as function of the level of competition M for several different levels of opacity μ. In line with Proposition 2, aggregate entry and thus aggregate revenue does not change with competition if competition is fierce enough ($M > \bar{M}$). On the other hand, while the presence of some players ($M < \bar{M}$) unambiguously increases welfare in the competitive equilibrium, with high levels of competition, welfare is decreasing and tends to zero. We prove this formally in the following proposition.

Proposition 4. ***Welfare and competition.*** *If $M > \bar{M}$, aggregate entry and the aggregate revenue of players stays constant as we increase M. However, welfare becomes decoupled from revenue, and converges to zero from above as $M \to \infty$:*

$$W(\bar{M}) > \lim_{M\to\infty} W(M) = 0. \tag{15}$$

The reason for welfare tending to zero in this case is due to learning expenditure, see Figure 2. As the mass of players in the market grows, they start worrying about crowding, and thus their relative type θ, inducing them to learn about it. A rat race ensues with increasing amounts invested

Figure 2: **Real outcomes as a function of competition**



The four panels show the effect of competition M on aggregate entry, aggregate revenue, total amount spent on learning by all players and welfare of all players. For all panels the thin solid line is the social optimum. The light thin solid line shows the competitive outcome with full opacity ($\mu = \infty$) and the light thin dashed line that for full transparency ($\mu = 0$). The thick (blue) lines are the competitive outcomes for intermediate values of opacity μ: the dashed line is that for $\mu = 0.2$, while the solid line for $\mu = 1$. The planner's solution is a (red) line. The other parameters are $\beta = 4$, $\alpha = 2$, $\kappa = 0$. While aggregate entry and revenue are constant above a given level of competition, welfare in the competitive equilibrium always tends to zero with increasing competition M for any level of positive opacity $\mu > 0$. Note that total learning expenditure is zero in the planner's solution and in the benchmark competitive equilibria for $\mu = 0$ and $\mu = \infty$.

in learning and reduced welfare. Thus increasing competition among players leads to a drop in welfare not because of overcrowding as was the case for full opacity (Lemma 5) but because of increased spending on learning. In fact, competition does not affect crowding at all, it only improves the type of entrants (see Appendix B).

Note, unlike in the no learning benchmark in Lemma 5, welfare is strictly positive under competitive solution for any parameter values. This shows that learning is useful under the competitive solution. In fact, learning is a substitute for coordination. When learning is prohibitively expensive, each player enters with a probability that is too large. When players learn, they can partially condition their entry on their type allowing for less aggregate entry. Indeed, as we later show in Proposition 5, crowding decreases when markets are less opaque. The resulting coordination is inefficient, because social and private incentives for entry do not coincide. Still, it is always better than no learning at all. In contrast, the planner does not have to learn to coordinate as it can choose all players' actions.

3.3.2 Opacity

We now turn to the welfare effect of opacity. Since the a policymaker might be able to make the market more opaque or transparent, as explained in Section 2.4, the results help us understand what kind of markets benefit from more transparency. We show that only markets with fierce competition unambiguously benefit from more transparency. Less fiercely contested markets might even benefit from more opacity, partially because it makes players spend less on learning.

Welfare depends on aggregate revenue which itself is a function of aggregate entry, and thus crowding, in the market. As a first step, in Proposition 5 we show that increasing opacity μ unambiguously increases crowding in the market.

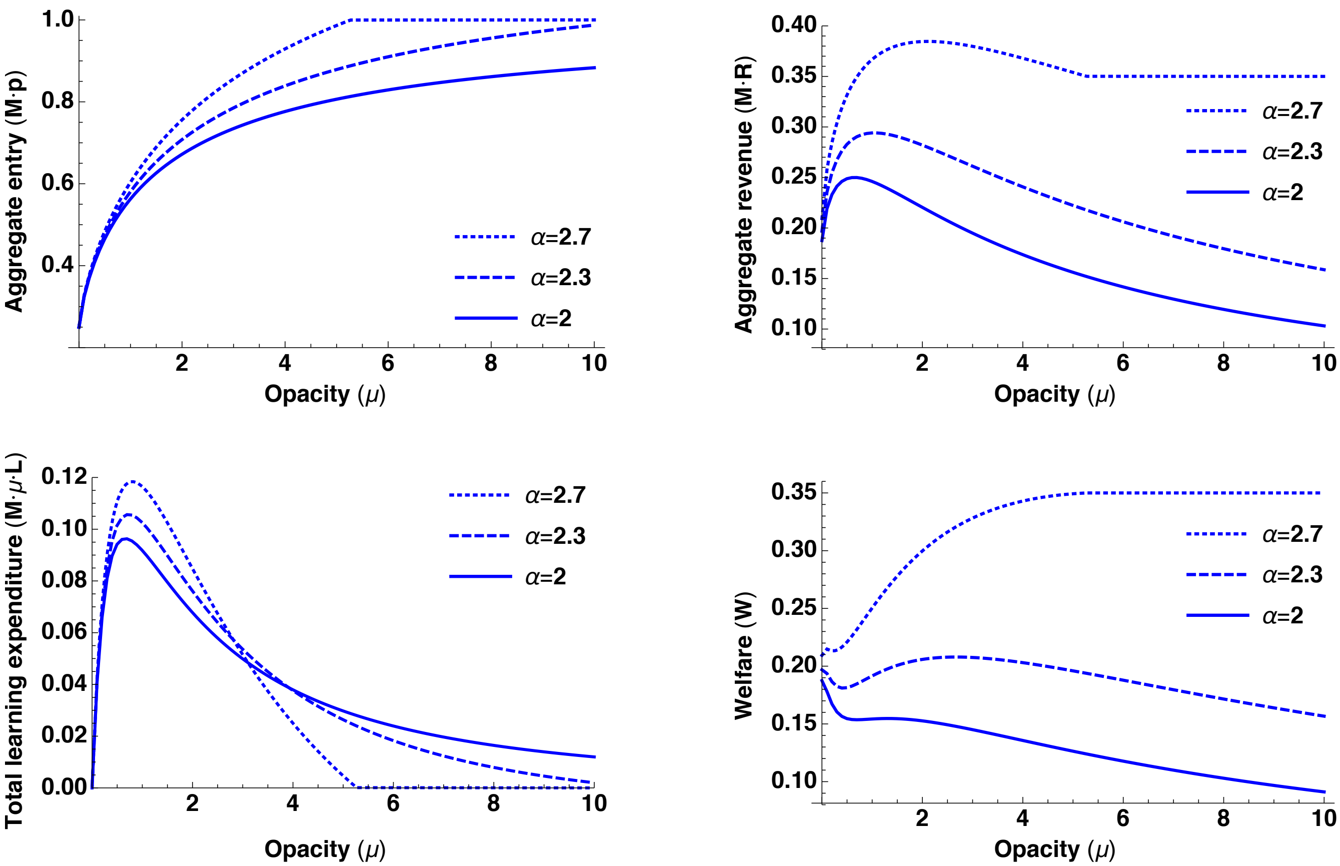
Proposition 5. ***Comparative statics of crowding.*** *If there is a sufficient mass $M > \max[\bar{M}, \bar{M}_s]$ of players, the relative amount of competitive aggregate entry to social aggregate entry $\frac{\bar{M}}{\bar{M}_s}$, i.e. crowding, is increasing in opacity μ.*

More costly information leads to more crowding, because the game is closer to a tragedy of commons problem as explained in Section 3.2. Figure 3 summarizes how aggregate entry, revenue, learning expenditure and welfare change as function of opacity μ. One might expect that in a market prone to crowding, making the market transparent (i.e. decreasing μ) increases welfare. However, as Figure 3 shows, welfare is not necessarily decreasing in opacity μ. In Proposition 6, we formally show that there are levels of competition M when higher opacity increases welfare.

Proposition 6. Optimal opacity.

1. *If $M \in \left(\frac{1}{\beta}, \frac{\alpha+\beta}{(\beta-\alpha)\cdot\beta}\right)$ then maximal welfare in the competitive equilibrium is attained at $\mu > 0$.*

Figure 3: **Real outcomes as a function of market opacity**

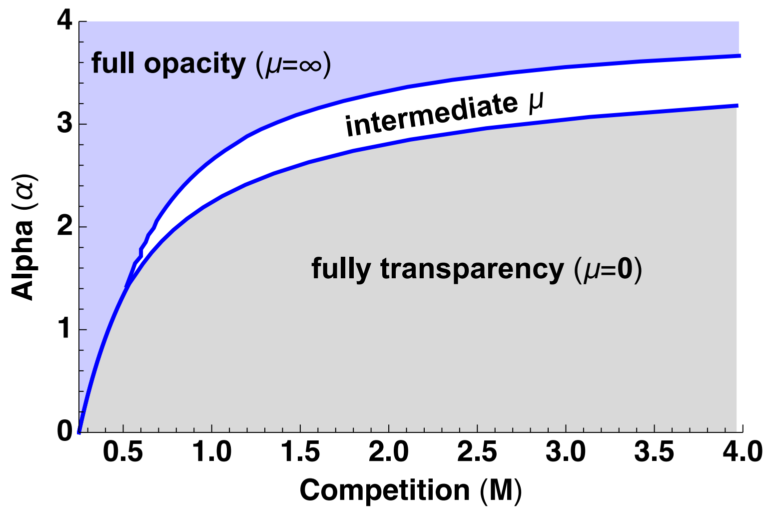


The four panels show the effect of opacity μ on aggregate entry, aggregate revenue, total amount spent on learning by all players and welfare of all players. We show the competitive equilibrium outcome for three levels of α. The solid line is that for $\alpha = 2$ and in this case full transparency ($\mu = 0$) is optimal. The dashed line is for $\alpha = 2.3$ in which case an interior level of opacity μ is optimal. For the dotted line $\alpha = 2.7$, opacity above a threshold or full opacity is optimal. The other parameters are $\beta = 4$, $M = 1$, $\kappa = 0$.

2. *If α is close enough to β, there exists an interval of μ in which welfare is increasing in opacity (i.e. $\frac{\partial W}{\partial \mu} > 0$) for any $M > \bar{M}$.*

The first result that full transparency is not necessarily optimal is driven by competitive under-entry for $\alpha > 0$. As discussed in Section 3.1, full information is not necessarily optimal in this case because there is insufficient entry. As opacity increases, entry gets closer to the efficient level, which increases welfare. While this above result is already apparent from Lemma 5, adding endogenous learning allows us to analyze the optimal level of opacity. Numerical results show that in many cases the optimal opacity of the market is inbetween full transparency and full opacity. We show the optimal level of opacity as a function of α and M in Figure 4.

Figure 4: **Optimal Opacity**



The figure shows the optimal level of opacity as a function of α and M with $\beta = 4$ and $\kappa = 0$. In the upper left area, full opacity ($\mu = \infty$) leads to the highest possible welfare, keeping all other parameters fixed. In the lower right area, full transparency ($\mu = 0$) is optimal. Between the two regions, in the light area, there is an interior level of opacity μ that leads to the highest welfare. From the figure it is clear that if the only difference between two markets is the level of competition, then less competitive markets benefit from opacity, while more competitive markets from transparency.

The intuition for the second part of Proposition 6 is much more subtle and is driven by the welfare loss from endogenous learning. Increasing opacity has two additional effects beyond increasing crowding: First, it increases the expenditure on learning for a fixed amount learning. Second, by increasing the marginal cost of learning it discourages players from learning. This second effect dominates for large μ. This is illustrated by Figure 3: total learning expenditure is hump-shaped in μ. Thus for intermediate values of μ, increasing opacity has a stronger beneficial effect on decreasing learning expenditure than the detrimental effect it has on increasing crowding. Strictly speaking, the Proposition only shows welfare is locally increasing in opacity but numerical simulations in Figure 3 show that full transparency is not a global optimum in this case either. E.g. for $\alpha = 2.3$ in the figure, an interior level of μ is optimal even though there is already excessive crowding as is apparent from the decreasing aggregate revenue of players.

Overall, the increasing opacity entails a trade-off between potentially decreasing total learning expenditure and making crowding worse. As shown in Figure 4, intermediate opacity is optimal if α is high and the competition in the market M is fierce because in this case learning incentives are very high and curtailing them has a benefit. Full opacity is not optimal, since it leads to huge crowding and zero welfare. On the other hand, full opacity is likely to be optimal at low levels of competition, since in these cases insufficient entry is the main issue and not allowing players to learn about their type increases welfare by moving aggregate entry closer to its socially optimal level. Overall, if the policymaker can influence the degree of opacity a competitive market (change μ), he should make markets with lots of competition more transparent, while markets with less competition and high α might benefit from less transparency.

4 Extensions

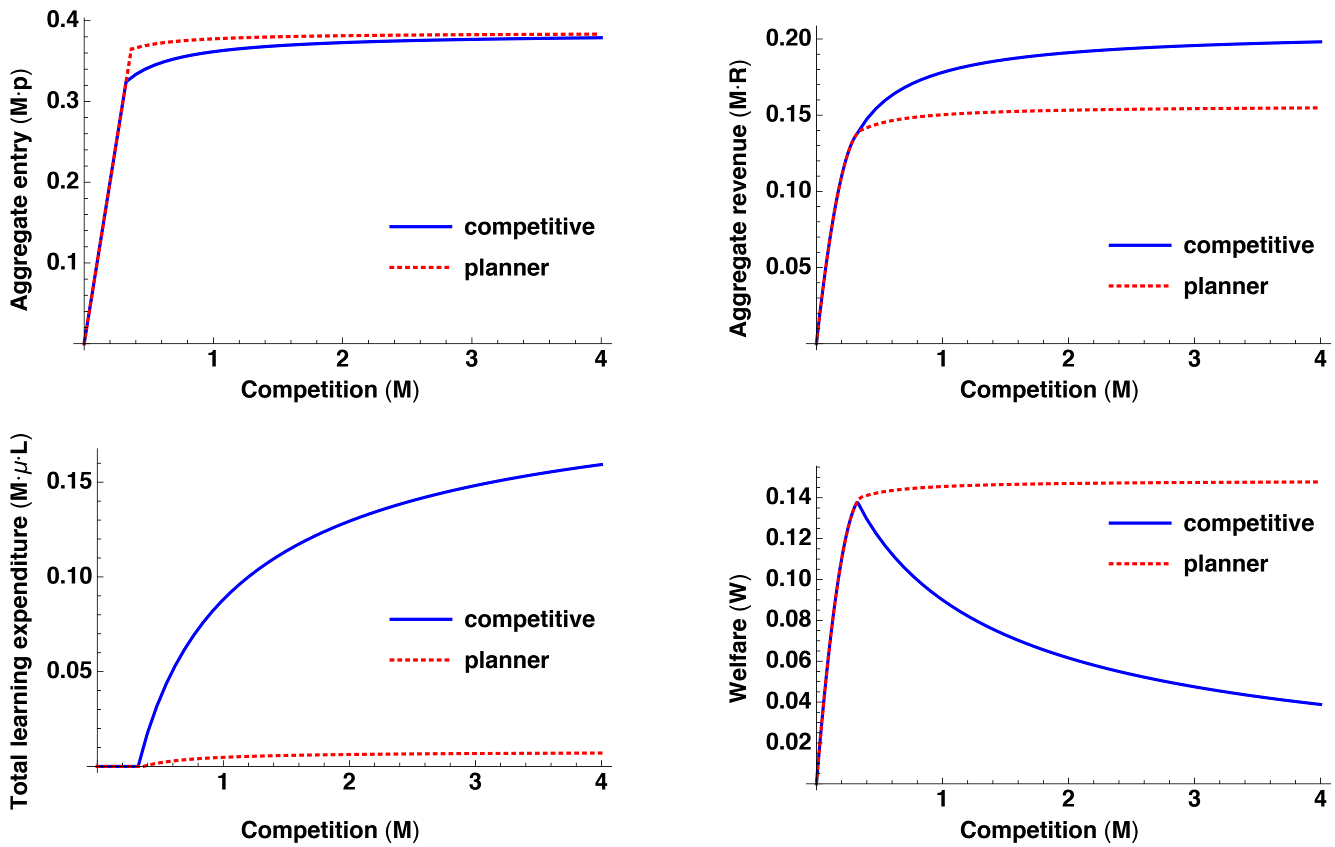
In our first extension we allow for $\kappa > 0$, i.e. the payoff to players with better types is higher irrespective of what other players do. We show that our previous results, attained in the more tractable baseline setting with $\kappa = 0$, qualitative hold in this more general setting as well. Thus the assumption that better entrants do not get a higher payoff, i.e. there is no social value to better types entering, is not driving our results.

In our second extension, we slightly reinterpret our measure of opacity μ: one can think of it not only as a measure of the market but also as an attribute of players. Some players might be "smarter" and thus have a lower marginal cost of learning than others. We show that increasing the share of smart players is similar to making the market more transparent, thus might not be optimal.

4.1 Socially more efficient types

Until now we analyzed the case where payoff of a player θ depends only on her rank among those who entered. As a result, the type-dependent part of utility was simply redistributive. That is, the planner was not interested in which type enters, only in aggregate entry. In this part, we consider the case of $\kappa > 0$, where better types are more efficient in both a social and a competitive sense. The model can only be solved numerically and the real outcomes are shown in Figure 5 as a function of competition. The formal calculation of the equilibrium can be found in Online Appendix E.

Figure 5: **Competitive and social optimum with socially more efficient types**



The four panels show aggregate entry, aggregate revenue, total amount spent on learning by all players and welfare of all players as a function of competition M. The solid line shows the competitive outcome, while the dashed line the planner's optimum. The parameters are: $\beta = 4$, $\alpha = 2$, $\mu = 0.5$, $\kappa = 0.5$. The main difference compared to Figure 2 using $\kappa = 0$ is that while aggregate entry is close to flat above a threshold M, it is not completely flat. Also, there is positive spending on learning even in the planner's equilibrium, though much less than in the competitive equilibrium.

In this case, the social planner also wants to differentiate between states, the socially optimal entry function $m_s(\theta)$ is no longer constant, it is also downward sloping. However, the incentive for competitive learning is even higher since private incentives include the rat race effect. Every player wants to know whether it is ahead of the others.

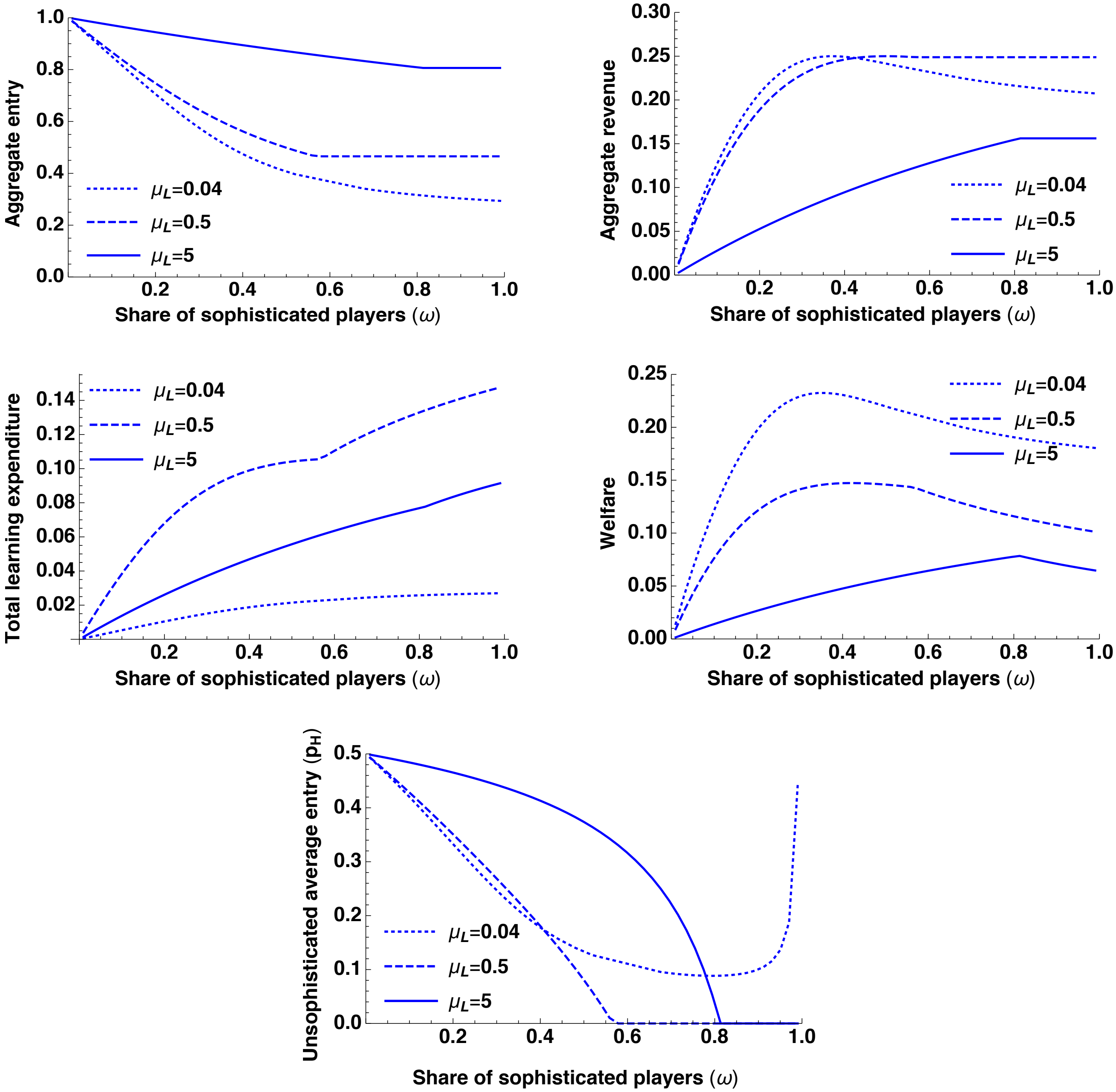
We can understand the effect of $\kappa > 0$ by comparing Figures 2 and 5. As competition increases, aggregate entry is slightly increasing in both the competitive and the social optimum because better type entry has increasing benefit as the mass of players increases and there are more potential good entrants with low θ. In both cases, aggregate revenue is increasing faster in competition M than aggregate entry because most of the additional efficiency comes from the better entrants, not simply more entry. The main insight of the simpler model still holds, players are also motivated by the rat race so there is excessive learning, though the welfare loss is partially offset by the welfare gain from the shifting type distribution of entry towards better players. Nevertheless, the revenue gains from better entrants, which is a side-effect of over-learning, cannot offset the loss from excessive learning, thus welfare still converges to zero.

4.2 Heterogenous players

In this section we consider an extension with heterogenous players. E.g. in a financial market, one can think of high-frequency traders and certain hedge funds as more sophisticated than pension funds. Instead of changing the opacity μ of the market, we analyze how changing the composition of players influences welfare. We show that increasing the share of sophisticated traders does not necessarily increase welfare because of the ensuing rat-race, similarly to our previous analysis. This allows us to draw the conclusion that e.g. the increasing presence of high-frequency traders in markets might be welfare-destroying.

We modify the model from Section 2 by considering two groups of players: $\omega \cdot M$ mass of players is sophisticated and faces a lower learning cost of μ_L, while $(1-\omega) \cdot M$ mass of players is unsophisticated and faces a higher learning cost of $\mu_H > \mu_L$. Both groups of players have types θ that are uniformly distributed over $[0, 1]$. We consider the symmetric equilibrium in which sophisticated players choose the same entry strategy of $m_L(\theta)$, while unsophisticated players choose the same $m_H(\theta)$. To simplify the problem, we assume that the unsophisticated cannot learn at all, i.e. $\mu_H \to \infty$, resulting in a constant m_H in θ as described in Lemma 3. Otherwise the solution would be a set of two joint differential equations which cannot be easily solved. We relegate details on how to compute the equilibrium in this case to Online Appendix E.

Figure 6: **Real outcomes with varying composition of players**



In this Figure we vary the portion ω of sophisticated players who can learn with low cost μ_L, while $1-\omega$ cannot learn at all, while the mass of players M is kept constant. The first four panels show the same outcomes as in our original model, see Figure 2. The fifth figure shows the share of unsophisticated players who enter. Parameters: $\beta = 4$, $\alpha = 2$, while μ takes three different values: $\mu_L = 0.04$ (dotted line), $\mu_L = 0.5$ (dashed line), $\mu_L = 5$ (solid line). In all cases the social planner would allow each player to enter with probability $\frac{1}{M \cdot (\beta - \alpha)} = \frac{1}{4}$, yielding aggregate entry of $\frac{1}{\beta-\alpha} = \frac{1}{2}$.

We solve the above set of equations numerically since it is analytically intractable. In Figure 6, we vary the portion ω of sophisticated players who can learn with cost μ_L. Thus, the overall mass of players M is kept constant but a growing fraction of players is sophisticated. At $\omega = 0$ only unsophisticated are present and thus they enter until revenue is zero (given that M is large enough), yielding zero welfare. As ω initially increases, welfare increases since the average player is more sophisticated and overcrowding is alleviated. There are two effects leading to decreasing welfare as ω increases further. First, if the sophisticated are very sophisticated (low μ_L) then having lots of sophisticated leads to under-entry for $\alpha > 0$, thus decreasing welfare. Second, and more interestingly, welfare can be decreasing in the share of sophisticated players ω even for high μ_L in the absence of under-entry. Consider for example the case of $\mu_L = 5$ in Figure 6. Aggregate entry is well above its socially optimal level of 0.5, still if ω is high, increasing ω is welfare destroying. The reason is that as ω increases, the unsophisticated are less likely to enter and above about $\omega = 0.82$ they are completely driven out of the market. Once the unsophisticated are not present, welfare is decreasing in ω as the sophisticated engage in a rat-race of learning, as illustrated by the fast increasing learning expenditure.

Figure 6 also highlights the intricate interplay between the entry and learning strategies of the sophisticated and the unsophisticated. First, the remaining unsophisticated are less likely to enter as the fraction ω of sophisticated increases because there is more and more aggregate entry at low θ, cream-skimming the market and leaving less revenues for unsophisticated players who enter indiscriminately. Second, unsophisticated players are completely driven out of the market for high ω when sophisticated players are also not perfectly sophisticated (if $\mu_L > 0$), thus they are competitors of the unsophisticated, cannibalizing their revenues and eventually driving them out. The intuition is similar to that in high-frequency trading where some players may stay out of the market because they are afraid of very fast players front-running them.

Interestingly, if the sophisticated players are sophisticated enough (μ_L close to zero), unsophisticated players will never be completely driven out of the market, see Figure 6. The reason is that perfectly sophisticated players follow cutoff strategies with the last entrant at the cutoff getting zero payoff and being indifferent. If only sophisticated players are present in the market, then an unsophisticated player with a uniform prior about its θ knows it can get positive payoff if its θ is smaller than the cutoff of the perfectly sophisticated players and gets zero payoff (equal to that of the last perfectly sophisticated to enter) with θ higher than the cutoff since there are no other entrants with higher θ in equilibrium. Figure 6 shows that in the case of $\mu_L = 0.04$ if there are only few unsophisticated players (ω close to one), all unsophisticated enter.

The above analysis also highlights how a not very well informed (unsophisticated) player should behave if she learns about an investment opportunity. She should enter with relatively high probability if she thinks players in the market are predominantly sophisticated but only if she believes that the sophisticated players are very sophisticated. On the other hand, she should not enter at all, if she thinks the other sophisticated players are not sufficiently sophisticated. She may also choose to enter if she thinks that players are predominantly unsophisticated.

5 Conclusions

We analyze a novel entry-game with endogenous information acquisition to study the welfare effects of opacity and competition. Since players in an opaque market are uncertain about their competitive advantage relative to others, they construct optimal signals to learn about it subject to an entropy cost. The opacity of the market is captured by the marginal cost of learning, while the extent of competition is modelled by the mass of players standing on the sidelines. In general, the individually optimal entry and learning decisions are socially suboptimal. Players always over-invest in learning and more opaque markets tend to be more crowded. Nevertheless, transparency is not always optimal as more opaqueness might still lead to higher welfare by discouraging costly learning without excessively increasing crowding. Opaqueness is more likely to be beneficial if competition is not too fierce, while transparency is preferred if competition is excessive. Fiercer competition leads to more wasteful learning leading to deteriorating welfare without affecting crowding. Also, a larger share of sophisticated players can be welfare destroying.

References

Abreu, Dilip, and Markus K. Brunnermeier, 2003, Bubbles and crashes, *Econometrica* 71, 173–204.

Colombo, Luca, Gianluca Femminis, and Alessandro Pavan, 2014, Information acquisition and welfare, *The Review of Economic Studies* 81, 1438–1483.

Glode, Vincent, Richard C. Green, and Richard Lowery, 2012, Financial expertise as an arms race, *Journal of Finance* 67, 1723–1759.

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

Hirshleifer, Jack, 1971, The private and social value of information and the reward to inventive activity, *American Economic Review* 61, 561–574.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp, 2016, A rational theory of mutual funds' attention allocation, *Econometrica* 84, 571–626.

Kondor, Peter, and Adam Zawadowski, 2016, Learning in crowded markets, The Paul Woolley Centre Paper Series.

Krueger, Anne O., 1974, The political economy of the rent-seeking society, *American Economic Review* 64, pp. 291–303.

Loury, Glenn C., 1979, Market structure and innovation, *Quarterly Journal of Economics* 93, 395–41.

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

Maćkowiak, Bartosz, and Mirko Wiederholt, 2009, Optimal sticky prices under rational inattention, *American Economic Review* 99, 769–803.

Matějka, Filip, and Alisdair McKay, 2015, Foundation for the multinomial logit model, *American Economic Review* 105, 272–98.

Moinas, Sophie, and Sebastien Pouget, 2013, The bubble game: An experimental study of speculation, *Econometrica* 81, 1507–1539 University of Toulouse working paper.

Myatt, David P., and Chris Wallace, 2012, Endogenous information acquisition in coordination games, *Review of Economic Studies* 79, 340–374.

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

——— , 2003, Implications of rational inattention, *Journal of Monetary Economics* 50, 665–90.

Stein, Jeremy C., 2009, Presidential address: Sophisticated investors and market efficiency, *Journal of Finance* 64, 1517–1548.

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

Woodford, Michael, 2008, Inattention as a source of randomized discrete adjustment, New York University working paper.

Yang, Ming, 2015a, Coordination with rational inattention, *Journal of Economic Theory* 158, 721–738.

——— , 2015b, Optimality of debt under flexible information acquisition, Duke University working paper.

A Proofs

Proof of Lemma 1

Proof. The proof follows that of Lemma 1 in Woodford (2008) which is identical to that of Lemma 1 in Yang (2015). For a detailed proof, see the working paper version of Yang (2015). □

Proof of Lemma 2

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

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

yielding $\bar{\theta} = \frac{1}{M \cdot \beta}$. Clearly if $\bar{\theta} > 1$ in the above, then even the last player finds it optimal to enter and $m(\theta) = 1$. This yields (7).

To derive the planner's choice, note that welfare is by definition aggregate revenue minus total cost of learning. We show that with $\kappa = 0$ aggregate revenue $M \cdot R$ only depends on the aggregate entry $M \cdot p$.

$$M \cdot R = M \cdot \int_0^1 m(\theta) \cdot \Delta u(\theta) \mathrm{d}\theta = M \cdot \int_0^1 m(\theta) \cdot \left(1 - M \cdot \beta \cdot \int_0^\theta m(\tilde{\theta}) \mathrm{d}\tilde{\theta} + M \cdot \alpha \cdot \int_\theta^1 m(\tilde{\theta}) \mathrm{d}\tilde{\theta}\right) \mathrm{d}\theta = \tag{A.2}$$

$$M \cdot p - M^2 \cdot \beta \cdot \int_0^1 m(\theta) \int_0^\theta m(\tilde{\theta}) \mathrm{d}\tilde{\theta} \mathrm{d}\theta + M^2 \cdot \alpha \cdot \int_0^1 m(\theta) \int_\theta^1 m(\tilde{\theta}) \mathrm{d}\tilde{\theta} \mathrm{d}\theta = \tag{A.3}$$

$$M \cdot p - M^2 \cdot \beta \cdot \int_0^1 \int_0^1 m(\theta) \cdot m(\tilde{\theta}) \cdot \mathbf{1}_{\{\theta > \tilde{\theta}\}} \mathrm{d}\tilde{\theta} \mathrm{d}\theta + M^2 \cdot \alpha \cdot \int_0^1 \int_0^1 m(\theta) \cdot m(\tilde{\theta}) \cdot \mathbf{1}_{\{\theta < \tilde{\theta}\}} \mathrm{d}\tilde{\theta} \mathrm{d}\theta \tag{A.4}$$

By symmetry, each double integral can be simplified by splitting it into two identical "halves" which then can be rejoined. We show this formally for the first one but the same can be shown for the second integral.

$$\int_0^1 \int_0^1 m(\theta) \cdot m(\tilde{\theta}) \cdot \mathbf{1}_{\{\theta < \tilde{\theta}\}} \mathrm{d}\tilde{\theta} \mathrm{d}\theta = \frac{1}{2} \cdot \int_0^1 \int_0^1 m(\theta) \cdot m(\tilde{\theta}) \cdot \mathbf{1}_{\{\theta < \tilde{\theta}\}} \mathrm{d}\tilde{\theta} \mathrm{d}\theta + \frac{1}{2} \cdot \int_0^1 \int_0^1 m(\theta) \cdot m(\tilde{\theta}) \cdot \mathbf{1}_{\{\theta > \tilde{\theta}\}} \mathrm{d}\tilde{\theta} \mathrm{d}\theta = \tag{A.5}$$

$$\frac{1}{2} \cdot \int_0^1 \int_0^1 m(\theta) \cdot m(\tilde{\theta}) \mathrm{d}\tilde{\theta} \mathrm{d}\theta = \frac{1}{2} \cdot \int_0^1 m(\theta) \mathrm{d}\theta \cdot \int_0^1 m(\tilde{\theta}) \mathrm{d}\tilde{\theta} = \frac{1}{2} \cdot p^2 \tag{A.6}$$

Thus the aggregate revenue $M \cdot R$ simplifies to a function of aggregate entry $M \cdot p$:

$$M \cdot R = M \cdot p - M^2 \cdot \beta \cdot \frac{1}{2} \cdot p^2 + M^2 \cdot \alpha \cdot \frac{1}{2} \cdot p^2 = (M \cdot p) - (\beta - \alpha) \cdot \frac{1}{2} \cdot (M \cdot p)^2 \tag{A.7}$$

All that remains is thus to pin down the optimal average (and thus aggregate) entry that maximizes (A.7). The first order condition for the maximum yields:

$$p = \frac{1}{M} \cdot \frac{1}{\beta - \alpha}. \tag{A.8}$$

Clearly if the resulting $p > 1$, then the optimal strategy is $m_s(\theta) = 1$, leading to (8). Because learning is free, we could choose many symmetric entry functions. For simplicity, let us choose the strategy in which all players with $\theta < \bar{\theta}$ enter, the others stay out.[9] □

Proof of Lemma 3

Proof. Under no information, in the competitive equilibrium every player 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:

$$\int_0^1 (M \cdot \alpha \cdot (1-\theta) \cdot p - M \cdot \beta \cdot \theta \cdot p + 1)\, \mathrm{d}\theta = 0 \tag{A.9}$$

yielding the unconditional entry probability in (9). If M is low and the implied entry is larger than one, then the revenue is not driven to zero and everyone enters for sure implying $p = 1$.

Only aggregate entry matters for the revenue and the entropy function penalizes any learning that leads to a non-constant conditional entry function $m(\theta)$. Thus in the social planner's optimum, the planner maximizes welfare by choosing a constant conditional entry function $m(\theta) = p$ that maximizes welfare:

$$W(p) = \int_0^1 p \cdot (M \cdot \alpha \cdot (1-\theta) \cdot p - M \cdot \beta \cdot \theta \cdot p + 1)\, \mathrm{d}\theta. \tag{A.10}$$

Taking derivative with respect to p and setting to zero, this implies the entry probability in (8). As before, if the implied entry probability is larger than one then everyone enters for sure $m(\theta) = 1$ implying $p_s = 1$. □

Proof of Lemma 4

Proof. This Lemma is the adaptation of Proposition 1 of Yang (2015) for our model (which itself is based on Lemma 2 in Woodford (2008)). For a detailed proof, see the working paper version of Yang (2015). For exhibitional purposes we restate the most important part of the proof that derives (10).

First note that if the strategy function of all other players is $\tilde{m}(\theta)$, then

$$\Delta u(\theta) = 1 - M \cdot \beta \cdot \int_0^\theta \tilde{m}(\tilde{\theta}) d\tilde{\theta} + M \cdot \alpha \cdot \int_\theta^1 \tilde{m}(\tilde{\theta}) d\tilde{\theta}. \tag{A.11}$$

We use a first order perturbation method to devise the first order condition. We set $m(\theta) + \chi \cdot \epsilon(\theta)$ as $m(\theta)$, while we keep the entry decision of the others $\tilde{m}$ fixed:

$$\int_0^1 ((m(\theta) + \chi \cdot \epsilon(\theta)) \cdot \Delta u(\tilde{m}, \theta) - \mu \cdot L(m(\theta) + \chi \cdot \epsilon(\theta)))\, \mathrm{d}\theta. \tag{A.12}$$

[9] In fact for the case $\kappa > 0$ this is the unique solution.

We then take derivative with respect to χ and then set $\chi = 0$ yielding the FOC:

$$\int_0^1 \epsilon(\theta) \cdot \left(\Delta u(\tilde{m}, \theta) - \mu \cdot \left[\log\left(\frac{m(\theta)}{1 - m(\theta)} \right) - \log\left(\frac{\int_0^1 m(\tilde{\theta}) d\tilde{\theta}}{1 - \int_0^1 m(\tilde{\theta}) d\tilde{\theta}} \right) \right] \right) \mathrm{d}\theta = 0. \tag{A.13}$$

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 and get (10).

□

Proof of Proposition 1

Proof. First, from part i) of Lemma 4 the symmetric solution is $m(\theta) = 1$ if and only if the following condition holds:

$$\int_0^1 e^{-\frac{\Delta u(\theta)}{\mu}} \mathrm{d}\theta = \int_0^1 e^{-\frac{1 - M \cdot \beta \cdot \theta + M \cdot \alpha \cdot (1-\theta)}{\mu}} \mathrm{d}\theta \leq 1 \tag{A.14}$$

where we simply substituted $\Delta u(\theta)$ using $m(\theta) = 1$. Simplifying and rearranging this yields $M \leq \bar{M}$ where $\bar{M}$ is defined by (11). Now we show that (11) has a unique solution $\bar{M}$ for all admissible parameters. To show this first observe that the second derivative of the function $\frac{M \cdot (\alpha+\beta)}{\mu} - e^{-\frac{1-\beta \cdot M}{\mu}} + e^{-\frac{1+\alpha \cdot M}{\mu}}$ in M is negative if and only if $\alpha^2 < \beta^2 e^{\frac{M(\alpha+\beta)}{\mu}}$ which holds for all $M > 0$ because by our assumptions $|\beta| > |\alpha|$, $\alpha + \beta > 0$ and $\mu > 0$. Thus, the above function defining $\bar{M}$ is continuous, concave, has the value 0 at $M = 0$. It is also increasing in M at $M = 0$ because the first derivative is positive if $\alpha + \beta > 0$, which holds by assumption. It is also true that the function diverges to $-\infty$ as $M \to \infty$ if $\beta > |\alpha|$, which again holds by assumption. Thus, it has a unique positive root $\bar{M}$. In this case Lemma 4 states that $m(\theta) = 1$ is a unique best response, thus this is also a unique symmetric equilibrium.

Second, $m(\theta) = 0$ is never a symmetric equilibrium since if it was, then the following would have to hold

$$\int_0^1 e^{\frac{\Delta u(\theta)}{\mu}} \mathrm{d}\theta = e^{\frac{1}{\mu}} \leq 1 \tag{A.15}$$

where we simply substituted $\Delta u(\theta)$ using $m(\theta) = 0$. This can never hold because $\mu > 0$.

Third, if $M > \bar{M}$ we must thus be in case iii) of Lemma 4 and therefore have a symmetric interior solution. Differentiating the first order condition (10) we arrive at the following differential equation:

$$(M \cdot \alpha + M \cdot \beta) \cdot \tilde{m}(\theta) = -\frac{\mu \cdot m'(\theta)}{m(\theta) \cdot (1 - m(\theta))}. \tag{A.16}$$

Thus, the competitive equilibrium strategy $m(\theta)$ in the symmetric equilibrium ($m = \tilde{m}$) has to solve the above differential equation with the original first order condition (e.g. evaluated at $\theta = 0$) as a boundary condition which is (13). The solution of (A.16) is

$$\frac{\frac{1}{m(\theta)} + \log\left(\frac{1-m(\theta)}{m(\theta)}\right)}{M(\alpha + \beta)} = C + \frac{\theta}{\mu} \tag{A.17}$$

for an appropriate constant C. Setting $\theta = 0$ above and subtracting from the above we can eliminate C and thus arrive at

$$\frac{1}{m(\theta)} + \log\left(\frac{1-m(\theta)}{m(\theta)}\right) - \frac{M(\alpha+\beta)}{\mu} \cdot \theta = \frac{1}{m(0)} + \log\left(\frac{1-m(0)}{m(0)}\right). \tag{A.18}$$

Taking logs and using the definition of the Lambert function (upper branch if $z > 0$) yields (12). Note that we have only proved the existence of a symmetric equilibrium, as the above is a solution to (10) which – according to Lemma 4 – describes the unique best response to the others playing the same interior strategy. Note that we do not prove that (A.17) is a unique solution to the differential (A.16) but based on numerical simulations this is likely to be the case. □

Proof of Lemma 5

Proof. In neither cases are there any learning costs, as either learning is free ($\mu = 0$) or because there is no learning (if $\mu = \infty$). Thus in all cases $W = M \cdot R$ and the expressions in the Lemma directly follow from plugging in the aggregate entry levels (from Lemmas 2 and 3 into (A.7). □

Proof of Proposition 2

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 first order condition (10) at $\theta = 0$ and $\theta = 1$.

$$p = \frac{\mu\left(\log\left(\frac{m(0)}{1-m(0)}\right) - \log\left(\frac{m(1)}{1-m(1)}\right)\right)}{M(\alpha+\beta)} \tag{A.19}$$

Second, the boundary condition (10) at $\theta = 0$

$$\alpha M p + 1 = \mu\left(\log\left(\frac{m(0)}{1-m(0)}\right) - \log\left(\frac{p}{1-p}\right)\right). \tag{A.20}$$

Third, the implicit (A.18) for $m(\theta)$ evaluated at $\theta = 1$

$$\log\left(\frac{m(0)}{1-m(0)}\right) - \log\left(\frac{m(1)}{1-m(1)}\right) = \frac{M(\alpha+\beta)}{\mu} + \frac{1}{m(0)} - \frac{1}{m(1)} \tag{A.21}$$

Substituting

$$x_0 = \log\left(\frac{m(0)}{1-m(0)}\right) \tag{A.22}$$

and

$$x_1 = \log\left(\frac{m(1)}{1-m(1)}\right) \tag{A.23}$$

the system of three equations can be written as:

$$p = \frac{\mu(x_0 - x_1)}{M(\alpha+\beta)} \tag{A.24}$$

$$\alpha M p + 1 = \mu \left(x_0 - \log \left(\frac{p}{1-p} \right) \right) \tag{A.25}$$

$$x_0 - x_1 = \frac{M(\alpha+\beta)}{\mu} + e^{-x_0} - e^{-x_1}. \tag{A.26}$$

Note that $p > 0$ by definition (3) which implies $x_0 > x_1$ by (A.24). Substituting out p from (A.24), (A.25), (A.26) we arrive at a system of two equations:

$$F = \mu \left(x_0 - \log \left(\frac{\mu(x_0 - x_1)}{M(\alpha+\beta) + \mu(x_1 - x_0)} \right) \right) - \left(\frac{\alpha\mu(x_0 - x_1)}{\alpha+\beta} + 1 \right) = 0 \tag{A.27}$$

$$G = \frac{M(\alpha+\beta)}{\mu} - (x_0 - x_1) + e^{-x_0} - e^{-x_1} = 0 \tag{A.28}$$

To prove $M \cdot p$ is constant, it is sufficient to prove $\frac{\partial(M \cdot p)}{\partial M} = 0$ which from (A.24) is equivalent to

$$\frac{\partial x_0}{\partial M} = \frac{\partial x_1}{\partial M} \tag{A.29}$$

We apply Cramer's rule both for x_0 and x_1 to the system of equations (A.27) and (A.28):

$$\frac{\partial x_1}{\partial M} = \frac{\begin{vmatrix} \frac{\partial F}{\partial x_0} & -\frac{\partial F}{\partial M} \\ \frac{\partial G}{\partial x_0} & -\frac{\partial G}{\partial M} \end{vmatrix}}{\begin{vmatrix} \frac{\partial F}{\partial x_0} & \frac{\partial F}{\partial x_1} \\ \frac{\partial G}{\partial x_0} & \frac{\partial G}{\partial x_1} \end{vmatrix}} \tag{A.30}$$

$$\frac{\partial x_0}{\partial M} = \frac{\begin{vmatrix} -\frac{\partial F}{\partial M} & \frac{\partial F}{\partial x_1} \\ -\frac{\partial G}{\partial M} & \frac{\partial G}{\partial x_1} \end{vmatrix}}{\begin{vmatrix} \frac{\partial F}{\partial x_0} & \frac{\partial F}{\partial x_1} \\ \frac{\partial G}{\partial x_0} & \frac{\partial G}{\partial x_1} \end{vmatrix}} \tag{A.31}$$

First, we show that the denominator,

$$\frac{\partial F}{\partial x_0}\frac{\partial G}{\partial x_1} - \frac{\partial F}{\partial x_1}\frac{\partial G}{\partial x_0} \tag{A.32}$$

is always positive. For this, note that $\frac{\partial F}{\partial x_0} + \frac{\partial F}{\partial x_1} = \mu$. Hence, we can rewrite the denominator as

$$\frac{\partial F}{\partial x_0} \left(\frac{\partial G}{\partial x_0} + \frac{\partial G}{\partial x_1} \right) - \mu \cdot \frac{\partial G}{\partial x_0}. \tag{A.33}$$

As $\frac{\partial G}{\partial x_0} = -\frac{1}{m(0)} < 0$ and (using $x_0 > x_1$),

$$\frac{\partial G}{\partial x_0} + \frac{\partial G}{\partial x_1} = e^{-x_1} - e^{-x_0} > 0, \tag{A.34}$$

it is sufficient to show that $\frac{\partial F}{\partial x_0} > 0$. This is implied by the observations $\lim_{M\to 0} \frac{\partial F}{\partial x_0} = \frac{\beta\mu}{\alpha+\beta} > 0$ and

$$\frac{\partial \frac{\partial F}{\partial x_0}}{\partial M} = \frac{\mu^2(\alpha+\beta)}{(M(\alpha+\beta)+\mu(x_1-x_0))^2} > 0. \tag{A.35}$$

Second, we show that the numerators of the Cramer rule for the two derivatives are equal, yielding the sufficient condition

$$\frac{(\alpha+\beta)e^{-x_0-x_1}\left(e^{x_0+x_1}(M(\alpha+\beta)+\mu(x_1-x_0)) - \mu e^{x_0} + \mu e^{x_1}\right)}{M(\alpha+\beta)+\mu(x_1-x_0)} = 0. \tag{A.36}$$

It follows from (A.26) that the denominator is non-zero if $x_0 \neq x_1$. Thus it is sufficient to prove that

$$\frac{M(\alpha+\beta)}{\mu} + (x_1 - x_0) + \frac{1}{e^{x_0}} - \frac{1}{e^{x_1}} = 0, \tag{A.37}$$

which is exactly the function $G = 0$ defined in (A.28). Thus the identity holds and we have proved that $M \cdot p$ is constant in M for interior solutions. □

Lemma A.1. Technical properties of $\bar{M}$. *Let us define*

$$A \equiv e^{\frac{\bar{M}(\alpha+\beta)}{\mu}} - 2e^{\frac{\alpha\bar{M}+1}{\mu}} + 1 \tag{A.38}$$

and

$$B \equiv \beta\left(e^{\frac{\alpha\bar{M}+1}{\mu}} - e^{\frac{\bar{M}(\alpha+\beta)}{\mu}}\right) + \alpha\left(e^{\frac{\alpha\bar{M}+1}{\mu}} - 1\right). \tag{A.39}$$

For any set of feasible parameters, $A > 0$, $B < 0$, and

$$B \cdot \bar{M} + e^{\frac{\bar{M}(\alpha+\beta)}{\mu}} - 1 < 0. \tag{A.40}$$

Furthermore,

$$\frac{d\bar{M}}{d\beta} = -\left(\frac{A}{B}\alpha + 1\right)\bar{M}(\beta-\alpha) \tag{A.41}$$

$$\frac{d\bar{M}}{d\alpha} = \left(\frac{A}{B}\beta + 1\right)\bar{M}(\beta-\alpha) \tag{A.42}$$

$$\frac{d\bar{M}}{d\mu} = \left(\bar{M} + \frac{e^{\bar{M}\frac{\alpha+\beta}{\mu}} - 1}{B}\right)\frac{1}{\mu} \tag{A.43}$$

.

Proof of Lemma A.1[10]

[10] We thank Gábor Lippner for helping us with this proof.

Proof. First note that rearranging (11) defining $\bar{M}$ yields:

$$e^{\frac{1}{\mu}} = \frac{e^{\beta\cdot\frac{\bar{M}}{\mu}} - e^{-\alpha\cdot\frac{\bar{M}}{\mu}}}{\frac{\bar{M}}{\mu}\cdot(\alpha+\beta)}. \tag{A.44}$$

Let us start by showing that $A > 0$. Using (A.44), (A.38) can be rewritten as:

$$A = e^{(\alpha+\beta)\cdot\frac{\bar{M}}{\mu}} - 2\cdot e^{\frac{1}{\mu}}\cdot e^{\alpha\frac{\bar{M}}{\mu}} + 1 = e^{(\alpha+\beta)\cdot\frac{\bar{M}}{\mu}} - 2\cdot\frac{e^{(\alpha+\beta)\cdot\frac{\bar{M}}{\mu}-1}}{(\alpha+\beta)\cdot\frac{\bar{M}}{\mu}} + 1 = e^z - \frac{2}{z}\cdot(e^z-1)+1, \tag{A.45}$$

where we denote $z = (\alpha+\beta)\cdot\frac{\bar{M}}{\mu}$. Note that $z > 0$ by assumption thus it suffices to show that $z\cdot A > 0$ for all $z > 0$. This follows from the fact that $\lim_{z\to 0^+} A\cdot z = 0$ and that the first derivative of $A\cdot z$ is $1 - e^z(1-z)$ which is strictly positive for $z > 0$. The latter statement is obvious for $z \geq 1$ and for $z < 1$ it reduces to showing $e^z < \frac{1}{1-z}$. Using the Taylor series of e^z around $z = 0$ it holds that $e^z < 1+z+z^2$ for $z < 1$ since $\sum_{k=3}^{\infty}\frac{1}{k!} < \frac{1}{2}$. It is thus sufficient to prove $1+z+z^2 < \frac{1}{1-z}$ which holds for all $z \in (0,1)$. Thus we have shown $A > 0$.

In order to prove $B < 0$ it suffices to prove that $\frac{\bar{M}}{\mu}\cdot e^{-\alpha\cdot\frac{\bar{M}}{\mu}}\cdot B < 0$ since $\frac{\bar{M}}{\mu}\cdot e^{-\alpha\cdot\frac{\bar{M}}{\mu}} > 0$. Using (A.39) one can thus write:

$$\frac{\bar{M}}{\mu}\cdot e^{-\alpha\cdot\frac{\bar{M}}{\mu}}\cdot B = (\alpha+\beta)\cdot\frac{\bar{M}}{\mu}\cdot e^{\frac{1}{\mu}} - \beta\cdot\frac{\bar{M}}{\mu}\cdot e^{\beta\frac{\bar{M}}{\mu}} - \alpha\cdot\frac{\bar{M}}{\mu}\cdot e^{-\alpha\cdot\frac{\bar{M}}{\mu}} = e^{\beta\cdot\frac{\bar{M}}{\mu}} - e^{-\alpha\cdot\frac{\bar{M}}{\mu}} - \beta\cdot\frac{\bar{M}}{\mu}\cdot e^{\beta\cdot\frac{\bar{M}}{\mu}} - \alpha\cdot\frac{\bar{M}}{\mu}e^{-\alpha\cdot\frac{\bar{M}}{\mu}} \tag{A.46}$$

where we substituted $e^{\frac{1}{\mu}}$ from (A.44). Thus we only have to show:

$$e^{\beta\cdot\frac{\bar{M}}{\mu}} - \beta\cdot\frac{\bar{M}}{\mu}\cdot e^{\beta\cdot\frac{\bar{M}}{\mu}} < e^{-\alpha\cdot\frac{\bar{M}}{\mu}} - \left(-\alpha\cdot\frac{\bar{M}}{\mu}\right)\cdot e^{-\alpha\cdot\frac{\bar{M}}{\mu}}. \tag{A.47}$$

Denote $f(t) = e^t - t\cdot e^t$, and the above simplifies to

$$f\left(\beta\cdot\frac{\bar{M}}{\mu}\right) < f\left(-\alpha\cdot\frac{\bar{M}}{\mu}\right). \tag{A.48}$$

Since $\beta\cdot\frac{\bar{M}}{\mu} > -\alpha\cdot\frac{\bar{M}}{\mu}$ follows from $\beta+\alpha > 0$. If $\alpha \leq 0$ then (A) follows from the fact that $f(t)$ is a monotone decreasing function if $t > 0$ since $f'(t) = -t\cdot e^t < 0$ (and it is monotone increasing if $t < 0$). If $\alpha > 0$ then we use that by our assumptions that $\beta > |\alpha|$. Thus $f\left(-\alpha\cdot\frac{\bar{M}}{\mu}\right) > f\left(-\beta\cdot\frac{\bar{M}}{\mu}\right)$ since $f(t)$ is a monotone increasing function if $t < 0$. Thus in this case it suffices to show that $f\left(\beta\cdot\frac{\bar{M}}{\mu}\right) < f\left(-\beta\cdot\frac{\bar{M}}{\mu}\right)$ which means all we need to show that is $f(t) < f(-t)$ for all $t > 0$, i.e. $e^{2t}\cdot(t-1)+t+1 > 0$ for $t > 0$ which is obvious for $t \geq 1$ and for $t \in (0,1)$ reduces to showing $\frac{t+1}{1-t} > e^{2t}$. An upper bound on e^{2t} using the Taylor expansion is $e^{2t} < 1+2t+2t^2+\frac{4t^3}{3}+2t^4$ for $t \in (0,1)$, which follows from approximating all fourth and higher order terms from above using $\sum_{k=4}^{\infty}\frac{2^k}{k!} < 2$. Thus it suffices to show $\frac{t+1}{1-t} > 1+2t+2t^2+\frac{4t^3}{3}+2t^4$ which simplifies to $\frac{2}{3}+2t^2 > \frac{2}{3}t$ and is true for all $t \in (0,1)$. Thus we have shown $B < 0$.

In order to prove $B \cdot \bar{M} + e^{\frac{\bar{M}(\alpha+\beta)}{\mu}} - 1 < 0$ we first multiply it by $\frac{1}{\mu} \cdot e^{-\alpha \cdot \frac{\bar{M}}{\mu}} > 0$ and thus have to show

$$e^{-\alpha \cdot \frac{\bar{M}}{\mu}} \cdot \frac{\bar{M}}{\mu} \cdot B + \frac{1}{\mu} \cdot \left(e^{\beta \cdot \frac{\bar{M}}{\mu}} - e^{-\alpha \cdot \frac{\bar{M}}{\mu}} \right) < 0 \tag{A.49}$$

using (A.46) this can be rewritten as

$$\left(e^{\beta \frac{\bar{M}}{\mu}} - e^{-\alpha \frac{\bar{M}}{\mu}} \right) (1 + 1/\mu) - \beta \frac{\bar{M}}{\mu} e^{\beta \frac{\bar{M}}{\mu}} - \alpha \frac{\bar{M}}{\mu} e^{-\alpha \frac{\bar{M}}{\mu}} < 0 \tag{A.50}$$

Let $s = \beta \frac{\bar{M}}{\mu}$ and $t = -\alpha \frac{\bar{M}}{\mu}$, by our assumptions $s > t$. Using (A.44) $\frac{1}{\mu} = \log e^{\frac{1}{\mu}} = \log \frac{e^s - e^t}{s-t}$, substituting this into (A.50), one needs to show:

$$1 + \log \left(\frac{e^s - e^t}{s - t} \right) < \frac{se^s - te^t}{e^s - e^t} = t + \frac{(s-t)e^s}{e^s - e^t} = t + (s-t) \frac{e^{s-t}}{e^{s-t} - 1} \tag{A.51}$$

Note that $1 + \log \left(\frac{e^s - e^t}{s-t} \right) = 1 + t + \log \left(\frac{e^{s-t} - 1}{s-t} \right)$, thus denoting $z = s - t > 0$, it suffices to prove

$$1 + \log(e^z - 1) - \log z - z \frac{e^z}{e^z - 1} < 0. \tag{A.52}$$

At $z \to 0+$ the left hand side evaluates to zero, thus it suffices to show that it is monotonically decreasing in z. Taking the derivative, we need to show

$$\frac{-1 - e^{2z} + e^z(2 + z^2)}{z(e^z - 1)^2} < 0 \tag{A.53}$$

Since the denominator is positive, we need to show $-1 - e^{2z} + e^z(2 + z^2) < 0$ for $z > 0$. Note that at $z = 0$ this evaluates to 0, thus it suffices to show that $-1 - e^{2z} + e^z(2 + z^2)$ is monotone decreasing. Taking the derivative again, one has to show $e^z \left(z^2 + 2z - 2e^z + 2 \right) < 0$ which simply follows from the fact that the Taylor expansion of e^z implies $e^z > 1 + z + \frac{z^2}{2}$ for $z > 0$, as we can neglect the other positive terms. Thus we have shown that $B \cdot \bar{M} + e^{\frac{\bar{M}(\alpha+\beta)}{\mu}} - 1 < 0$.

For the derivatives of $\bar{M}$ recall that $\bar{M}$ is defined implicitly by (11):

$$F = \frac{\bar{M} \cdot (\alpha + \beta)}{\mu} - e^{-\frac{1 - \beta \cdot \bar{M}}{\mu}} + e^{-\frac{1 + \alpha \cdot \bar{M}}{\mu}} = 0. \tag{A.54}$$

Thus, $\frac{d\bar{M}}{d.}$ for any parameter "." becomes: $-\frac{\frac{\partial F}{\partial .}}{\frac{\partial F}{\partial \bar{M}}}$ by the implicit function theorem. Then, (A.42)-(A.43) are given by simple algebra. ◻

Proof of Proposition 3

Proof. $\bar{M}_s = \frac{1}{\beta - \alpha}$ follows from (8). Using it the derivative of interest $\frac{\partial \frac{\bar{M}}{\bar{M}_s}}{\partial .}$ for any parameter "." becomes:

$$\frac{d \frac{\bar{M}}{\bar{M}_s}}{d.} = \frac{\partial \bar{M}}{\partial .} \cdot (\beta - \alpha) + \frac{\partial (\beta - \alpha)}{\partial .} \cdot \bar{M}. \tag{A.55}$$

Basic algebra and Lemma A.1 yields

$$\frac{d\frac{\bar{M}}{\bar{M}_s}}{d\alpha} = \frac{A}{B} \cdot \beta \cdot \bar{M} < 0 \tag{A.56}$$

proving the first statement, and

$$\frac{d\frac{\bar{M}}{\bar{M}_s}}{d\beta} = -\frac{A}{B} \cdot \alpha \cdot \bar{M}, \tag{A.57}$$

the sign of which is the same as the sign of α, proving the second statement in the proposition. □

Proof of Proposition 4

Proof. By Proposition 1 when $M \leq \bar{M}$ all players enter with probability 1. Hence, all equilibrium objects are the same for the planner and in the decentralized solution. In particular, average entry of a player is $p = 1$ thus expected aggregate entry is M. Aggregate revenue and welfare in this full entry case are

$$M \cdot R = W = M \cdot R_s = W_s = M \cdot \int_0^1 (M \cdot \alpha \cdot (1-\theta) - M \cdot \beta \cdot \theta + 1) \, \mathrm{d}\theta = M - \frac{M^2 (\beta - \alpha)}{2}. \tag{A.58}$$

To arrive at a formula for $W(M)$ for $M > \bar{M}$, we first express aggregate learning from (2) and rearrange by using the fact that p is constant in θ and that $p = \int_0^1 m(\theta) \, \mathrm{d}\theta$:

$$M \cdot L = M \cdot \left(p \log \left[\frac{1}{p}\right] + (1-p) \log \left[\frac{1}{1-p}\right] \right) - M \cdot \int_0^1 \left(m(\theta) \log \left[\frac{1}{m(\theta)}\right] + (1-m(\theta)) \log \left[\frac{1}{1-m(\theta)}\right] \right) \mathrm{d}\theta =$$

$$M \cdot \int_0^1 p \log \left[\frac{1}{p}\right] + (1-p) \log \left[\frac{1}{1-p}\right] - m(\theta) \log \left[\frac{1}{m(\theta)}\right] - (1-m(\theta)) \log \left[\frac{1}{1-m(\theta)}\right] \mathrm{d}\theta =$$

$$M \cdot \int_0^1 m(\theta) \cdot \log \left[\frac{1}{p}\right] + (1-m(\theta)) \cdot \log \left[\frac{1}{1-p}\right] - \log \left[\frac{1}{1-m(\theta)}\right] + m(\theta) \log \left[\frac{m(\theta)}{1-m(\theta)}\right] \mathrm{d}\theta =$$

$$M \cdot \int_0^1 m(\theta) \cdot \left(\log \left(\frac{m(\theta)}{1-m(\theta)} \right) - \log \left(\frac{p}{1-p} \right) \right) \mathrm{d}\theta - M \cdot \int_0^1 \log \left(\frac{1-p}{1-m(\theta)} \right) \mathrm{d}\theta. \tag{A.59}$$

Multiplying the above by μ and replacing the interior part of the first integral using the first order condition expressed in (10) yields:

$$M \cdot \mu \cdot L = M \cdot \int_0^1 m(\theta) \cdot \left[M \cdot \alpha \cdot \int_\theta^1 \tilde{m}(\tilde{\theta}) d\tilde{\theta} - M \cdot \beta \cdot \int_0^\theta \tilde{m}(\tilde{\theta}) d\tilde{\theta} + 1 \right] \mathrm{d}\theta - M \cdot \int_0^1 \mu \cdot \log \left(\frac{1-p}{1-m(\theta)} \right) \mathrm{d}\theta. \tag{A.60}$$

Thus the first integral is exactly the definition of aggregate revenue. Since $M \cdot p$ is constant if $M \geq \bar{M}$ (see Proposition 2), so is aggregate revenue $M \cdot R$. Rearranging yields the below expression for W:

$$W(M) = M \cdot \mu \cdot \int_0^1 \log \left(\frac{1-p}{1-m(\theta)} \right) \mathrm{d}\theta \tag{A.61}$$

We now show that welfare converges to zero for $M \to \infty$. Since both $m(\theta) \to 0$ and $p \to 0$ in this case, one can use the first order approximation that

$$\lim_{m(\theta),p\to 0} \log\left(\frac{1-p}{1-m(\theta)}\right) = \lim_{p\to 0} \log(1-p) - \lim_{m(\theta)\to 0} \log(1-m(\theta)) \approx m(\theta) - p \tag{A.62}$$

Since $m \to 0$ it holds that $\frac{\log(\frac{1}{m})}{(\frac{1}{m})} \to 0$ and $\log(1-m) \to 0$ thus the implicit equation (A.17) for $m(\theta)$ can be approximated to first order by

$$\frac{1}{m(\theta)} - M(\alpha+\beta)\left(C + \frac{\theta}{\mu}\right) = 0. \tag{A.63}$$

A closed form solution can be obtained in this limit case:

$$m(\theta) = \frac{\mu}{M(\alpha+\beta)(C\mu+\theta)} \tag{A.64}$$

for an appropriate C. By the definition of the average entry p this implies

$$M \cdot p = M \cdot \int_0^1 m(\theta)\mathrm{d}\theta = \frac{\mu}{\alpha+\beta} \cdot \log\left(\frac{1}{C\mu} + 1\right). \tag{A.65}$$

From Proposition 2 we know that $M \cdot p$ is a constant for any $M > \bar{M}$, thus (A.65) yields:

$$C = \frac{1}{\mu \cdot \left(e^{\frac{\bar{M}\cdot(\alpha+\beta)}{\mu}} - 1\right)}. \tag{A.66}$$

Now we turn back to showing that welfare converges to zero for $M \to \infty$. Using the first order approximations of $m(\theta)$ and p as expressed in (A.64) and (A.65), one can thus write:

$$\lim_{M\to\infty} W(M) = \lim_{M\to\infty} M \cdot \mu \cdot \int_0^1 \log\left(\frac{1-p}{1-m(\theta)}\right)\mathrm{d}\theta = \lim_{M\to\infty} \mu \cdot \int_0^1 M \cdot m(\theta) - M \cdot p \;\mathrm{d}\theta = \tag{A.67}$$

$$\mu \cdot \int_0^1 \left(\frac{\mu}{(\alpha+\beta)(C\mu+\theta)} - \frac{\mu}{\alpha+\beta} \cdot \log\left(\frac{1}{C\mu}+1\right)\right)\mathrm{d}\theta = \frac{\mu^2}{\alpha+\beta} \cdot \left[\int_0^1 \frac{1}{C\mu+\theta}\mathrm{d}\theta - \log\left(\frac{1}{C\mu}+1\right)\right] = 0$$

Thus $W(M) \to 0$ and this convergence happens from above, since the payoff per player $\frac{V=W}{M}$ cannot be negative, otherwise players would choose not to enter. □

Proof of Proposition 5

Proof. Following the logic of the proof of Proposition 3 basic algebra and Lemma A.1 yields:

$$\frac{d\frac{\bar{M}}{\bar{M}_s}}{d\mu} = \frac{B \cdot \bar{M} + e^{\frac{\bar{M}(\alpha+\beta)}{\mu}} - 1}{B} \cdot \frac{\beta-\alpha}{\mu} > 0. \tag{A.68}$$

□

Proof of Proposition 6

Proof. Part 1 of the Proposition follows directly from Lemma 5.

Part 2 follows from using the expression for W from (A.61) and differentiating with respect to μ:

$$\frac{\partial W}{\partial \mu} = M \cdot \int_0^1 \log\left(\frac{1-p}{1-m(\theta)}\right) d\theta - M \cdot \mu \cdot \frac{1}{1-p}\frac{\partial p}{\partial \mu} + M \cdot \mu \cdot \int_0^1 \frac{1}{1-m(\theta)}\frac{\partial m(\theta)}{\partial \mu} d\theta \tag{A.69}$$

Also, differentiating (A.18) with respect to θ:

$$-\frac{1}{m^2(\theta)(1-m(\theta))}\frac{\partial m(\theta)}{\partial \mu} = -\frac{1}{m^2(0)(1-m(0))}\frac{\partial m(0)}{\partial \mu} - \frac{M(\alpha+\beta)}{\mu^2} \cdot \theta \tag{A.70}$$

When $\theta = 1$:

$$-\frac{1}{m^2(1)(1-m(1))}\frac{\partial m(1)}{\partial \mu} = -\frac{1}{m^2(0)(1-m(0))}\frac{\partial m(0)}{\partial \mu} - \frac{M(\alpha+\beta)}{\mu^2} \tag{A.71}$$

(A.70) minus (A.71) yields:

$$-\frac{1}{m^2(\theta)(1-m(\theta))}\frac{\partial m(\theta)}{\partial \mu} = -\frac{1}{m^2(1)(1-m(1))}\frac{\partial m(1)}{\partial \mu} + \frac{M(\alpha+\beta)}{\mu^2} \cdot (1-\theta) \tag{A.72}$$

Next we solve $\partial m(0)/\partial \mu$ and $\partial m(1)/\partial \mu$. (A.19) is equivalent to:

$$\log(\frac{m(0)}{1-m(0)}) - \log(\frac{m(1)}{1-m(1)}) = \frac{M \cdot p \cdot (\alpha+\beta)}{\mu} = \frac{\overline{M}(\alpha+\beta)}{\mu} \tag{A.73}$$

The latter equality used the fact that $M \cdot p$ is constant in M once the solution m is in interior. Define

$$Y = \frac{\overline{M}(\alpha+\beta)}{\mu}. \tag{A.74}$$

Differentiate (A.73) with respect to μ:

$$\frac{1}{m(0)(1-m(0))}\frac{\partial m(0)}{\partial \mu} - \frac{1}{m(1)(1-m(1))}\frac{\partial m(1)}{\partial \mu} = \frac{\partial Y}{\partial \mu} \tag{A.75}$$

From (A.71) and (A.71):

$$\frac{1}{m^2(0)(1-m(0))}\frac{\partial m(0)}{\partial \mu} = \frac{1}{m(0)-m(1)}\left(m(1) \cdot \frac{M(\alpha+\beta)}{\mu^2} + \frac{\partial Y}{\partial \mu}\right) \tag{A.76}$$

$$\frac{1}{m^2(1)(1-m(1))}\frac{\partial m(1)}{\partial \mu} = \frac{1}{m(0)-m(1)}\left(m(0) \cdot \frac{M(\alpha+\beta)}{\mu^2} + \frac{\partial Y}{\partial \mu}\right) \tag{A.77}$$

Calculating $\frac{m(0)-p}{m(0)-m(1)} \cdot$(A.76) $+\frac{p-m(1)}{m(0)-m(1)} \cdot$(A.77) and then substituting (A.70) and (A.72):

$$\frac{m(0)-p}{m(0)-m(1)} \cdot \frac{1}{m^2(0)(1-m(0))} \frac{\partial m(0)}{\partial \mu} + \frac{p-m(1)}{m(0)-m(1)} \cdot \frac{1}{m^2(1)(1-m(1))} \frac{\partial m(1)}{\partial \mu}$$
$$= \frac{1}{m(0)-m(1)} \left(p \cdot \frac{M(\alpha+\beta)}{\mu^2} + \frac{\partial Y}{\partial \mu} \right) = \frac{1}{m(0)-m(1)} \left(\frac{Y}{\mu} + \frac{\partial Y}{\partial \mu} \right) \tag{A.78}$$

From (A.70), (A.72) and (A.78):

$$\frac{1}{m^2(\theta)(1-m(\theta))} \frac{\partial m(\theta)}{\partial \mu}$$
$$= \frac{1}{m(0)-m(1)} \left(\frac{Y}{\mu} + \frac{\partial Y}{\partial \mu} \right) + \frac{m(0)-p}{m(0)-m(1)} \cdot \frac{M(\alpha+\beta)}{\mu^2} \cdot \theta - \frac{p-m(1)}{m(0)-m(1)} \cdot \frac{M(\alpha+\beta)}{\mu^2} \cdot (1-\theta) \tag{A.79}$$

We begin from the last term of (A.69), substituting (A.79):

$$\int_0^1 \frac{1}{1-m(\theta)} \frac{\partial m(\theta)}{\partial \mu} d\theta$$
$$= \frac{1}{m(0)-m(1)} \left(\frac{Y}{\mu} + \frac{\partial Y}{\partial \mu} \right) \int_0^1 m^2(\theta) d\theta + \frac{m(0)-p}{m(0)-m(1)} \cdot \frac{M(\alpha+\beta)}{\mu^2} \int_0^1 \theta m^2(\theta) d\theta$$
$$- \frac{p-m(1)}{m(0)-m(1)} \cdot \frac{M(\alpha+\beta)}{\mu^2} \int_0^1 (1-\theta) m^2(\theta) d\theta \tag{A.80}$$

From (A.16):

$$m'(\theta) = -\frac{M(\alpha+\beta)}{\mu} m^2(\theta)(1-m(\theta)) \tag{A.81}$$

Applying integration by parts:

$$\frac{M(\alpha+\beta)}{\mu^2} \int_0^1 \theta m^2(\theta) d\theta = -\frac{1}{\mu} \int_0^1 \theta \cdot \frac{m'(\theta)}{1-m(\theta)} d\theta$$
$$= \frac{1}{\mu} \int_0^1 \theta \, d(\log(1-m(\theta))) = \frac{1}{\mu} \left[\log(1-m(1)) - \int_0^1 \log(1-m(\theta)) d\theta \right] \tag{A.82}$$

Similarly:

$$\frac{M(\alpha+\beta)}{\mu^2} \int_0^1 (1-\theta) m^2(\theta) d\theta = \frac{1}{\mu} \left[\int_0^1 \log(1-m(\theta)) d\theta - \log(1-m(0)) \right] \tag{A.83}$$

Substitute (A.82) and (A.83) into (A.80):

$$\int_0^1 \frac{1}{1-m(\theta)} \frac{\partial m(\theta)}{\partial \mu} d\theta = \frac{1}{m(0)-m(1)} \left(\frac{Y}{\mu} + \frac{\partial Y}{\partial \mu} \right) \int_0^1 m^2(\theta) d\theta - \frac{1}{\mu} \int_0^1 \log(1-m(\theta)) d\theta$$
$$+ \frac{1}{\mu} \frac{m(0)-p}{m(0)-m(1)} \log(1-m(1)) + \frac{1}{\mu} \frac{p-m(1)}{m(0)-m(1)} \log(1-m(0)) \tag{A.84}$$

Next we study the second term of (A.69), using (A.79):

$$\begin{aligned}\frac{\partial p}{\partial \mu} = \int_0^1 \frac{\partial m(\theta)}{\partial \mu} &= \frac{1}{m(0)-m(1)}\left(\frac{Y}{\mu}+\frac{\partial Y}{\partial \mu}\right)\int_0^1 m^2(\theta)(1-m(\theta))d\theta \\ &+\frac{m(0)-p}{m(0)-m(1)}\cdot\frac{M(\alpha+\beta)}{\mu^2}\int_0^1 \theta m^2(\theta)(1-m(\theta))d\theta \\ &-\frac{p-m(1)}{m(0)-m(1)}\cdot\frac{M(\alpha+\beta)}{\mu^2}\int_0^1 (1-\theta)m^2(\theta)(1-m(\theta))d\theta \end{aligned} \tag{A.85}$$

Applying integration by parts and using (A.16):

$$\frac{M(\alpha+\beta)}{\mu^2}\int_0^1 \theta m^2(\theta)(1-m(\theta))d\theta = -\frac{1}{\mu}\int_0^1 \theta\cdot m'(\theta)d\theta = \frac{1}{\mu}\left[\int_0^1 m(\theta)d\theta - m(1)\right] = \frac{1}{\mu}(p-m(1)) \tag{A.86}$$

Similarly, using (A.16) and applying integration by parts:

$$\frac{M(\alpha+\beta)}{\mu^2}\int_0^1 (1-\theta)m^2(\theta)(1-m(\theta))d\theta = \frac{1}{\mu}\left[m(0)-\int_0^1 m(\theta)d\theta\right] = \frac{1}{\mu}(m(0)-p) \tag{A.87}$$

Substitute (A.86) and (A.87) into (A.85):

$$\frac{\partial p}{\partial \mu} = \frac{1}{m(0)-m(1)}\left(\frac{Y}{\mu}+\frac{\partial Y}{\partial \mu}\right)\int_0^1 m^2(\theta)(1-m(\theta))d\theta \tag{A.88}$$

Substitute (A.84) and (A.88) into (A.69):

$$\begin{aligned}\frac{1}{M}\frac{\partial W}{\partial \mu} &= \frac{m(0)-p}{m(0)-m(1)}\log(1-m(1)) + \frac{p-m(1)}{m(0)-m(1)}\log(1-m(0)) - \log(1-p) \\ &+2\left(\log(1-p)-\int_0^1 \log(1-m(\theta))d\theta\right) \\ &+\frac{1}{m(0)-m(1)}\left(Y+\mu\cdot\frac{\partial Y}{\partial \mu}\right)\left(\int_0^1 m^2(\theta)d\theta - \frac{1}{1-p}\int_0^1 m^2(\theta)(1-m(\theta))d\theta\right)\end{aligned} \tag{A.89}$$

In the last section, we prove that when $\alpha/\beta \to 1$, there exists a range of μ such that $\partial W/\partial \mu > 0$ for any M.

First, the function $\log(1+x) - x + \frac{1}{2\cdot(1-a)}\cdot x^2 \geq 0$ for $x \in [-a,\infty)$. Let $a = \frac{m(0)-p}{1-p}$. Substitute in $x = \frac{p-m(1)}{1-p}$ and $x = -\frac{m(0)-p}{1-p}$:

$$\log(1-m(1)) \geq \log(1-p) + \frac{p-m(1)}{1-p} - \frac{1}{2(1-a)}\left(\frac{p-m(1)}{1-p}\right)^2 \tag{A.90}$$

$$\log(1-m(0)) \geq \log(1-p) - \frac{m(0)-p}{1-p} - \frac{1}{2(1-a)}\left(\frac{m(0)-p}{1-p}\right)^2 \tag{A.91}$$

$\frac{m(0)-p}{m(0)-m(1)} \cdot (19) + \frac{p-m(1)}{m(0)-m(1)} \cdot (20)$:

$$\frac{m(0)-p}{m(0)-m(1)} \cdot \log(1-m(1)) + \frac{p-m(1)}{m(0)-m(1)} \cdot \log(1-m(0)) - \log(1-p) \geq -\frac{1}{2}\frac{(p-m(1))(m(0)-p)}{(1-p)^2} \cdot \frac{1}{1-a}$$
$$\geq -\frac{1}{8}\frac{(m(0)-m(1))^2}{(1-p)^2} \cdot \frac{1-p}{1-m(0)} \geq -\frac{1}{8}\frac{(m(0)-m(1))^2}{(1-p)^2} \cdot \frac{m(0)(1-m(1))}{m(1)(1-m(0))} = -\frac{1}{8}\frac{(m(0)-m(1))^2}{(1-p)^2} \cdot e^Y, \tag{A.92}$$

where the last equality comes from (A.73) and (A.74).

Second, $\log(1-x)$ is a concave function of x. Taking a Taylor expansion and using Jensen's inequality for all third order and higher terms:

$$\log(1-p) - \int_0^1 \log(1-m(\theta))d\theta \geq \frac{1}{2}\min_{x\in[m_1,m_0]}\frac{1}{(1-x)^2}\int_0^1 (m^2(\theta)-p^2)d\theta = \frac{1}{2}\frac{1}{(1-m(1))^2}\int_0^1 (m(\theta)-p)^2 d\theta \tag{A.93}$$

Let θ_p be the value such that $m(\theta_p) = p$ is satisfied, using (A.16)

$$\int_0^1 (m(\theta)-p)^2 d\theta = \int_0^{\theta_p} (m(\theta)-m(\theta_p))^2 d\theta + \int_{\theta_p}^1 (m(\theta_p)-m(\theta))^2 d\theta \geq \min_{\theta\in[0,1]} \left(m'(\theta)\right)^2 \left[\int_0^{\theta_p} (\theta-\theta_p)^2 d\theta + \int_{\theta_p}^1 (\theta_p-\theta)^2 d\theta\right]$$
$$\geq \left[\frac{M(\alpha+\beta)}{\mu} m^2(1)(1-m(0))\right]^2 \cdot \frac{1}{3}\left(\theta_p^3 + (1-\theta_p)^3\right) \geq \frac{1}{12}\frac{M^2(\alpha+\beta)^2}{\mu^2} m^4(1)(1-m(0))^2 \tag{A.94}$$

From (A.20) and (A.22):

$$\frac{1}{m(1)} - \frac{1}{m(0)} = \frac{M(\alpha+\beta)}{\mu} - \frac{Mp(\alpha+\beta)}{\mu} = \frac{M(\alpha+\beta)}{\mu} \cdot (1-p) \tag{A.95}$$

Substitute (A.95) into (A.94):

$$\int_0^1 (m(\theta)-p)^2 d\theta \geq \frac{1}{12}\frac{M^2(\alpha+\beta)^2}{\mu^2} m^4(1)(1-m(0))^2 = \frac{1}{12}\frac{(m(1)-m(0))^2}{(1-p)^2} \cdot \frac{m^2(1)(1-m(0))^2}{m^2(0)} \tag{A.96}$$

Substitute (A.96) into (A.93):

$$\log(1-p) - \int_0^1 \log(1-m(\theta))d\theta \geq \frac{1}{24}\frac{(m(1)-m(0))^2}{(1-p)^2} \cdot \frac{m^2(1)(1-m(0))^2}{m^2(0)(1-m(1))^2}$$
$$= \frac{1}{24}\frac{(m(1)-m(0))^2}{(1-p)^2} \exp\left[-2 \cdot \left(\log(\frac{m(0)}{1-m(0)}) - \log(\frac{m(1)}{1-m(1)})\right)\right]$$
$$= \frac{1}{24}\frac{(m(1)-m(0))^2}{(1-p)^2} \cdot e^{-2Y} \tag{A.97}$$

where the last step used (A.19) and (A.74).

Third, we provide a lower bound for the third term in (A.89). Define $g(\theta) = m(\theta) - m(1)$, where $g(\theta)$ is non-negative for $\theta \in [0,1]$ and decreasing.

$$\int_0^1 m^2(\theta)d\theta - \frac{1}{1-p}\int_0^1 m^2(\theta)(1-m(\theta))d\theta = \frac{1}{1-p}\left[\int_0^1 m^3(\theta)d\theta - \int_0^1 m(\theta)d\theta \cdot \int_0^1 m^2(\theta)d\theta\right]$$

$$= \frac{1}{1-p}\cdot\left[\int_0^1 (m(1)+g(\theta))^3 d\theta - \int_0^1 (m(1)+g(\theta))d\theta \cdot \int_0^1 (m(1)+g(\theta))^2 d\theta\right]$$

$$= \frac{1}{1-p}\cdot\left[m(1)^3 + 3m^2(1)\int_0^1 g(\theta)d\theta + 3m(1)\int_0^1 g^2(\theta)d\theta + \int_0^1 g^3(\theta)d\theta\right]$$

$$+\frac{1}{1-p}\cdot\left[-m(1)^3 - 3m^2(1)\int_0^1 g(\theta)d\theta - m(1)\int_0^1 g^2(\theta)d\theta - 2m(1)\left(\int_0^1 g(\theta)d\theta\right)^2 - \int_0^1 g(\theta)d\theta\int_0^1 g^2(\theta)d\theta\right]$$

$$= \frac{1}{1-p}\cdot\left(2m(1)\left[\int_0^1 g^2(\theta)d\theta - \left(\int_0^1 g(\theta)d\theta\right)^2\right] + \left[\int_0^1 g^3(\theta)d\theta - \int_0^1 g(\theta)d\theta\int_0^1 g^2(\theta)d\theta\right]\right) \tag{A.98}$$

For the first term in (A.98)

$$\int_0^1 g^2(\theta)d\theta - \left(\int_0^1 g(\theta)d\theta\right)^2 = \int_0^1\left(g(\theta) - \int_0^1 g(\theta)d\theta\right)^2 = \int_0^1 (m(\theta) - m(1) - (p - m(1)))^2 = \int_0^1 (m(\theta)-p)^2 \tag{A.99}$$

Now define θ_p such that $m(\theta_p) = p$, and then take the second order Taylor expansion of $m(\theta)$ around θ_p. (A.99) can be approximated by

$$\int_0^1 (m(\theta)-p)^2 \geq \min_{\theta\in[0,1]}\left(m'(\theta)\right)^2\left[\int_0^{\theta_p}(\theta-\theta_p)^2 d\theta + \int_{\theta_p}^1 (\theta_p - \theta)^2 d\theta\right] \geq \frac{1}{12}\min_{\theta\in[0,1]}\left(m'(\theta)\right)^2 \tag{A.100}$$

Since both $g(\theta)$ and $g^2(\theta)$ are decreasing functions of θ. From Chebyshev's sum inequality,

$$\int_0^1 g^3(\theta)d\theta \geq \int_0^1 g(\theta)d\theta\int_0^1 g^2(\theta)d\theta \tag{A.101}$$

Plugging (A.99), (A.100), (A.101) into (A.98) one arrives at the following lower bound for the third term in (A.89):

$$\begin{aligned}&\int_0^1 m^2(\theta)d\theta - \frac{1}{1-p}\int_0^1 m^2(\theta)(1-m(\theta))d\theta \geq \frac{1}{1-p}\cdot\frac{1}{6}\cdot m(1)\cdot\min_{\theta\in[0,1]}\left(m'(\theta)\right)^2\\ =&\frac{1}{6}\cdot\frac{1}{1-p}\cdot m(1)\cdot\frac{(m(1)-m(0))^2}{(1-p)^2}\cdot\frac{m^2(1)(1-m(0))^2}{m^2(0)}\end{aligned} \tag{A.102}$$

where we used the same steps as in (A.94), (A.95), and (A.96).

From (A.95):

$$\frac{m(0)-m(1)}{m(0)m(1)} = \frac{M(\alpha+\beta)}{\mu}\cdot(1-p) = \frac{M\cdot p\cdot(\alpha+\beta)}{\mu}\cdot\frac{1-p}{p} = \frac{\overline{M}\cdot(\alpha+\beta)}{\mu}\cdot\frac{1-p}{p} = \frac{Y(1-p)}{p}, \tag{A.103}$$

rearrange to get

$$\frac{Y}{m(0)-m(1)} = \frac{1}{m(0)m(1)}\frac{p}{1-p} \tag{A.104}$$

Combining (A.104) and (A.102)

$$\begin{aligned}
&\frac{1}{m(0)-m(1)}\left(Y+\mu\cdot\frac{\partial Y}{\partial \mu}\right)\left(\int_0^1 m^2(\theta)d\theta - \frac{1}{1-p}\int_0^1 m^2(\theta)(1-m(\theta))d\theta\right)\\
\geq&\left(1+\frac{\mu}{Y}\cdot\frac{\partial Y}{\partial \mu}\right)\cdot\frac{Y}{m(0)-m(1)}\cdot\frac{1}{6}\frac{1}{1-p}m(1)\cdot\frac{(m(1)-m(0))^2}{(1-p)^2}\cdot\frac{m^2(1)(1-m(0))^2}{m^2(0)}\\
=&\frac{1}{6}\left(1+\frac{\mu}{Y}\cdot\frac{\partial Y}{\partial \mu}\right)\frac{(m(1)-m(0))^2}{(1-p)^2}\cdot\frac{pm^2(1)(1-m(0))^2}{(1-p)^2m(0)^3}\\
\geq&\frac{1}{6}\left(1+\frac{\mu}{Y}\cdot\frac{\partial Y}{\partial \mu}\right)\frac{(m(1)-m(0))^2}{(1-p)^2}\cdot\frac{m^3(1)(1-m(0))^3}{m(0)^3(1-m(1))^3} = \frac{1}{6}\left(1+\frac{\mu}{Y}\cdot\frac{\partial Y}{\partial \mu}\right)\frac{(m(1)-m(0))^2}{(1-p)^2}\cdot e^{-3Y}
\end{aligned} \tag{A.105}$$

(A.92), (A.97) and (A.105), using (A.89), imply that $\partial W/\partial \mu > 0$ for any M if the following sufficient condition is satisfied:

$$-\frac{1}{8}e^{Y} + \frac{1}{12}e^{-2Y} + \frac{1}{6}e^{-3Y}\left(1+\frac{\mu}{Y}\cdot\frac{\partial Y}{\partial \mu}\right) > 0 \tag{A.106}$$

This is true if $Y < 0.03$ and

$$1+\frac{\mu}{Y}\cdot\frac{\partial Y}{\partial \mu} > \frac{1}{3} \tag{A.107}$$

(A.107) is equivalent to:

$$\frac{\partial \frac{1}{\mu}}{\partial Y} > \frac{3}{2}\cdot\frac{1}{Y}\cdot\frac{1}{\mu} \tag{A.108}$$

Note that (11) can be rewritten as:

$$e^{\frac{1}{\mu}} = \frac{1}{Y}\left(e^{\frac{\beta}{\alpha+\beta}Y} - e^{-\frac{\alpha}{\alpha+\beta}Y}\right), \tag{A.109}$$

taking logs, this can be rewritten to express $\frac{1}{\mu}$ in terms of Y and Z:

$$\frac{1}{\mu} = \log\left(\frac{1}{Y}\left(e^{\frac{\beta}{\alpha+\beta}Y} - e^{-\frac{\alpha}{\alpha+\beta}Y}\right)\right) = Z + \frac{\beta-\alpha}{2(\alpha+\beta)}Y \tag{A.110}$$

where Z is defined by

$$Z = \log\left(\frac{e^{Y/2}-e^{-Y/2}}{Y}\right), \tag{A.111}$$

and $\lim_{Y\to 0} Z = 0$.

(A.108) holds if and only if:

$$\frac{dZ}{dY} - \frac{3}{2}\frac{Z}{Y} > \frac{1}{4} \cdot \frac{\beta - \alpha}{\alpha + \beta}. \tag{A.112}$$

Note that $\lim_{\mu\to\infty} Y = 0$ and taking a Taylor expansion of $\frac{dZ}{dY} - \frac{3}{2}\frac{Z}{Y}$ around $Y = 0$ yields

$$\lim_{\mu\to\infty} \frac{1}{Y} \cdot \left(\frac{dZ}{dY} - \frac{3}{2} \cdot \frac{Z}{Y}\right) = \frac{1}{48}. \tag{A.113}$$

Rearranging (A.114) we need to show that

$$4 \cdot \frac{1}{Y} \cdot \left(\frac{dZ}{dY} - \frac{3}{2} \cdot \frac{Z}{Y}\right) > \frac{1}{Y} \cdot \frac{\beta - \alpha}{\alpha + \beta}. \tag{A.114}$$

Using (A.113), for high enough μ, the right hand side can be bounded from below by e.g. $\frac{1}{24}$. Now for the fixed Y associated with this μ, one can find an *alpha* close enough to β such that the right hand side is smaller than $\frac{1}{24}$. For high enough M we have $M > \bar{M}$ (for the $\bar{M}$ associated with the chosen μ) thus we are in an interior equilibrium that (A.61) can indeed be used.

Thus we have shown that for high enough μ, $\frac{\partial W}{\partial \mu} > 0$ for all M above the threshold $\bar{M}$.

□