The Great Industry Gamble: Market Structure Dynamics with Moral Hazard*

Aron Toth†
University of Warwick

October 29, 2008

Abstract
I investigate the dynamic effect of moral hazard on market structure in a general framework. In my model, the evolution of market structure determines the severity of moral hazard and, in turn, moral hazard fuels market structure dynamics through a survival contest. In the absence of scale economies, I show that the presence of moral hazard results in a convergence towards market concentration, regardless of the intensity of competition. On the other hand, the dynamics leading to concentration reduces moral hazard even when prices do not increase with concentration (e.g. with Bertrand competition). As a result, counterintuitively, a firm in a more concentrated market can be less profitable. The equilibrium is a completely endogenous Markov process with a unique steady state distribution that follows a lognormal pattern. The model displays ongoing turbulence in the steady state and explains two well known empirical regularities in a novel way. Specifically, I show that moral hazard in and of itself is sufficient to produce correlated exit and entry rates and also shake-outs from time to time. The main policy implication is that market concentration can be effective against moral hazard and as such, welfare increasing. The model helps to explain the puzzling market transformation of industries such as banking, health care and audit.

1 Introduction
Consumers often cannot verify certain key product characteristics before consumption (experience and credence goods).1 Firms, therefore, have a natural incentive to exploit their informational advantage and cheat on consumers - this is a classical example of moral hazard. Interestingly, many markets where the unobserved product characteristics (quality) are crucial have a tendency towards concentration over time, despite the fact that scale economies are often not substantial. Moreover, there is a common perception that concentration has a positive effect on quality in these industries, even when the intensity of price competition does not ease with concentration. These tendencies are particularly prominent in previously regulated markets where competition of some sort has been introduced. For instance, in banking, health care and audit, concentration has been increasing steadily since market liberalization, while price competition appears to be fiercer than ever. Traditional arguments (e.g. scale economies), however, often fail to offer a credible explanation for these peculiar industry dynamics.

The goal of this paper is to show how these market characteristics arise endogenously in a simple infinitely repeated oligopoly model with moral hazard. The dynamic relationship between moral hazard and market structure is explored in the absence of scale economies. I argue that market evolution determines the severity of moral hazard while, in turn, moral hazard drives market structure through a survival contest. I show that the presence of this basic driving force does not depend on the nature of market competition.

---

*I am greatly indebted to my supervisors Amrita Dhillon, Claudio Mezzetti and Margaret Slade. I also thank Marcus Asplund, Daniel Sgroi and various seminar participants at the University of Warwick (2006), the Econometric Society Meeting (2006 Chennai), the Royal Economic Society Annual Conference (2007), the European Association for Research in Industrial Economics Annual Conference (2007) and the Bank of England (2008) for helpful comments and suggestions. Financial support from the Economics Department at the University of Warwick is gratefully acknowledged. All remaining errors are mine.

†Department of Economics, University of Warwick, Coventry CV4 7AL, UK; Email: a.toth@warwick.ac.uk.

1These key product characteristics are usually labelled as quality in the literature and hereafter I will follow this convention. I would like to stress that what I consider quality here is the set of key characteristics in a sense that they essentially determine the value of the product. Quality, therefore is different from "amenities", which are rather designed to give the appearance of quality. So for example, in a hospital, quality would be the effectiveness of clinical treatment (unobserved), whereas waiting times, complaint resolutions, etc (observed) would be less crucial characteristics.
In my model, a firm can invest in (unobserved) quality and this investment increases its probability of survival through the reduction of moral hazard. The contest is a result of strategic gambling on survival: firms try to outlive rivals so that they can be among the few to benefit from a future, lucrative, concentrated market structure. However, the presence of many rivals reduces the chances of winning this survival battle. Hence firms care little about the future and they invest less in quality, thereby exacerbating the moral hazard problem. This triggers a shake-out when the market is shared by many. The failure of many firms, on the other hand, raises the stakes for survival, since the market gets closer to favourable, concentrated states. Surviving firms invest more and hence alleviate the moral hazard problem even when prices do not increase with concentration (e.g. with Bertrand competition). Counterintuitively, this means that a firm in a more concentrated market can be less profitable along the equilibrium path. As the chances of survival increase, remaining firms fail with smaller and smaller probability and it becomes more and more difficult to outlast rivals. The convergence towards concentration slows down and the contest turns into a ruthless struggle.

This survival contest is present independently of the intensity of market competition. Lax competition induces more investment, reducing moral hazard and as such supports a less concentrated market in the long run. Intense market rivalry, on the other hand, speeds up the convergence towards concentration and results in fewer firms in equilibrium.

The paper also contributes to the literature on stochastic industry dynamics. I introduce stochastic entry into the model in a novel way so that the equilibrium yields a completely endogenous and "fully" stochastic Markov process. The steady state distribution is unique, follows a lognormal pattern and displays ongoing turbulence. The model explains two well known empirical regularities. Specifically, I show that moral hazard in and of itself is sufficient to produce correlated exit and entry rates and also shake-outs from time to time.

The paper is organized as follows. In the remainder of this introduction, I describe some stylized facts and the related literature. In Section 2, I discuss the main assumptions and outline the baseline model without entry. I introduce entry in Section 3 and discuss the steady state dynamics of the model. Finally, I conclude in Section 4. Appendix A contains an extension of the model, which explicitly incorporates consumers into the game discussed in the main body of the paper. All the lengthy proofs are relegated to Appendix B.

1.1 Stylized facts and empirical evidence

The industry dynamics described in this paper can be observed in markets which were previously regulated. In particular, once regulatory restrictions are lifted and competition is unleashed, the market goes through a rapid transformation. The initial phase is characterised by poor quality and heavy shake-out in the form of failures, mergers and acquisitions. Yet, as the market is driven towards concentration, failure rates decrease, while quality often improves substantially even when cut-throat market competition appears to remain. Conventional arguments for an industry transformation of this sort (e.g. scale economies) do not always seem to apply. Three industries of this kind have received particular attention in the last couple of decades: banking, health care and audit.

A. The Banking Industry

In the banking industry, I consider quality to be the risk that a bank chooses to take by the design of its investment portfolio. This risk (the probability of insolvency) is never directly observed by the depositors (moral hazard). However, once the bank’s financial position is in doubt, depositors withdraw their money and the bank often fails.2

From the 80’s the deregulation and liberalization process of the financial systems started to gather momentum worldwide. As a result, competition started to challenge banks which led to a failure, merger and acquisition wave of an unexpected magnitude almost all over the world. The most striking example is the US where the number of banks has decreased by 40% since 1984, but market concentration has also increased steadily in almost all developed and developing countries in the last couple of decades. (Rare exceptions are Finland and France.) This spectacular concentration process has long been a puzzle in the academic literature, since empirical research has more often than not failed to find significant economies of scale (Bikker, 2004). Furthermore, it is well documented in numerous empirical studies that, as a direct consequence of the liberalization process, banks started to engage in riskier activities (Keeley, 1999; Allen and Gale, 2000; Brewer III and Jackson III, 2006). However, as banking markets steadily tended towards a concentrated structure, less and less risk was observed to be taken (Maudo & Guevara, 2004; Beck et al, 2006).

2 One may ask: why would the depositor care about risk of insolvency if deposits are insured? First, deposit insurance is never designed to give full protection for standard moral hazard reasons. Second, it’s not only the deposit that a client can lose in the case of bankruptcy. With time, consumers build up valuable relationship with the bank, which gets lost if the bank goes bankrupt. In this particular industry, failure can also be thought of as an event independent of consumer’s attitude towards risk (quality); that is, in banking it doesn’t really matter if consumers care about risk or not, since banks can fail as a direct result of their investment decisions.
As a consequence, this perceived negative relationship between risk and market concentration (the "concentration-stability" paradigm) has widely been appreciated by competition authorities: despite the apparent lack of efficiency gains from mergers, antitrust policy has been particularly lenient in the financial industry in the last couple of decades (Boyd and Graham, 1996; Boyd and Nicolo, 2005).

B. The Health Care Industry

Moral hazard in health care is very apparent: a patient is almost never in the position to observe and verify the quality of clinical treatment she receives. In fact, measuring quality of health care is a notoriously difficult task even for researchers. Hence, like in banking, for decades a puzzling policy question in health care has been whether the total welfare effect of competition is positive or negative.

While competition undoubtedly reduces health care costs in the long run, its effect on quality is ambiguous. The controversies of quality measurement in health care are particularly severe since the quality of clinical treatment is very poorly observed. It is a difficult task to find hard empirical evidence on the effect of competition on quality. Not surprisingly, the empirical findings on the relationship between competition and quality is very mixed and largely dependent on the way quality is defined and measured (Wong, 2004; Pauly, 2004; Romano and Mutter, 2004). To my knowledge, there is only one article which avoids the problem of measuring quality explicitly and yet convincingly succeeds in giving some important insight into the relationship of market concentration and quality. Jin (2005) investigates health maintenance organizations’ (HMO) voluntary disclosure of quality. The market of HMOs is highly competitive. Since the early 1990s, these organizations have been widely criticized for the presumed low service quality, and the market observed many failures as well as mergers and acquisitions (Jin, 2005). As a response, an independent and non-profit agency, the National Committee of Quality Assurance (NCQA) started to accredit these institutions on a voluntary basis. By 1998 around half of the HMOs were accredited. Jin (2005) finds that the proportion of HMOs which chose to voluntarily disclose information on quality via accreditation is significantly higher in more concentrated markets. This certainly appears to be at odds with theory as well as common sense, unless less concentrated markets indeed provide lower quality.

Similarly to banking, recent empirical studies provide only minimal evidence of scale economies in managed care: such economies have been found to exist only at relatively low levels (Glied, 2000). The findings from these articles are in line with merger studies, because mergers have minor effects on health care costs (Glied, 2000). As a consequence of an ongoing consolidation in the HMO market, 79% of MSAs were highly concentrated (HHI $>3000$) by 2004 (AMA 2005).

C. The Audit Industry

The market for audit services is another typical example of credence goods. The certification of the accuracy of financial statements by a third party is based on a credence claim: the integrity (quality) of an audit is never directly observed. The independent auditor faces a dilemma. On the one hand, if it is more lenient about the accounting standards, more willing to bend the truth, more companies will want to subscribe to its audit services. On the other hand, the more false the picture the auditor agrees to certify (that is, the more corrupt the auditor), the more probable it is that the actual performance of its clients reveals the auditor’s misconduct (for example, one of the audited firms unexpectedly fails), and hence the higher the threat of litigation and reputational damages and, as a consequence, of failure. Audits are primarily produced for shareholders and potential investors who rely on the credibility of the audit.

\[^3\] Risk taken by financial institutions is of utmost importance because bank failures could easily lead to systemic banking crises, which usually have tremendous social costs. In the USA, for example, between 1984 and 1991 more than 1400 savings and loans company and 1300 banks failed, resulting in clean up costs of 3.2% of the GDP. In many cases the direct cost of systemic crises have been around 10-20% of the GDP, with occasional magnitude of 40-55% of GDP (Chile, Argentina), (Caprio et al, 1996).

\[^4\] In theory, in competitive markets informational asymmetries can be overcome by certain market mechanisms such as voluntary quality disclosure via independent quality verification agencies. A high quality producer has a natural incentive to disclose information on its quality through these agencies, distinguishing itself from other firms which produce inferior quality. For further details see the references in Jin (2005).

\[^5\] In their article, a systemic banking crisis is defined by emergency measures taken to assist the banking system, and/or non-performing assets that reached 10% of the total assets, and/or fiscal rescue operations that exceeded 2% of the GDP. It may also be worthwhile noting that in their sample 47 crises are included.

\[^6\] I should emphasize that the theoretical and empirical evidence on the negative relationship between risk and concentration is best described as mixed (e.g. Beck et al, 2006). This is largely due to the fact that financial risk is very complex and can be measured in numerous ways. However, the very fact that merger policy appears to have adopted the "concentration-stability" view suggest that this relationship may indeed be prevalent.

\[^7\] For broader (product) market definition of combined HMO and Preferred Provider Organization (PPO) market, the corresponding percentage would be 67%.
and ultimately, the credibility of the auditor. Once credibility is seriously questioned, the product (audit) is worthless
and the auditor quite possibly goes bankrupt. Essentially, this is the stylised story of Arthur Andersen’s failure in
2002. By the time of its indictment for misconduct, Arthur Andersen had already lost nearly all of its clients.

Despite fierce rivalry, we observe an extremely concentrated audit market today although this was not always the
case. In the USA, for instance, the audit industry came into existence as a result of the Securities Acts of 1933 and
1934. At the beginning, competition among auditors was severely impaired: advertising, competitive bidding and
soliciting clients were all strictly prohibited. As a result, hundreds of audit firms were operating undisturbed for three
decades. Beginning from the 70s, however, rules restricting auditors from advertising and competitive bidding were
loosened unleashing fierce competition. Ever since, the industry has been going through a remarkable concentration
process. By the 1980s, eight firms dominated the American audit market and by 1998, this was down to five as a result
of a series of mergers in 1980s and 1990s among the big 8. Then, the collapse of Arthur Andersen further reduced the
number of dominant firms to four. Whereas only 22% of the smaller companies (revenue less than $100 million) were
audited by the big 4, the large company market is very concentrated: 96% of the companies with revenues over $500
million were big 4 clients in 2006 (GAO-08-163). Unfortunately, unlike in the banking and health care markets, data
on audit costs are not publicly available, therefore, direct tests for economies of scale in the industry are simply hard
to come by. However, after a careful inspection of the predominantly labour intensive audit technology, it’s di¢ cult
to imagine why substantial economies of scale should exist in this industry that could credibly determine such a high
level of concentration.

Recent events such as the bankruptcy of Arthur Andersen raise similar policy issues to banking and health care:
is concentration good or bad, does a concentrated market promote quality or, on the contrary, does it actually harm
the integrity of auditing? (Bar-Yosef and Sarath, 2005.)

In this paper, I build a model which aims to rationalise the following stylized facts: i) Steady convergence towards
concentration in the absence of scale economies, ii) High failure rates in fragmented markets, and iii) Product quality
increasing with market concentration. The model shows that instead of technological reasons, the presence of moral
hazard can be the key driving force behind the puzzling market evolution observed in the banking, health care and
audit industries.

1.2 Related literature

In the context of unobserved quality, asymmetric information has been investigated extensively under static competitive
conditions. For instance, Allen (1984) presents a model of moral hazard where highly sophisticated consumers infer
quality from prices and an equilibrium exists where firms always produce high quality in Bertrand competition. Hörner
(2002) shows that high quality can be sustained in perfect competition under adverse selection and moral hazard.
More recently, Rob and Fishman (2005) argue that local monopolists provide higher quality as their size grows, when
information about quality flows among consumers by word of mouth. However, only Allen (1984) introduces strategic
interaction among firms and none of these articles addresses market structure dynamics explicitly.

Kranton (2003), Bar-Isaac (2005) and Dana and Fong (2008) consider models where competition and market
structure influence profits earned and, therefore, the incentive to invest in quality. Their main concern is how and
to what extent different competitive conditions can undermine the incentive to invest in quality. In these papers the
dynamics of market structure is not analysed since they focus on one direction of the relationship between market
structure and moral hazard: market structure affects moral hazard. These papers ignore the other important aspect of
the relationship: moral hazard can affect market structure and, in turn, the strategic behaviour of firms. In particular,
if consumers can prompt failure, then it seems reasonable to think that firms recognise that they are in a gambling
game for survival and hence they must ask: what do I gain if others fail? This is the driving force that the current
study focuses on.

In sum, my paper differs from previous studies because it explicitly models the full extent of the dynamic relation-
ship between market structure and investment in quality (moral hazard), while allowing for all of the possible forms
of strategic interactions among firms considered in the literature. The paper identifies a key driving force that has
been overlooked so far: a survival contest.

The paper is also related to the literature on industry dynamics and stochastic entry. Although stochastic entry
is rather the norm than the exception in the literature of industry dynamics, the entry process has almost never
been "genuinely" stochastic in these models. For instance, in Jovanovic (1982), Hopenhayn (1992), and Amir and
Lambson (2003) entry is only stochastic because the state evolves stochastically. In their model, entry in equilibrium
is a deterministic function of the state and hence affects incumbents’ actions deterministically. To my knowledge, only
Ericson and Pakes (1995) constructs a game where the entry process is truly stochastic in equilibrium; however, their model is not tractable analytically.

I depart from the literature on stochastic entry by developing a model where entry is completely stochastic. In particular, the number of entrants in a given period is a stochastic function of the state of the game and therefore entry affects the state stochastically. In the present model, the entry process in equilibrium is stochastic as a result of ex ante heterogeneous entrants. In particular, before entry entrants observe their own fixed entry cost which is random and private information. This generates a non-degenerate distribution of the number of entrants at each state, as opposed to a degenerate one as in many previous studies. The game therefore yields a Markov chain with a completely endogenous stochastic structure. This is a realistic feature of the model which differentiates it from most of the literature on industry dynamics.

2 The baseline model

Denote the number of firms in the market by \( n \in \{0, 1, 2, \ldots\} \). Each individual firm \( i = 1, \ldots, n \) produces a (homogenous or heterogenous) good, the quality of which can be of two types, high or low. Neither consumers nor firms observe quality before consumption; it is commonly observed after consumption. Consumers obtain zero utility from a low quality good and positive utility from a high quality good.\(^8\) Firms can invest in quality (or exert effort) in every period and this investment is private information; that is, there is moral hazard. The probability of producing high quality in a given period depends only on how much the firm invests in that period. There is no knowledge accumulation or increase in efficiency as a result of past investments which could potentially introduce scale economies. Firm \( i \) can invest in quality \( x_i \in [0, 1] \) at the cost of \( g(x_i) \) each period and produces high quality at the end of that period with probability \( x_i \). I assume that a firm that produces low quality fails and leaves the market.

Assumption 1. If a firm produces low quality, it fails.

Assumption 1 can be justified by incorporating consumers explicitly into the game and showing that if they ever experience low quality from a firm, consumers rationally stop buying from that firm in equilibrium. The details of this extended game can be found in Appendix A. This is very standard consumer behaviour in the unobserved quality literature (e.g. Allen 1984, Hörner 2002).\(^9\)

As a consequence of Assumption 1, a firm fails with probability \( 1 - x_i \) at the end of the period and hence in any given period firms that have produced only good quality are in the market. Therefore, including firm \( i \), the number of firms (the market structure) at the beginning of the next period will be \( n - k \) with probability \( x_i \Pr(k|x_{-i}) \), where \( x_{-i} = |x_j|_{j \neq i} \) and \( k = 0, \ldots, n - 1 \) is the number of firms other than firm \( i \) failing at the end of the period. \( \Pr(k|x_{-i}) \) is the probability mass function of the convolution of \( n - 1 \) Bernoulli distributions with "success" probabilities \( x_{-i} \).

At the beginning of each period, firms engage in price or quantity competition. For simplicity, I do not explicitly model this market game. Denote the symmetric equilibrium profit for a firm from the market game by \( \pi(n) \), when \( n \) firms are in the market.

Assumption 2. \( \pi(n) \leq \pi(n - 1) \), for \( n > 2 \) and \( \pi(2) < \pi(1) \) with \( \pi(\cdot) \geq 0 \).

The per period gross profit is weakly decreasing in the number of firms and non-negative for \( n > 1 \); the per period gross monopoly profit is strictly positive. Importantly, Assumption 2 does not impose any restriction on the nature of the market competition. It is general enough to account for homogeneous-product as well as differentiated Bertrand or Cournot competition. This specification essentially captures all the cases typically considered in the literature.

Having realised per period gross profits, firms invest in quality. Hence, the per period net profit for firm \( i \) is \( \pi(n) - g(x_i) \). I make the following assumptions on the cost of investment.

Assumption 3A. \( g(x_i), g'(x_i), g''(x_i) > 0 \) for \( x_i > 0 \) and \( g(0) = g'(0) = 0, g(1) < \infty. \)

\(^8\)The good, therefore, can be heterogenous only along observable characteristics.

\(^9\)There is one crucial difference between the consumers’ strategy in the models of Allen (1984) and Hörner (2002) and in the current game. In their models, the equilibrium configuration is such that either consumers are aware of industry specifics so that they can calculate the price that ensures high quality (Allen 1984) or they simply hold the "right beliefs" about the lowest price necessary for high quality (Hörner 2002). In other words, price has a signalling role in these models. This requires high level of consumer knowledge and sophistication. In my modelling approach I follow Rob and Fishman (2005) and construct an equilibrium in which consumers’ conduct does not link quality and price.
The cost of investment is a strictly increasing convex function. The assumption of bounded investment costs is not essential, the main results are unaffected if I relax it. Its only purpose is to ensure that a corner solution ("high quality equilibrium") may exist so that I can compare my results to the findings of previous studies in a straightforward manner. As we will see, without entry, this can also be a way to ensure that the market structure converges to a steady state. Were the cost to tend to infinity, \( \lim_{x_i \to 1} g(x_i) = \infty \), then not even a monopoly would find it optimal to invest the maximum amount \( (x_i = 1) \) with finite gross profits. As a consequence, even a monopoly would fail sooner or later and the market would vanish in the long run. With entry, however, we do not need the assumption of bounded costs in order for the market structure to converge to a steady state, as we will see in Section 3.

In most markets \( g(x_i) \) has a natural interpretation. In the banking industry \( g(x_i) \) can be understood as an opportunity cost: the lower risk (higher quality) the bank takes, the more potential profit it sacrifices. The market game in banking, therefore, could be thought of as competition for deposits. Suppose there are investment opportunities for banks in every period ranked by their risk. Having acquired deposits at the equilibrium interest rate in the stage game, the bank can invest in the riskiest project available and can ensure a profit of \( \pi(n) \) in that period. However, investing in less risky assets, the bank willingly decreases its potential profit in order to increase the probability of survival into the next period and acquires a per period net payoff \( \pi(n) - g(x_i) \). This net profit can even be negative if the bank offers loans at a lower rate than it acquires deposits.\(^{10}\)

For expository purposes, I first present the model without entry.

**Assumption 4A. There is no entry.**

Assumption 1, 2, 3A and 4A are maintained throughout this section. In Section 3 I will relax Assumption 4A and allow for entry.

Time is discrete and infinite. The common discount factor is \( \beta \in (0, 1] \). I consider the quality game in the framework of an infinitely repeated game.\(^{11}\) The summary of timing in this (reduced) game is as follows:

1. Given \( n \), gross profits, \( \pi(n) \), are realised.
2. Firms simultaneously choose investment in quality, \( x_i, i = 1, ..., n \).
3. Failures and exits occur.

The strategies are assumed to be Markov and I focus on symmetric Markov Perfect Equilibrium (MPE) throughout the analysis. That is, strategies depend only on payoff relevant information. The payoff relevant information can be conveniently "condensed" into a state variable, which is the number of firms \( n \).

With \( n \) firms in the market, the dynamic programme of firm \( i \) is simply

\[
v(n;x_{-i}) = \pi(n) + \max_{0 \leq x_i \leq 1} \left\{-g(x_i) + \beta x_i \sum_{k=0}^{n-1} V(n-k) \Pr(k|x_{-i}) \right\}
\]

(1)

where \( n \) is the current state and \( V(n-k) \) is the value of the firm in a symmetric equilibrium in state \( n-k \). After differentiating and imposing symmetry (i.e. \( x_i = x \) for all \( i \)), the first order conditions (FOC) of this programme are

\[
\beta \sum_{k=0}^{n-1} V(n-k) \left( \begin{array}{c} n-1 \\ k \end{array} \right) (1-x)^k x^{n-k-1} - g'(x) + \lambda - \mu = 0, \quad \lambda x = 0, \quad \mu(x-1) = 0
\]

(2)

where \( \lambda, \mu \) are Lagrange multipliers, and I used the fact that if \( x_j = x \) for all \( j \neq i \), then \( \Pr(k|x_{-i}) = \left( \begin{array}{c} n-1 \\ k \end{array} \right) (1-x)^k x^{n-k-1} \).

The simple nature of the first order conditions is the result of two facts. First, the convolution of Bernoulli trials is simply the binomial distribution, due to symmetry. Second, the Bernoulli probabilities are linear. The latter may

\(^{10}\) Bounded costs in Assumption 3, therefore, are not unreasonable. In the banking industry, for instance, if the risk of the available investment opportunities is bounded, there is certainly a maximum amount that a bank can sacrifice in order to secure survival unless one is willing to entertain the idea of negative credit rates.

\(^{11}\) Since the results on moral hazard in the reputation literature are often sensitive to the length of the horizon, perhaps I should emphasise that the findings of the current section readily go through with finite horizon as well; in fact, two periods would be sufficient.
appear overly simplistic. However, given the non-linear costs, it is not. In other words, it is not more restrictive than assuming non-linear survival probabilities but constant marginal costs, a commonly applied modelling structure (e.g. Rob and Fisherman 2007, Ericson and Pakes 1995). In the current model the marginal benefit is constant (given rivals’ strategy) and the marginal cost is not, whilst in the other models it’s exactly the opposite. In both cases the trade off between a unit cost and a unit benefit is non-linear and the first order conditions yield non-linear strategies in the state variable.

Denote the solution to the first order conditions (2) by \( x(n) \). Given Assumption 2 and 3, it is easy to see that \( x(n) \neq 0 \). There are two possible solutions: one corner and one interior. It is \( x(n) = 1 \) as long as

\[
\pi(n) \geq g'(1) \frac{1-\beta}{\beta} + g(1)
\]  

Alternatively, \( 0 < x(n) < 1 \) where \( x(n) \) solves (2) with \( \lambda = \mu = 0 \).

Before we proceed, let’s have a closer look at the corner solution, which represents the high quality equilibrium, the traditional focus of the literature on reputation. If (3) holds, then firms will invest the maximum amount in quality, there will be no failures at the end of the period and the market structure will not change. The necessary condition for this to happen is that firms generate enough revenue from the market game to cover the cost of producing the highest quality. If they don’t discount the future, that is, if they are infinitely patient \( (\beta = 1) \), then non-negative net profits \( (\pi(n) - g(1) \geq 0) \) are sufficient for the market structure being at a standstill. However, the more impatient the firms, the higher the net profit they must make, the higher the price premium the consumers must pay for the highest quality. In the limit, as \( \beta \to 0 \), in the static game, no firm would ever produce high quality, regardless of profits. In other words, as the impatience of firms grows, the bigger is the incentive to cheat on consumers, since firms care less about survival. Inequality (3) is consistent with many of the previous findings in the literature on reputation (e.g. Allen 1984).

I first establish the existence of a symmetric pure strategy equilibrium.

**Proposition 1** In the game without entry, there exists a symmetric Markov Perfect Equilibrium in pure strategies.

**Proof.** See Appendix B. ■

The next Proposition follows from the first order conditions directly:

**Proposition 2** In a symmetric equilibrium of the game without entry, in the long run the number of firms, \( n^* \), is no bigger than \( \tilde{n} \) \( (n^* \leq \tilde{n}) \), where

\[
\tilde{n} = \begin{cases} 
0 & \text{if } \pi(1) < g'(1) \frac{1-\beta}{\beta} + g(1) \\
\max\{n : \pi(n) \geq g'(1) \frac{1-\beta}{\beta} + g(1)\} & \text{otherwise}
\end{cases}
\]

It is immediate from (3) that the higher the cost of investment in quality, the smaller the number of firms viable in the long run. Also, perhaps paradoxically, more intense competition leads to more concentrated markets since depressed gross profits discourage investment in quality. Greater patience, however, as captured by the discount factor, supports the existence of more firms in a long run equilibrium.

The following Proposition states the main result. It shows that equilibrium investment rises with concentration even if gross profits \( \pi(n) \) don’t. As a consequence, paradoxically a firm with only a few rivals may realise less profit along the equilibrium path than a firm with many competitors.

**Proposition 3** In a symmetric equilibrium of the game without entry, more firms in the market implies strictly less investment in quality, that is, \( x(n+1) < x(n) < 1 \), unless \( n + 1 \leq \tilde{n} \), in which case \( x(n+1) = x(n) = 1 \).
Proposition 3 states that, unless gross profits are sufficiently high, each firm's investment will always be strictly less when there are \( n + 1 \) firms than with \( n \) firms. However, if gross profits in a market with \( n + 1 \) firms are high enough (inequality (3) holds), then firms in this market (and any other market with fewer firms) will invest the maximum amount. In this "high quality equilibrium" no firm fails and the market never gets more concentrated. In other words, if the market structure is such that \( n > \bar{n} \), then firms will fail with positive probability \( (n < 1) \) and the market will steadily converge towards \( \bar{n} \). Therefore, without entry \( n^* \leq \bar{n} \) are the only viable market structures in the long run.

Importantly, note that it is sufficient for the result in Proposition 3 to hold that gross profits are only weakly decreasing with the number of firms. That is, even if \( \pi(n + 1) = \pi(n) \) for \( n > 1 \), the strict inequality between equilibrium investments holds if \( n > \bar{n} \). The intuition for this is that the model dynamics generates a survival contest, independently of current profits. To see this, consider homogeneous-product Bertrand competition, where \( \pi(n + 1) = \pi(n) = 0 \) for all \( n > 1 \). In this case, investment in quality (moral hazard) is driven by future expectations of market structure, namely by the possibility that a firm may become a monopolist. However, the further away the current market structure from the monopoly state (the bigger \( n \)), the less likely a firm is to be the lucky one to become a monopolist, since many firms compete for the "prize". Therefore, the expected value of the continuation payoff is smaller, leading to less investment. As the market becomes concentrated, the stakes rise since the probability that the firm will be the one which survives into the monopoly stage increases. As a consequence, firms care more about future and invest more. Thus, as the market is driven towards concentration, the survival contest becomes more intense, firms grow more desperate to survive and as a result will fail with smaller probability. In general a firm in a more concentrated market will invest more, even if current gross profits are no more than that of less concentrated markets \( (\pi(n + 1) = \pi(n)) \), resulting, paradoxically, in firms in more concentrated markets being less profitable.

Competition has two conflicting effects on the investment in quality. On the one hand, it depresses profits and undermines the incentive to invest. On the other hand, it allows consumers to switch and hence to inflict maximum punishment on firms. Maximum punishment in and of itself, however, fails to give firms sufficient incentives to invest in quality if the market is too fragmented. Interestingly, as market concentration increases, the consequences of maximum punishment become more severe even if profits do not increase with concentration. This leads to firms investing more and, consequently, less moral hazard.

In the banking industry, for instance, numerous studies have found a negative relationship between concentration and deposit rates (see Brewer III and Jackson III, 2006, and references therein). Carrying on interpreting the market game as a competition for deposits, this observation is equivalent to a decreasing \( \pi(n) \). However, Brewer III and Jackson III (2006) find empirical evidence that although banks in concentrated markets pay lower deposit rates, they also take less risk. In other words, deposit rates are lower in more concentrated markets and these depressed deposit rates allow banks to invest in less risky assets (i.e. assets with smaller returns). In the present framework, this "sacrifice of potential profits" is \( g(x) \), the opportunity cost of investing in less risky assets. Proposition 3, therefore, confirms the rationale behind the recent empirical findings of Brewer III and Jackson III (2006). In sum, while market power may increase gross profits, these profits are (partly) used to finance investment in quality (i.e. less risky assets).

The simple structure of the baseline model allows me to fully characterise the equilibria by comparative static analysis. The following Proposition summarizes these results.

**Proposition 4** (Comparative statics) In a symmetric equilibrium without entry, the higher the per period gross profit \( \pi(n) \), the monopoly profit \( \pi(1) \), or the discount factor \( \beta \), the higher amount firms invest in quality, the lower the moral hazard. That is,

\[
\frac{dx(n)}{d\pi(n)} > 0, \quad \frac{dx(n)}{d\pi(1)} > 0, \quad \frac{dx(n)}{d\beta} > 0
\]

**Proof.** See Appendix B.
the more oligopolists will invest in quality, reducing the chance that the market actually ever becomes a monopoly. Regardless of how low the value of gross profit $\pi(n)$ is, as the value of monopoly tends to infinity, the oligopolists’ investment will approach the maximum investment in quality and the market turns into a monopoly with probability approaching zero. In other words, as the monopoly profit tends to infinity ($\pi(1) \to \infty$), any market structure can be sustained almost surely, regardless of the intensity of market competition. That is, firms engage in fierce competition to survive and as a consequence, none will fail with probability approaching one, preserving current market structure almost surely. Note that if the monopoly profit is sufficiently large, then firms are willing to endure severe losses in the oligopoly phase, even if there is little hope that the monopoly stage will ever be reached. The following corollary, which is a direct consequence of Proposition 4, summarises this observation.

**Corollary 1** In a symmetric equilibrium of the game without entry, as the value of the monopoly firm approaches infinity, the optimal investment in quality tends to the maximum value. That is, $\lim_{\pi(1)\to\infty} x(n) = 1$. As a consequence, for any $n$, the expected number of failures tends to zero: $\lim_{\pi(1)\to\infty} n(1 - x(n)) = 0$.

The survival contest interpretation of the game, the "contest for a prize" has some features in common with patent races. In both models, the higher the prize, the more firms invest. However, there is at least one crucial difference between the current model and a standard patent race. The survival contest identified in the current model is not a race, in fact it’s quite the opposite. In a simple symmetric patent race more investment in the probability of discovery would imply a quicker end of the game since the probability that one of the firms wins is increasing with investments. In the current model, however, higher level of investment increases the probability of firm survivals and as such reduces the probability that any of the firms succeeds in winning the prize (e.g. becomes a monopolist in the Bertrand case). In other words, the probability that the "game ends" and the prize will be won is decreasing with the value of the prize, which is quite the opposite to what we observe in patent race models. This implies that the survival contest is a much more ruthless struggle, since firms have to endure investment costs for a much longer period.

### 2.1 Discussion of the baseline model

Market structure dynamics plays the key role in the determination of investment in quality, which, in turn, fuels the evolution of market dynamics. Competition and unforgiving consumers can depress prices to the level where investment is insufficient to prevent failure ($x < 1$). This mechanism triggers failures setting the market in motion. Market dynamics are driven by firms’ expectations of future market structure, regardless of price dynamics. Firms want to make it into some profitable market structure, but the further they are from this state, the less probable that they will ever reach it. As a result, they have little to gain from survival and consequently they care less about the future and invest little. The closer the "winners" get to a profitable stage, however, the more desperate they become to survive, because their stake in survival is much higher, as it is more likely they will be among the few who actually arrive at the lucrative phase. In other words, some win, others lose, and failing firms indirectly raise the stakes for the survivors. Surviving firms, therefore, end up in a fierce contest to stay on and invest more in quality. Paradoxically, this reduces the chance that the desired stage will ever be reached. As a result, fragmented markets display heavy shake-out and fewer and fewer failures occur as the market gets concentrated.

The main results of this section are therefore quite different from previous findings in the literature of reputation and competition: for example in Allen (1984) and Hörner (2002), high quality production is sustainable in competitive markets when quality is imperfectly observed. The main driving force behind their result is that consumers know the firms’ technology, and hence they can deduce the unobserved quality produced by a firm from the observable price that it charges. The maximum punishment strategy of unforgiving consumers has supreme bite in their models due to the signalling role of prices. I construct an equilibrium in which prices do not have this signalling role and consumers form their expectations based only on firms’ past quality performance. This consumer behaviour is consistent with unsophisticated consumers who don’t know the firms’ technology, but it is also consistent with knowledgeable consumers in a pooling equilibrium when firms’ prices convey no information.

I show that consumers’ maximum punishment strategy (Assumption 1) in and of itself fails to give the right incentives for firms to produce high quality in general. Market evolution, however, will lead to a high quality equilibrium in the long run. The very nature of strategic gambling for survival (i.e. firms optimally choosing $x < 1$) ensures steady convergence towards high quality.

The model in this section is also related to that of Rob and Fishman (2005). In their model, firms are local monopolists and therefore there is no strategic interaction among them. Reputation about firms spreads among
consumers by word of mouth, securing a continuous customer base expansion for surviving firms. This expanding
demand boosts period monopoly profits, which results in an investment pattern increasing with firm age. Therefore
in their article, as in the model of this section, high quality is a product of some sort of evolution.

3 Entry

Entry can be introduced into the model in several ways. A natural starting point is free and non-stochastic entry. This straightforward entry process, however, leads to a much simplified industry dynamics and in the limit to a
degenerate steady state distribution of the market structure. In order to analyse a more complex and interesting
industry dynamics, I introduce instead stochastic entry into the model.

To simplify the exposition, in what follows I will focus only on interior solutions in which \( x_i \in (0, 1) \). For this
purpose, Assumption 3A is replaced by the following modified assumption.

**Assumption 3B.** \( g(x_i), g'(x_i), g''(x_i) > 0 \) for \( x_i > 0 \) and \( g(0) = g'(0) = 0 \), \( \lim_{x \to 0} g(x) = \infty. \)

In this section, in each of an infinite number of periods a stage game is played which is itself composed by a
sequence of rounds. To be precise, in the stage game played in each period there are \( N \) entry rounds followed by
one production round. In each entry round only one entrant is allowed to enter; in other words, in each period there
are \( N \) potential entrants who enter sequentially. In the production round, each firm simultaneously decides on the
investment in quality. Time "stops" within each period, there is discounting only across periods.

Let the fixed cost of entry be \( F = F + \varepsilon \), where \( F > 0 \) and \( \varepsilon \) is iid with \( E(\varepsilon) = 0 \) and CDF \( \rho(\cdot) \) with support \([ -F, \infty ]\). Each entrant knows her own fixed cost of entry before entering. An entrant’s fixed cost is private information, but all entrants know the distribution of fixed costs.

Thus, entrants are heterogenous. This property complicates the space of strategic interactions since, in addition
to the game between incumbents and entrants, it allows for a non-trivial game among the entrants themselves. In
particular, entrants are not symmetric, so, despite the sequential structure, they are unable to foresee the entry process
with certainty within a period. As a result, upon entry an entrant must contemplate the fact that further entry in
that period can make her entry unprofitable ex post. In previous studies, the symmetry of entrants (before entering)
and the sequential structure often led to entry being a deterministic function of the state of the game in equilibrium.
In this model, on the contrary, the \( n \) firms in the market in round \( l \) (that is, the incumbents and the entrants who
have entered before round \( l + 1 \) will expect an entrant to enter in the next round with probability \( \rho_{l+1,n} \), which is to
be determined endogenously. In sum, heterogeneity before entry induces a non-degenerate distribution of the number
of entrants at each state.

Assumption 4A is, therefore, modified as follows:

**Assumption 4B.** In each period there are \( N \) potential entrants, which enter sequentially at the beginning of the
period. The \( l \)-th entrant enters with probability \( \rho_{l,n} \) when there are \( n \) firms in the market at the end of the previous
(production or entry) round.

Assumption 1, 2, 3B, 4B are maintained throughout this section. The summary of timing within a period is as follows:

1. The first entrant observes the number of firms in the market after failures occurred in the last period. If there
were \( n \) firms producing at the end of last period and \( k \) failed, then the first entrant enters with probability \( \rho_{l,n-k} \).
Subsequent entrants observe the number of firms in the last entry round and enter sequentially. The \( l \)-th entrant enters
with probability \( \rho_{l,n} \) if there are \( n \) firms in the \( l-1 \)-th entry round.
2. After the final entry round, the production round starts. Given the number \( n \) of active firms (entrants plus
incumbents), gross profits, \( \pi(n) \), are realised.
3. Firms simultaneously choose investment in quality, \( x_i, i = 1, \ldots, n. \)
4. Failures and exits occur.

The strategies are assumed to be Markov and I focus on symmetric Markov Perfect Equilibrium (MPE). Let the
equilibrium value function at the end of the \( l \)-th entry round be \( W_l(n) \). (That is, \( n \) includes the \( l \)-th entrant if it has entered the market.) Then,

\[
W_l(n) = W_{l+1}(n + 1)\rho_{l+1,n} + W_{l+1}(n)(1 - \rho_{l+1,n})
\]

(4)
Note that at the end of the last entry round, since there is no further entry during that period, \( W_N(n) = V(n) \), where \( V(n) \) is the equilibrium value function of the firms at the beginning of the production round. The incumbents’ value function in symmetric equilibrium then is defined by

\[
V^e(n) = \pi(n) - g(x^e(n)) + \beta x^e(n) \sum_{k=0}^{n-1} [W_1(n-k+1)\rho_{1,n-k} + W_1(n-k)(1-\rho_{1,n-k})] \left( \frac{n-1}{k} \right) (1-x^e(n))^k (x^e(n))^{n-1-k} \quad (5)
\]

where \( x^e(n) \) is the optimal level of investment, i.e. the solution to the dynamic program (5) when there are \( n \) firms in the market. The \( e \) superscript stands for entry. First, I establish the strong monotonicity of the value functions.

**Proposition 5** \( V^e(n) \) and \( W_1(n) \) are strictly decreasing in \( n \).

**Proof.** See Appendix B. ■

Now I can prove the strong monotonicity of the equilibrium investments with stochastic entry. The next proposition is just the analogue of Proposition 3, the main result of the previous section.

**Proposition 6** Firms’ investment level is a strictly increasing function of market concentration, that is \( x^e(n + 1) < x^e(n) \).

**Proof.** The proof is very similar to that of Proposition 3. ■

It is important to note that the results in Proposition 5 and 6 do not depend on the distribution of entry probabilities \( \rho(\cdot) \). Therefore, the strong monotonicity of the value functions and the investment schedule is completely invariant to the stochastic structure of the entry process.

I am now in a position to fully characterise the entry process. The \( l \)th entrant enters if \( W_l(n + 1) - F > 0 \), that is, it enters with probability \( \rho_{l,n} = \rho(W_l(n + 1) - F) \) when there are \( n \) firms in the market at the end of the previous (production or entry) round. It is not difficult to see that \( 0 < W_l(n) < \infty \), and therefore \( 0 < \rho_{l,n} < 1 \). Consequently, entry is a continuous (stochastic) threat in all states.

Since \( W_l(n) < W_{l+1}(n) \), from (4) and Proposition 5 it is also simple to see that \( \rho_{l,n} < \rho_{l+1,n} \) and \( \rho_{l,n} > \rho_{l,n+1} \). In other words, within a period, \( \text{ceteris paribus} \), entrants enter with greater probability if closer to the production round and when the market is more concentrated. Both results are very intuitive. An entrant closer to the production round enters because with higher probability because the uncertainty regarding the remaining potential entries in that period is smaller. Also, more concentrated markets naturally attract more entry. These may seem to be countervailing effects. However, from (4) and Proposition 5 it is immediate that \( \rho_{l,n} > \rho_{l+1,n+1} \). Therefore, if entry occurs in a given entry round, it will unambiguously lower the probability of further entry. These observations can be summarised in the following proposition.

**Proposition 7** \( \rho_{l,n} < \rho_{l+1,n}, \quad \rho_{l,n} > \rho_{l,n+1}, \quad \rho_{l,n} > \rho_{l+1,n+1}, \quad 0 < \rho_{l,n} < 1, \quad \lim_{n \to \infty} \rho_{l,n} = 0. \)

**Proof.** The inequalities follow from Proposition 5 and the limit from Proposition 5 and Assumption 2. ■

The model describes a very intuitive industry dynamics. When there are many firms in the market, the equilibrium value of the firm is low. Thus, firms will invest very little and hence will fail with considerable probability. At the same time, the low value of the incumbent firms will attract little entry. As a result, the market is driven towards concentration. However, as the market gets concentrated, firms invest more and fail with smaller probability. In addition, since the value of being in the market increases, entry will occur with a greater probability preventing the market to become too concentrated. As a result, concentration can never fully resolve the moral hazard problem, since while concentration eases the problem of moral hazard, entry prevents the market to become too concentrated. Consequently, the model exhibits continuous turbulence in the form of failures and entry even in the limit.

Now I turn to the analysis of the long run behaviour of the model.
3.1 Steady State dynamics

The idea behind the analysis of the steady state dynamics is to use the equilibrium properties of the discrete time game above to derive transition probabilities when each period of the game is infinitely small. In essence, I am going to transform the discrete time Markov transition kernel into a continuous time analogue and will prove that the game produces a birth and death process.

Let’s introduce explicitly the length of a period $dt$. Note that the results in the previous section were obtained in the special case when $dt = 1$. Since I am going to relate the dynamics to birth and death processes, it seems convenient to redefine the probability of failure as $1 - x^e(n) = \theta_n$. Therefore, in a period a firm will fail with probability $\theta_n dt$ and will survive with probability $1 - \theta_n dt$. Also, the $l$th entrant enters with probability $\rho_{l,n} dt$ when there are $n$ firms in the market. There are two facts that I am going to use extensively in what follows:

\[
(dt)^i = o(dt) \quad \text{for} \quad i \geq 2 \quad \text{and} \quad (1 - o(dt))^k = \sum_{i=0}^{k} \binom{k}{i} (-o(dt))^i = 1 - kadt + o(dt)
\]

where $o(dt)$ is a term vanishing (in probability) at the rate of at least second order. Formally, \(\lim_{dt \to 0} o(dt)/dt = 0\).

Denote the random variable of the number of failures as $\omega(t)$. Using (6) and (7) we can show that the transition probabilities in the production round are

\[
\Pr(\omega(t + dt) = k|n) = \binom{n}{k} (\theta_n dt)^k (1 - \theta_n dt)^{n-k} = \binom{n}{k} (\theta_n dt)^k (1 - (n-k)\theta_n dt + o(dt))
\]

\[
= \begin{cases} 
1 - n\theta_n dt + o(dt) & \text{if } k = 0 \\
\theta_n dt + o(dt) & \text{if } k = 1 \\
o(dt) & \text{if } k > 1 
\end{cases}
\]

Similarly, letting the random variable $\psi(t)$ denote the number of entering firms, the transition probabilities in the entry process are

\[
\Pr(\psi(t + dt) = j|n - k) = \sum_{i=1}^{N} (1 - \rho_{i,n-k} dt) = 1 - \sum_{i} \rho_{i,n-k} dt + o(dt)
\]

\[
\Pr(\psi(t + dt) = j|n - k) = \prod_{i=1}^{N} (1 - \rho_{i,n-k} dt) \prod_{i > j} (1 - \rho_{i,n-k+1} dt) = \sum_{i} \rho_{i,n-k} dt + o(dt) \quad \text{if } j = 1 \\
o(dt) & \text{if } j > 1 
\]

Let the random variable $Z(t)$ denote the number of firms in the market at time $t$. Then the change in $Z(t)$ is simply the convolution of $\omega(t)$ and $\psi(t)$. In particular, $dZ(t)/dt = \psi(t) - \omega(t)$. The transition probabilities in a period when there are $n$ firms in the market are as follows

\[
\Pr(Z(t + dt) = n + 1|Z(t) = n) = [1 - n\theta_n dt + o(dt)] \left[ \sum_{i} \rho_{i,n} dt + o(dt) \right] + o(dt)
\]

\[
\Pr(Z(t + dt) = n - 1|Z(t) = n) = [n\theta_n dt + o(dt)] \left[ 1 - \sum_{i} \rho_{i,n} dt + o(dt) \right] + o(dt)
\]

\[
\Pr(Z(t + dt) = n|Z(t) = n) = [1 - n\theta_n dt + o(dt)] \left[ 1 - \sum_{i} \rho_{i,n} dt + o(dt) \right] + o(dt)
\]

\[
= 1 - n\theta_n dt - \sum_{i} \rho_{i,n} dt + o(dt)
\]

\[
\Pr(Z(t + dt) = n + m|Z(t) = n) = o(dt) \quad \text{if } m \neq -1, 0, 1
\]

with initial conditions $\Pr(Z(t + dt) = 0|Z(t) = 0) = 1 - \sum_{i} \rho_{i,0} dt + o(dt)$ and $\Pr(Z(t + dt) = 1|Z(t) = 0) = \sum_{i} \rho_{i,0} dt + o(dt)$.

Denote the intensity parameters of the Markov process as $\mu_n = n\theta_n$ and $\lambda_n = \sum_{i} \rho_{i,n}$. The parameter $\mu_n$ can be interpreted as the average number of failures ("deaths") in a period $dt$ with $n$ firms, while the parameter $\lambda_n$ can be interpreted as the average number of entries ("births") in a period with $n$ firms. Although the entry process is
Proposition 8 There exists a unique stationary steady state distribution \( P \), with probability mass function \( P_n = \frac{\lambda_{n-1}}{\mu_n} P_{n-1} \) for \( n \geq 1 \) with \( P_0 = \left(1 + \sum_{n=1}^{\infty} \prod_{i=0}^{n-1} \frac{\lambda_i}{\mu_{i+1}}\right)^{-1}, \) and \( E(n) < \infty \).

Proof. See Appendix B. ■

We can think of \( \mu_n P_n \) as the flow of exits from the state with \( n \) firms, and \( \lambda_{n-1} P_{n-1} \) as the flow of new entries when there are \( n - 1 \) firms. In the steady state distribution these two flows are equal. Equivalently, the likelihood ratio \( P_n / P_{n-1} \) is equal to the ratio of new entries from state \( n - 1 \) over the exits from state \( n \).

Proposition 9 The most frequently visited state in steady state, that is, the mode of the steady state distribution is \( n^* = \max\{n : \lambda_{n-1} > \mu_n\} \).

Proof. By Proposition 6 \( \mu_n \) is strictly increasing in \( n \), while by Proposition 7 \( \lambda_n \) is strictly decreasing in \( n \). Furthermore, \( \lim_{n \to -\infty} \lambda_n = 0 \) and \( \lim_{n \to -\infty} \mu_n = \infty \). Since \( P_n / P_{n-1} = \lambda_{n-1} / \mu_n \), the probability mass function is increasing while \( \lambda_{n-1} \geq \mu_n \) and decreasing afterwards. As a result, the probability mass function reaches its maximum for the highest \( n \) for which \( \lambda_{n-1} > \mu_n \).

The average number of exits from a state is an increasing function of the number of firms, while the average number of entries is a decreasing function of the number of firms. This implies that the steady state likelihood ratio \( P_n / P_{n-1} \) decreases with the number of firms \( n \). It follows that the most likely state is the state with the highest number of firms \( n \) for which \( P_n / P_{n-1} > 1 \). This makes intuitive sense, since such a state \( n^* \) is the only state with the following property: the (average) number of firms entering from \( n^* - 1 \) is larger than the number of firm exiting from \( n^* \), while the number of firms entering from \( n^* \) is smaller than the number of firms exiting from \( n^* + 1 \).

3.2 Discussion of the steady state

Some properties of the distribution can be readily deduced. For instance, as a consequence of the strong monotonicity of the entry and failure parameters, the shape of the distribution can be of two types. If there exists no \( n \) such that \( \lambda_{n-1} > \mu_n \), then the probability mass function is strictly decreasing and will look like an exponential distribution. In this case, the mode of the distribution, that is, the most frequently visited state is state zero (no firm in the market).

If, on the contrary, \( \lambda_{n-1} > \mu_n \) for some \( n \), then the probability mass function is strictly increasing for small \( n \), and strictly decreasing afterwards. In other words, it will look like an \( F \) distribution. These two observations could be nicely described by the lognormal distribution: a high variance case will yield a shape similar to the exponential distribution, whereas a lower variance would result in a shape similar to the \( F \) distribution as depicted in Figure 1.

Using the recursive nature of the steady state probabilities \( P_n / P_{n-1} = \lambda_{n-1} / \mu_n \) or \( P_n - P_{n-1} = (\lambda_{n-1} - \mu_n) / \mu_n \), we can say something about the variance too. The distribution is tight (the variance is small) if the ratio \( \lambda_{n-1} / \mu_n \) is steep around \( n^* \). In other words, if there is a big jump in the willingness to enter or in the failure rates around the most frequently visited state, then the variance will most probably be low. This may occur if the variance of the distribution of the fixed cost is relatively small or the cost function \( g(\cdot) \) is relatively convex around \( n^* \).

While the symmetric framework prohibits the model to address many interesting empirical phenomena (e.g. variance of firm growth rates, size distribution, etc.), it is still able to rationalise two important stylized facts, namely the presence of shake-out and the strong correlation between entry and exit rates. Both are well known empirical regularities.

It has long been observed that most industries are characterised by shake-outs. In particular, in the period immediately after the birth of an industry there are typically only few firms. This is followed by a period of steady increase in the number of incumbents. After a while, a wave of failures, mergers and acquisitions occurs. So far the theoretical literature has explained shake-outs through the introduction of some innovation process (Sutton 2006). In
the current model, a shake-out is simply the result of the presence of moral hazard, while the technology is constant over time. The model exhibits shake-outs from time to time: the market can end up with many firms with positive probability and then firms fail in great numbers (high $\mu_n$).

Entry and exit rates tend to be highly correlated in many industries. So far the literature has focused on the following possible driving forces: demand fluctuations, displacement of existing technologies, displacement of existing products, fluctuations in relative efficiency (Sutton 2006). The present model shows that correlation between entry and exit rates can also be a direct consequence of moral hazard. This is immediate from Proposition 9, after observing that the most frequently visited state, the mode of the steady state distribution, is the market structure where entry and failure rates are approximately equal.

4 Conclusion

In this paper, I have developed a model in which strategic rivalry leads firms to underinvest in imperfectly observed quality (there is moral hazard) when the market is too fragmented. Consumers follow a strategy which inflicts maximum punishment, abandoning the firm should it deliver low quality. This is the standard reputational mechanism through which asymmetric information problems are supposed to get rectified. In my model, however, this very mechanism in and of itself is not sufficient in general to give firms the necessary incentives to produce high quality. Rivalry triggers a survival contest and failing firms drive a fragmented market towards a concentrated structure. I show that the moral hazard problem is effectively eased along this dynamics even when prices do not increase with concentration. In particular, since market concentration increases the prospects for firms to reach profitable states, the industry dynamics leads to higher investments in quality, reducing moral hazard. This results in a convergence towards concentration in fragmented markets, although at an ever decreasing speed. Entry, of course, can potentially block this monotonic market evolution. Introducing stochastic entry, the model displays shake-outs from time to time and correlated exit and entry rates. These two empirical regularities are often rationalised using technological arguments, whereas in my model they are simply the results of moral hazard.

The model explains why, even in the absence of scale economies, we may observe massive failure rates and steady convergence towards concentration in fragmented markets of some experience and credence goods. It also explains why more concentrated markets produce higher qualities in general, regardless of the intensity of market competition and entry. This study, therefore, contributes to our understanding of the peculiar market transformation and dynamics in industries such as banking, health care and audit. A notable policy implication is that in some markets concentration
may be necessary to alleviate moral hazard. As a consequence, promoting competition in markets where imperfectly observed product characteristics are of utmost importance may well be counterproductive.
Appendix A The baseline model with consumers: an extension

In this appendix, I discuss a straightforward way to include consumers into the game analysed in the body of the paper. The main objective is to show that consumers’ maximum punishment strategy is rational and can emerge as an equilibrium outcome of an extended game, similarly to Allen (1984), Hörner (2002) and Rob and Fishman (2005). Assumption 1 and 2 are satisfied in the equilibrium of the game.

As in Salop (1979), there is a continuum of consumers, whose measure is normalised to one, located uniformly along a circle. Each consumer buys one product each period. The product can be of two qualities, high or low. Quality is not observed before purchase but after consumption. The value of this unobserved quality to each consumer is identical, $u > 0$ if quality is high and zero if quality is low. Products are horizontally differentiated. Denote the distance between a consumer and firm $i$ by $s_i \in (0, 1)$.

The *ex post* net benefit for a buyer to consume the good from firm $i$ at time $t$ and price $p_t^i \in \mathbb{R}_+$ is $U_t^i = u - p_t^i - cs_t^i$ if quality is high, and $U_t^i = -p_t^i - cs_t^i$ if quality is low. The parameter $c \geq 0$ measures the travel cost, and is an index of the degree of product differentiation. If $c = 0$, then goods are homogeneous, whereas if $c > 0$ products are differentiated. This utility function implies identical consumer attitudes towards (unobserved) quality, since every consumer gets negative utility from consuming a low quality product at a positive price, regardless of her distance from firms. Were the quality known before purchase, no consumer would buy low quality products.

Assumption 1 and 2 are satisfied in the equilibrium of the game. In this appendix, I discuss a straightforward way to include consumers into the game analysed in the body of the paper. The main objective is to show that consumers’ maximum punishment strategy is rational and can emerge as an equilibrium outcome of an extended game, similarly to Allen (1984), Hörner (2002) and Rob and Fishman (2005).

In this appendix, I discuss a straightforward way to include consumers into the game analysed in the body of the paper. The main objective is to show that consumers’ maximum punishment strategy is rational and can emerge as an equilibrium outcome of an extended game, similarly to Allen (1984), Hörner (2002) and Rob and Fishman (2005). Assumption 1 and 2 are satisfied in the equilibrium of the game.

In each period, decisions and actions take place in the following order:

1. Firms decide to quit or stay. Incumbents relocate symmetrically.
2. Firms choose prices simultaneously.
3. Consumers choose firms on the basis of last period quality, travel distance and prices.
4. Firms choose investment in quality simultaneously.
5. Consumers and firms observe current period quality.

Markov Strategies

I will look at Markov strategies that only depend on last period qualities. Define the (last period) history of qualities as: $H^t = x_{tN'}^{-1} \{ u^{t-1}, 0 \}$. A consumer chooses which firm to buy from, as a function of the history of qualities, her distance from firms, and current prices. Her strategy is described by the mapping $\mu : H^t \times [x_{iN'}(0, 1)] \times [x_{iN'} \mathbb{R}_+] \rightarrow N'$. Having observed last period qualities, firms decide to quit or stay, then incumbents choose prices. After having observed its own gross profits from the price game, a firm sets the level of investment in quality. Thus, a firm $i$’s strategy consists of three mappings: $\tau_i : H^t \rightarrow \{ \text{Quit}, \text{Stay} \}$, $p_t : H^t \rightarrow \mathbb{R}_+$ and $x_t : H^t \times \mathbb{R}_+ \rightarrow [0, 1]$.

Symmetric Markov Perfect Equilibrium

The following strategies and beliefs constitute a Symmetric Subgame Perfect Nash Equilibrium in Markov strategies:

Equilibrium Strategy of Consumers:

- Do not buy from a firm, which has produced bad quality.$^{13}$

---

$^{12}$It is possible to extend the model to allow firms to choose their location. Economides (1984) shows that with free location choice, there exists a symmetric equilibrium in locations and prices.

$^{13}$This is Assumption 1 in the main text.
- Buy from a firm which minimizes total costs, $p_i t + cs_t$.

**Equilibrium Beliefs of Consumers:**

- If a firm has produced bad quality, it will always produce bad quality with probability one.
- In period $t$, firms produce good quality with probability $q_t^i = q_t = x_t^i$, if they haven’t produced bad quality before.

**Equilibrium Strategy of Firms:**

- Quit if you produced bad quality last period.
- If you produced good quality last period, choose $p_t^i$ such that $p_t^i = r(p_{t-1}^i, \mu_t^i) \in \arg \max_{p_t^i} \pi(p_t^i; p_{t-1}^i, \mu_t^i)$.
- If you produced good quality last period, choose $x_t^i$ which maximizes the dynamic programme of (1), where $\pi(n) \equiv \hat{\pi}(r(p_t^i, \mu_t^i), p_t^i, \mu_t^i)$.

It is simple to see why $p_t^i, x_t^i$ are optimal for firm $i$, and also why firm $i$ is no better off staying in the market if it has produced bad quality last period, given consumers’ and other firms’ strategy. Looking at (1) and Proposition 4, observe that there is no dynamic link between price and investment in quality: given other firms’ and consumers’ strategy (i.e. that they do not trade off price and quality), posting a price other than $p_t^i$ would just reduce gross profit, which would in turn induce lower investment in quality, lowering the overall value of the firm.

On the other hand, given the firms’ strategies, consumers can clearly do no better by switching to another firm because, by the symmetric nature of the equilibrium, consumers would be no better off if $c = 0$, and would be worse off if $c > 0$. Also, consumer can do no better by deviating to staying loyal rather than switching from the firm which has produced bad quality. Finally, note that a Nash equilibrium in Markov strategies is necessarily subgame perfect.

**Appendix B Proofs**

The following Lemmas will be useful in what follows.

**Lemma B.1** $xg'(x) > g(x)$ for $x > 0$.

**Proof.** Recall that $g''(\cdot) > 0$ by Assumption 3. Hence, the following holds for any two distinct points $v, x$ in the domain: $g(v) > g(x) + g'(x)(v - x)$. Letting $v = 0$ and recalling that $g(0) = 0$ gives the inequality result. ■

**Lemma B.2** In a symmetric equilibrium, the value function is strictly decreasing: $V(n) < V(n - 1)$.

**Proof.** The proof is by induction. First, I show that $V(2) < V(1)$. Then supposing $V(n) < ... < V(1)$, I’ll show that $V(n + 1) < V(n)$ which proves the claim. To prove $V(2) < V(1)$ observe that

$$V(2) - V(1) = \pi(2) - g(x(2)) + \beta x(2)[V(2)x(2) + V(1)(1 - x(2))] - V(1)$$

$$\leq \pi(2) - g(x(2)) + \beta x(2)V(2) - \beta x(2)^2[V(2) - V(1)] - V(1)$$

$$= \frac{\pi(2) - g(x(2)) + \beta x(2)V(1) - V(1)}{1 - \beta x(2)^2} < \frac{\pi(1) - g(x(2)) + \beta x(2)V(1) - V(1)}{1 - \beta x(2)^2} \leq 0$$

where the first inequality follows from Assumption 2 and the second from the fact that $\pi(1) - g(x(2)) + \beta x(2)V(1) \leq V(1)$ since $x(2)$ is the maximiser of the duopoly’s dynamic programme, rather than the monopoly’s.
Next, suppose $V(n) < \ldots < V(1)$. Then, I’ll show that $V(n + 1) < V(n)$. Let $x \equiv x(n + 1)$ be the value which maximises the firm’s dynamic program in a symmetric equilibrium when there are $n + 1$ firms in the market. Then,

$$V(n + 1) = \pi(n + 1) - g(x) + \beta \sum_{k=0}^{n} V(n + 1 - k) \binom{n}{k} (1 - x)^k x^{n+1-k}$$

Therefore,

$$V(n + 1) \leq \pi(n) - g(x) + \beta \sum_{k=0}^{n-1} V(n - k) \binom{n-1}{k} (1 - x)^k x^{n-k} + \beta \sum_{k=0}^{n} V(n + 1 - k) \binom{n}{k} (1 - x)^k x^{n+1-k}$$

$$- \beta \sum_{k=0}^{n-1} V(n - k) \binom{n-1}{k} (1 - x)^k x^{n-k}$$

where the inequality follows from Assumption 2 again. Now, observe that

$$\pi(n) - g(x) + \beta \sum_{k=0}^{n-1} V(n - k) \binom{n-1}{k} (1 - x)^k x^{n-k} \leq V(n)$$

since $x$ maximises the value function when there are $n + 1$ firms, rather than $n$. Thus, from (B.1) we get

$$V(n + 1) - V(n) \leq \beta \sum_{i=1}^{n-1} V(n - i) \binom{n}{i+1} (1 - x)^i x^{n-i}$$

$$- \beta \sum_{i=0}^{n-1} V(n - i) \binom{n-1}{i} (1 - x)^i x^{n-i}$$

where I switched indexes setting $k = i + 1$ in the first term and $k = i$ in the second. I use the convention $\binom{n}{l} = 0$ for $l < 0$ and $l > n$ throughout the derivations. Next, using Pascal’s identity

$$\binom{n}{i+1} - \binom{n-1}{i} = \binom{n-1}{i+1}$$

the inequality can be rearranged as

$$V(n + 1) - V(n) \leq \beta \sum_{i=1}^{n-1} V(n - i) \left[ \binom{n-1}{i+1} + \binom{n-1}{i} \right] (1 - x)^i x^{n-i}$$

$$- \beta \sum_{i=0}^{n-1} V(n - i) \binom{n-1}{i} (1 - x)^i x^{n-i}$$

$$= \beta \sum_{i=1}^{n-1} V(n - i) \binom{n-1}{i+1} (1 - x)^i x^{n-i}$$

$$- \beta \sum_{i=0}^{n-1} V(n - i) \binom{n-1}{i} (1 - x)^i x^{n+1-i}$$

The last element in the first sum after equality can be dropped since it’s zero. Then, switching the index back again, set $i = k$ in the first sum and $i = k$ in the second. This yields:

$$V(n + 1) - V(n) \leq \beta \sum_{k=0}^{n-1} V(n + 1 - k) \binom{n-1}{k} (1 - x)^k x^{n+1-k}$$

$$- \beta \sum_{k=0}^{n-1} V(n - k) \binom{n-1}{k} (1 - x)^k x^{n+1-k}$$

$$= \beta \sum_{k=0}^{n-1} \left[ V(n + 1 - k) - V(n - k) \right] \binom{n-1}{k} (1 - x)^k x^{n+1-k}$$
Rearranging yields
\[ V(n + 1) - V(n) \leq \frac{\beta \sum_{k=1}^{n-1} [V(n + 1 - k) - V(n - k)](n-1)k x^{n+1-k}}{1 - \beta x^{n+1}} < 0 \]
where the last inequality holds by the induction hypothesis. ■

**Corollary B.1** Suppose \( x(n) < 1 \). Then, in a symmetric equilibrium, the value of the firm when there are \( n \) firms in the market is \( V(n) = \pi(n) - g(x(n)) + x(n)g'(x(n)) > \pi(n) \). Furthermore, \( 0 < V(n) < \infty \).

**Proof.** The first part of the claim is the result of the first order condition (2) being substituted into the value function; the inequality follows from Lemma B.1. The second part of the Corollary, \( 0 < V(n) \) holds because \( 0 \leq \pi(n) \) by Assumption 2 and the strictly increasing function \(-g(x) + xg'(x) > 0\) for all \( x > 0 \) by Lemma B.1. \( V(n) < \infty \) since \( V(n) < V(1) \) for all \( n > 1 \) by Lemma B.2 and \( V(1) \leq \pi(1)/(1 - \beta) < \infty \). ■

**PROOF OF PROPOSITION 1**

By inspection of (1), it is immediate that Blackwell’s sufficiency conditions, namely monotonicity and discounting, hold. Therefore, the value function is unique. Also, the second order condition is simply \(-g''(x) < 0\), which holds by Assumption 3. This implies that the reaction functions are unique. It remains to show that these reaction functions define a symmetric equilibrium of the game in pure strategies. This follows from the the symmetric reaction functions being continuous and downward sloping. Let’s write the first order condition to the programme (1) in the following way
\[ f_i(x_i; x_j, x_l, n) = \beta x_j \sum_{k=0}^{n-2} V(n-k)c_k + \beta(1-x_j) \sum_{k=0}^{n-2} V(n-1-k)c_k - g'(x_i) = 0 \]
where \( x_i = [x_h]_{h \neq i,j} \) and \( c_k = \Pr(k|x_i) \) is the probability mass function of the convolution of \( n - 2 \) Bernoulli trials and \( \sum_{k=0}^{n-2} c_k = 1 \). Continuity is obvious. Furthermore, the reaction functions are downward sloping because, by the Implicit Function Theorem,
\[ \frac{\partial x_i}{\partial x_j} = -\frac{\partial f_i(x_i; x_j, x_l, n)}{\partial x_j} / \frac{\partial f_i(x_i; x_j, x_l, n)}{\partial x_i} = \frac{\beta \sum_{k=0}^{n-2} [V(n-k) - V(n-1-k)]c_k}{g''(x_i)} < 0 \]
where the inequality follows from Lemma B.2. Note that the reactions functions are linear (\( \partial x_i/\partial^2 x_j = 0 \)), hence the symmetric equilibrium is unique if \( |\partial x_i/\partial x_j| \neq 1 \).

**PROOF OF PROPOSITION 3**

The proof consists of two parts. In the first part I show that the left hand side of the first order condition decreases with \( x \). In the second, I prove that it also decreases stochastically with \( n \). As a result, higher \( n \) must imply lower \( x \) in equilibrium.

Let’s define the left hand side of the FOC (2) as follows:
\[ f(x; n, \beta) = \beta \sum_{k=0}^{n-1} V(n-k)\left(\frac{n-1}{k}\right)(1-x)^k x^{n-k-1} - g'(x) \]  \hspace{1cm} (B.3)

First, I show that \( \partial f(x; n, \beta)/\partial x < 0 \). To see this, differentiate (B.3)
\[ \frac{\partial f(x; n, \beta)}{\partial x} = \beta \frac{\partial}{\partial x} \sum_{k=0}^{n-1} V(n-k)\left(\frac{n-1}{k}\right)(1-x)^k x^{n-k-1-k} - g''(x) \]
The last term is negative by Assumption 3. The first term is proportional to
\[
\frac{\partial}{\partial x} \sum_{k=0}^{n-1} V(n-k) \binom{n-1}{k} (1-x)^{k} x^{n-1-k} = \sum_{k=0}^{n-1} V(n-k) \binom{n-1}{k} (n-1-k)(1-x)^k x^{n-2-k} \\
- \sum_{k=0}^{n-1} V(n-k) \binom{n-1}{k} k(1-x)^{k-1} x^{n-1-k}
\]

Dropping the \((n-1)\)th element in the first sum and the first element in the second sum, which are both zero, we obtain

\[
\frac{\partial}{\partial x} \sum_{k=0}^{n-1} V(n-k) \binom{n-1}{k} (1-x)^{k} x^{n-1-k} = (n-1) \sum_{i=0}^{n-2} V(n-i) \frac{(n-2)!}{i!(n-2-i)!} (1-x)^i x^{n-2-i} \\
- (n-1) \sum_{i=1}^{n-2} V(n-1-i) \frac{(n-2)!}{i!(n-2-i)!} (1-x)^i x^{n-2-i}
\]

Substitute \(i = k\) in the first sum and \(i = k-1\) in the second sum. Then

\[
\frac{\partial}{\partial x} \sum_{k=0}^{n-1} V(n-k) \binom{n-1}{k} (1-x)^{k} x^{n-1-k} = (n-1) \sum_{i=0}^{n-2} [V(n-i) - V(n-1-i)] \binom{n-2}{i} (1-x)^i x^{n-2-i} < 0
\]

because from Lemma B.2 \(V(n-i) < V(n-1-i)\) for all \(i = 0, \ldots, n-2\). Therefore, we’ve established that

\[
\frac{\partial f(x; n, \beta)}{\partial x} < 0 \quad \text{(B.4)}
\]

Next, I’ll show that \(f(\cdot)\) is (stochastically) decreasing in \(n\). In particular, for any \(x \in (0, 1)\), \(f(x; n+1, \beta) < f(x; n, \beta)\). For this, it is sufficient to show that

\[
\sum_{k=0}^{n} V(n+1-k) \binom{n}{k} (1-x)^k x^{n-k} \quad < \quad \sum_{i=0}^{n-1} V(n-i) \binom{n-1}{i} (1-x)^i x^{n-1-i}
\]

\[
= \sum_{k=1}^{n} V(n+1-k) \binom{n-1}{k-1} (1-x)^{k-1} x^{n-k} \quad \text{(B.5)}
\]

where the equality follows from setting \(i = k-1\). In other words, the expected value of \(V(n+1) < \ldots < V(1)\) with CDF \(B(v; n, 1-x)\) is smaller than the expected value of \(V(n) < \ldots < V(1)\) with CDF \(B(v; n-1, 1-x)\), where:

\[
B(v; n, 1-x) = \sum_{k=0}^{v} b(k; n, 1-x) = \sum_{k=0}^{v} \binom{n}{k} (1-x)^k x^{n-k}
\]

\[
B(v; n-1, 1-x) = \sum_{k=1}^{v} b(k; n-1, 1-x) = \sum_{k=1}^{v} \binom{n-1}{k-1} (1-x)^{k-1} x^{n-k}
\]

Now, since \(V(\cdot)\) is strictly increasing in \(k\) (Lemma B.2), by the First Order Stochastic Dominance Theorem (FOSD), (B.5) holds if and only if \(B(v; n-1, 1-x)\) with support \(v = 1, \ldots, n\) (first order) stochastically dominates \(B(v; n, 1-x)\) with support \(v = 0, \ldots, n\); that is,
\[
B(v; n, 1 - x) \geq B(v; n - 1, 1 - x)
\]  \hspace{1cm} (B.6)

for \( v = 0, \ldots, n \) and with strict inequality for some \( v \). Inequality (B.6) is shown to hold by induction. It can be easily seen that

\[
B(0; n, 1 - x) - B(0; n - 1, 1 - x) = x^n
\]

\[
B(1; n, 1 - x) - B(1; n - 1, 1 - x) = (n - 1)(1 - x)x^{n-1}
\]

Therefore, the inducional hypothesis is

\[
B(v; n, 1 - x) - B(v; n - 1, 1 - x) = \left(\frac{n - 1}{v}\right)(1 - x)^v x^{n-v}
\]  \hspace{1cm} (B.7)

Observe that

\[
B(v + 1; n, 1 - x) = B(v; n, 1 - x) + b(v + 1; n, 1 - x)
\]

Therefore,

\[
B(v + 1; n, 1 - x) - B(v + 1; n - 1, 1 - x) = B(v; n, 1 - x) - B(v; n - 1, 1 - x)
\]

\[
+ b(v + 1; n, 1 - x) - b(v + 1; n - 1, 1 - x)
\]

Thus, using the inducional hypothesis (B.7), equation above can be written as

\[
B(v + 1; n, 1 - x) - B(v + 1; n - 1, 1 - x) = \left(\frac{n - 1}{v}\right)(1 - x)^v x^{n-v} + \left(\frac{n}{v + 1}\right)(1 - x)^{v+1} x^{n-v-1}
\]

where in the last line I used Pascal’s identity (B.2). Therefore, the inducional hypothesis is proven, and we can conclude that for all \( v \)

\[
B(v; n, 1 - x) - B(v; n - 1, 1 - x) = \left(\frac{n - 1}{v}\right)(1 - x)^v x^{n-v} \geq 0
\]

The inequality is strict for all \( v < n \). Therefore, by the theorem of FOSD, inequality (B.5) holds, and we can conclude that since \( f(x; n, \beta) \) is decreasing in both \( x \) and \( n \). The first order condition (B.3), therefore, implies that in a symmetric equilibrium the more firms are in the market, the less they invest; in other words, \( x(n + 1) < x(n) \) as required.

**PROOF OF PROPOSITION 4**

Recall the first order condition (B.3)

\[
f(x; n, \beta) = \beta \sum_{k=0}^{n-1} V(n - k) \left(\frac{n - 1}{k}\right)(1 - x)^k x^{n-k-1} - g'(x) = 0
\]

Then,
\[
\frac{\partial f(x; n, \beta)}{\partial \pi(n)} = \beta x^{n-1} \frac{\partial V(n)}{\partial \pi(n)} = \frac{\beta x^{n-1}}{1 - \beta x^n} > 0,
\]
\[
\frac{\partial f(x; n, \beta)}{\partial \pi(1)} = \beta \sum_{k=0}^{n-1} \frac{\partial V(n-k)}{\partial \pi(1)} \left( \frac{n-1}{k} \right) (1-x)^k x^{n-k} > 0
\]

The second inequality follows from a simple induction argument. Clearly,
\[
\frac{\partial V(1)}{\partial \pi(1)} = \frac{1}{1 - \beta x_1} > 0
\]

Now suppose \(\partial V(1)/\partial \pi(1), \ldots, \partial V(n-1)/\partial \pi(1) > 0\). Then,
\[
\frac{\partial V(n)}{\partial \pi(1)} = \frac{\beta}{1 - \beta x^n} \sum_{k=1}^{n-1} \frac{\partial V(n-k)}{\partial \pi(1)} \left( \frac{n-1}{k} \right) (1-x)^k x^{n-k} > 0
\]

where the inequality holds by the induction hypothesis. So
\[
\sum_{k=0}^{n-1} \frac{\partial V(n-k)}{\partial \pi(1)} \left( \frac{n-1}{k} \right) (1-x)^k x^{n-k} > 0
\]

A very similar argument shows that
\[
\frac{\partial f(x; n, \beta)}{\partial \beta} > 0,
\]

Then, using the Implicit Function Theorem, (B.4) and the inequalities derived above, we have
\[
\frac{dx(n)}{d\pi(n)} = -\frac{\partial f(x; n, \beta)}{\partial \pi(n)} \cdot \frac{\partial f(x; n, \beta)}{dx} > 0, \quad \frac{dx(n)}{d\pi(1)} = -\frac{\partial f(x; n, \beta)}{\partial \pi(1)} \cdot \frac{\partial f(x; n, \beta)}{dx} > 0,
\]
\[
\frac{dx(n)}{d\beta} = -\frac{\partial f(x; n, \beta)}{\partial \beta} \cdot \frac{\partial f(x; n, \beta)}{dx} > 0
\]

as required.

**PROOF OF PROPOSITION 5**

The proof is as follows. First, I consider the finite version of the dynamic programme and by a straightforward inducitonal argument I establish that \(V^*_t(n-1) - V^*_t(n) \geq 0\) and \(W_{t,t}(n-1) - W_{t,t}(n) \geq 0\) for all \(t, l\) and \(n\), where \(V^*_t(\cdot)\) and \(W_{t,t}(\cdot)\) are value functions when there are \(t\) periods remaining. Then, by standard dynamic programming argument \(\lim_{t \to \infty} V^*_t(n) = V^*(n)\) so it is proven that \(V^*(n-1) - V^*(n) \geq 0\). Similarly \(W_{t}(n-1) - W_{t}(n) \geq 0\). Lastly, I will show that these inequalities must be strict.

In what follows I drop the superscript and the argument of \(x^*(n)\) in order to ease notation so \(x = x^*(n)\). By Assumption 2, it is trivial to see that \(V^*_0(n-1) - V^*_0(n) \geq 0\) for all \(n\) since \(V^*_0(n) = \pi(n)\). Now noting that \(V^*_0(n) = W_{X,0}(n)\), from (4) a straightforward backwards induction shows that \(W_{t,0}(n-1) - W_{t,0}(n) \geq 0\) for all \(l\) and \(n\). So suppose \(V^*_t-1(n-1) - V^*_t-1(n) \geq 0\) and \(W_{t,t-1}(n-1) - W_{t,t-1}(n) \geq 0\) for all \(l\) and \(n\). Then, from (5) and Assumption 2 again
\[
V^*_t(n) \leq \pi(n-1) - g(x) + \beta x \sum_{k=0}^{n-1} [W_{1,t-1}(n-k+1)\rho_{1,n-k} + W_{1,t-1}(n-k)(1-\rho_{1,n-k})] \left( \frac{n-1}{k} \right) (1-x)^k x^{n-1-k} \tag{B.8}
\]

Now observe that if \(x\) is the maximiser of \(V^*_t(n)\), then
Therefore, (B.8) can be rearranged easily as

\[ [V^e_t(n) - V^e_t(n-1)]/\beta x \leq \sum_{k=0}^{n-2} [W_{1,t-1}(n-k)\rho_{1,n-k} + W_{1,t-1}(n-k-1)(1-\rho_{1,n-k-1})] \left( \frac{n-2}{k} \right) (1-x)^k x^{n-2-k} \]

where after the last equality sign in the first sum I substituted \( k = i + 1 \) and in the second \( k = i \). Using Pascal’s identity

\[ \binom{n-1}{i+1} - \binom{n-2}{i} = \binom{n-2}{i+1} \]

and noting that by convention \( \binom{n}{i} = 0 \) for \( i < 0 \) and \( i > n \), the inequality can be rearranged again as

\[ [V^e_t(n) - V^e_t(n-1)]/\beta x \leq \sum_{i=0}^{n-2} [W_{1,t-1}(n-i)\rho_{1,n-i} + W_{1,t-1}(n-i-1)(1-\rho_{1,n-i-1})] \left( \frac{n-2}{i+1} \right) (1-x)^i x^{n-2-i} \]

Switching the indices back again, this leads us to

\[ V^e_t(n) - V^e_t(n-1) \leq \beta \sum_{k=0}^{n-2} ([W_{1,t-1}(n-k+1)\rho_{1,n-k} + W_{1,t-1}(n-k)(1-\rho_{1,n-k})] - [W_{1,t-1}(n-k)\rho_{1,n-k-1} + W_{1,t-1}(n-k-1)(1-\rho_{1,n-k-1})]) \left( \frac{n-2}{k} \right) (1-x)^k x^{n-k} \]

where the last inequality follows from the induction hypothesis: since \( W_{1,t-1}(n-k+1) \leq W_{1,t-1}(n-k) \leq W_{1,t-1}(n-k-1) \) for all \( 0 \leq k \leq n-2 \), all the elements of the sum on the right hand side of (B.9) are non-positive. Noting that \( V^e_t(n) = W_{X,t}(n) \), from (4) a straightforward application of backwards induction shows that \( W_{1,t}(n-1) - W_{1,t}(n) \geq 0 \). Therefore, for all \( t, l \) and \( n, V^e_t(n-1) - V^e_t(n) \geq 0 \) and \( W_{1,t}(n-1) - W_{1,t}(n) \geq 0 \). Taking
the limit \( \lim_{t \to \infty} V^t(n) = V^c(n) \), it follows that \( V^c(n-1) - V^c(n) \geq 0 \) and similarly, \( W_l(n-1) - W_l(n) \geq 0 \) for all \( n \) and \( l \). Weak monotonicity, therefore is established. In what follows, I show that these two weak inequalities are actually strict and hence the strong monotonicity of \( V^c(n) \) and \( W_l(n) \).

Observe that inequality (B.9) holds at the limit too, that is, we can consider it without the time subscripts. Furthermore, from Assumption 2 for \( n = 2 \) inequality (B.8) is actually strict, therefore, (B.9) can be written as

\[
V^c(2) - V^c(1) < \beta x^2 [W_l(1) \rho_{1,2} + W_l(2)(1 - \rho_{1,2})] - [W_l(2) \rho_{1,1} + W_l(1)(1 - \rho_{1,1})] \leq 0
\]

The second inequality follows from what we have established earlier, that is \( W_l(n) - W_l(n) \geq 0 \) for all \( l \). But then \( W_{N-1}(1) = V^c(2) \rho_{1,1} + V^c(1)(1 - \rho_{1,1}) > W_{N-1}(2) = V^c(3) \rho_{N,2} + V^c(2)(1 - \rho_{N,2}) \) since \( W_l(n) = V^c(n) \). Noting again \( W_l(2) - W_l(3) \geq 0 \) for all \( l \), this implies through backwards induction that \( W_l(1) > W_l(2) \). As a consequence, there is at least one negative element on the right hand side of (B.9), which just means the second inequality in (B.9) is strict. Therefore, \( V^c(n) < V^c(n - 1) \) for all \( n \). Following the logic of the argument above it immediately follows that \( W_l(n) < W_l(n - 1) \) for all \( l \) and \( n \).

**Proof of Proposition 8**

The process has a transition matrix

\[
G = \begin{pmatrix}
1 - \lambda_0 & \lambda_0 & 0 & 0 & \cdots \\
\mu_1 & 1 - \lambda_1 - \mu_1 & \lambda_1 & 0 & \cdots \\
0 & \mu_2 & 1 - \lambda_2 - \mu_2 & \lambda_2 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots 
\end{pmatrix}
\]

Stationarity implies that the average flow into a state is equal to the average flow out of that state. The stationary distribution, therefore, satisfies \( P'G = P \) where \( P' = [P_0, P_1, \ldots] \). Thus

\[
(1 - \lambda_0)P_0 + \mu_1P_1 = P_0
\]

\[
\lambda_{n-1}P_{n-1} + (1 - \lambda_n - \mu_n)P_n + \mu_{n+1}P_{n+1} = P_n \text{ for } n \geq 1
\]

and solving this system of equations recursively yields

\[
P_n = \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} P_0 \text{ for } n > 0
\]

For existence the probability mass function must sum up to one, that is

\[
\sum_{n=0}^{\infty} P_n = P_0 + \sum_{n=1}^{\infty} P_n = P_0 + P_0 \sum_{n=1}^{\infty} \prod_{i=0}^{n-1} \frac{\lambda_i}{\mu_{i+1}} = 1
\]

Therefore, for the above to hold it is necessary and sufficient to show that

\[
\sum_{n=1}^{\infty} \prod_{i=0}^{n-1} \frac{\lambda_i}{\mu_{i+1}} < \infty
\]

From Proposition 7 \( \lim_{n \to \infty} \lambda_n = 0 \) and from Proposition 6 \( \lim_{n \to \infty} \mu_n = \infty \). Applying, for instance, the ratio test it is immediate that the sequence converges (very rapidly, in fact). Then,

\[
P_0 = \left( 1 + \sum_{n=1}^{\infty} \prod_{i=0}^{n-1} \frac{\lambda_i}{\mu_{i+1}} \right)^{-1}
\]

Uniqueness follows from the facts that the transition probability matrix is standard and the chain is irreducible. (Grimmett and Stirzaker 2001, Theorem 6.9.21) The chain is standard (although not uniform) because \( \lim_{dt \to 0} \Pr(Z(t+dt) = i | Z(t) = j) = 1 \) if \( i = j \) and zero otherwise as it is clear from the definition of the transition probabilities in (8). The chain is clearly irreducible since any state can be visited from any state with strictly positive probability.

Furthermore,
\[ E(n) = \sum_{n=1}^{\infty} nP_n < \infty \]

since applying for example the ratio test again

\[ \lim_{n \to \infty} \frac{(n+1)P_{n+1}}{nP_n} = \lim_{n \to \infty} \frac{(n+1)\lambda_n}{n\mu_{n+1}} = \lim_{n \to \infty} \frac{\lambda_n}{n\theta_{n+1}} = 0 \]
References


Bikker, J., 2004: Competition and Efficiency in a Unified European Banking Market


Dana, J. D. and Fong, Y., F. 2008: Product Quality, Reputation, and Tacit Collusion, mimeo, Kellogg School of Management, 2008


General Accounting Office 2008: Audits of Public Companies: Continued Concentration in Audit Market for Large Public Companies Does Not Call for Immediate Action, GAO-08-163


26


Salop, S. 1979: Monopolistic Competition with Outside Goods, Bell Journal of Economics 10, 141-156
