INSTITUTO DE ECONOMÍA

DOCUMENTO de TRABAJO

436
2013

Reputation-Driven Industry Dynamics

Bernardita Vial; Felipe Zurita.

www.economia.puc.cl • ISSN (edición impresa) 0716-7334 • ISSN (edición electrónica) 0717-7593

Reputation-driven industry dynamics*

by Bernardita Vial† and Felipe Zurita‡

December 3, 2012

Abstract

This paper studies the entry-exit dynamics of an experience good industry. Consumers observe noisy signals of past firm behavior and hold common beliefs regarding their types, or reputations. There is a small chance that firms may independently and unobservably be exogenously replaced. The market is perfectly competitive: entry is free, and all participants are price-takers. Entrants have an endogenous reputation μ_E. In the steady-state equilibrium, μ_E is the lowest reputation among active firms: firms that have done poorly leave the market, and some re-enter under a new name. This endogenous replacement of names drives the industry dynamics. In particular, exit probabilities are higher for younger firms, for inept firms, and for firms with worse reputations. Competent firms have stochastically larger reputations than inept firms both in the population as a whole and within each cohort, and thus are able to live longer and charge higher prices.

JEL Classification: C7, D8, L1
Keywords: reputation, industry dynamics, free entry, exit and entry rates

1 Introduction

By now there is a large empirical literature that studies the dynamics of firms within an industry. Among the most salient patterns that have consistently been found are[1]: (1) The presence of sizeable entry and exit rates even in industries that are scarcely growing,

*We are grateful to Axel Anderson, Francesc Dilme, Juan Dubra, Juan Escobar, Claudio Fernández, Hugo Hopenhayn, David K. Levine, César Martinelli, and Federico Weinschelbaum for their helpful comments. We are also grateful to conference participants at the Association for Public Economic Theory, European and Latin American meetings of the Econometric Society, *Jornadas de Economía del Banco Central del Uruguay, Jornadas Latinoamericanas de Teoría Económica*, Midwest Economic Theory Meetings, *Sociedad de Economía de Chile*, Spain-Italy-Netherlands Meeting on Game Theory, World Congress of the Game Theory Society, and seminar participants at *Universidad de Santiago*, *Universidad de Chile*, and *Pontificia Universidad Católica de Chile*. This paper shares content with an earlier manuscript entitled "On Reputational Rents as an Incentive Mechanism in Competitive Markets." Financial support from FONDECYT grant #1121096 is gratefully acknowledged.

†*Instituto de Economía, Pontificia Universidad Católica de Chile*. E-mail:bvial@uc.cl

‡*Instituto de Economía, Pontificia Universidad Católica de Chile*. E-mail:fzurita@uc.cl

[1]The empirical literature has also given a great deal of attention to firm growth and firm size. We will abstract from this issue by assuming that all fims have a capacity constraint of une unit.

with large heterogeneity across industries (Dunne et al., 1988); (2) younger firms are –ceteris paribus– more likely to exit and also (3) more likely to charge lower prices (Foster et al., 2008).

A recent strand of the literature adds a number of regularities related to firms' reputations. McDevitt (2011) focuses on an industry where firms with widely different track records compete with each other, and where exit, entry, and name changes occur frequently. Similarly, Cabral and Hortacsu (2010) study the reputational mechanism of eBay, an online auctioneer. These studies confirm previous findings, and add that: (4) the firms that are more likely to change names or exit are those with worse or shorter track records (McDevitt, 2011); and (5) the probability that a given seller will exit the market increases as its reputation worsens (Cabral and Hortacsu, 2010).

This paper presents a model of an industry in which the dynamics are driven entirely by reputation. The model is fully consistent with the five patterns above, and yields novel theoretical predictions as well. As in those industries analyzed by the recent empirical literature, in our model the active firms are not only heterogeneous in age (as defined by the number of periods that have elapsed since the firm began operating under its current name), but also within any age cohort they are heterogeneous in both records (histories) and prices.

Ours is an adverse selection model with imperfect public monitoring. Reputations are the common belief regarding a firm's type. There is perfect competition in the sense of Gretsky et al. (1999): firms are price-takers, and there is free entry. The incumbents' reputation is the Bayesian update of a common prior given an observed history of (imperfect, public) signals; the entrants' reputation is also a consistent belief. Incumbents have the option to re-enter the market under a new name. In the non-revealing, pure strategy equilibrium that is our focus, they choose to exercise this option whenever their reputation falls below a threshold, so that (5) holds. This threshold coincides with the entrants' reputation. In the steady state, there are exit and entry flows while the industry as a whole is stagnant, as in (1). Competent firms' reputation stochastically dominates that of inept firms; still, full separation is never achieved. Thus, the option to change names is used more intensively among the inept, so that (4) holds. As a consequence, each time the option is exercised the reputation distribution of any given cohort shifts to the right, and thus older cohorts have stochastically better reputations. In turn, this also implies that the probability of exiting the market is decreasing in age, as in (2). Moreover, since prices are increasing in reputation, (3) holds. Even though the reputation of older firms first-order stochastically dominates that of younger ones, there is always heterogeneity both within and between cohorts; in fact, the reputation distributions for all cohorts have full support.

The theoretical literature has investigated a number of possible explanations for the five patterns above. One strand asks whether such dynamics can be the result of individual productivity shocks in a perfectly competitive market for a homogeneous good (the seminal paper of Hopenhayn, 1992, stands out). A related strand looks at the combination of productivity shocks and financial frictions (Cooley and Quadrini, 2001, Albuquerque and Hopenhayn, 2004, Clementi and Hopenhayn, 2006) or labor market frictions (Hopenhayn and Rogerson, 1993). While (1) and (2) are consistent with this view, the law of one price is at odds with (3). Also, the empirical concepts of reputation and track records do not have a theoretical counterpart in this setting. The same is true in Fishman and Rob (2003), a paper in which the dynamics of the industry are driven by consumer inertia in a context of search costs and older firms sell more because they have a larger customer base.

On the other hand, there is a large body of theoretical literature that looks at the creation and maintenance of firms' reputations in markets for experience goods (e.g., Klein and Leffler, 1981, Fudenberg and Levine, 1989, and Mailath and Samuelson, 2001, to name just a few; Mailath and Samuelson, 2006, and Bar-Isaac and Tadelis, 2008, present comprehensive expositions of the literature.) This literature discusses primarily the monopoly case. In spite of this, some papers still manage to look at entry and exit decisions. For instance, Bar-Isaac (2003) assumes that the firm has the option to leave the market. When the firm knows its own type, in equilibrium the high-quality firm never leaves, while the low-quality firm plays a strictly mixed strategy at low levels of reputation–i.e., below some threshold. The mixed strategy is such that the post-exit reputation of any firm that has crossed the threshold becomes the threshold. Having a strictly positive probability of exiting, the low-quality type eventually leaves; this implies that there is complete separation in the long run. Board and Meyer-ter Vehn (2010) extend this analysis by incorporating moral hazard and the possibility of entry, and focus their analysis on the investment and exit decisions over the life cycle of the firm. In this equilibrium, the entry-level reputation coincides with the threshold as well.

Within the strand of the literature that looks at reputation dynamics in competitive markets, some papers focus on markets in which the information flow to potential customers is quite limited, and fundamentally different from that to customers–namely, private monitoring; Hörner (2002) and Fishman and Rob (2005) stand out. Instead, we want to examine markets where information–albeit imperfect–flows constantly to potential customers as well; for instance, the eBay feedback system (Cabral and Hortacsu, 2010), or the complaint record of plumbing firms (McDevitt, 2011). Indeed, the internet-related technological progress turns an increasing number of markets fall into this category by providing means of communication among customers; think for instance of the travel industry with TripAdvisor, Expedia, etc.

Tadelis (1999) is one of the first papers to formally analyze competition under imperfect public monitoring. It presents an adverse-selection model with a continuum of firms. However, the author focuses on an equilibrium where firms leave the market after one bad outcome; this means that active firms either don't have any history (they are new), or they must have impeccable records. Tadelis (2002) develops a similar model, under moral hazard. While this kind of model can explain certain stylized facts of industry dynamics, like the differences in pricing and probability of exit between cohorts, it cannot explain the observed heterogeneity in these variables after controlling by age: all firms of the same age must have the same records and reputation. In particular, it cannot account for observations (4) and (5) beyond age.

Our model recasts Mailath and Samuelson (2001)'s in a Walrasian environment, obtaining heterogeneous reputations even in the steady state (as in Vial, 2010, but considering entry). In our model, the entry-level reputation and the reputation distributions are endogenous. These variables turn out to be important determinants of the industry dynamics, as the rate of endogenous exit (hence the exit-entry flow) is tied to them. Hence, our paper complements recent literature on reputation under competition that features heterogeneous reputations, where some papers assume the entry-level reputation to be exogenous (e.g., Ordonez, forthcoming) while others obtain it independently from the reputation distributions because of their focus on mixed strategies (e.g., Atkeson et al., 2012). When analyzing industry dynamics, the difference between mixed and pure strategy equilibria becomes important: in the former case, there are incumbents among those with the entrants' reputation while in the latter only entrants carry the entry-level reputation. In each case the resulting age distributions of firms are therefore different.

The entrants' reputation μ_E is a consistent consumer belief. This consistency condition ties together consumers' belief updating, the reputation distributions, μ_E and the firms' strategies in a non-trivial fashion. The reputation of entering firms must coincide with the fraction of competent firms among them. However, the mass of competent firms that choose to change their names and re-enter the market depends precisely on the level of reputation with which they would re-enter–and so does the reputation distribution. We prove that such a consistent entry-level reputation exists, and moreover, that it is unique (Theorem 1). Within this equilibrium, the determinants of the entry-level reputation μ_E are purely informational.[2] We find that the entry-level reputation is increasing in the exogenous replacement rate (Theorem 3): industries in which competence is more transient (for instance, because of a high rate of technological development), the entry-level reputation will be more demanding.

We also find that highly reputable firms are less likely to cross the μ_E barrier during any given time interval (Theorem 4), so that "better" names last longer, in a stochastic sense; the same is true of competent firms, both in the population as a whole (Theorem 2) and within each cohort (Theorem 5). As a consequence, as time goes by each generation or cohort of firms improves its reputation (Theorem 6).

The rest of the paper is organized as follows: Section 2 presents the model. Section 3 introduces the equilibrium concept. Section 4 discusses existence and uniqueness issues; it is technical and can be skipped without loss of continuity. Section 5 examines the relationship between the replacement rate and the turnover ratio. Section 6, the core of the paper, analyzes the dynamics of the industry in the steady state.

2 The model

2.1 Preliminaries

We consider an infinitely repeated game in which, at every date $t = 0, 1, 2, ...$, a market for a given service opens. Firms are long-run players, while consumers are not. Instead, at every stage there is a different generation of short-lived consumers.

The service is an experience good as per Nelson (1970): its quality is ex ante unobservable to buyers. We assume that there is no communication among consumers. Since consumers only live for one period, the information each one obtains as a result of consuming the service is not transferred to the next generation, but lost altogether. Hence, quality is also unobservable ex post. Nevertheless, after consumption takes place, an imperfect signal r of the quality each active firm provided is publicly observed.

Each generation of consumers is of mass 1. In contrast, there is an unlimited supply of potential firms. Each individual may consume or produce at most one unit per period. Hence, while all consumers may purchase, not all firms will be able to sell. We call "active at t" a firm that produces at time t, and "inactive" a firm that does not. We will assume that consumers are homogeneous, and that their willingness to pay is high enough so that they all buy; as a consequence, the mass of active firms will–in equilibrium–be 1.

[2] In contrast to Bar-Isaac (2003), Board and Meyer-ter Vehn (2010) and Atkeson et al. (2012), in our model the determination of μ_E is independent from the zero-profit condition.

There are two types of firm: competent (C) and inept (I). Competent firms are those that can only produce a high-quality variety of the service, while inept firms can only produce a low-quality one. The total mass of competent firms is denoted by θ, constant over time and less than 1.

Each active firm is subject to the possibility of dying. A dead firm is replaced immediately by a newly born firm. While consumers are aware of this replacement process, they do not observe it. The process is assumed to be i.i.d. across time and firms. λ denotes the probability of dying, and ϕ the probability that a dead firm is replaced by a competent one. This replacement process ensures that throughout any history there is never almost certainty about any firm's type. In effect, Cripps et al. (2004) shows that the adverse selection model with imperfect monitoring needs a mechanism for replenishing uncertainty about types in order for doubts about players' types to persist in the long run. Different mechanisms have been studied. One of them is given by information frictions, such as limited memory (Liu and Skrzypacz, 2009), coarse observability (e.g., in Ekmekci, 2011, in which consumers observe discrete ratings rather than full histories), or costly observation of records (Liu, 2011). A second, related approach is that of Tadelis (1999) and Tadelis (2002), where consumers forget certain aspects of a history (what he calls "reputation reduction"), with the same effect. Another mechanism is provided by "trembles", as in Levine and Martinelli (1998). The approach we follow is the one advanced in Mailath and Samuelson (2001). By adding an unobservable replacement process, consumers are never certain of who they are dealing with. The replacement of exiting firms may be plainly exogenous (our choice) or endogenous as it is in the literature that studies the possibility of trading names.[3] If instead of a replacement process we had chosen a process of unobservable type changes, the dynamics of the industry–which is the focus of our paper–would be exactly the same.

On the other hand, a mass $\eta \ll 1$ of competent firms is also born each period among inactive firms. Note that for the total mass of competent firms to be constant over time, it is necessary that $\phi = \frac{\lambda\theta - \eta}{\lambda}$.

Types are privately observed. Thus, this is a pure adverse selection model. Hence, from the consumers' perspective, the probability of receiving a high-quality service is the same as the probability of facing a competent firm.

Firm names play a key role in our model. Consumers only observe each firms' name and the history of public signals since the last spell of uninterrupted use of that name. They don't know if that name has always belonged to that firm, or if it was first used by some of its predecessors. In that sense, a firm's reputation is really the reputation of the name they are currently using.

When a name is not used in one period, it is forgotten by consumers–together with its associated signal history. Then, should a firm become inactive for one period or longer, consumers will forget its history. The next time the firm becomes active, it will have to do so under a new name. Hence, maintaining a name requires remaining active without interruption. In addition, the firm may also choose to change its name at the beginning of each stage; we assume that this is done by exiting the market and reentering immediately.

Thus, a firm may have different names at different times according to the (endogenous) name-changing process. Moreover, the same name may pass from one firm to another

[3] See, for example, Tadelis (1999) and Mailath and Samuelson (2001). A different strand of this literature looks at the case in which trading names is observable, as in Wang (2011) and Hakenes and Peitz (2007).

under the (exogenous) replacement (or birth-death) process, in which the latter inherits the former's history.

We assume consumers have common priors. The timeline for the stage game is shown in Figure (1).

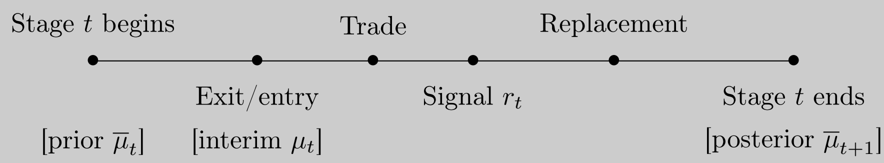


Figure 1: Time line for the date-t stage game

Consumers' beliefs refer to the probability that a given name belongs to a competent firm, conditional on all available information. We refer to this consumer belief as the name's reputation. At the beginning of each stage, each incumbent is endowed with a **prior** reputation $\overline{\mu}_t$. Then, should the incumbent decide to produce (be active) at t, it must choose whether to change its name to a new one–which will carry the reputation associated with a name with no history μ_E–or keep its old name and prior reputation $\overline{\mu}_t$. We will refer to this as the firm's **interim** reputation, and denote it by μ_t. This is the reputation the firm will have when the market opens. After trading, the signal r_t will be publicly observed, and the replacement process takes place–yet it is unobserved by consumers; the Bayesian update of the interim reputation given r_t, which takes into account the possibility of having being replaced, will be the firm's **posterior** reputation. This posterior will be the next period's prior, and so it will be denoted by $\overline{\mu}_{t+1}$.

2.2 Signals

The signal r refers to any piece of information that is publicly available to consumers. For instance, if the firms were schools, r could be the score percentile on a standardized test; if the firms were academic journals, r could be their impact factor; if the firms were health care providers, r could be their medical malpractice track records; if the firms were car makers, r could be the consumer reports, and so on.

The signal lies in the open unit interval: $r \in (0,1)$. When a firm provides high quality, its signal is distributed according to the c.d.f. F_H; when it provides low quality, it is distributed according to the c.d.f. F_L. The p.d.f.'s are denoted by f_H and f_L, respectively. We assume that a higher signal makes it more likely that a firm provided high quality:

Assumption 1 (Monotone likelihood ratio)**.** *The likelihood ratio* $R(r) \equiv \frac{f_H(r)}{f_L(r)}$ *is a monotonically increasing bijection from* $(0,1)$ *to* $(0,\infty)$.

This assumption implies that $F_H(r) \leq F_L(r)$ for all r, this is to say, the signal conditional on H first-order stochastically dominates the signal conditional on L.

2.3 Firms

At every stage t, each firm chooses whether to produce or not (i.e., remain active in the case of active firms, or enter in the case of inactive firms). We denote by $q_t \in \{0, 1\}$ the production level. In addition, those active firms that choose to produce must decide whether to keep their previous name or change it (by exiting and reentering immediately at no cost). We denote by $n_t = 1$ the decision to keep the name, and by $n_t = 0$ the decision to change it. Those firms that were inactive do not have this choice.

The date-t profits π_t are given by:

$$\pi(\mu_t) = \begin{cases} p(\mu_t) - c & \text{if } q_t = 1 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where c is the production cost, and $p(\mu)$ the competitive price for a service that is of high quality with probability μ. μ reflects consumers' beliefs, and is endogenously determined in equilibrium, taking into consideration the equilibrium naming and production policies.

The production and naming decisions jointly maximize the expected, discounted profits:

$$(1 - \delta) \sum_{t=0}^{\infty} (1 - \lambda)^t \delta^t E[\pi_t], \tag{2}$$

where $\delta \in (0, 1)$ is a discount factor. The expectation depends on the firm's type, as this affects the signal distribution.

Firms have heterogeneous and ever-changing reputations. Let $\overline{G}_t$ denote the cdf of prior reputations at the beginning of stage t of those firms that were active at $t - 1$. Stage t begins with firms' exit and entry decision. G_t denotes the cdf of interim reputations of those firms that chose to be active at t. Thus, G_t and $\overline{G}_t$ differ because some active firms chose to exit, some to re-enter, and some inactive firms decided to enter (the last two with a reputation μ_E).

The superscripts C and I denote the corresponding subpopulations of competent and inept firms. Lemma (3) below proves that there is a unique steady-state pair of reputation distributions $\overline{G}^C$ and $\overline{G}^I$.

2.4 Consumers

Consumers are homogeneous. Their willingness to pay for a high-quality unit of the service is β (with $\beta > 0$), while for a low-quality one it is normalized to zero. Hence, when facing a reputation-μ firm, the expected utility of buying is:

$$E[u] = \mu\beta - p, \tag{3}$$

Consumers are indifferent between two firms with reputations μ and μ' if $p(\mu) - p(\mu') = \beta(\mu - \mu')$.

2.5 Equilibrium: definition

By equilibrium we mean stationary Markov perfect equilibrium. Each firm's state variable is the (commonly known) prior reputation $\overline{\mu}_t$, and its (privately known) type $\tau \in \{C, I\}$. Recall that if a firm does change its name, its previous history becomes unobservable to consumers; as a consequence, its reputation is not $\overline{\mu}_t$. Rather, consumers do not distiguish among entering firms, regardless of whether they were active or not in the previous period. Thus, all entrants will have the same reputation level, denoted by μ_E. Hence, the firm's reputation is given by:

$$\mu_t = n_t \overline{\mu}_t + (1 - n_t)\,\mu_E$$

Definition 1. An equilibrium is:

- A naming policy function $n(\overline{\mu}, \tau)$;
- A production policy function $q(\overline{\mu}, \tau)$;
- A price function $p(\mu)$;
- A belief system, namely:
  - A probability that a firm is competent
    - † $\overline{\mu}_t = \varphi(\mu_{t-1}, r_{t-1})$, the Bayesian update of the prior μ_{t-1} given the signal r_{t-1}, for firms that were active in the previous period and kept their names
    - † An entry-level reputation μ_E, for firms that enter under a new name, and
  - A pair of steady-state, population-wide reputation distributions $\left(\overline{G}^C, \overline{G}^I\right)$ for active firms

such that:

E1. Beliefs are consistent,

E2. n and q maximize (2),

E3. Consumers' choices maximize (3), and

E4. The market clears.

As is common in this kind of model, many equilibria are supported by different off-equilibrium-path beliefs. For instance, there is a trivial equilibrium with no reputation-building: All consumers believe only inept firms are active at all times, so that every active firm has a null reputation and is indifferent as to whether to produce or not. Instead, we look at a reputational equilibrium–an equilbrium where the reputation of each firm is affected by its signals. Specifically, we are assuming that signals and only signals affect the reputation of firms that remain active; in particular, the naming decision is uninformative. As competent and inept firms with reputations higher than μ_E will keep their name (see Section 3 below), the event that a firm keeps its name in the equilibrium path is indeed uninformative–there is no assumption so far. Firms with priors below μ_E will change their names. Should one such firm–contrary to the equilibrium strategy–keep its name, we assume that consumers won't change their prior. For instance, they interpret this deviation as an uninformative "tremble".[4]

[4] A few remarks are in order. First, any out-of-equilibrium belief that assigns a deviator a reputation

3 Reputational equilibrium

Price function. Consumers must be indifferent among providers. Thus, E3, together with the assumption that consumers are homogeneous, implies that the price as a function of μ is given by:

$$p(\mu) = \alpha + \beta\mu \tag{4}$$

where $\alpha \leq 0$ for consumers to buy. More reputable firms are paid higher prices.

Remark 1. With heterogeneous consumers, none of the main results would change. For instance, if consumers varied in their willingness to pay, the equilibrium assignment would be positively assortative, as in Vial (2010), and the price function's slope would depend also on the supply side, namely, the population-wide reputation distribution.

Naming policy. Writing the firms' optimization problem (2) for a type τ firm in recursive form, we get:

$$\begin{aligned} v(\overline{\mu}, \tau) = (1-\delta) \max_{n,q} \{ & q(n(p(\overline{\mu}) - c) + (1-n)(p(\mu_E) - c)) + \\ & \left. \delta(1-\lambda) \int_0^1 v(q\varphi(n\overline{\mu} + (1-n)\mu_E, r) + (1-q)\mu_E, \tau)\, dF_\tau \right\} \end{aligned} \tag{5}$$

Notice that if the firm chooses not to produce at a given date, the state in the following period would be (μ_E, τ) should the firm continue to exist.

Lemma 1 (naming policy)**.** *The optimal naming policy is to choose the highest between* $\overline{\mu}$ *and* μ_E*:*

$$n^*(\overline{\mu}, \tau) = \begin{cases} 1 & \text{if } \overline{\mu} \geq \mu_E \\ 0 & \text{otherwise} \end{cases}$$

Proof. Since the price is increasing in reputation, so is the flow utility. On the other hand, the law of motion φ is also increasing in this argument since it is the Bayesian update of the a priori μ. Then, a firm that wants to produce prefers to do so at the highest available reputation. $\square$

Observe that both types find it optimal to act in the same way. Then, firms choose to change their name as soon as their reputation falls below the threshold μ_E. The value function becomes:

$$\begin{aligned} v(\mu, \tau) = (1-\delta) \max_{q} \{ & q(p(\mu) - c) + \\ & \left. \delta(1-\lambda) \int_0^1 v(q\varphi(\mu, r) + (1-q)\mu_E, \tau)\, dF_\tau \right\} \end{aligned} \tag{6}$$

where

$$\mu = \max\{\overline{\mu}, \mu_E\} \tag{7}$$

strictly smaller than μ_E also supports these equilibrium strategies. This is so because any such belief makes it optimal for low-reputation firms to change their names. Second, there is no equilibrium in which low-reputation firms that keep their names get a reputation higher than μ_E. If they did, then the inept firms would also want to keep their names–and at these reputation levels the fraction of competents is smaller than μ_E. Third, a mixed-strategy equilibrium in which some firms with a prior reputation below μ_E keep their names and get a reputation *equal* to μ_E is conceivable. This is the sort of equilibrium analyzed in Bar-Isaac (2003). This equilibrium is not essentially different from the one we focus on, since it is still the case that the interim reputation is the maximum between the prior and the entry-level reputation.

Production policy. Since there is an unlimited number of potential entrants, competition will drive prices down so that a positive mass of firms choose to stay out.

Lemma 2 (production policy)**.** *The optimal production policy is:*

$$q^*(\overline{\mu}, C) = 1 \qquad \forall \overline{\mu}$$

$$q^*(\overline{\mu}, I) = \begin{cases} 1 & \text{if } \overline{\mu} > \mu_E \\ 0 \text{ or } 1 & \text{if } \overline{\mu} \leq \mu_E \end{cases}$$

This is to say, all competent firms will always be active, and so will inept firms with a reputation above the entry-level. The other inept firms will be indifferent as to whether to produce or not.

Proof. The value function (6) is increasing in μ for both types. On the other hand,

$$v(\mu, C) > v(\mu, I)$$

because the signals for a competent firm are stochastically larger than those of an inept one. It follows that the free-entry condition applies only to inept firms:

$$v(\mu_E, I) = 0$$

□

It follows that market clearing is obtained when a mass $(1-\theta)$ of inept firms are active. The equilibrium price function (i.e., α) is pinned down by the no-entry condition for inept firms:

$$v(\mu_E, I) = (1-\delta)\left\{(\alpha + \beta\mu_E - c) + \delta(1-\lambda)\int_0^1 v(\varphi(\mu_E, r), I)\, dF_L\right\} = 0$$

Beliefs. Each firm's reputation is formed as the Bayesian update of a common prior upon observation of the public history. As usual, we require beliefs to be consistent with the equilibrium strategies.

At the begining of the game ($t = 0$), the mass of competent firms is θ. From then on, a firm with a given name dies with probability λ and passes its name to a newly born firm, which is competent with probability ϕ. This process is hidden to consumers. Hence, from the consumers' viewpoint, names are associated with underlying firm types that may change according to the transition matrix in Table (1). This process, though, is taken into consideration in the Bayesian update of the name's prior μ given the signal r as follows:

$$\varphi(\mu, r) = \lambda\phi + (1-\lambda)\frac{f_H(r)\mu}{f_H(r)\mu + f_L(r)(1-\mu)}. \tag{8}$$

In effect, the probability that the firm under that name is competent at t ($\Pr(\tau_t = C \mid \mu_{t-1}, r_{t-1}) \equiv \varphi(\mu_{t-1}, r_{t-1})$) is made up of the probability that there was a previous firm that died and was replaced by a newly born competent firm ($\lambda\phi$), and the probability that the firm survived $(1-\lambda)$ times the conditional posterior probability of it being competent at t after a signal r_{t-1}.

		Type at t (τ_t)	
		C	I
Type at $t+1$	C	$1-\lambda+\lambda\phi$	$\lambda\phi$
(τ_{t+1})	I	$\lambda(1-\phi)$	$1-\lambda\phi$

Table 1: Transition matrix for types under fixed names

We now turn to the consistent entry-level reputation μ_E. It must coincide with the fraction of competent firms among the group of entrants:

$$\mu_E = \frac{\text{Competent entrants}}{\text{Entrants}}. \tag{9}$$

Before the exit-entry process, the mass of competent firms is $\theta - \eta$,[5] the sum of:

- active competents that survived: $\theta(1-\lambda)$
- newly born competents that replace active competents that died: $\theta\lambda\phi$
- newly born competents that replace active inepts that died: $(1-\theta)\lambda\phi$

According to the optimal production and naming policies, the competent entrants are then the sum of all competents whose reputation fell below μ_E, $(\theta-\eta)\overline{G}^C(\mu_E)$, because they change their names, and all newly born competent among inactive firms, η.

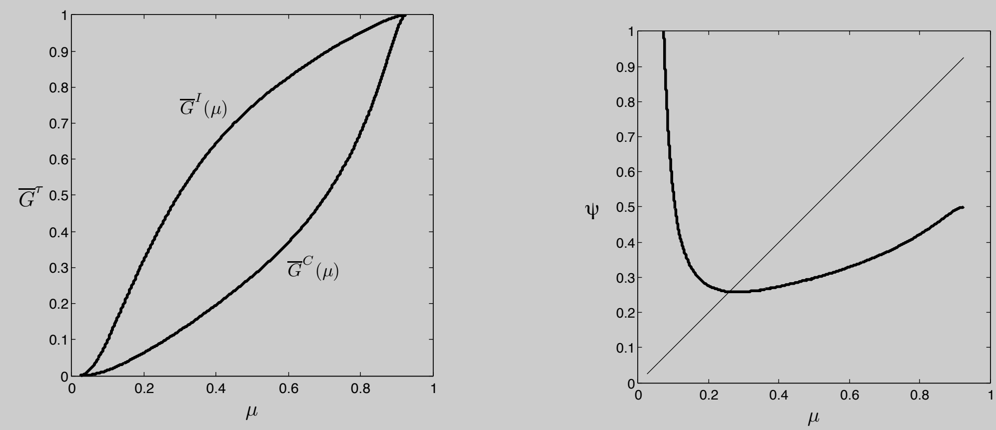
The inept firms whose reputation fell below μ_E either change their names (i.e., exit and re-enter), or simply exit to be replaced by other, previously inactive, firms. Regardless, since market clearing (E4) requires the mass of active firms to be 1 at all times, then the mass of entrants must equal the mass of exiting firms. Thus, the mass of entrants is $\overline{G}(\mu_E)$.

The righ-hand side of Equation (9) can be written as:

$$\psi(\mu_E) \equiv \frac{(\theta-\eta)\overline{G}^C(\mu_E)+\eta}{\overline{G}(\mu_E)}. \tag{10}$$

Then, a consistent entry-level reputation μ_E is a fixed point of ψ. Theorem 1 shows that there is only one such μ_E. Figure 2 depicts a numerical example of steady-state reputation distributions, and the corresponding ψ function; the uniqueness of μ_E is apparent.

[5]Recall that $\phi = \theta - \eta/\lambda$.



(a) $\overline{G}^{C}(\mu)$ and $\overline{G}^{I}(\mu)$ (b) $\psi(\mu)$

Figure 2: Reputation distributions and consistent entry-level reputation

Notes: $r|_H \sim Be\,(r \mid 3, 2)$ and $r|_L \sim Be\,(r \mid 2, 3)$. The parameters are $\lambda = 0.1$, $\theta = 0.5$ and $\eta = 0.025$. The resulting μ_E is 0.26.

In equilibrium, the net inflows of firms must be zero. If there is entry, there must be exiting. When competent firms are born among inactive firms (i.e., $\eta > 0$), the entry-level reputation μ_E must be strictly larger than $\lambda\phi$, so that a positive mass of low-reputation firms chooses to leave the market and stay out. The newly born competents will enter.

On the other hand, when no competent firms are born among inactive firms (i.e., $\eta = 0$), the only firms that would enter are those that exited because of having a reputation lower than the entry level threshold. Each individual firm's reputation would change by erasing its history, but the mean reputation of the group cannot change, as consumers are aware of the fact that the entrants are the same ones that just exited. Then, this adverse selection argument shows that the entry-level reputation μ_E cannot be larger than $\lambda\phi$: A consistent entry level reputation is one so low that nobody wants to change their name.

4 Existence and uniqueness

Section 3, assuming the existence of a non-revealing reputational equilibrium, characterized the equilibrium strategies. This section shows that there is a unique pair of steady-state reputation distributions and a unique entry-level reputation which are consistent with those equilibrium strategies.

We proceed in two steps. First, the entry-level reputation is assumed to be an exogenous parameter $y \in (0, 1)$. Under this assumption, we explain how Bayes' law defines the stochastic change over time of individual reputations, as signals accumulate. We also establish the difference-equation system that defines the dynamic process of population distributions, for a fixed, arbitrary, entry-level reputation y. Lemma 3 shows that there is a unique steady-state distribution pair for competent and inept firms. This existence and uniqueness result is important because we want to focus on steady-state equilibria. The analysis also allows us to characterize those distributions: They are comparable under first-order stochastic dominance, are continuous and have full support.

Second, the entry-level reputation y is endogenized by requiring it to be consistent: $y = \mu_E$. Indeed, consistency implies that the fraction of competent firms among those active firms whose histories grant them a given reputation μ is precisely μ, and similarly, that the fraction of competent firms among entrants is precisely μ_E. These two properties turn out to be closely related. In the steady state new firms will enter (and some old ones will exit) if and only if new competent firms are born among inactive firms.

This section is technical, and can be skipped without loss of continuity.

4.1 Fixed, exogenous y

Rewriting Equation (8) in the likelihood ratio form we have:

$$\varphi(\mu, r) = \lambda\phi + (1-\lambda)\frac{R(r)\mu}{R(r)\mu + 1 - \mu}. \tag{11}$$

Define the functions $\tilde{r}(x, \mu)$ and $\tilde{\mu}(x, r)$ from $x = \varphi(\mu, r)$ as:[6]

$$\begin{aligned} r &= \tilde{r}(x,\mu) \Leftrightarrow x = \lambda\phi + (1-\lambda)\frac{R(\tilde{r}(x,\mu))\mu}{R(\tilde{r}(x,\mu))\mu + 1 - \mu}, && (12) \\ \mu &= \tilde{\mu}(x,r) \Leftrightarrow x = \lambda\phi + (1-\lambda)\frac{R(r)\tilde{\mu}(x,r)}{R(r)\tilde{\mu}(x,r) + 1 - \tilde{\mu}(x,r)}. && (13) \end{aligned}$$

The first function, $\tilde{r}(x, \mu)$, says what the signal value should be for a firm of current reputation μ to have a reputation x the next period. The second function, $\tilde{\mu}(x, r)$, indicates what the reputation was in the previous period of a firm with a signal r that currently enjoys a reputation x. Similarly, $\varphi(\mu, r)$ is the reputation in the next period of a firm that started off with a reputation μ and whose signal was r.

Appendix A shows that the end-of-period reputation cdf for each type of firm is thus given by:

$$\begin{pmatrix} \overline{G}^C_{t+1}(x) \\ \overline{G}^I_{t+1}(x) \end{pmatrix} \equiv \begin{pmatrix} \frac{\theta(1-\lambda+\lambda\phi)}{\theta-\eta} & \frac{(1-\theta)\lambda\phi}{\theta-\eta} \\ \frac{\theta\lambda(1-\phi)}{1-\theta+\eta} & \frac{(1-\theta)(1-\lambda\phi)}{1-\theta+\eta} \end{pmatrix} \begin{pmatrix} \int_0^{\tilde{r}(x,y)} G^C_t(\tilde{\mu}(x,r))\,dF_H \\ \int_0^{\tilde{r}(x,y)} G^I_t(\tilde{\mu}(x,r))\,dF_L \end{pmatrix} \tag{14}$$

The exit-entry process changes these distributions in two ways. First, firms whose reputation fell below the threshold y replace their reputation with y, so that the evolution of μ_t is actually defined by $\mu_t = \max\{\varphi(\mu_{t-1}, r_{t-1}), y\}$. All competent firms remain active, since competent firms with a low reputation clean up their names and re-enter immediately. Second, there is a mass η of newly born competent firms.

[6] Assumption (1) ensures that φ is strictly increasing in r, so that higher signal values always improve a firm's reputation. Moreover, the image of φ is $(\lambda\phi, \lambda\phi + 1 - \lambda)$ because the likelihood ratio is surjective: $\varphi((\lambda\phi,\ 1-\lambda+\lambda\phi) \times (0,1)) = (\lambda\phi, \lambda\phi + 1 - \lambda)$. On the other hand, it is readily seen that φ is strictly increasing in μ, and continuous. Hence, being onto, it is a bijection, and these implicit functions are well-defined.

These cdf's are transformed by the entry-exit process as follows:

$$\begin{aligned} G^{C}_{t+1}(x) &= \begin{cases} 0 & \text{if } x < y \\ \frac{1}{\theta}\left(\eta + (\theta - \eta)\,\overline{G}^{C}_{t+1}(x)\right) & \text{if } x \geq y \end{cases} \\ G^{I}_{t+1}(x) &= \begin{cases} 0 & \text{if } x < y \\ \frac{1}{1-\theta}\left(-\eta + (1 - \theta + \eta)\,\overline{G}^{I}_{t+1}(x)\right) & \text{if } x \geq y \end{cases} \end{aligned} \tag{15}$$

Note that when $\eta = 0$, G^{C}_{t+1} and G^{I}_{t+1} are simply the truncated versions of $\overline{G}^{C}_{t+1}$ and $\overline{G}^{I}_{t+1}$, respectively, because the only change is the replacement of the lowest reputations by y.[7] On the other hand, when $\eta > 0$, the mass of newly born competent firms enters to replace an equal mass of currently inept firms.[8]

Replacing in (15) and rearranging, we get:

$$\begin{aligned} \begin{pmatrix} \overline{G}^{C}_{t+1}(x) \\ \overline{G}^{I}_{t+1}(x) \end{pmatrix} &= \begin{pmatrix} \frac{(1-\lambda+\lambda\phi)}{\theta-\eta}\eta & -\frac{\lambda\phi}{\theta-\eta}\eta \\ \frac{\lambda(1-\phi)}{1-\theta+\eta}\eta & -\frac{(1-\lambda\phi)}{1-\theta+\eta}\eta \end{pmatrix} \begin{pmatrix} F_H(\tilde{r}(x,y)) \\ F_L(\tilde{r}(x,y)) \end{pmatrix} \\ &+ \begin{pmatrix} (1-\lambda+\lambda\phi) & \frac{\lambda\phi(1-\theta+\eta)}{\theta-\eta} \\ \frac{\lambda(1-\phi)(\theta-\eta)}{1-\theta+\eta} & (1-\lambda\phi) \end{pmatrix} \begin{pmatrix} \int_0^{\tilde{r}(x,y)} \overline{G}^{C}_{t}(\tilde{\mu}(x,r))\,dF_H \\ \int_0^{\tilde{r}(x,y)} \overline{G}^{I}_{t}(\tilde{\mu}(x,r))\,dF_L \end{pmatrix} \end{aligned} \tag{16}$$

Define the right-hand side of Equation (16) as the operator T in $B[0,1] \times B[0,1]$–the product space of continuous bounded functions. The steady-state reputation distributions $\overline{G}^{C}$ and $\overline{G}^{I}$ are a fixed point of T. Since T depends parametrically on y, so do $\overline{G}^{C}$ and $\overline{G}^{I}$.

Lemma 3. *T has a unique fixed point for any given $y \in [\lambda\phi, 1 - \lambda + \lambda\phi]$. Moreover:*

1. *$\overline{G}^{C}$ and $\overline{G}^{I}$ are absolutely continuous with support $[\lambda\phi, 1 - \lambda + \lambda\phi]$*
2. *G^{C} and G^{I} have the common support $[\max\{y, \lambda\phi\}, 1 - \lambda + \lambda\phi]$*

Proof. In Appendix B. □

The steady-state distributions are continuous in the parameter y, as they are the fixed points of a contraction.[9]

In equilibrium, then, there will be a unique pair of steady-state reputation distributions. Their absolute continuity follows from the absolute continuity of the signal distributions. The corresponding densities will be denoted by $\overline{g}^{C}$ and $\overline{g}^{I}$, respectively.

[7] The reader may wonder about the consistency of the supporting beliefs in this case. It is odd that consumers assign a reputation y to firms that in equilibrium achieved a lower reputation in the previous round. After all, the entry-level reputation should be the mean reputation for entrants. Indeed, Lemma 4 below shows that any consistent entry-level reputation is such that no firm ever wants to exit the market, namely, $\mu_E \leq \lambda\theta$.

[8] Observe that when $\eta > 0$, it is necessary that the mass of exiting inept firms be sufficiently large so that they make enough room for the entrants, in order for G^{I}_{t+1} to be nonnegative. This is a condition over μ_E. Consistency will require this to be the case in equilibrium. However, this restriction has no bearing on the convergence result we are about to describe.

[9] See De la Fuente, 2000, Chapter 2, Theorem 7.18.

4.2 Consistent y

When useful, we will stress the fact that $\overline{G}^{\tau}$ depends parametrically on y by writing $\overline{G}^{\tau}(x, y)$; this is the fraction of type$-\tau$ firms with reputation no greater than x when the distributions $\overline{G}^{C}$ and $\overline{G}^{I}$ have been generated under the cutoff value y. We write $\psi(x, y)$ accordingly as:

$$\psi(x, y) \equiv \frac{(\theta - \eta)\,\overline{G}^{C}(x, y) + \eta}{\overline{G}(x, y)}, \tag{17}$$

This function is well defined for $x > \lambda\phi$. Consistency requires that μ_E be a fixed point of ψ in both arguments, i.e., $\mu_E = \psi(\mu_E, \mu_E)$.

Observe that consistency also requires that the expected probability of being competent in any information set be equal to the fraction of competent firms within that set. In particular, the probability of being competent conditional on the firm's reputation being x should be exactly x:

$$x = \frac{(\theta - \eta)\,\overline{g}^{C}(x, \mu_E)}{\overline{g}(x, \mu_E)} \tag{18}$$

When new competent firms are born among inactive firms ($\eta > 0$), the entry-level reputation must be high enough so that old firms wish to exit and make room for the newcomers. Then, any consistent entry-level reputation must be larger than $\lambda\phi$. This can be appreciated in Figure 2. Theorem 1 below asserts that there is actually a unique such μ_E.

Theorem 1. *If $\eta > 0$, there is a unique consistent entry-level reputation μ_E. Moreover, $\mu_E \in (\lambda\phi, \theta)$.*

Proof. In Appendix C. □

Thus, when there is type change among inactive firms there is a continuous process of endogenous name renovation. The next theorem establishes that the reputation of competent firms is stochastically larger than that of inept firms, and hence this name renovation is biased towards inept firms.

Theorem 2. *Both before and after the exit-entry process takes place the reputation of competent firms first-order stochastically dominates the reputation of inept firms. That is, for all $x \in (\lambda\phi, 1 - \lambda + \lambda\phi)$, $\overline{G}^{C}(x, \mu_E) \leq \overline{G}^{I}(x, \mu_E)$ and $G^{C}(x, \mu_E) \leq G^{I}(x, \mu_E)$.*

Proof. In Appendix D. □

Using Equation (18), we obtain:

$$E[\mu \mid \mu < \mu_E] = \frac{(\theta - \eta)\,\overline{G}^{C}(\mu_E, \mu_E)}{\overline{G}(\mu_E, \mu_E)} \tag{19}$$

Thus, the consistency condition $\mu_E = \psi(\mu_E, \mu_E)$ can be written as:

$$\mu_E = E[\mu \mid \mu < \mu_E] + \frac{\eta}{\overline{G}(\mu_E, \mu_E)}. \tag{20}$$

That is, the entrants' reputation is the mean reputation that the firms that have just left the market would have had if they didn't exit, increased by the fact that a mass η of inept firms that exit are replaced by the newly created competent firms that enter.

From Equation (20) it is clear that in the case of $\eta = 0$, where no new competent firms are born among inactive firms, μ_E cannot be larger than $\lambda\theta$. Indeed, Equation (20) becomes $\mu_E = E[\mu \mid \mu < \mu_E]$. However, $E[\mu \mid \mu < \mu_E] < \mu_E$ for all $\mu_E > \lambda\theta$; this is to say, unless the event $\mu < \mu_E$ is null, Equation (20) is inconsistent. It follows that the consistent entry-level reputation μ_E satisfies $\mu_E \leq \lambda\theta$. Without loss of generality, we assume that $\mu_E = \lambda\theta$. When $\eta = 0$, then, there is no actual exiting or entry; as the entrant's reputation would be worse than any active firm's reputation, no matter how bad the track records may be, no firm would ever want to change its name. This proves that:

Lemma 4. *If $\eta = 0$, the entry-level reputation μ_E is no greater than $\lambda\theta$ in equilibrium, and there is no actual entry.*

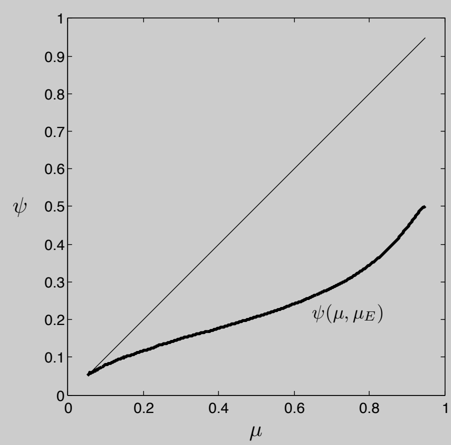


Figure 3: Consistent entry-level reputation when $\eta = 0$

Notes: $r|_H \sim Be(r \mid 3, 2)$ and $r|_L \sim Be(r \mid 2, 3)$. The parameters are $\lambda = 0.1$, $\theta = 0.5$ and $\eta = 0$. The resulting μ_E is 0.05.

Summing up, this section proved that there is a unique steady-state reputation distribution for each type of firm G^C and G^I, and a unique consistent entry-level reputation μ_E. When no competent firms are born among inactive firms ($\eta = 0$), this μ_E is the lowest possible, and as a consequence all firms want to keep their names at all times and there are no exit-entry flows. Any reputation is better than μ_E. Nevertheless, the threat of entry is not without consequences; rather, it serves the purpose of keeping prices down. On the other hand, when new competent firms are born among inactive firms ($\eta > 0$), μ_E is such that there are exit-entry flows, and at all times a positive mass of firms chooses to change their names.

5 Exogenous and endogenous replacement rates

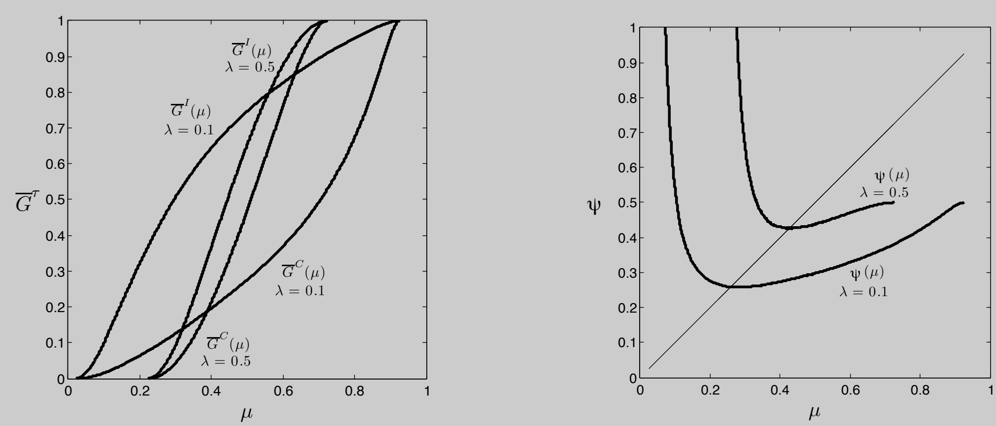
The empirical literature on industry dynamics finds that there is considerable heterogeneity among industries in terms of entry and exit rates. On the other hand, these rates are sizeable even in industries that are neither growing nor shrinking. Typically, within an industry the gross entry and exit rates are similar to each other, but at the same time they are orders of magnitude larger than the net rates (Dunne et al., 1988). Our focus is on a steady state, where the net exit rate is zero. Still, there is a constant renewal (exit and entry), given by $\overline{G}\left(\mu_E\right)$–the fraction of incumbents that leave the market, or turnover ratio.

In our model there are two replacement processes: that of firms, and that of names. The first one, although empirically unobservable, affects the latter. Theorem 3 discusses how the replacement rate λ affects the entry-level reputation μ_E:

Theorem 3 (entry-level reputation and replacement)**.** *The entry-level reputation μ_E is increasing in λ, the exogenous replacement probability.*

Proof. See Appendix E. □

A higher level of λ implies a decreased informativeness of histories, as the past becomes less useful in predicting a firm's current type. As a consequence, the reputation distributions of competent and inept firms move closer to each other. By this mechanism, the fraction of competent firms among those below the threshold μ_E increases, so that μ_E increases. This is depicted in Figure 4.

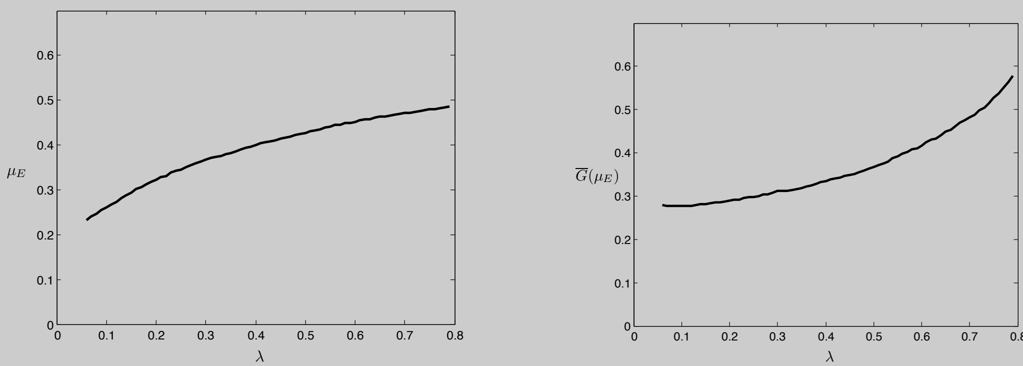


(a) $\overline{G}^C(\mu)$ and $\overline{G}^I(\mu)$ for different values of λ (b) $\psi(\mu)$ for different values of λ

Figure 4: Distributions and consistent entry-level reputation for different values of λ

Notes: $r|_H \sim Be\left(r \mid 3,2\right)$ and $r|_L \sim Be\left(r \mid 2,3\right)$. The parameters are $\theta = 0.5$ and $\eta = 0.025$.

As for the turnover ratio, it not only depends on μ_E but also on the population-wide reputation distribution–which shifts when λ increases. Figure 5 shows that in our example, when the replacement rate is small it has a large effect on μ_E but a negligible effect on the turnover ratio $\overline{G}(\mu_E)$, and that the opposite occurs when the replacement rate is large. Notice that a larger μ_E means that less information transpires to consumers, as a higher fraction of histories are erased by the name-changing process. Hence, the exogeneous decrease in the informativeness of histories given by the higher λ has the effect of a further endogenous decrease through this channel.



(a) μ_E for different values of λ

(b) $\overline{G}(\mu_E)$ for different values of λ

Figure 5: Entry-level reputation and turnover for different values of λ

Notes: $r|_H \sim Be\,(r \mid 3, 2)$ and $r|_L \sim Be\,(r \mid 2, 3)$. The parameters are $\theta = 0.5$ and $\eta = 0.025$.

6 Industry dynamics

Our model has a number of predictions, some of which have been investigated empirically. Cabral and Hortacsu (2010), in a study of eBay auctions, find that the probability that a firm exits the market increases as its reputation declines (as defined by eBay's reputation mechanism.) In turn, McDevitt (2011) studies the plumbing services market in Illinois and finds that, all else being equal, the firms that are more likely to change names or exit are those with worse track records–a variable that resembles the history of public signals in our model. Theorem 4 establishes that this is exactly what we should expect.

Regarding the cross section, the literature finds that younger firms are more likely to exit, and charge lower prices, than older firms (Foster et al., 2008). In the same vein, McDevitt (2011) finds that the exit probability is monotonically decreasing with age. Theorem 6 asserts that these findings are consistent with our model. It should be noted, though, that in our model we track the age of names, not of firms. This is so because consumers do not observe the birth-death process, and so they don't distinguish among those firms that have acted under the same name. Hence, firms are the appropriate empirical counterparts of names in our model.

We start by looking at the cross-sectional variation in exit rates:

Theorem 4 (Name dynamics)**.** *For any active firm, the probability of exiting the market in the next period is:*

1. *Higher for inept firms than for competent firms;*
2. *decreasing in the firm's reputation (both conditional and unconditional on type); and also*
3. *decreasing in the current signal (both conditional and unconditional on type).*

Proof. In Appendix F. □

The exit probability is the probability that the reputation falls below the threshold μ_E. Part (1) of the theorem follows from the fact that competent firms' signals are stochastically larger than inept firms'; part (2) of the theorem follows from the fact that according to Bayes' rule, the posterior probability of an event is increasing in the prior; and finally, part (3) of the theorem follows from the monotone likelihood assumption. These results are consistent with the empirical findings reported in Cabral and Hortacsu (2010) and McDevitt (2011), that relate exit probabilities with reputation and track records, respectively.

We now turn to the cross-sectional differences in reputation distributions of names of different age in the steady state. This is to say, we compare across cohorts. Notice that in the steady state, the group of age a looks exactly the same as the group of age 0 in a periods in the future. In this sense, studying the cross-sectional variation (across cohorts) is equivalent to studying the evolution over time of a given cohort.

Per Lemma 4, if no new competents were born among inactive firms ($\eta = 0$), then there would be no entry nor exiting, this is to say, no dynamics. In this case, it is not possible to distinguish among cohorts, as all names are introduced at the same time. Age is irrelevant, as it is completely detached from reputation. This contrasts with Hörner's model (Hörner, 2002), where age and reputation are biunivocally related, as only firms with perfect records survive.

On the other hand, as soon as there is a positive mass of competents born outside the market ($\eta > 0$), a flow of entry/exit emerges.

Let a denote the age of a name that was introduced a periods ago in the market and has been kept throughout (this, regardless of whether the firm that carries it has died and been replaced in that lapsus or not). All such names conform the cohort a. Let $\overline{G}_a^\tau$ denote the prior reputation distribution of the set of firms of cohort a and type τ, and $\overline{m}_a^\tau$ its mass; similarly, G_a^τ and m_a^τ denote the interim reputation distributions and its mass after exit.

At any date, a new cohort of mass $\overline{G}\left(\mu_E\right)$ enters. Out of them, a fraction μ_E, is competent: $m_0^C = \mu_E \overline{G}\left(\mu_E\right)$. As all new names carry the same reputation μ_E, we have:

$$G_0^C(x) = G_0^I(x) = \begin{cases} 0 & \text{if } x < \mu_E \\ 1 & \text{if } x \geq \mu_E \end{cases}$$

As time goes by, at each period two changes occur: (i) The birth-death process shifts the masses of competent and inept firms within the cohort according to the transition matrix 1:

$$\begin{pmatrix} \overline{m}_{a+1}^{C} \\ \overline{m}_{a+1}^{I} \end{pmatrix} = \begin{pmatrix} 1-\lambda+\lambda\phi & \lambda\phi \\ \lambda(1-\phi) & 1-\lambda\phi \end{pmatrix} \begin{pmatrix} m_a^C \\ m_a^I \end{pmatrix} \tag{21}$$

and (ii) The mass of surviving names in each subpopulation τ shrinks by a factor of $\left(1-\overline{G}_{a+1}^{\tau}(\mu_E)\right)$, as those firms that exit are not replaced by other firms from the same cohort. Hence:

$$m_{a+1}^{\tau} = \overline{m}_{a+1}^{\tau}\left(1-\overline{G}_{a+1}^{\tau}(\mu_E)\right) \tag{22}$$

In turn, the evolution of the prior reputation distributions in a given cohort at different ages is given by:

$$\begin{pmatrix} \overline{G}_{a+1}^{C}(x) \\ \overline{G}_{a+1}^{I}(x) \end{pmatrix} = \begin{pmatrix} \frac{(1-\lambda+\lambda\phi)m_a^C}{\overline{m}_{a+1}^{C}} & \frac{\lambda\phi m_a^I}{\overline{m}_{a+1}^{C}} \\ \frac{\lambda(1-\phi)m_a^C}{\overline{m}_{a+1}^{I}} & \frac{(1-\lambda\phi)m_a^I}{\overline{m}_{a+1}^{I}} \end{pmatrix} \begin{pmatrix} \int_0^{\tilde{r}(x,\mu_E)} G_a^C(\tilde{\mu}(x,r))\,dF_H \\ \int_0^{\tilde{r}(x,\mu_E)} G_a^I(\tilde{\mu}(x,r))\,dF_L \end{pmatrix} \tag{23}$$

The distributions of interim reputations relate to the priors' as follows:

$$G_a^{\tau}(x) = \begin{cases} 0 & \text{if } x < \mu_E \\ \frac{\overline{G}_a^{\tau}(x)-\overline{G}_a^{\tau}(\mu_E)}{1-\overline{G}_a^{\tau}(\mu_E)} & \text{if } x \geq \mu_E \end{cases} \tag{24}$$

Equation 23 is analogous to Equation 14. The only difference is in the weights. In the population of active firms as a whole the total mass and the ratio of competent to inept are constant over time. In contrast, not only each cohort is losing mass over time, but also each type does so at different rates. Similarly, Equation 24 resembles Equation 15; they differ in that within each cohort there is only exiting and no entrance.

Theorem 5 (Reputation across types by cohort)**.** *Within each cohort, the (prior, interim) reputation of competent firms first-order stochastically dominates the (prior, interim) reputation of inept firms:*

$$(\forall a \in \mathbb{N})\,(\forall x \in [\lambda\phi, 1-\lambda+\lambda\phi]), \qquad \overline{G}_a^{C}(x) \leq \overline{G}_a^{I}(x) \text{ and } G_a^C(x) \leq G_a^I(x)$$

Proof. See Appendix G ◻

Theorem 5 implies that conditional on age, inept firms are more likely to exit the market than competent firms. As a consequence, the rate at wich the mass of competents decreases is smaller than that of the inepts in Equation (22), as $\overline{G}_{a+1}^{C}(\mu_E) < \overline{G}_{a+1}^{I}(\mu_E)$. In turn, this implies that the reputation distribution for each cohort as a whole improves over time as each cohort ages:

Theorem 6 (Reputation across cohorts)**.** *The (prior, interim) reputation of age-$a+1$ firms first-order stochastically dominates the (prior, interim) reputation of age-a firms both, conditional and unconditional on types:*

$$(\forall a \in \mathbb{N})\,(\forall x \in [\lambda\phi, 1-\lambda+\lambda\phi]), \qquad \begin{array}{l} \overline{G}_{a+1}^{\tau}(x) \leq \overline{G}_a^{\tau}(x) \text{ and } \overline{G}_{a+1}(x) \leq \overline{G}_a(x) \\ G_{a+1}^{\tau}(x) \leq G_a^{\tau}(x) \text{ and } G_{a+1}(x) \leq G_a(x) \end{array}$$

Proof. See Appendix H. ◻

This "cleansing" effect is an imperfect, stochastic version of what happens in Hörner (2002). Figure 6 illustrates this in our example: older firms (i.e., names) have stochastically better reputations than younger ones, and yet not even the limit distributions are degenerate. This result is consistent with the empirical findings in McDevitt (2011) and Foster et al. (2008), where the exit probability is monotonically decreasing in age. Moreover, since better-reputation firms charge higher prices, this result is also consistent with the finding that older firms charge higher prices.

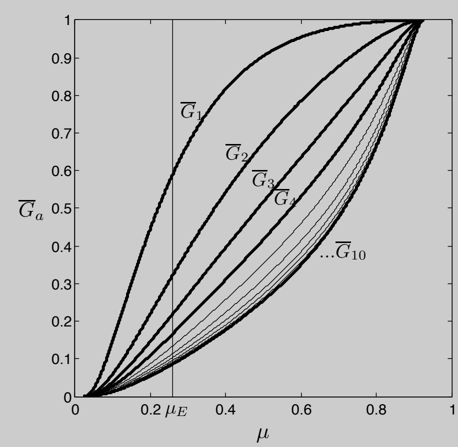


Figure 6: Reputation distributions for different ages, $\overline{G}_a(x)$

Notes: $r|_H \sim Be(r \mid 3, 2)$ and $r|_L \sim Be(r \mid 2, 3)$. The parameters are $\lambda = 0.1$, $\theta = 0.5$ and $\eta = 0.025$. μ_E is 0.26.

Remark 2. Theorem 6 implies that the mean reputation is monotonically increasing in age. The variance, however, is not. Figure 7 illustrates this in our example. As all age-0 firms (i.e., entrants) have the same reputation μ_E, the variance starts at 0. As time goes by and each firm gets a different realization of the signal process, variance increases. The cleansing effect that excludes from each cohort a disproportionate fraction of inept firms, however, brings the reputation of the surviving firms closer together.

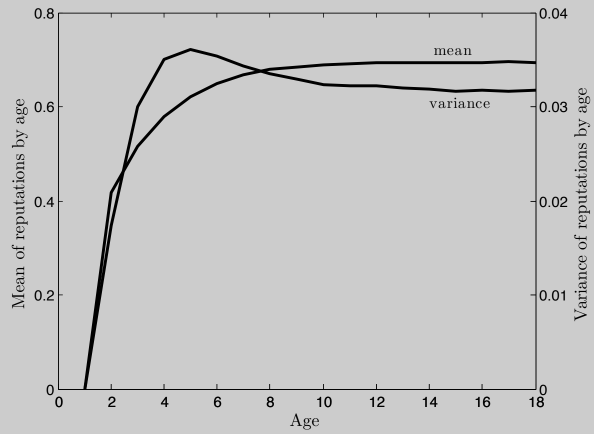


Figure 7: Mean and variance of reputations by age

Notes: $r|_H \sim Be\,(r \mid 3, 2)$ and $r|_L \sim Be\,(r \mid 2, 3)$. The parameters are $\lambda = 0.1$, $\theta = 0.5$ and $\eta = 0.025$. μ_E is 0.26.

References

Albuquerque, R., Hopenhayn, H. A., 2004. Optimal lending contracts and firm dynamics. The Review of Economic Studies 71 (2), pp. 285–315.
URL `http://www.jstor.org/stable/3700627`

Atkeson, A., Hellwig, C., Ordonez, G., March 2012. Optimal regulation in the presence of reputation concerns. Working Paper 17898, National Bureau of Economic Research.
URL `http://www.nber.org/papers/w17898`

Bar-Isaac, H., 2003. Reputation and survival: Learning in a dynamic signalling model. The Review of Economic Studies 70 (2), pp. 231–251.
URL `http://www.jstor.org/stable/3648633`

Bar-Isaac, H., Tadelis, S., 2008. Seller reputation. Foundations and Trends® in Microeconomics 4, 273–351.

Board, S., Meyer-ter Vehn, M., 2010. A reputational theory of firm dynamics.

Cabral, L., Hortacsu, A., 2010. The dynamics of seller reputation: Evidence from ebay. The Journal of Industrial Economics 58 (1), 54–78.

Clementi, G. L., Hopenhayn, H. A., 2006. A theory of financing constraints and firm dynamics. The Quarterly Journal of Economics 121 (1), pp. 229–265.
URL `http://www.jstor.org/stable/25098789`

Cooley, T. F., Quadrini, V., 2001. Financial markets and firm dynamics. The American Economic Review 91 (5), pp. 1286–1310.
URL http://www.jstor.org/stable/2677926

Cripps, M. W., Mailath, G. J., Samuelson, L., 2004. Imperfect monitoring and impermanent reputations. Econometrica 72 (2), pp. 407–432.
URL http://www.jstor.org/stable/3598908

De la Fuente, A., 2000. Mathematical methods and models for economists. Cambridge University Press.

Dunne, T., Roberts, M. J., Samuelson, L., 1988. Patterns of firm entry and exit in u.s. manufacturing industries. The RAND Journal of Economics 19 (4), pp. 495–515.
URL http://www.jstor.org/stable/2555454

Ekmekci, M., March 2011. Sustainable reputations with rating systems. Journal of Economic Theory 146 (2), 479–503.

Fishman, A., Rob, R., 2003. Consumer inertia, firm growth and industry dynamics. Journal of Economic Theory 109 (1), 24 – 38.
URL http://www.sciencedirect.com/science/article/pii/S0022053102000352

Fishman, A., Rob, R., 2005. Is bigger better? customer base expansion through word-of-mouth reputation. Journal of Political Economy 113 (5), 1146–1175.

Foster, L., Haltiwanger, J., Syverson, C., 2008. Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? The American Economic Review 98 (1), pp. 394–425.
URL http://www.jstor.org/stable/29729976

Fudenberg, D., Levine, D. K., 1989. Reputation and equilibrium selection in games with a patient player. Econometrica 57 (4), 759–778.
URL http://www.jstor.org/stable/1913771

Gretsky, N., Ostroy, J., Zame, W., 1999. Perfect Competition in the Continuous Assignment Model. Journal of Economic Theory 88 (1), 60–118.

Hakenes, H., Peitz, M., 2007. Observable reputation trading. International Economic Review 48 (2), 693–730.
URL http://dx.doi.org/10.1111/j.1468-2354.2007.00442.x

Hopenhayn, H., Rogerson, R., 1993. Job turnover and policy evaluation: A general equilibrium analysis. Journal of Political Economy 101 (5), pp. 915–938.
URL http://www.jstor.org/stable/2138602

Hopenhayn, H. A., 1992. Entry, exit, and firm dynamics in long run equilibrium. Econometrica 60 (5), pp. 1127–1150.
URL http://www.jstor.org/stable/2951541

Hörner, J., 2002. Reputation and competition. The American Economic Review 92 (3), 644–663.
URL http://www.jstor.org/stable/3083359

Klein, B., Leffler, K. B., 1981. The role of market forces in assuring contractual performance. The Journal of Political Economy 89 (4), 615–641.
URL http://www.jstor.org/stable/1833028

Levine, D., Martinelli, C., 1998. Reputation with noisy precommitment. Journal of Economic Theory 78 (1), 55–75.

Liu, Q., 2011. Information acquisition and reputation dynamics. The Review of Economic Studies 78 (4), 1400–1425.

Liu, Q., Skrzypacz, A., Jul 2009. Limited records and reputation. Research Papers 2030, Stanford University, Graduate School of Business.
URL http://ideas.repec.org/p/ecl/stabus/2030.html

Mailath, G., Samuelson, L., 2006. Repeated games and reputations: long-run relationships. Oxford University Press, USA.

Mailath, G. J., Samuelson, L., 2001. Who wants a good reputation? The Review of Economic Studies 68 (2), 415–441.
URL http://www.jstor.org/stable/2695935

McDevitt, R. C., 2011. Names and reputations: An empirical analysis. AEJ: Microeconomics 3 (3), 193–209.

Nelson, P., 1970. Information and consumer behavior. The Journal of Political Economy 78 (2), 311–329.
URL http://www.jstor.org/stable/1830691

Ordonez, G. L., ???? Fragility of reputation and clustering of risk-taking. Theoretical Economics.
URL http://ideas.repec.org/a/the/publsh/1207.html

Tadelis, S., 1999. What's in a name? reputation as a tradeable asset. The American Economic Review 89 (3), 548–563.
URL http://www.jstor.org/stable/117032

Tadelis, S., 2002. The market for reputations as an incentive mechanism. The Journal of Political Economy 110 (4), 854–882.
URL http://www.jstor.org/stable/3078538

Vial, B., 2010. Walrasian equilibrium and reputation under imperfect public monitoring. The BE Journal of Theoretical Economics, Advances 10 (1).

Wang, T., November 2011. The dynamics of names: A model of reputation. International Economic Review 52 (4), 1039–1058.

Appendices

A Computation of the reputation distributions

This section derives Equation (14), which gives the distributions of reputation for competent and inept types, $\Pr\left(\overline{\mu}_{t+1} \leq x \mid \tau_{t+1}\right) = \overline{G}_{t+1}^{\tau_{t+1}}(x)$.

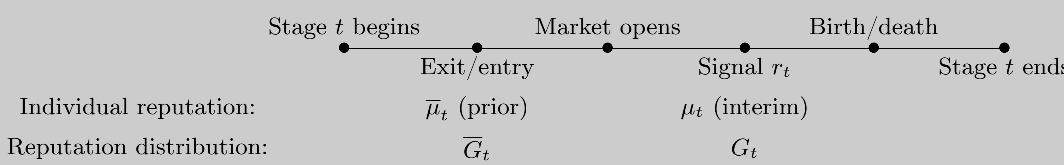


Figure 8: Time line for the date-t stage game

Let $H_{\overline{\mu}_{t+1},\mu_t,\tau_{t+1},\tau_t,r_t}$ denote the joint cdf of the random variables $\left(\overline{\mu}_{t+1}, \mu_t, r_t, \tau_t, \tau_{t+1}\right)$ after the time-t exit-entry process and before the time-$t+1$ exit-entry process. At the beginning of this time interval, the firms' types are τ_t and their reputation levels are μ_t; after r_t is observed and the birth-death process occurs, firms' types are τ_{t+1} and consumers' beliefs are updated to $\overline{\mu}_{t+1}$. H is a mixed distribution: while μ_t, $\overline{\mu}_{t+1}$ and r_t are continuous random variables, τ_t and τ_{t+1} are discrete. Moreover, μ_t has a point mass at y. The marginal cdf of H over μ_t is denoted by $G_t(\mu_t)$; notice that it is discontinuous at y. The marginal cdf of H over r_t conditional on τ_t is denoted by F_{τ_t}, and assumed to have a density. Note that the distribution of r_t only depends on the type τ_t. Hence, the distributions of reputations and signals relate to H as follows:

$$\begin{aligned} \overline{G}_{t+1}^{\tau}(x) &\equiv H_{\overline{\mu}_{t+1}|\tau_{t+1}}(x \mid \tau) \quad & \overline{G}_{t+1}(x) &\equiv H_{\overline{\mu}_{t+1}}(x) \\ G_t^{\tau}(x) &\equiv H_{\mu_t|\tau_t}(x \mid \tau) & G_t(x) &\equiv H_{\mu_t}(x) \\ F_{\tau}(r) &\equiv H_{r_t|\tau_t}(r \mid \tau) \end{aligned}$$

On the other hand, the exogenous birth-death process among active firms at t is independent of any other variable, and is fully described by the joint distribution in Table (2):

		Before (τ_t) C	I	Marginal
After	C	$\theta(1-\lambda+\lambda\phi)$	$\lambda\phi(1-\theta)$	$\theta-\eta$
(τ_{t+1})	I	$\lambda\theta(1-\phi)$	$(1-\lambda\phi)(1-\theta)$	$1-\theta+\eta$
	Marginal	θ	$1-\theta$	

Table 2: Joint distribution of types among active firms at t

Note: Recall that $\lambda(\theta-\phi)=\eta$.

The marginal distribution over $\overline{\mu}_{t+1}$ (before time $t+1$ exit-entry process), conditional on $\tau_{t+1}=\tau'$, is the expectation of $H_{\overline{\mu}_{t+1}|\mu_t,r_t,\tau_t,\tau_{t+1}}$ over μ_t, r_t, and τ_t.

$$\begin{aligned} \overline{G}_{t+1}^{\tau'}(x) &= \Pr\left(\overline{\mu}_{t+1} \leq x \mid \tau_{t+1}=\tau'\right) \\ &= H_{\overline{\mu}_{t+1}|\tau_{t+1}}\left(x \mid \tau'\right) \\ &= E_{\tau_t|\tau_{t+1}}\left[E_{r_t|\tau_t,\tau_{t+1}}\left[E_{\mu_t|r_t,\tau_t,\tau_{t+1}}\left[H_{\overline{\mu}_{t+1}|\mu_t,r_t,\tau_t,\tau_{t+1}}\left(x \mid \mu, r, \tau, \tau'\right)\right]\right]\right] \end{aligned}$$

Notice that there is a functional relation between $\overline{\mu}_{t+1}$ and (μ_t, r_t), as established by Equation (11):

$$\overline{\mu}_{t+1} = \varphi(\mu_t, r_t)$$

This implies that $\overline{\mu}_{t+1}$ and (τ_t, τ_{t+1}) are conditionally independent given (μ_t, r_t). Hence, $H_{\overline{\mu}_{t+1}|\mu_t,r_t,\tau_t,\tau_{t+1}}(x \mid \mu, r, \tau, \tau') = H_{\overline{\mu}_{t+1}|\mu_t,r_t}(x \mid \mu, r)$. Moreover, given (μ_t, r_t), $\overline{\mu}_{t+1}$ becomes determinate, so that:

$$H_{\overline{\mu}_{t+1}|\mu_t,r_t}(x \mid \mu, r) = \mathbf{1}_{\{\varphi(\mu,r)\le x\}}$$

Thus, we can write:

$$\overline{G}^{\tau'}_{t+1}(x) = E_{\tau_t|\tau_{t+1}}\left[E_{r_t|\tau_t,\tau_{t+1}}\left[E_{\mu_t|r_t,\tau_t,\tau_{t+1}}\left[\mathbf{1}_{\{\varphi(\mu,r)\le x\}}\left(x \mid \mu, r, \tau, \tau'\right)\right]\right]\right]$$

There are different ways in which we can compute this expectation, depending on the order of integration we choose. If we start with the expectation over the current type, we get:

$$\overline{G}^{\tau'}_{t+1}(x) = \sum_{\tau_t} \Pr\left(\tau_t \mid \tau_{t+1} = \tau'\right) E_{r_t|\tau_t,\tau_{t+1}}\left[E_{\mu_t|r_t,\tau_t,\tau_{t+1}}\left[\mathbf{1}_{\{\varphi(\mu,r)\le x\}}\left(x \mid \mu, r, \tau, \tau'\right)\right]\right]$$

Recall from Equations (13) and (12) that we write:

$$\mu = \tilde{\mu}(x, r) \text{ and } r = \tilde{r}(x, \mu)$$

for the implicit functions for μ and r from $x = \varphi(\mu, r)$. Hence:

$$E_{\mu_t|r_t,\tau_t,\tau_{t+1}}\left[\mathbf{1}_{\{\varphi(\mu,r)\le x\}}\left(x \mid \mu, r, \tau, \tau'\right)\right] = E_{\mu_t|r_t,\tau_t,\tau_{t+1}}\left[\mathbf{1}_{\{\mu\le\tilde{\mu}(x,r)\}}\left(x \mid \mu, r, \tau, \tau'\right)\right] = G^{\tau}_t(\tilde{\mu}(x, r))$$

so that we get:

$$\begin{aligned}\overline{G}^{\tau'}_{t+1}(x) &= \Pr\left(\tau_t = C \mid \tau_{t+1} = \tau'\right) E_{r_t|\tau_t,\tau_{t+1}}\left[G^C_t(\tilde{\mu}(x, r)) \mid r, C, \tau'\right] \\ &\quad + \Pr\left(\tau_t = I \mid \tau_{t+1} = \tau'\right) E_{r_t|\tau_t,\tau_{t+1}}\left[G^I_t(\tilde{\mu}(x, r)) \mid r, I, \tau'\right]\end{aligned}$$

On the other hand, the expectation $E_{r_t|\tau_t,\tau_{t+1}}\left[G^{\tau}_t(\tilde{\mu}(x, r)) \mid r, \tau, \tau'\right]$ must be computed bearing in mind that r induces a discontinuous distribution over $G^C_t(\tilde{\mu}(x, r_t))$, with point mass at y. Thus, we split it into the two events separated by the point of discontinuity:

$$\begin{aligned}E_{r_t|\tau_t,\tau_{t+1}}\left[G^C_t(\tilde{\mu}(x, r_t)) \mid r, C, \tau'\right] &= \Pr\left(r_t > \tilde{r}(x, y) \mid r, C, \tau'\right) E_{r_t|\tau_t,\tau_{t+1}}\left[G^C_t(\tilde{\mu}(x, r_t)) \mid r_t > \tilde{r}(x, y), C, \tau'\right] \\ &\quad + \Pr\left(r_t \le \tilde{r}(x, y) \mid r, C, \tau'\right) E_{r_t|\tau_t,\tau_{t+1}}\left[G^C_t(\tilde{\mu}(x, r_t)) \mid r_t \le \tilde{r}(x, y), C, \tau'\right]\end{aligned}$$

where $E_{r_t|\tau_t,\tau_{t+1}}\left[G^C_t(\tilde{\mu}(x, r_t)) \mid r_t > \tilde{r}(x, y), C, \tau'\right] = 0$, so that:

$$E_{r_t|\tau_t,\tau_{t+1}}\left[G^C_t(\tilde{\mu}(x, r_t)) \mid r_t, C, \tau'\right] = \int_0^{\tilde{r}(x,y)} G^C_t(\tilde{\mu}(x, r_t))\, dF_H$$

Similarly,

$$E_{r_t|\tau_t,\tau_{t+1}}\left[G^I_t(\tilde{\mu}(x, r_t)) \mid r_t, I, \tau'\right] = \int_0^{\tilde{r}(x,y)} G^I_t(\tilde{\mu}(x, r_t))\, dF_L$$

Plugging into the equation for $\overline{G}^C_{t+1}(x)$, we get:

$$\overline{G}^C_{t+1}(x) = \frac{\theta(1-\lambda+\lambda\phi)}{\theta-\eta} \int_0^{\tilde{r}(x,y)} G^C_t(\tilde{\mu}(x, r_t))\, dF_H + \frac{\lambda\theta(1-\phi)}{\theta-\eta} \int_0^{\tilde{r}(x,y)} G^I_t(\tilde{\mu}(x, r_t))\, dF_L \tag{25}$$

An analogous equation is obtained for $G^I_{t+1}(x)$:

$$\overline{G}^I_{t+1}(x) = \frac{\lambda\theta(1-\phi)}{1-\theta+\eta} \int_0^{\tilde{r}(x,y)} G^C_t(\tilde{\mu}(x, r_t))\, dF_H + \frac{(1-\theta)(1-\lambda\phi)}{1-\theta+\eta} \int_0^{\tilde{r}(x,y)} G^I_t(\tilde{\mu}(x, r_t))\, dF_L \tag{26}$$

B Proof of Lemma 3

To lighten notation we supress the reference to y as an argument of T and the distributions.

We start by establishing that the operator T defined by Equation (**??**) is a contraction mapping in the set of pairs of continuous, normalized[10] functions $(\overline{G}^C, \overline{G}^I)$ endowed with the following metric:

$$\rho\left(\left(\overline{G}^C, \overline{G}^I\right), \left(\overline{G}'^C, \overline{G}'^I\right)\right) = \max\left\{\rho_\infty\left(\overline{G}^C, \overline{G}'^C\right), \rho_\infty\left(\overline{G}^I, \overline{G}'^I\right)\right\},$$

where:

$$\rho\left(\overline{G}^\tau, \overline{G}'^\tau\right) = \sup_{x\in[\lambda\phi, 1-\lambda+\lambda\phi]}\left|\overline{G}^\tau(x) - \overline{G}'^\tau(x)\right|$$

for $\tau \in \{C, I\}$. The supremum is taken over $x \in [\lambda\phi, 1-\lambda+\lambda\phi]$ since the supports of $\overline{G}$ and $\overline{G}'$ are always contained in this interval.

Recall that the operator T transforms the pair $\left(\overline{G}_t^C, \overline{G}_t^I\right)$ into a pair $\left(\overline{G}_{t+1}^C, \overline{G}_{t+1}^I\right)$ according to:

$$\begin{aligned}\begin{pmatrix}\overline{G}_{t+1}^C(x) \\ \overline{G}_{t+1}^I(x)\end{pmatrix} &= \begin{pmatrix}\frac{(1-\lambda+\lambda\phi)}{\theta-\eta}\eta & -\frac{\lambda\phi}{\theta-\eta}\eta \\ \frac{\lambda(1-\phi)}{1-\theta+\eta}\eta & -\frac{(1-\lambda\phi)}{1-\theta+\eta}\eta\end{pmatrix}\begin{pmatrix}F_H(\tilde{r}(x,y)) \\ F_L(\tilde{r}(x,y))\end{pmatrix} \\ &+ \begin{pmatrix}(1-\lambda+\lambda\phi) & \frac{\lambda\phi(1-\theta+\eta)}{\theta-\eta} \\ \frac{\lambda(1-\phi)(\theta-\eta)}{1-\theta+\eta} & (1-\lambda\phi)\end{pmatrix}\begin{pmatrix}\int_0^{\tilde{r}(x,y)}\overline{G}_t^C(\tilde{\mu}(x,r))\,dF_H \\ \int_0^{\tilde{r}(x,y)}\overline{G}_t^I(\tilde{\mu}(x,r))\,dF_L\end{pmatrix} \end{aligned} \tag{27}$$

and where $y \in [0,1]$ is treated as a constant.

First notice that there are no firms with reputation either below y or above $1-\lambda+\lambda\phi$ after entry-exit decisions are made; therefore, we have:

$$\begin{aligned}\overline{G}_t^\tau(\mu) &= \overline{G}_t'^\tau(\mu) = 0 \text{ if } \mu < y \text{ and} \\ \overline{G}_t^\tau(\mu) &= \overline{G}_t'^\tau(\mu) = 1 \text{ if } \mu > 1-\lambda+\lambda\phi\end{aligned}$$

for any distribution of reputations. Moreover, the definitions of $\tilde{r}(x,\mu)$ and $\tilde{\mu}(x,r)$ imply that:

1. $\tilde{\mu}(x, \tilde{r}(x,y)) = y$. This is to say, the previous reputation of a firm who obtained a signal $\tilde{r}(x,y)$ that changed its reputation from y to x was y;
2. $\tilde{\mu}(x,r) < y \Leftrightarrow r > \tilde{r}(x,y)$: Those firms who have a reputation x today and had a reputation lower than y the previous period are those who obtained signals of at least $\tilde{r}(x,y)$; and
3. $\tilde{\mu}(x,r) > 1-\lambda+\lambda\phi \Leftrightarrow r < \tilde{r}(x, 1-\lambda+\lambda\phi)$: Those firms who had a higher reputation than $1-\lambda+\lambda\phi$ the previous period and have a reputation x today are those whose signals were lower than $\tilde{r}(x, 1-\lambda+\lambda\phi)$.

Hence,

$$\begin{aligned}\overline{G}_t^\tau(\tilde{\mu}(x,r_t)) &= \overline{G}_t'^\tau(\tilde{\mu}(x,r_t)) = 0 \quad \text{if } r > \tilde{r}(x,y) \quad \text{and} \\ \overline{G}_t^\tau(\tilde{\mu}(x,r_t)) &= \overline{G}_t'^\tau(\tilde{\mu}(x,r_t)) = 1 \quad \text{if } r < \tilde{r}(x, 1-\lambda+\lambda\phi).\end{aligned}$$

Therefore, the distance between $\overline{G}_{t+1}^C$ and $\overline{G}_{t+1}'^C$:

$$\rho_\infty\left(\overline{G}_{t+1}^C, \overline{G}_{t+1}'^C\right) = \sup_x\left|\overline{G}_{t+1}^C(x) - \overline{G}_{t+1}'^C(x)\right|$$

can be rewritten as:

$$\begin{aligned}\rho_\infty\left(\overline{G}_{t+1}^C, \overline{G}_{t+1}'^C\right) = \sup_x\Bigg|&(1-\lambda+\lambda\phi)\int_{\tilde{r}(x,1-\lambda+\lambda\phi)}^{\tilde{r}(x,y)}\left(\overline{G}_t^C(\tilde{\mu}(x,r_t)) - \overline{G}_t'^C(\tilde{\mu}(x,r_t))\right)dF_H \\ &+\lambda\phi\frac{(1-\theta+\eta)}{\theta-\eta}\int_{\tilde{r}(x,1-\lambda+\lambda\phi)}^{\tilde{r}(x,y)}\left(\overline{G}_t^I(\tilde{\mu}(x,r_t)) - \overline{G}_t'^I(\tilde{\mu}(x,r_t))\right)dF_L\Bigg|.\end{aligned}$$

Using the properties of the sup and $|\cdot|$ operators, we obtain:

$$\begin{aligned}\rho_\infty\left(\overline{G}_{t+1}^C, \overline{G}_{t+1}'^C\right) \leq (1-\lambda+\lambda\phi)\sup_x&\left|\int_{\tilde{r}(x,1-\lambda+\lambda\phi)}^{\tilde{r}(x,y)}\left(\overline{G}_t^C(\tilde{\mu}(x,r_t)) - \overline{G}_t'^C(\tilde{\mu}(x,r_t))\right)dF_H\right| \\ + \lambda\phi\frac{1-\theta+\eta}{\theta-\eta}\sup_x&\left|\int_{\tilde{r}(x,1-\lambda+\lambda\phi)}^{\tilde{r}(x,y)}\left(\overline{G}_t^I(\tilde{\mu}(x,r_t)) - \overline{G}_t'^I(\tilde{\mu}(x,r_t))\right)dF_L\right|\end{aligned}$$

[10]The function $\overline{G}^\tau$ is normalized if $\overline{G}^\tau(\lambda\phi) = 0$ and $\overline{G}^\tau(1-\lambda+\lambda\phi) = 1$.

Moving the absolute value inside the integrals:

$$\rho_\infty\left(\overline{G}_{t+1}^C,\overline{G}_{t+1}^{\prime C}\right)\le(1-\lambda+\lambda\phi)\sup\int_{\tilde r(x,1-\lambda+\lambda\phi)}^{\tilde r(x,y)}\left|\overline{G}_t^C\left(\tilde\mu\left(x,r_t\right)\right)-\overline{G}_t^{\prime C}\left(\tilde\mu\left(x,r_t\right)\right)\right|dF_H$$
$$+\lambda\phi\frac{1-\theta+\eta}{\theta-\eta}\sup_x\int_{\tilde r(x,1-\lambda+\lambda\phi)}^{\tilde r(x,y)}\left|\overline{G}_t^I\left(\tilde\mu\left(x,r_t\right)\right)-\overline{G}_t^{\prime I}\left(\tilde\mu\left(x,r_t\right)\right)\right|dF_L$$

By the definition of $\rho_\infty\left(G_t^\tau,H_t^\tau\right)$, we also know that:

$$\rho_\infty\left(\overline{G}_{t+1}^C,\overline{G}_{t+1}^{\prime C}\right)\le(1-\lambda+\lambda\phi)\sup_x\int_{\tilde r(x,1-\lambda+\lambda\phi)}^{\tilde r(x,y)}\rho_\infty\left(\overline{G}_t^C,\overline{G}_t^{\prime C}\right)dF_H$$
$$+\lambda\phi\frac{1-\theta+\eta}{\theta-\eta}\sup_x\int_{\tilde r(x,1-\lambda+\lambda\phi)}^{\tilde r(x,y)}\rho_\infty\left(\overline{G}_t^I,\overline{G}_t^{\prime I}\right)dF_L$$

Taking the distances outside the integrals, we get:

$$=(1-\lambda+\lambda\phi)\,\rho_\infty\left(\overline{G}_t^C,\overline{G}_t^{\prime C}\right)\sup_x\int_{\tilde r(x,1-\lambda+\lambda\phi)}^{\tilde r(x,y)}dF_H$$
$$+\lambda\phi\frac{1-\theta+\eta}{\theta-\eta}\rho_\infty\left(\overline{G}_t^I,\overline{G}_t^{\prime I}\right)\sup_x\int_{\tilde r(x,1-\lambda+\lambda\phi)}^{\tilde r(x,y)}dF_L$$

But $\int_{\tilde r(x,1-\lambda+\lambda\phi)}^{\tilde r(x,y)}dF=F\left(\tilde r\left(x,y\right)\right)-F\left(\tilde r\left(x,1-\lambda+\lambda\phi\right)\right)<1$ for all $x\in[\lambda\phi,1-\lambda+\lambda\phi]$. Let us define β as follows:

$$\beta=\max\left\{\sup_{x\in[\lambda\phi,1-\lambda+\lambda\phi]}\left(F_H\left(\tilde r\left(x,y\right)\right)-F_H\left(\tilde r\left(x,1-\lambda+\lambda\phi\right)\right)\right),\right.$$
$$\left.\sup_{x\in[\lambda\phi,1-\lambda+\lambda\phi]}\left(F_L\left(\tilde r\left(x,y\right)\right)-F_L\left(\tilde r\left(x,1-\lambda+\lambda\phi\right)\right)\right)\right\}.$$

Observe that $\beta\in(0,1)$. Then:

$$\rho_\infty\left(\overline{G}_{t+1}^C,\overline{G}_{t+1}^{\prime C}\right)\le\beta\left(1-\lambda+\lambda\phi\right)\rho_\infty\left(\overline{G}_t^C,\overline{G}_t^{\prime C}\right)$$
$$+\beta\lambda\phi\frac{1-\theta+\eta}{\theta-\eta}\rho_\infty\left(\overline{G}_t^I,\overline{G}_t^{\prime I}\right)$$
$$\le\beta\rho\left(\left(\overline{G}_t^C,\overline{G}_t^I\right),\left(\overline{G}_t^{\prime C},\overline{G}_t^{\prime I}\right)\right)$$

This is so because the sum of the coefficients is smaller than 1. Following a similar procedure, it can be shown that the distance between $\overline{G}_{t+1}^I$ and $\overline{G}_{t+1}^{\prime I}$ satisfies:

$$\rho_\infty\left(\overline{G}_{t+1}^I,\overline{G}_{t+1}^{\prime I}\right)\le\beta\rho\left(\left(\overline{G}_t^C,\overline{G}_t^I\right),\left(\overline{G}_t^{\prime C},\overline{G}_t^{\prime I}\right)\right).$$

We thus conclude that:

$$\rho\left(\left(\overline{G}_{t+1}^C,\overline{G}_{t+1}^{\prime C}\right),\left(\overline{G}_{t+1}^I,\overline{G}_{t+1}^{\prime I}\right)\right)\le\beta\rho\left(\left(\overline{G}_t^C,\overline{G}_t^I\right),\left(\overline{G}_t^{\prime C},\overline{G}_t^{\prime I}\right)\right),$$

i.e., T is a contraction mapping.

On the other hand, the set of continuous, bounded real functions endowed with the sup norm is complete. Moreover, the subset of normalized functions is closed,[11] thereby complete. Hence, by Banach's Fixed Point Theorem T has a unique fixed point, which is a pair of continuous and normalized functions. □

Remark 3. If we had used $\lambda\phi$ instead of y as the lower bound of reputations in the proof, we would have obtained a higher modulus β–but we would still be able to prove that T is a contraction mapping–.

Remark 4. y not only affects the modulus of the contraction, but also the limiting distribution.

Remark 5. If y is consistent, $\overline{G}^C(x)$ and $\overline{G}^I(x)$ are increasing functions because $G^C(x)$ and $G^I(x)$ are non-negative in the whole domain, while $\tilde r(x,y)$ is increasing in x. Thus, $\overline{G}^C$ and $\overline{G}^I$ are not only normalized and continuous, but also increasing–i.e., distribution functions.

Thus, the proof of the existence of a steady-state distribution of reputations is not affected by the endogenous entry-exit process as long as y is fixed, but the shape of the steady-state distribution is. The absolute continuity of F_H an F_L implies that $\overline{G}^C$ and $\overline{G}^I$ are absolutely continuous, with common support $[\lambda\phi,1-\lambda+\lambda\phi]$.

[11]See Lemma 1 in Vial, 2010 for a proof.

C Proof of Proposition 1

Define the function $\sigma(\mu) \equiv \psi(\mu, \mu)$, with $\mu > \lambda\phi$, as:

$$\psi(x,y) = \frac{(\theta-\eta)\,\overline{G}^C(x,y) + \eta}{(\theta-\eta)\,\overline{G}^C(x,y) + (1-\theta+\eta)\,\overline{G}^I(x,y)}.$$

with $x = y = \mu$, i.e., Equation (17) evaluated at the diagonal. We need to prove that $\sigma(\mu)$ has a unique fixed point. We begin by observing that:

Lemma 5. *σ has at least one fixed point.*

Proof. The function $f(x) \equiv \overline{G}^I(x,x)$ is continuous, with $f(\lambda\phi) = 0$ and $f(1-\lambda+\lambda\phi) = 1$. Hence, by the intermediate value theorem there is at least one $\xi \in (\lambda\phi, 1-\lambda+\lambda\phi)$ such that $f(\xi) = \frac{\eta}{1-\theta+\eta}$, and so $\sigma(\xi) = 1$. We also know that σ is continuous in its domain, and that $\sigma(\xi) - \xi = 1 - \xi > 0$ and $\sigma(1-\lambda+\lambda\phi) - (1-\lambda+\lambda\phi) = \theta - (1-\lambda+\lambda\phi) < 0$ (this follows from the assumption that $\lambda < 1-\theta$). By the intermediate value theorem, then, there is at least one $\mu \in (\xi, 1-\lambda+\lambda\phi)$ such that $\sigma(\mu) - \mu = 0$. Hence, there is at least one $\mu_E \in (\lambda\phi, 1-\lambda+\lambda\phi)$ such that $\psi(\mu_E, \mu_E) = \mu_E$. □

The next step is to establish uniqueness.

Lemma 6. *$\sigma'(\mu_E) = 0$ if μ_E is a fixed point. Hence, the fixed point is unique.*

Proof. Indeed,

$$\sigma'(\mu_E)\,d\mu_E = \frac{\partial\psi}{\partial x}dx + \frac{\partial\psi}{\partial y}dy$$

evaluated at $x = y = \mu_E$ and $dx = dy = d\mu_E$.

The first term corresponds to:

$$\frac{\partial\psi}{\partial x}(x,y) = \left(\frac{(\theta-\eta)\,\overline{g}^C(x,y) + (1-\theta+\eta)\,\overline{g}^I(x,y)}{(\theta-\eta)\,\overline{G}^C(x,y) + (1-\theta+\eta)\,\overline{G}^I(x,y)}\right)\left(\frac{(\theta-\eta)\,\overline{g}^C(x,y)}{(\theta-\eta)\,\overline{g}^C(x,y) + (1-\theta+\eta)\,\overline{g}^I(x,y)} - \psi(x,y)\right)$$

However, consitency requires that $\frac{a\overline{g}^C(x,y)}{a\overline{g}^C(x,y)+(1-a)\overline{g}^I(x,y)} = x$, namely, the fraction of competent among those firms with reputation x is exactly x. Moreover, at a fixed point $\psi(x,y) = x$. Hence, $\frac{\partial\psi}{\partial x} = 0$ at $x = y = \mu_E$. In words, the entrants' reputation $\psi(x,y)$ increases when the exit reputation level increases if and only if the firms that leave and reenter after this change have a higher reputation than those that are already replacing their names. At the fixed point, however, those firms have exactly the same mean reputation, so moving the cutuff point will have no effect on the entrants' reputation.

In turn, the second term corresponds to

$$\begin{aligned}\frac{\partial\psi}{\partial y}(x,y) &= \frac{\partial\psi(x,y)}{\partial\overline{G}^C}\frac{\partial\overline{G}^C}{\partial y}(x,y) + \frac{\partial\psi(x,y)}{\partial\overline{G}^I}\frac{\partial\overline{G}^I}{\partial y}(x,y),\\ &= \frac{1}{\overline{G}(x,y)}\left(a(1-\psi(x,y))\frac{\partial\overline{G}^C}{\partial y}(x,y) - (1-a)\,\psi(x,y)\frac{\partial\overline{G}^I}{\partial y}(x,y)\right) \qquad (28)\end{aligned}$$

and $\psi(x,y) = x$ if $x = y = \mu_E$ is a fixed point.

After a change of variables, the operator T in Equation 16 can be rewritten as:

$$\begin{aligned}T\begin{pmatrix}\overline{G}_t^C\\ \overline{G}_t^I\end{pmatrix}(x,y) &= \eta\begin{pmatrix}\frac{(1-\lambda+\lambda\phi)}{\theta-\eta} & -\frac{\lambda\phi}{\theta-\eta}\\ \frac{\lambda(1-\phi)}{1-\theta+\eta} & -\frac{(1-\lambda\phi)}{1-\theta+\eta}\end{pmatrix}\begin{pmatrix}F_H(\tilde r(x,y))\\ F_L(\tilde r(x,y))\end{pmatrix}\\ &- \begin{pmatrix}(1-\lambda+\lambda\phi) & \frac{\lambda\phi(1-\theta+\eta)}{\theta-\eta}\\ \frac{\lambda(1-\phi)(\theta-\eta)}{1-\theta+\eta} & (1-\lambda\phi)\end{pmatrix}\begin{pmatrix}\int_y^{\tilde\mu(x,0)}\overline{G}_t^C(\mu,y)\,f_H(\tilde r(x,\mu))\frac{\partial\tilde r(x,\mu)}{\partial\mu}d\mu\\ \int_y^{\tilde\mu(x,0)}\overline{G}_t^I(\mu,y)\,f_L(\tilde r(x,\mu))\frac{\partial\tilde r(x,\mu)}{\partial\mu}d\mu\end{pmatrix}\\ &\equiv \begin{pmatrix}T^C(x,y)\\ T^I(x,y)\end{pmatrix} \qquad (29)\end{aligned}$$

For a given y, the (unique pair of) steady-state distriutions $\overline{G}^C$ and $\overline{G}^I$ satisfy $\begin{pmatrix} \overline{G}^C \\ \overline{G}^I \end{pmatrix}(x,y) = T\begin{pmatrix} \overline{G}^C \\ \overline{G}^I \end{pmatrix}(x,y)$. In particular, let us consider as the initial steady-state distributions those obtained for $y=\mu_E$ (i.e., for a consistent entry-level reputation). Changing y affect $\overline{G}^C$ and $\overline{G}^I$ in two ways: (i) as y affects the operator itself, it also affects its fixed point (direct effect); and (ii) as the distributions also affect the operator, they also affect its fixed point (indirect effect):

$$\begin{pmatrix} \frac{\partial \overline{G}^C}{\partial y} \\ \frac{\partial \overline{G}^I}{\partial y} \end{pmatrix}(x,y) = \begin{pmatrix} \left.\frac{\partial T^C}{\partial y}\right|_{\overline{G}} \\ \left.\frac{\partial T^I}{\partial y}\right|_{\overline{G}} \end{pmatrix}(x,y) + \begin{pmatrix} \int_y^{\tilde{\mu}(x,0)} \frac{\partial \overline{G}^C}{\partial y}(\mu,y) f_H(\tilde{r}(x,\mu)) \frac{\partial \tilde{r}(x,\mu)}{\partial \mu} d\mu + \int_y^{\tilde{\mu}(x,0)} \frac{\partial \overline{G}^I}{\partial y}(\mu,y) f_L(\tilde{r}(x,\mu)) \frac{\partial \tilde{r}(x,\mu)}{\partial \mu} d\mu \\ \int_y^{\tilde{\mu}(x,0)} \frac{\partial \overline{G}^C}{\partial y}(\mu,y) f_H(\tilde{r}(x,\mu)) \frac{\partial \tilde{r}(x,\mu)}{\partial \mu} d\mu + \int_y^{\tilde{\mu}(x,0)} \frac{\partial \overline{G}^I}{\partial y}(\mu,y) f_L(\tilde{r}(x,\mu)) \frac{\partial \tilde{r}(x,\mu)}{\partial \mu} d\mu \end{pmatrix} \tag{30}$$

The direct effect is obtained from direct computation of the derivatives on Equation 29 (when evaluated at the steady-state distributions obtained for $y=\mu_E$):

$$\begin{pmatrix} \left.\frac{\partial T^C}{\partial y}\right|_{\overline{G}} \\ \left.\frac{\partial T^I}{\partial y}\right|_{\overline{G}} \end{pmatrix}(x,y) = \frac{\partial \tilde{r}(x,y)}{\partial y} \begin{pmatrix} \frac{\theta(1-\lambda+\lambda\phi)}{\theta-\eta} & \frac{\lambda\phi(1-\theta)}{\theta-\eta} \\ \frac{\theta\lambda(1-\phi)}{1-\theta+\eta} & \frac{(1-\lambda\phi)(1-\theta)}{1-\theta+\eta} \end{pmatrix} \begin{pmatrix} G^C(y,y) f_H(\tilde{r}(x,y)) \\ G^I(y,y) f_L(\tilde{r}(x,y)) \end{pmatrix} \tag{31}$$

Equation 30 defines an integral equations system for $\frac{\partial \overline{G}^C}{\partial y}$ and $\frac{\partial \overline{G}^C}{\partial y}$. However, instead or looking for a solution of this integral equations system, we can iterate the operator T to show that 28 equals zero when evaluated at a fixed point: since the steady-state distributions are the fixed point of a contraction mapping in a complete metric space, they can be obtained as the limit of the sequence defined by repeatedly iterating T starting from any feasible pair $\overline{G}_0^C$ and $\overline{G}_0^I$.

Let denote the t-th iteration of T as $\overline{G}_t^C$ and $\overline{G}_t^I$, and define $\psi_t(x,y)$ as $\frac{\theta G_t^C(x,y')}{\overline{G}_t(x,y')}$. If we start iterating T from the initial steady-state distributions (obtained with $y=\mu_E$), then $y=\psi_0(y,y)=\frac{\theta G_0^C(y,y)}{\overline{G}_0(y,y)}$. Notice that iterating T with the same y gives a constant sequence, since $\overline{G}_0^C$ and $\overline{G}_0^I$ are already fixed points of T for $y=\mu_E$. But after iterating T with a different level of y we obtain $\overline{G}_1^\tau \neq \overline{G}_0^\tau$; moreover, for an infinitesimally small change in y, the difference between them is exactly $\frac{\partial T^\tau}{\partial y}dy$. Hence, we can define $\overline{G}_1^\tau$ as:

$$\overline{G}_1^\tau(x,y') = \overline{G}_0^\tau(x,y) + \left.\frac{\partial T^\tau}{\partial y}\right|_{\overline{G}_0}(x,y)\,dy$$

When evaluating Equation 31 in G_0^C y G_0^I, and taking into account that $y=\frac{\theta G_0^C(y,y)}{\overline{G}_0(y,y)}$, we obtain:

$$\begin{pmatrix} \left.\frac{\partial T^C}{\partial y}\right|_{\overline{G}_0} \\ \left.\frac{\partial T^I}{\partial y}\right|_{\overline{G}_0} \end{pmatrix}(x,y) = \overline{G}_0(y,y)\frac{\partial \tilde{r}(x,y)}{\partial y} \begin{pmatrix} \frac{1}{\theta-\eta}\left((1-\lambda+\lambda\phi)\, y f_H(\tilde{r}(x,y)) + \lambda\phi(1-y) f_L(\tilde{r}(x,y))\right) \\ \frac{1}{1-\theta+\eta}\left(\lambda(1-\phi)\, y f_H(\tilde{r}(x,y)) + (1-\lambda\phi)(1-y) f_L(\tilde{r}(x,y))\right) \end{pmatrix} \tag{32}$$

Moreover, from Equations 8 and 12we obtain:

$$(1-\lambda+\lambda\phi)\, y f_H(\tilde{r}(x,y)) + (\lambda\phi)(1-y) f_L(\tilde{r}(x,y)) = x\left(y f_H(\tilde{r}(x,y)) + (1-y) f_L(\tilde{r}(x,y))\right) \tag{33}$$

Hence, the direct effect can be rewritten as:

$$\begin{pmatrix} \left.\frac{\partial T^C}{\partial y}\right|_{\overline{G}_0} \\ \left.\frac{\partial T^I}{\partial y}\right|_{\overline{G}_0} \end{pmatrix}(x,y) = \overline{G}_0(y,y)\frac{\partial \tilde{r}(x,y)}{\partial y}\left(y f_H(\tilde{r}(x,y)) + (1-y) f_L(\tilde{r}(x,y))\right) \begin{pmatrix} \frac{x}{\theta-\eta} \\ \frac{1-x}{1-\theta+\eta} \end{pmatrix}$$

and the pair $\left(\overline{G}_1^C, \overline{G}_1^I\right)$ can be defined as:

$$\begin{pmatrix} \overline{G}_1^C \\ \overline{G}_1^I \end{pmatrix}(x,y') = \begin{pmatrix} \overline{G}_0^C \\ \overline{G}_0^I \end{pmatrix}(x,y) + \overline{G}_0(y,y)\,\upsilon_0(x,y') \begin{pmatrix} \frac{x}{\theta-\eta} \\ \frac{1-x}{1-\theta+\eta} \end{pmatrix} \tag{34}$$

with $\upsilon_0(x,y') \equiv \frac{\partial \tilde{r}(x,y)}{\partial y}(yf_H(\tilde{r}(x,y)) + (1-y) f_L(\tilde{r}(x,y)))\, dy$, while ψ_1 can be obtained as:

$$\psi_1(x,y') = \frac{(\theta-\eta)\overline{G}_0^C(x,y) + x\overline{G}_0(y,y)\,\upsilon_0(x,y) + \eta}{\overline{G}_0(x,y) + \overline{G}_0(y,y)\,\upsilon_0(x,y)}$$

After evaluating at $x = y$ and rearranging, we obtain:

$$\psi_1(y,y') = \frac{\psi_0(y,y)(1+\upsilon_0(x,y)\,dy)}{1+\upsilon_0(x,y)\,dy} = \psi_0(y,y)$$

Repeatedly iterating the operator T we obtain a sequence of distributions defined as:

$$\begin{pmatrix} \overline{G}_t^C \\ \overline{G}_t^I \end{pmatrix}(x,y') = \begin{pmatrix} \overline{G}_{t-1}^C \\ \overline{G}_{t-1}^I \end{pmatrix}(x,y') + \overline{G}_0(y,y)\,\upsilon_{t-1}(x,y')\begin{pmatrix} \frac{x}{\theta-\eta} \\ \frac{1-x}{1-\theta+\eta} \end{pmatrix} \tag{35}$$

where $\upsilon_t(x,y') \equiv -\int_{y'}^{\tilde{\mu}(x,0)} \upsilon_{t-1}(\mu,y')(\mu f_H(\tilde{r}(x,\mu)) + (1-\mu) f_L(\tilde{r}(x,\mu)))\frac{\partial \tilde{r}(x,\mu)}{\partial \mu} d\mu$ for all $t > 0$. Hence, we conclude that $\psi_t(y,y') = \psi_0(y,y)$ for all $t > 0$, and hence $\frac{\partial \psi}{\partial y}(y,y) = 0$ when $y = \mu_E$. □

Finally, we conclude that $\mu_E < \theta$:

Lemma 7. *The consistent entry-level reputation is lower than the fraction of competents among active firms after the exit-entry process takes place: $\mu_E < \theta$.*

Proof. Notice that $\psi(\mu_E,\mu_E) = \mu_E$ and $\psi(1-\lambda+\lambda\phi,\mu_E) = \theta$. Moreover,

$$\frac{\partial \psi}{\partial x}(x,\mu_E) = \left(\frac{(\theta-\eta)\,\overline{g}^C(x,\mu_E) + (1-\theta+\eta)\,\overline{g}^I(x,\mu_E)}{\overline{G}(x,\mu_E)}\right)(x - \psi(x,\mu_E)),$$

is strictly positive in the interval $(\mu_E, 1-\lambda+\lambda\phi)$. Hence, $\mu_E < \theta$. □

D Proof of Theorem 2

Lemma 8. *For all $x \in (\lambda\phi, 1-\lambda+\lambda\phi)$, $\overline{G}^C(x,\mu_E) < \overline{G}^I(x,\mu_E)$.*

Proof. Notice first that $\mu_E < \theta$ implies that $\frac{(\theta-\eta)\overline{G}^C(\mu_E,\mu_E)}{\overline{G}(\mu_E,\mu_E)} < \theta - \frac{\eta}{\overline{G}(\mu_E,\mu_E)} < \theta-\eta$. But $\frac{(\theta-\eta)\overline{G}^C(\mu_E,\mu_E)}{\overline{G}(\mu_E,\mu_E)} < \theta-\eta$ if and only if $\overline{G}^C(\mu_E,\mu_E) < \overline{G}^I(\mu_E,\mu_E)$.

Consistency also requires that for all $x \in (\lambda\phi, 1-\lambda+\lambda\phi)$, $\frac{(\theta-\eta)\overline{g}^C(x,\mu_E)}{\overline{g}(x,\mu_E)} = x$. Hence, $\overline{g}^C(x,\mu_E) = \frac{x}{\theta-\eta}\overline{g}(x,\mu_E)$ and $\overline{g}^I(x,\mu_E) = \frac{1-x}{1-\theta+\eta}\overline{g}(x,\mu_E)$. Therefore, $\overline{G}^I(x,\mu_E) - \overline{G}^C(x,\mu_E) > 0$ at $x = \mu_E$, and this difference attains its maximum at $x = \theta-\eta$:

$$\frac{\partial\left(\overline{G}^I(x,\mu_E) - \overline{G}^C(x,\mu_E)\right)}{\partial x} = \overline{g}^I(x,\mu_E) - \overline{g}^C(x,\mu_E) = \left(\frac{\theta-\eta-x}{(\theta-\eta)(1-\theta+\eta)}\right)\overline{g}(x,\mu_E)$$

Notice that $\frac{\partial\left(\overline{G}^I(x,\mu_E) - \overline{G}^C(x,\mu_E)\right)}{\partial x} < 0$ for all $x > \theta-\eta$, and since $\overline{G}^I(x,\mu_E) - \overline{G}^C(x,\mu_E) = 0$ at $x = 1-\lambda+\lambda\phi$, we conclude that $\overline{G}^I(x,\mu_E) > \overline{G}^C(x,\mu_E)$ for all $x \in (\lambda\phi, 1-\lambda+\lambda\phi)$. □

Lemma 9. *For all $x \in (\mu_E, 1-\lambda+\lambda\phi)$, $G^C(x,\mu_E) < G^I(x,\mu_E)$.*

Proof. As $\mu_E < \theta$, we know that $G^C(\mu_E,\mu_E) < G^I(\mu_E,\mu_E)$, since $\frac{\theta G^C(\mu_E,\mu_E)}{G(\mu_E,\mu_E)} < \theta$ if and only if $G^C(\mu_E,\mu_E) < G^I(\mu_E,\mu_E)$.

Consistency also requires that for all $x \in (\mu_E, 1-\lambda+\lambda\phi)$, $\frac{\theta g^C(x,\mu_E)}{g(x,\mu_E)} = x$. Hence, $G^I(x,\mu_E) - G^C(x,\mu_E) > 0$ at $x = \mu_E$, and this difference attains its highest value at $x = \theta$:

$$\frac{\partial\left(G^I(x,\mu_E) - G^C(x,\mu_E)\right)}{\partial x} = g^I(x,\mu_E) - g^C(x,\mu_E) = \frac{\theta-x}{\theta(1-\theta)}g(x,\mu_E)$$

Notice that $\frac{\partial\left(G^I(x,\mu_E) - G^C(x,\mu_E)\right)}{\partial x} < 0$ for all $x > \theta$, and since $G^I(x,\mu_E) - G^C(x,\mu_E) = 0$ at $x = 1-\lambda+\lambda\phi$, we conclude that $G^I(x,\mu_E) > G^C(x,\mu_E)$ for all $x \in (\mu_E, 1-\lambda+\lambda\phi)$. □

E Proof of Theorem 3

Consider the function $\sigma(\mu) \equiv \psi(\mu, \mu)$, with $\mu > \lambda\phi$ as defined in Appendix C; i.e., Equation 17 evaluated at the diagonal, with:

$$\psi(x,y) = \frac{(\theta-\eta)\overline{G}^C(x,y) + \eta}{(\theta-\eta)\overline{G}^C(x,y) + (1-\theta+\eta)\overline{G}^I(x,y)}.$$

We know that $\mu_E = \sigma(\mu_E)$. Totally differentiating this expression, we get:

$$d\mu_E = \frac{(\theta-\eta)(1-\psi)}{\overline{G}(\mu_E,\mu_E)}\frac{\partial\overline{G}^C(\mu_E,\mu_E)}{\partial\lambda}d\lambda - \frac{\psi(1-\theta+\eta)}{\overline{G}(\mu_E,\mu_E)}\frac{\partial\overline{G}^I(\mu_E,\mu_E)}{\partial\lambda}d\lambda + \sigma'(\mu_E)\,d\mu_E$$

and therefore:

$$\frac{d\mu_E}{d\lambda} = \frac{\frac{(\theta-\eta)(1-\psi)}{\overline{G}}\frac{\partial\overline{G}^C}{\partial\lambda} - \frac{\psi(1-\theta+\eta)}{\overline{G}}\frac{\partial\overline{G}^I}{\partial\lambda}}{1-\sigma'(\mu_E)} \tag{36}$$

The factor $\frac{1}{1-\sigma'(\mu_E)}$ emphasizes that as λ has a direct effect on $\overline{G}^C$ and $\overline{G}^I$, it must affect the entry-level reputation, which in turn affects the reputation distributions G^C and G^I and hence ψ. But as we proved before, $\sigma'(\mu_E) = 0$ and therefore the indirect effect vanishes. Hence, we must restrict attention to the direct effect.

As in Appendix C, $\frac{\partial\overline{G}^C}{\partial\lambda}$ and $\frac{\partial\overline{G}^I}{\partial\lambda}$ are the solution of an integral-equation system defined as:

$$\begin{pmatrix} \frac{\partial\overline{G}^C}{\partial\lambda} \\ \frac{\partial\overline{G}^I}{\partial\lambda} \end{pmatrix}(x,y) = \begin{pmatrix} \left.\frac{\partial T^C}{\partial\lambda}\right|_{\overline{G}} \\ \left.\frac{\partial T^I}{\partial\lambda}\right|_{\overline{G}} \end{pmatrix}(x,y)$$
$$+\begin{pmatrix} \int_y^{\tilde{\mu}(x,0)}\frac{\partial\overline{G}^C}{\partial\lambda}(\mu,y)\,f_H(\tilde{r}(x,\mu))\frac{\partial\tilde{r}(x,\mu)}{\partial\mu}d\mu + \int_y^{\tilde{\mu}(x,0)}\frac{\partial\overline{G}^I}{\partial\lambda}(\mu,y)\,f_L(\tilde{r}(x,\mu))\frac{\partial\tilde{r}(x,\mu)}{\partial\mu}d\mu \\ \int_y^{\tilde{\mu}(x,0)}\frac{\partial\overline{G}^C}{\partial\lambda}(\mu,y)\,f_H(\tilde{r}(x,\mu))\frac{\partial\tilde{r}(x,\mu)}{\partial\mu}d\mu + \int_y^{\tilde{\mu}(x,0)}\frac{\partial\overline{G}^I}{\partial\lambda}(\mu,y)\,f_L(\tilde{r}(x,\mu))\frac{\partial\tilde{r}(x,\mu)}{\partial\mu}d\mu \end{pmatrix} \tag{37}$$

We know that:

- $\phi = \frac{\lambda\theta-\eta}{\lambda}$ (by definition);
- $\frac{\partial\tilde{r}(x,y)}{\partial x} = \frac{R}{R'}\frac{1-\lambda}{(x-\lambda\phi)(1-\lambda+\lambda\phi-x)}$, while $\frac{\partial\tilde{r}(x,y)}{\partial\lambda} = -\frac{\theta-\eta-x}{1-\lambda}\frac{\partial\tilde{r}(x,y)}{\partial x}$
- $\frac{\partial\tilde{\mu}(x,r)}{\partial x} = \frac{(1-R(r))\tilde{\mu}(x,r)-1}{((x-1+\lambda-\lambda\phi)R(r)-(x-\lambda\phi))}$, while $\frac{\partial\tilde{\mu}(x,r)}{\partial\lambda} = -\frac{\theta-\eta-x}{1-\lambda}\frac{\partial\tilde{\mu}(x,r)}{\partial x}$
- $\overline{g}^C_{t+1}$ y $\overline{g}^I_{t+1}$ can be obtained by taking the derivative of $\overline{G}^C_{t+1}$ y $\overline{G}^I_{t+1}$ with respect to x:

$$\overline{g}^C_{t+1}(x,y) = \frac{\partial\tilde{r}(x,y)}{\partial x}\left(\frac{\theta(1-\lambda+\lambda\phi)}{\theta-\eta}G^C_t(y,y)\,f_H(\tilde{r}(x,y)) + \frac{(1-\theta)\lambda\phi}{\theta-\eta}G^I_t(y,y)\,f_L(\tilde{r}(x,y))\right)$$
$$+\frac{\theta(1-\lambda+\lambda\phi)}{\theta-\eta}\int_0^{\tilde{r}(x,y)} g^C_t(\tilde{\mu}(x,r),y)\frac{\partial\tilde{\mu}(x,r)}{\partial x}dF_H + \frac{(1-\theta)\lambda\phi}{\theta-\eta}\int_0^{\tilde{r}(x,y)} g^I_t(\tilde{\mu}(x,r),y)\frac{\partial\tilde{\mu}(x,r)}{\partial x}dF_L$$

and

$$\overline{g}^I_{t+1}(x,y) = \frac{\partial\tilde{r}(x,y)}{\partial x}\left(\frac{\theta\lambda(1-\phi)}{1-\theta+\eta}G^C_t(y,y)\,f_H(\tilde{r}(x,y)) + \frac{(1-\theta)(1-\lambda\phi)}{1-\theta+\eta}G^I_t(y,y)\,f_L(\tilde{r}(x,y))\right)$$
$$+\frac{\theta\lambda(1-\phi)}{1-\theta+\eta}\int_0^{\tilde{r}(x,y)} g^C_t(\tilde{\mu}(x,r),y)\frac{\partial\tilde{\mu}(x,r)}{\partial x}dF_H + \frac{(1-\theta)(1-\lambda\phi)}{1-\theta+\eta}\int_0^{\tilde{r}(x,y)} g^I_t(\tilde{\mu}(x,r),y)\frac{\partial\tilde{\mu}(x,r)}{\partial x}dF_L$$

Hence, the direct effect of a change in λ can be written as:

$$\begin{pmatrix} \left.\frac{\partial T^C}{\partial\lambda}\right|_{\overline{G}} \\ \left.\frac{\partial T^I}{\partial\lambda}\right|_{\overline{G}} \end{pmatrix}(x,y) = -\left(\int_0^{\tilde{r}(x,y)} G^I(\tilde{\mu}(x,r),y)\,dF_L - \int_0^{\tilde{r}(x,y)} G^C(\tilde{\mu}(x,r),y)\,dF_H\right)\begin{pmatrix} \frac{\partial\left(\frac{\theta(1-\lambda+\lambda\phi)}{\theta-\eta}\right)}{\partial\lambda} \\ \frac{\partial\left(\frac{\theta\lambda(1-\phi)}{1-\theta+\eta}\right)}{\partial\lambda} \end{pmatrix}$$
$$-\frac{\theta-\eta-x}{1-\lambda}\begin{pmatrix} \overline{g}^C(x,y) \\ \overline{g}^I(x) \end{pmatrix} \tag{38}$$

where $\frac{\partial\left(\frac{\theta(1-\lambda+\lambda\phi)}{\theta-\eta}\right)}{\partial\lambda} = -\frac{\theta(1-\theta)}{\theta-\eta}$ and $\frac{\partial\left(\frac{\theta\lambda(1-\phi)}{1-\theta+\eta}\right)}{\partial\lambda} = \frac{\theta(1-\theta)}{1-\theta+\eta}$. But we also know that steady-state distributions satisfy:

$$\overline{G}^I(x,y) - \overline{G}^C(x,y) = \frac{\theta(1-\theta)(1-\lambda)}{(\theta-\eta)(1-\theta+\eta)}\left(\int_0^{\tilde{r}(x,y)} G^I(\tilde{\mu}(x,r),y)\,dF_L - \int_0^{\tilde{r}(x,y)} G^C(\tilde{\mu}(x,r),y)\,dF_H\right)$$

and that consistency requires that $\frac{(\theta-\eta)\overline{g}^C(x,y)}{\overline{g}(x,y)} = x$. Thus, the direct effect can be rewritten as:

$$\begin{pmatrix} \left.\frac{\partial T^C}{\partial\lambda}\right|_{\overline{G}} \\ \left.\frac{\partial T^I}{\partial\lambda}\right|_{\overline{G}} \end{pmatrix}(x,y) = -\frac{(\theta-\eta)(1-\theta+\eta)}{1-\lambda}\left(\overline{G}^I(x,y) - \overline{G}^C(x,y)\right)\begin{pmatrix} -\frac{1}{\theta-\eta} \\ \frac{1}{1-\theta+\eta}\end{pmatrix} - \frac{\theta-\eta-x}{1-\lambda}\overline{g}(x,y)\begin{pmatrix} \frac{x}{\theta-\eta} \\ \frac{1-x}{1-\theta+\eta}\end{pmatrix} \tag{39}$$

Also as in Appendix C we repeatedly iterate the operator T to find the sign of 36 starting from $\overline{G}_0^C$ y $\overline{G}_0^I$ –the steady-state distributions obtained with $y = \mu_E$ and the initial value of λ.

Let denote the t-th iteration of T as $\overline{G}_t^C$ and $\overline{G}_t^I$, and define $\psi_t(x,y)$ as $\frac{\theta G_t^C(x,y)}{\overline{G}_t(x,y)}$. Then $y = \psi_0(y,y) = \frac{\theta G_0^C(y,y)}{\overline{G}_0(y,y)}$. After iterating T with a different level of λ we obtain $\overline{G}_1^\tau \neq \overline{G}_0^\tau$; moreover, for an infinitesimally small change in λ, the difference between them is exactly $\frac{\partial T^\tau}{\partial\lambda}d\lambda$. Hence, we can define $\overline{G}_1^\tau$ as:

$$\overline{G}_1^\tau(x,y) = \overline{G}_0^\tau(x,y) + \left.\frac{\partial T^\tau}{\partial\lambda}\right|_{\overline{G}_0}(x,y)\,d\lambda$$

It follows that:

$$\begin{pmatrix} \overline{G}_1^C \\ \overline{G}_1^I \end{pmatrix}(x,y) = \begin{pmatrix} \overline{G}_0^C \\ \overline{G}_0^I \end{pmatrix}(x,y) + \alpha_0(x,y)\begin{pmatrix} \frac{1}{\theta-\eta} \\ -\frac{1}{1-\theta+\eta}\end{pmatrix} + \beta_0(x,y)\begin{pmatrix} \frac{x}{\theta-\eta} \\ \frac{1-x}{1-\theta+\eta}\end{pmatrix}$$

where $\alpha_0(x,y) \equiv \frac{(\theta-\eta)(1-\theta+\eta)}{1-\lambda}\left(\overline{G}_0^I(x,y) - \overline{G}_0^C(x,y)\right)d\lambda > 0$ and $\beta_0(x,y) \equiv -\frac{\theta-\eta-x}{1-\lambda}\overline{g}_0(x,y)\,d\lambda < 0$. When we evaluate this expression in the new value of λ we obtain:

$$\overline{G}_1(x,y) \quad = \overline{G}_0(x,y) + \beta_0(x,y)$$

and hence:

$$\psi_1(y,y) = \frac{(\theta-\eta)\overline{G}_0^C(y,y) + \alpha_0(y,y) + y\beta_0(y,y) + \eta}{\overline{G}_1(y,y)}$$

where $\frac{(\theta-\eta)\overline{G}_0^C(y,y)+\eta}{\overline{G}_0(y,y)} = \psi_0(y,y) = y$. Rearranging, we obtain:

$$\psi_1(y,y) = \psi_0(y,y) + \frac{\alpha_0(x,y)}{\overline{G}_1(x,y)}$$

Iterating the operator T (considering the new value of λ) we obtain:

$$\begin{pmatrix} \overline{G}_t^C \\ \overline{G}_t^I \end{pmatrix}(x,y) = \begin{pmatrix} \overline{G}_{t-1}^C \\ \overline{G}_{t-1}^I \end{pmatrix}(x,y) + \alpha_{t-1}(x,y)\begin{pmatrix} \frac{1}{\theta-\eta} \\ -\frac{1}{1-\theta+\eta}\end{pmatrix} + \beta_{t-1}(x,y)\begin{pmatrix} \frac{x}{\theta-\eta} \\ \frac{1-x}{1-\theta+\eta}\end{pmatrix}$$

with:

$$\overline{G}_t(x,y) = \overline{G}_{t-1}(x,y) + \beta_{t-1}(x,y)$$

where

$$\alpha_t(x,y) \equiv -\int_y^{\tilde{\mu}(x,0)} \alpha_{t-1}(\mu,y)\left((1-\lambda+\lambda\phi-x)f_H(\tilde{r}(x,\mu)) + (x-\lambda\phi)f_L(\tilde{r}(x,\mu))\right)\frac{\partial\tilde{r}(x,\mu)}{\partial\mu}d\mu > 0$$

and

$$\beta_t(x,y) \equiv \int_y^{\tilde{\mu}(x,0)} \left(\alpha_{t-1}(\mu,y)\left(f_L(\tilde{r}(x,\mu)) - f_H(\tilde{r}(x,\mu))\right) - \beta_{t-1}(\mu,y)\left(\mu f_H(\tilde{r}(x,\mu)) + (1-\mu)f_L(\tilde{r}(x,\mu))\right)\right)\frac{\partial\tilde{r}(x,\mu)}{\partial\mu}d\mu$$

for all $t > 0$. Thus, we conclude that for all the elements of the sequence, $\psi_t(y,y)$ can be written as:

$$\psi_t(y,y) = \frac{\psi_{t-1}(y,y)\overline{G}_{t-1}(y,y) + \alpha_{t-1}(y,y) + y\beta_{t-1}(y,y)}{\overline{G}_{t-1}(y,y) + \beta_{t-1}(y,y)}$$

Moreover:

$$\begin{pmatrix} \overline{G}_t^C \\ \overline{G}_t^I \end{pmatrix}(x,y) = \begin{pmatrix} \overline{G}_0^C \\ \overline{G}_0^I \end{pmatrix}(x,y) + \sum_{j=0}^{t-1} \alpha_j(x,y) \begin{pmatrix} \frac{1}{\theta-\eta} \\ -\frac{1}{1-\theta+\eta} \end{pmatrix} + \sum_{j=0}^{t-1} \beta_j(x,y) \begin{pmatrix} \frac{x}{\theta-\eta} \\ \frac{1-x}{1-\theta+\eta} \end{pmatrix}$$

with:

$$\overline{G}_t(x,y) = \overline{G}_0(x,y) + \sum_{j=0}^{t-1} \beta_j(x,y)$$

Since $\psi_0(y,y) = y$, we finally obtain:

$$\psi_t(y,y) = \psi_0(y,y) + \frac{\sum_{j=0}^{t-1} \alpha_j(y,y)}{\overline{G}_t(y,y)} > \psi_0(y,y)$$

and hence we conclude that for all $t > 0$, $\psi_t(y,y) > \psi_0(y,y)$, and therefore $\frac{d\mu_E}{d\lambda} > 0$.

F Proof of Lemma 4

Observe that $\Pr(\text{exit at } t+1|\mu_t,\tau_t) = F_{\tau_t}(\tilde{r}(\mu_E,\mu_t))$. The first-order stochastic dominance assumption implies directly that $\Pr(\text{exit at } t{+}1|\mu_t,C) = F_H(\tilde{r}(\mu_E,\mu_t)) \leq F_L(\tilde{r}(\mu_E,\mu_t)) = \Pr(\text{exit at } t+1|\mu_t,I)$. On the other hand, we have:

$$\frac{\partial \Pr(\text{exit at } t+1|\mu_t,\tau_t)}{\partial \mu_t} = f_{\tau_t}(\tilde{r}(\mu_E,\mu_t)) \frac{\partial \tilde{r}(\mu_E,\mu_t)}{\partial \mu_t} < 0$$

since $\frac{\partial \tilde{r}(\mu_E,\mu_t)}{\partial \mu_t} < 0$. Also, we have that $\Pr(\text{exit at } t+1|\mu_t) = \mu_t F_H(\tilde{r}(\mu_E,\mu_t)) + (1-\mu_t) F_L(\tilde{r}(\mu_E,\mu_t))$. As a consequence,

$$\frac{\partial \Pr(\text{exit at } t+1|\mu_t)}{\partial \mu_t} = (F_H - F_L) + (\mu_t f_H + (1-\mu_t) f_L) \frac{\partial \tilde{r}(\mu_E,\mu_t)}{\partial \mu_t} < 0$$

Regarding signals, we have that $\Pr(\text{exit at } t+1|r_t,\tau_t) = G^{\tau_t}(\tilde{\mu}(\mu_E,r_t))$. Hence,

$$\frac{\partial \Pr(\text{exit at } t+1|r_t,\tau_t)}{\partial r_t} = \frac{\partial G^{\tau_t}}{\partial \tilde{\mu}} \frac{\partial \tilde{\mu}(\mu_E,r_t)}{\partial r_t} < 0$$

since $\frac{\partial \tilde{\mu}(x,r_t)}{\partial r_t} < 0$. Also, the unconditional (on type) probability of exit is given by: $\Pr(\text{exit at } t{+}1|r_t) = \big((\theta - \lambda(\theta-\phi)) G^C(\tilde{\mu}(\mu_E,r_t)) + (1-\theta-\lambda(\theta-\phi)) G^I(\tilde{\mu}(\mu_E,r_t))\big)$, so that:

$$\frac{\partial \Pr(\text{exit at } t{+}1|r_t)}{\partial r_t} = \left((\theta - \lambda(\theta-\phi)) \frac{\partial G_C}{\partial \tilde{\mu}} + (1-\theta-\lambda(\theta-\phi)) \frac{\partial G_I}{\partial \tilde{\mu}}\right) \frac{\partial \tilde{\mu}(x,r_t)}{\partial r_t} < 0$$

G Proof of Theorem 5

Using Equations 23 and 21 we can write the difference $\overline{G}_{a+1}^C(x) - \overline{G}_{a+1}^I(x)$ as:

$$\overline{G}_{a+1}^C(x) - \overline{G}_{a+1}^I(x) = (1-\lambda) \frac{m_a^C m_a^I}{\overline{m}_{a+1}^C \overline{m}_{a+1}^I} \left(\int_0^{\tilde{r}(x,\mu_E)} G_a^C(\tilde{\mu}(x,r))\, dF_H - \int_0^{\tilde{r}(x,\mu_E)} G_a^I(\tilde{\mu}(x,r))\, dF_L \right)$$

However, $\int_0^{\tilde{r}(x,\mu_E)} G_a^C(\tilde{\mu}(x,r))\, dF_H < \int_0^{\tilde{r}(x,\mu_E)} G_a^C(\tilde{\mu}(x,r))\, dF_L$ for all $x \in (\lambda\phi, 1-\lambda+\lambda\phi)$ because $G_a^C(\tilde{\mu}(x,r))$ is decreasing in r and $F_H(r) < F_L(r)$ for all $r \in (0,1)$. Hence:

$$\overline{G}_{a+1}^C(x) - \overline{G}_{a+1}^I(x) < (1-\lambda) \frac{m_a^C m_a^I}{\overline{m}_{a+1}^C \overline{m}_{a+1}^I} \int_0^{\tilde{r}(x,\mu_E)} \Big(G_a^C(\tilde{\mu}(x,r)) - G_a^I(\tilde{\mu}(x,r)) \Big) dF_L$$

Hence, if $G_a^C(\tilde{\mu}(x,r)) \leq G_a^I(\tilde{\mu}(x,r))$, then $\overline{G}_{a+1}^C(x) < \overline{G}_{a+1}^I(x)$ for all $x \in (\lambda\phi, 1-\lambda+\lambda\phi)$.

When $a = 0$, $G_a^C(\tilde{\mu}(x,r)) - G_a^I(\tilde{\mu}(x,r)) = 0$, so that $\overline{G}_1^C(x) < \overline{G}_1^I(x)$ for all $x \in (\lambda\phi, 1-\lambda+\lambda\phi)$. On the other hand, within the population of incumbents of age at least 1 ($a > 0$), if the prior reputation conditional on an event A first-order stochastically dominates the prior reputation conditional on an event B, then the posteriors are also ordered by first-order stochastic dominance in the same way:

$$\Pr(\overline{\mu} < x|A) \leq \Pr(\overline{\mu} < x|B) \Rightarrow \Pr(\mu < x|A) \leq \Pr(\mu < x|B) \tag{40}$$

To see this, simply observe that $\frac{\Pr(\overline{\mu}<x|A)-\Pr(\overline{\mu}<\mu_E|A)}{1-\Pr(\overline{\mu}<\mu_E|A)} \leq \frac{\Pr(\overline{\mu}<x|B)-\Pr(\overline{\mu}<\mu_E|B)}{1-\Pr(\overline{\mu}<\mu_E|B)} \iff \Pr(\mu < x|A) \leq \frac{\Pr(\mu<x|B)(1-\Pr(\overline{\mu}<\mu_E|A))+\Pr(\overline{\mu}<\mu_E|A)-\Pr(\overline{\mu}<\mu_E|B)}{1-\Pr(\overline{\mu}<\mu_E|B)}$; since $0 < \Pr(\overline{\mu} < \mu_E|A) < \Pr(\overline{\mu} < \mu_E|B) < 1$, $\frac{\Pr(\mu<x|B)(1-\Pr(\overline{\mu}<\mu_E|A))+\Pr(\overline{\mu}<\mu_E|A)-\Pr(\overline{\mu}<\mu_E|B)}{1-\Pr(\overline{\mu}<\mu_E|B)} \leq \Pr(\mu < x|B)$, so that $\Pr(\mu < x|A) \leq \Pr(\mu < x|B)$. In particular,

$$\overline{G}^C_{a+1}(x) \leq \overline{G}^I_{a+1}(x) \Rightarrow G^C_{a+1}(x) \leq G^I_{a+1}(x). \tag{41}$$

Hence, the prior reputation of a competent of any cohort first-order stochastically dominate that of the inept of the same cohort.

H Proof of Theorem 6

Let μ_a denote the interim reputation of a firm of age a. Consistency requires that $E[\mu_a] = \frac{m^C_a}{m^C_a+m^I_a}$, the fraction of competents in cohort a. Using this and collecting terms, we can re-write Equation 23 as:

$$\begin{pmatrix} \overline{G}^C_{a+1}(x) \\ \overline{G}^I_{a+1}(x) \end{pmatrix} = \begin{pmatrix} \gamma^C_a & 1-\gamma^C_a \\ \gamma^I_a & 1-\gamma^I_a \end{pmatrix} \begin{pmatrix} \int_0^{\tilde{r}(x,\mu_E)} G^C_a(\tilde{\mu}(x,r))\,dF_H \\ \int_0^{\tilde{r}(x,\mu_E)} G^I_a(\tilde{\mu}(x,r))\,dF_L \end{pmatrix}$$

where

$$\begin{aligned} \gamma^C_a &= \frac{(1-\lambda+\lambda\phi)}{(1-\lambda+\lambda\phi)+\lambda\phi\frac{1-E[\mu_a]}{E[\mu_a]}} \\ \gamma^I_a &= \frac{\lambda(1-\phi)}{\lambda(1-\phi)+(1-\lambda\phi)\frac{1-E[\mu_a]}{E[\mu_a]}} \end{aligned}$$

The age$-a+1$ population-wide distribution can thus be written as:

$$\overline{G}_{a+1}(x) = E[\mu_a]\int_0^{\tilde{r}(x,\mu_E)} G^C_a(\tilde{\mu}(x,r))\,dF_H + (1-E[\mu_a])\int_0^{\tilde{r}(x,\mu_E)} G^I_a(\tilde{\mu}(x,r))\,dF_L.$$

From the proof of Theorem 5 (Equation 40) we know that:

$$\overline{G}^\tau_{a+1}(x) \leq \overline{G}^\tau_a(x) \Rightarrow G^\tau_{a+1}(x) \leq G^\tau_a(x)$$

and

$$\overline{G}_{a+1}(x) \leq \overline{G}_a(x) \Rightarrow G_{a+1}(x) \leq G_a(x)$$

We first prove that:

Lemma 10.

$$\left(\overline{G}^\tau_{a+1}(x) \leq \overline{G}^\tau_a(x) \text{ and } \overline{G}_{a+1}(x) \leq \overline{G}_a(x)\right) \Rightarrow \left(\overline{G}^\tau_{a+2}(x) \leq \overline{G}^\tau_{a+1}(x) \text{ and } \overline{G}_{a+2}(x) \leq \overline{G}_{a+1}(x)\right)$$

Proof. By hypothesis, $G^C_{a+1}(\tilde{\mu}(x,r)) \leq G^C_a(\tilde{\mu}(x,r))$ and $G^I_{a+1}(\tilde{\mu}(x,r)) \leq G^I_a(\tilde{\mu}(x,r))$. Hence, $\overline{G}^\tau_{a+2}(x) \leq \gamma^\tau_{a+1} \int_0^{\tilde{r}(x,\mu_E)} G^C_a(\tilde{\mu}(x,r))\,dF_H + \left(1-\gamma^\tau_{a+1}\right)\int_0^{\tilde{r}(x,\mu_E)} G^I_a(\tilde{\mu}(x,r))\,dF_L$. The right hand side can be written as: $\overline{G}^\tau_{a+1}(x) + \left(\gamma^\tau_{a+1}-\gamma^\tau_a\right)\left(\int_0^{\tilde{r}(x,\mu_E)} G^C_a(\tilde{\mu}(x,r))\,dF_H - \int_0^{\tilde{r}(x,\mu_E)} G^I_a(\tilde{\mu}(x,r))\,dF_L\right)$. Hence, $\left(\gamma^\tau_{a+1}-\gamma^\tau_a\right) > 0$ is a sufficient condition for $\overline{G}^\tau_{a+2}(x) \leq \overline{G}^\tau_{a+1}(x)$. Indeed, by hypothesis, $G_{a+1}(x) \leq G_a(x)$, which implies that $E[\mu_{a+1}] > E[\mu_a]$, and hence $\left(\gamma^\tau_{a+1}-\gamma^\tau_a\right) > 0$ for both types.

The proof for the unconditional case is analogous. □

We complete the induction argument by showing that:

Lemma 11. $G_1^\tau(x) \leq G_0^\tau(x)$ *and* $G_1(x) \leq G_0(x)$

Proof. Follows immediatly from the fact that in $a = 0$ both distributions are degenerate at μ_E, while in $a = 1$ both distributions are non-degenerate, with support starting at μ_E. □

Finally, the initial prior distributions are also ordered in the same fashion:

Lemma 12. $\overline{G}_2^\tau(x) \leq \overline{G}_1^\tau(x)$ *and* $\overline{G}_2(x) \leq \overline{G}_1(x)$

Proof. The proof follows a similar argument than that of Lemma 10 above. □