Reputation and product recalls

Boyan Jovanovic*

September 30, 2016

Abstract

We model intangible capital as a property of equilibrium play between a seller and buyers. A seller continually meets short-lived buyers. A low-effort signal occurs periodically. Its arrival triggers a reduction in prices that then gradually rise until the next low-effort signal and so on repeatedly. We then fit the model to data on product recalls and on stock-price reductions following such recalls. We estimate the model by constrained maximum likelihood. We then use the parameter estimates to compare welfare across several types of equilibria that the model gives rise to. We conclude that contract incompleteness leads to large welfare losses in this area.

1 Introduction

Reputations are an "intangible" capital. McGrattan and Prescott (2000) argue that as much as 40 percent of GNP is intangible capital. Evidence by sector is in Hall (2001). Created by advertising or simply by having a good product with word-of-mouth information diffusion among customers. This paper adds to the literature by structurally estimating a model of reputation building. It uses product-recalls and stock-price evidence to fit the model and to draw welfare conclusions.

In the model an agent – a seller – faces a continuum of risk-neutral, short-lived buyers whom he meets one at a time. The seller's effort is not observed

*NYU Economics. I thanks Sai Ma and Dong Wei for help with the research, the NSF and C.V. Starr Center for financial assistance, and Chris Flinn, Ricardo Lagos and Ennio Stacchetti for comments.

and payment for its services is made up front. No output contingent contracts are available and there are no repeat meetings or long-term contacts. Buyers are short-term and do not return.

Periodically a low-effort signal occurs. The seller can postpone the arrival of the signal by exerting higher effort, and so when the signal transpires it is bad news for the seller. When the seller is a publicly owned firm, the signal will cause the value of the firm to fall. Expectations then re-initialize and after the signal causes them to fall, they fall, prices then gradually rise until the next low-effort signal arrives, and so on.

The signal is interpreted as a product recall which is known to lead to a stock-price reduction of publicly traded firms. The model equilibrium fits the data on product recalls and stock-price reductions well – the maximum likelihood estimate is compared to data and the two are close. Of course, product recalls are quite common, but still rare enough to produce a negative stock-price impact.

Analytically we make the most progress in a "mixed equilibrium" in which there is a hazard of switching from the ratchet equilibrium to the grim trigger equilibrium. This version is also estimated by maximum likelihood.

Even though the model is quite special, the close fit to data suggests that the welfare implications of the model should be taken seriously. The estimates imply that equilibrium delivers only a very small fraction of maximal welfare, a little over three percent. This is because output is also only slightly over three percent of its socially optimal level. This is conditional on a particular type of equilibrium being played, one we refer to as the "ratchet" equilibrium. Grim trigger equilibria (of which there are two) yields higher welfare, but they do not fit the data. Yet another equilibrium is one that mixes between these equilibria that also yield higher welfare. The "ratchet" equilibrium is similar to that in Rob and Fishman (2005), whereas the grim trigger equilibrium is quite standard. Mixing between the two is done in an innocuous and standard way – via an extrinsic "sunspot" shock but that also reflects the firm's effort, and so one may conceive of it as reflecting earnings announcements, analyst reports, and so on.

Intangible capital in the model is a name for the additional output, value and welfare that some equilibria deliver. Intangible capital is a reflection of how the seller's receipts and effort reflect the history of the game between the seller and his customers including especially the public signals which are endogenous. Zero intangible capital is associated with the zero-effort equilibrium. Maximal capital is associated with the high grim-trigger equilibrium.

Both of equilibria always exist. The best commonly used term for this type of capital is "goodwill" that is thought to reflect buyers' attitudes towards a seller. Of course the term "intangible capital" is thought to include other things that this model leaves out.

Section 2 lays out the model and several of its equilibria. Section 3 discusses the data and the estimation. Section 4 discusses welfare in the various equilibria and Section 5 concludes. Some proofs and data details are in the Appendix.[1]

2 Model

A long-lived seller faces an infinite sequence of short-term, risk-neutral buyers who arrive continuously. The seller's effort is private information.

Seller's period payoff.—If the seller exerts effort x_t, it produces output y_t at the rate

$$dy_t = x_t dt + \sigma dW_t,$$

where W is Brownian motion. Effort is unobserved and the seller is paid in advance. Buyers are risk neutral and they bid up the price to p per expected unit of quality. The price is constant and reflects buyer's utility of consuming the product. The cost of effort is $x_t^2/2$. Thus the seller's period payoff is

$$px_t^* - \frac{1}{2}x_t^2$$

where p is price per unit of quality and the seller's revenue is px^* where x^* is the effort that the buyer expects him to exert.

Meetings and rent division.—Meetings are bilateral but that the seller has all the bargaining power and can extract from the buyer all the expected rents. A buyer's outside option is zero while the seller's outside option, v_t, will be endogenous. Many risk neutral buyers who pay the expected value of the seller's output for which their marginal utility is constant at p. The total price paid up front is

$$p_t = px_t^*.$$

Public signals.—A buyer's consumption experience is not observed by other prospective buyers. Periodically the public sees a signal which comes

[1] In reduced form the model is mathematically similar to the Bass (1969) model of information diffusion and Appendix 6 compares the two models formally.

sooner if x_t is low. The signal contains no other information about x_t. Reputation does not spread except through "bad news" which is a noisy signal that effort was low. The signal may take the form of a product recall that is widely publicized and that becomes a part of the seller's public history immediately.

Let the i'th signal arrive at date τ_i, and let the hazard rate of its arrival be $\lambda - x_t$, where $\lambda > 0$ is a parameter denoting maximum feasible effort so that

$$x_t \leq \lambda. \tag{1}$$

Let $t =$ the duration since the last signal. Normalize the date of arrival of the last signal at $\tau = 0$. Then t denotes time as well as time elapsed since the last signal. Conditional on an effort path (x_t), the waiting time until the next signal has CDF $F(t) = \Pr(\tau \leq t)$ given by

$$F(t) = 1 - \exp\left(-\int_0^t (\lambda - x_\tau)\, d\tau\right). \tag{2}$$

Higher effort postpones the arrival of the next signal. If x_t is bounded away from λ, $\lim_{t\to\infty} F(t) = 1$, i.e., the signal is sure to arrive eventually.

Equilibrium.—This game has several equilibria including the equilibrium in which $x_t = 0$ regardless of the history signals. We shall focus on two types of equilibrium that depend on past effort through the sequence of public signals $(\tau_i)_{i\in\mathbb{N}}$. The first is a standard grim-trigger equilibrium which will not fit the data but which is easy to explain. The second is a "ratchet" type of equilibrium which we shall then estimate.

2.1 Grim trigger equilibrium

A grim trigger equilibrium has expected effort $x^* > 0$ until a signal arrives, and then $x_t^* = 0$ for all $t > \tau$, where τ is the first time a bad signal is observed. The signal reduces the seller's value to zero permanently. In the pre-signal phase the seller's value v satisfies

$$rv = \max_{x\leq\lambda}\left\{px^* - \frac{1}{2}x^2 - (\lambda - x)\,v\right\}$$

with the FOC

$$v - x \geq 0$$

with equality if $v \leq \lambda$. This means that v satisfies

$$rv = \begin{cases} p\lambda - \frac{1}{2}\lambda^2 & \text{if } v \geq \lambda \\ pv - \frac{1}{2}v^2 - (\lambda - v)\, v & \text{if } v < \lambda \end{cases}$$

The solutions are

$$v = \begin{cases} \frac{1}{r}\left(p\lambda - \frac{1}{2}\lambda^2\right) \equiv x^{\mathrm{H}} & \text{if } v \geq \lambda \\ 2\,(r + \lambda - p) \equiv x^{\mathrm{L}} & \text{if } v < \lambda \end{cases} \tag{3}$$

The first requires that $v \geq \lambda$, which reads $\frac{1}{r}\left(p\lambda - \frac{1}{2}\lambda^2\right) \geq \lambda$, i.e.,

$$p \geq r + \frac{\lambda}{2}. \tag{4}$$

Since the second of these requires that $v < \lambda$, this means that $2\,(r + \lambda - p) < \lambda$, and i.e., again condition (4).but also that $v \geq 0$ so that $p \leq r + \lambda$. I.e., the low equilibrium exists if

$$r + \frac{\lambda}{2} \leq p \leq r + \lambda \tag{5}$$

Remark 1. *The equilibrium x^H exists whenever x^L does but not vice versa*

Thus when (4) holds, $x^{\mathrm{H}} = \lambda$ is an equilibrium – the "high" equilibrium – and when and $x^{\mathrm{L}} = 2\,(r + \lambda - p)$ are both trigger equilibria. On the contrary, (4) fails, no trigger equilibrium exists. The parameter estimates in Table 1 indicate that x^{H} could be an equilibrium (alternative to one that was estimated), but not x^{L}.

The grim-trigger equilibrium at x^{H} is not consistent with evidence because bad news never takes place – it has no "recalls" because the bad news hazard is zero. Equilibrium x^{L} has recalls, but equilibria x^{L} and x^{H} are both inconsistent with the evidence which says that upon having a product recall, a seller's value falls, but not to zero – sellers typically continue making profits. Our parameter estimates are based on the equilibrium described in Proposition 1 which is broadly consistent with that evidence. The estimates do imply, however, that (4) holds, so that even though a different equilibrium appeared to be in force, the environment could have given rise to a grim-trigger equilibrium too.

2.2 Ratchet equilibrium

We shall solve for an equilibrium in which, if the seller produces the bad signal, the market punishes him by "re-initializing" his expected lifetime value at v_0, leading to a value loss of $(v - v_0)$. That is, we define a Nash equilibrium in strategies $(x_t)_{t\geq 0}$ and prices such that the seller's price depends positively on the time elapsed since the last signal, and such that in equilibrium the public's expectations are correct, i.e.,

$$x_t = x_t^*. \tag{6}$$

The parameter v_0 is exogenous and will later be estimated. For now we note only that it cannot be less than the seller's outside option and so $v_0 \geq 0$.[2]

The HJB equation.—Let t denote time elapsed since the last signal, and suppose that t is the only variable determining effort. Let r be the rate of discount. Conditional on $(x_t^*)_0^\infty$, Bellman eq. yields the following equation for the lifetime value v

$$rv = \max_{x\leq\lambda}\left(px^* - \frac{x^2}{2} - (\lambda - x)(v - v_0) + \frac{dv}{dt}\right) \tag{7}$$

where we drop the t subscript to keep notation simple. The problem is concave in x and the first-order condition is

$$x = v - v_0. \tag{8}$$

We have assumed that the maximum is interior. We shall check ex post that $x_t < \lambda$ for all t.[3]

Substituting for x and for x^* in (7), it reads

$$rv = p(v - v_0) - \frac{(v - v_0)^2}{2} - (\lambda - (v - v_0))(v - v_0) + \frac{dv}{dt}.$$

But it is simpler to eliminate v instead; since $x = v - v_0$, we have the differential equation for x

$$\frac{dx}{dt} = rv_0 + (r + \lambda - p)x - \frac{1}{2}x^2, \tag{9}$$

[2] If v_0 is the resale value of the firm then $v_0 \geq 0$, because the owners of the firm are not forced to sell it and can just dispose of it freely.

[3] Implicit in (7) is the assumption that there are many more buyers than sellers and that the seller meets new buyers continually.

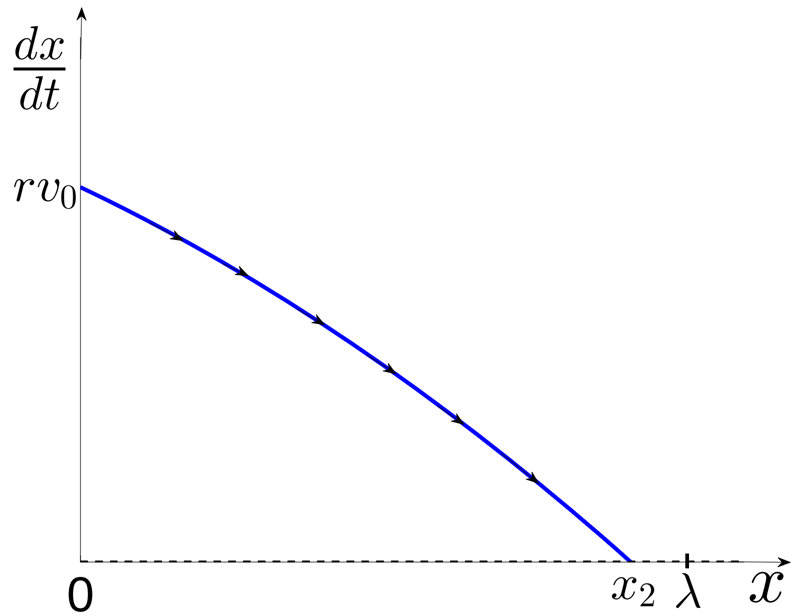

Figure 1: THE RHS OF (9) AS A FUNCTION OF x.

with the initial condition $x_0 = 0$. This is an equation of the Ricatti type.

Let x_1 and x_2 be the two roots of x of the equation $dx/dt = 0$, i.e., the two values of x at which the RHS of (9) is zero. If $v_0 > 0$, we have $x_1 < 0 < x_2$, and[4]

$$x_1 = r + \lambda - p - \sqrt{(r + \lambda - p)^2 + 2rv_0} < 0, \quad \text{and} \tag{10}$$

$$x_2 = r + \lambda - p + \sqrt{(r + \lambda - p)^2 + 2rv_0} > 0 \tag{11}$$

The dynamics are illustrated in Fig. 1. At $x = 0$, $dx/dt = v_0$. Thus if $v_0 > 0$,

$$\frac{dx}{dt} > 0, \text{ and } \lim_{t\to\infty} x_t = x_2. \tag{12}$$

The RHS of (9) peaks at the value $\hat{x} = r + \lambda - p$, which is where of dx/dt is largest. According to the estimates in Table 1, however, $\hat{x} = -0.34$, which means that $\frac{d^2x}{dt^2} < 0$, i.e., x grows at a decreasing rate.

[4] If $v_0 < 0$, $x_1 > 0$, in which case the FOCs would say that at $t = 0$ where $x = 0$, $dx/dt|_{t=0} < 0$. We shall assume $v_0 > 0$.

These statements hold iff $x_t < \lambda$ for all t. From (12) this is equivalent to the requirement that[5]

$$x_2 < \lambda, \tag{13}$$

This places a restriction on the admissible v_0:

$$0 < v_0 < \frac{p\lambda - \lambda^2/2}{r} - \lambda \tag{14}$$

The derivation of (14) is in Appendix 3.

First best rents which are

$$v_0^{\text{best}} = \int_0^\infty e^{-rt} \max\left(px - \frac{1}{2}x^2\right) dt = \frac{p^2}{2r}. \tag{15}$$

Claim 1.

$$v_0 < v_0^{best}$$

Proof. From $x_2 < \lambda$, we have $v_0 < \frac{p\lambda - \lambda^2/2}{r} - \lambda \leq \frac{p^2}{2r} - \lambda < \frac{p^2}{2r} = v_0^{\text{best}}$, where the second inequality comes from $\max_\lambda\{p\lambda - \lambda^2/2\} = \frac{p^2}{2}$. ∎

Thus any v_0 satisfying (14) is an equilibrium.

2.2.1 The equilibrium (x, v, F, G)

From the waiting-time distribution F in (2), we shall be able to derive the distribution of x at the moment that the signal arrives. This will enable the model to explain some data on the loss in market value of sellers that experience product recalls.

Denote the solution for $x_t \equiv \phi(t)$. Let $G(x)$ be the CDF of x_t at the time the signal arrives. Since ϕ is strictly increasing,

$$G(x) \equiv \Pr(x_\tau \leq x) = F\left(\phi^{-1}(x)\right).$$

Appendix 1 proves the following claim:

Proposition 1. *If (13) holds, the ODE in (9) has the solution*

$$x_t = \frac{x_2(1 - \exp\{-\frac{1}{2}(x_2 - x_1)t\})}{1 - \frac{x_2}{x_1}\exp\{-\frac{1}{2}(x_2 - x_1)t\}} \tag{16}$$

[5] The case where condition (13) is not met is discussed in Sec. 2.2.2 and in Appendix 4.

for all $t \geq 0$, *with* x_1 *and* x_2 *given in (10) and (11). Also,*

$$v_t = v_0 + x_t$$

$$F(t) = 1 - \left(\frac{1 - \frac{x_2}{x_1} \exp\{-\frac{1}{2}(x_2 - x_1)t\}}{1 - \frac{x_2}{x_1}} \right)^2 \exp\left(-(\lambda - x_2)t\right) \tag{17}$$

and

$$G(x) = 1 - \left(\frac{x_1}{x - x_1} \right)^{\frac{2(\lambda - x_1)}{(x_2 - x_1)}} \left(\frac{x - x_2}{x_2} \right)^{\frac{2(\lambda - x_2)}{(x_2 - x_1)}} \tag{18}$$

The solution for x_t depends on the parameters (p, v_0, r, λ) only through their effect (x_1, x_2). The solutions are depicted in the 4-quadrant Fig. 2, evaluated at the parameter estimates in Table 1. Panel 1 is an expanded version of 1, where we see the RHS of (9) crossing the zero axis at $x_1 = -.72$ and at $x_2 = .024$. These are the values given in (10) and (11) also evaluated at the estimates in Table 1.

Panel 2 shows that x indeed approaches x_2 as stated in (12). As x rises, v also rises. Additionally, v rises because the hazard rate $\lambda - x$declines, reducing the likelihood of a reversion to v_0.The third and fourth panels show the densities of F and G which are derived analytically in Appendix 2.

The distribution of % losses in value at recall.–Define the percentage loss at recall as z:

$$z = \frac{v - v_0}{v} = \frac{x}{x + v_0} = \frac{1}{1 + v_0/x}$$

which is an increasing function of x. The larger is z, the larger is the loss in value. Solving for x in terms of z yields $x = v_0 \frac{z}{1-z}$, and so the CDF of z is

$$\Psi(z) = G\left(v_0 \frac{z}{1 - z} \right) \qquad \text{for } z \in \left[0, \frac{1}{1 + \frac{v_0}{x_2}} \right]$$

and its density, $\psi(z)$, is illustrated in Fig. 3.Jarrell and Peltzman (1984) provide evidence that

$$E(z) \equiv \int_0^1 z\psi(z)\, dz \geq 0.027 \tag{19}$$

The estimate is conservative and so we shall use the inequality (19) to constrain the ML estimates.

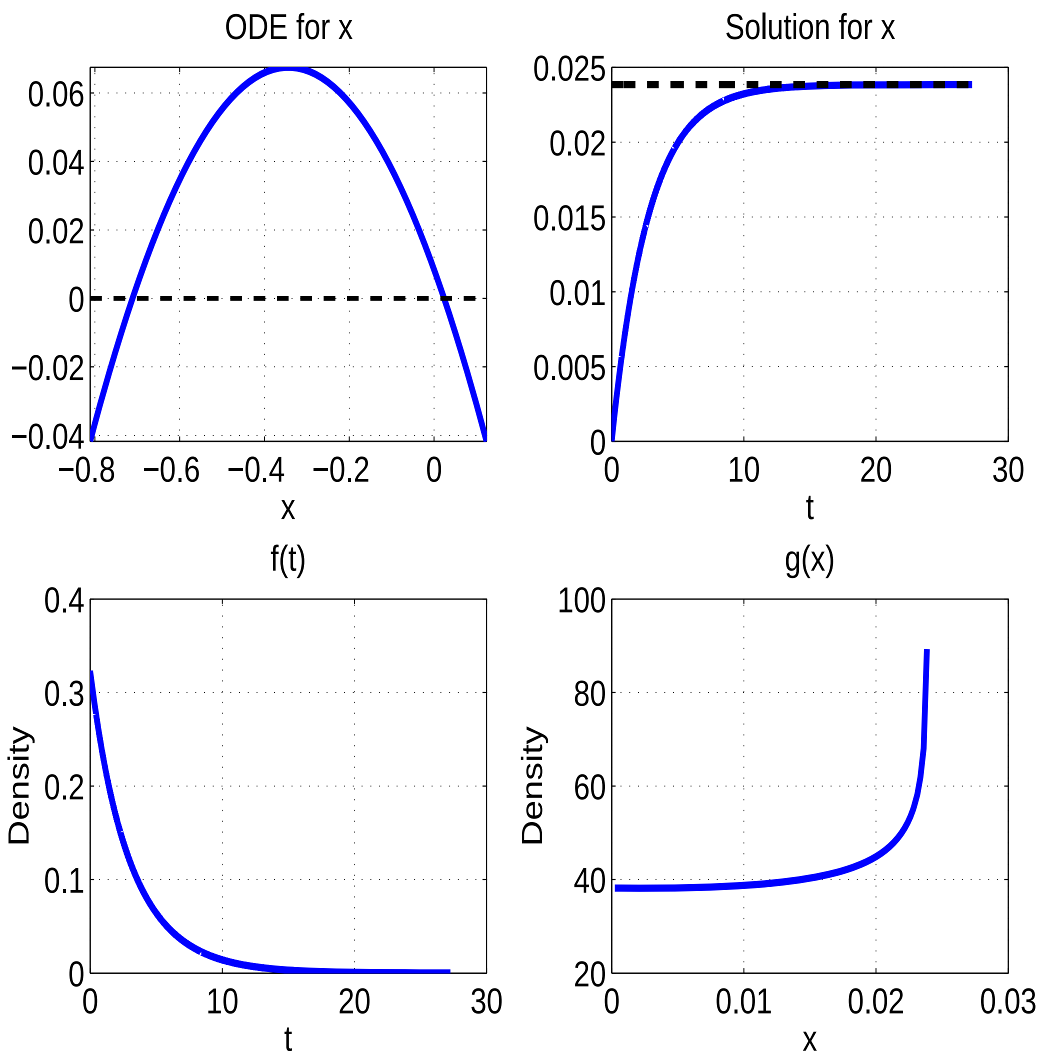

Figure 2: The solution evaluated at the parameter values in Table 1

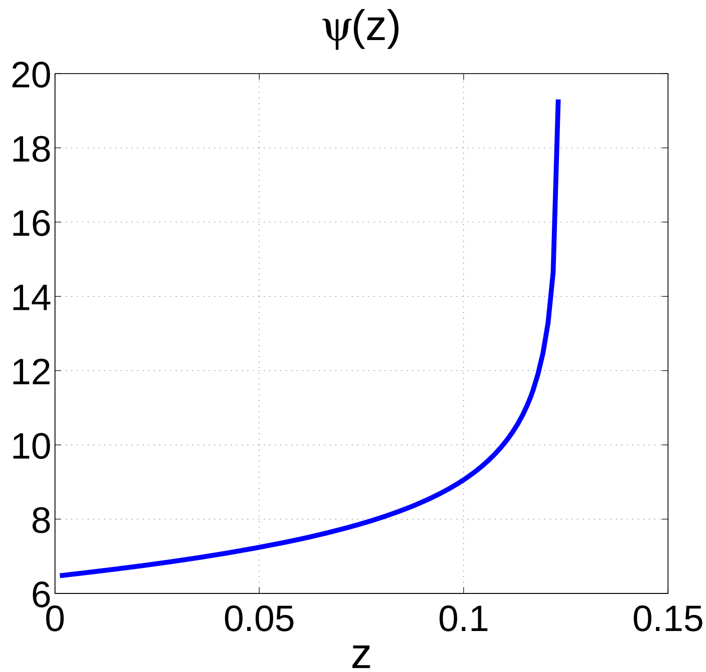

Figure 3: Density of losses z at recall, $h(z)$

2.2.2 The case $x_2 > \lambda$

This case is of special interest because, if an equilibrium exists that satisfies (16) on the region where $x_2 \leq \lambda$ it implies that x reaches λ in finite time, say t_λ, and if it remains at λ the probability of a signal would become zero. This would mean that $F(t)$ would be a defective distribution with the property that $F(t) = F(t_\lambda) < 1$ and that $v_t = v_{t_\lambda}$ for all $t \geq t_\lambda$.

No equilibrium of the type described here exists in this case, for any v_0. In this case x reaches λ in finite time and after it does so, v would have to remain at rest since the signal hazard then becomes zero. But this leads to a contradiction. Using (11), $x_2 > \lambda$ implies that

$$r - p + \sqrt{(r + \lambda - p)^2 + 2rv_0} > 0 \tag{20}$$

which holds automatically if $p < r$.

Case 1.— $p \leq r$. Then we can show that $v_0 < 0$. We have

$$\begin{aligned} x_t &= \frac{x_2(1 - \exp\{-\frac{1}{2}(x_2 - x_1)t\})}{1 - \frac{x_2}{x_1}\exp\{-\frac{1}{2}(x_2 - x_1)t\}} \\ v_t &= x_t + v_0 \end{aligned}$$

with $x_0 = 0$ and $v_0 = \frac{1}{r}(p\lambda - \frac{\lambda^2}{2}) - \lambda$. Then, if $p = r$, $v_0 = -\lambda^2/2r$, which contradicts the assumption that $v_0 \geq 0$.

Case 2.— $p > r$. Let us assume $x_2 > \lambda$ and, hence (20) can hold. Rearranging (20), we get $(r + \lambda - p)^2 + 2rv_0 > (p - r)^2$, (this step is needed only when $p > r$). i.e., $\lambda^2 + 2\lambda(r - p) + 2rv_0 > 0$, i.e.,

$$v_0 + \lambda > \frac{1}{r}(p\lambda - \frac{\lambda^2}{2}). \tag{21}$$

If equilibrium satisfying eq. (7) did exist, x_t would, according to (39), reach λ in finite time, say t_λ. From (8), we know that $v_t = v_0 + x_t$ for all $t < t_\lambda$, and $\lim_{t \uparrow t_\lambda} v_t = v_0 + \lambda$. If $v_t = v_{t_\lambda}$ for $t \geq t_\lambda$, (8) would continue to hold for $t \geq t_\lambda$, and hence $x_t = \lambda$. No further signals could then arise since their hazard rate is zero. But then

$$v_{t_\lambda} = v_0 + \lambda = \frac{1}{r}(p\lambda - \frac{\lambda^2}{2}), \tag{22}$$

which contradicts (21). Thus the type of equilibrium studied above cannot exist, since the approach of $x \to \lambda$ is not a Poisson event but occurs deterministically.

Since it does not involve the parameter v_0, the trigger equilibrium condition (4) is also independent of whether $\lambda \gtrless x_2$. Thus the trigger equilibrium can exist when the ratchet equilibrium fails to exist and vice versa. Let

$$T(v_0) \equiv \max_{(x_t)_{s=t}^{\infty}} \left\{ \int_0^\infty \left(\int_0^s e^{-rt} \left(px_t^*(\hat{v}_0) - \frac{x_t^2}{2} \right) dt + e^{-rs}\hat{v}_0 \right) f(s, (x_t)^s)\, ds \right\},$$

so that a stationary equilibrium is a number v_0 such that

$$v_0 = T(v_0).$$

A stationary solution $v_0 = 0 = T(0)$ always exists, and one can derive conditions under which a positive solution v_0 also exists by using the fact that v_0 is bounded and finding parameter values for which $T'(0) > 1$. I cannot solve for any positive solution analytically, however, and I therefore move on to another environment that I'll refer to as a "mixed" equilibrium.

2.3 Mixed equilibrium

The ratchet equilibrium exists only if $x_2 < \lambda$ which means that $1 - \exp\left(-\int_0^\infty (\lambda - x_t)\, dt\right) = 1$, i.e., every product eventually receives the bad signal and is eventually recalled. To have a positive fraction of products that never experience a recall, we now introduce a random time at which expectations shift to a high grim-trigger equilibrium in which there are never any recalls.

We introduce an effort-dependent hazard of switching from the ratchet equilibrium to one of the grim-trigger equilibria. A switch to a grim trigger equilibrium represents good news for the firm. The result is a mixture of "good news" and "bad news" signals in the form of competing hazards, both of which depend on the firm's effort x.

Interpretation of the good news signal.—Since it results in an upward jump in the value of the firm from v to v^{g}, we may look for examples of news shocks that deliver such positive stock-price hike. A favorable earnings announcement or some discovery or patent would be examples, although the model has no role for why such news should matter except that they change everyone's expectations.

In a ratchet equilibrium, let us introduce the hazard, θx, of receiving a "good news" signal taking subsequent expectations and play to a grim-trigger equilibrium (3). This signal is public and indicates a belief switch to $x^{\mathrm{g}} = v^{\mathrm{g}}$ for ever, where $g \in \{L, H\}$. We note that v^{g} is continuous in p around the point where $v = \lambda$, i.e., at the point where $p = r + \lambda/2$ both expressions for v^{g} are the same and equal to λ.

The HJB equation now is

$$rv = \max_x \left\{ px^* - \frac{x^2}{2} + (\lambda - x)(v_0 - v) + \theta x (v^{\mathrm{g}} - v) + \frac{dv}{dt} \right\} \tag{23}$$

with the FOC

$$\begin{aligned} x &= v - v_0 + \theta (v^{\mathrm{g}} - v) \\ &= (1-\theta) v + \theta v^{\mathrm{g}} - v_0. \end{aligned} \tag{24}$$

We see a qualitatively different relation between x and v depending on θ: For $\theta < 1$, x increases with v whereas when $\theta < 1$, x *decreases* with v. Solving for v in terms of x we get

$$v = \frac{1}{1-\theta} (x + v_0 - \theta v^{\mathrm{g}}). \tag{25}$$

Substituting (25) into (23), we have

$$r \frac{x + v_0 - \theta v^g}{1-\theta} = px - \frac{x^2}{2} + \lambda \left(v_0 - \frac{x + v_0 - \theta v^g}{1-\theta} \right) + x^2 + \frac{1}{1-\theta} \frac{dx}{dt}.$$

Rearranging, we get

$$\frac{dx}{dt} = -\frac{1-\theta}{2}x^2 + (\lambda + r - (1-\theta)\,p)\,x + (\theta\lambda + r)\,v_0 - \theta\,(\lambda + r)\,v^g \qquad (26)$$

Eq. (26) is once again a Ricatti eq., again with $\theta = 0$ giving us the ratchet eq., and $\theta = +\infty$ giving us a grim trigger equilibrium, provided that equilibrium exists which has yet to be shown.

Three regions for θ.—When $\theta < 1$, the RHS of (26) has an inverted-U shape, just as in Panel 1 of Fig 2, and we can expect the dynamics qualitatively the same as those we studied for the ratchet equilibrium, with $\theta = 0$ corresponding to the pure ratchet equilibrium. When $\theta = 1$ the RHS of 26) is linear in x and when $\theta > 1$ it is U-shaped. These two cases will give solutions for the dynamics of (x_t, v_t) that are qualitatively different. We focus on the case $\theta \leq 1$, because otherwise x is decreasing in v.

2.3.1 The case $\theta = 1$

This case solves analytically. It yields a constant x which, in turn, yields a constant recall hazard as well as a constant hazard θ of a switch to the trigger equilibrium.

The FOC for x, (24) now becomes

$$x = v^{\mathrm{g}} - v_0,$$

so that x is a constant. Then v_t is also a constant at v_0, which solves

$$rv_0 = p\,(v^{\mathrm{g}} - v_0) + \frac{(v^{\mathrm{g}} - v_0)^2}{2}$$

Equivalently, x solves

$$r\,(v^{\mathrm{g}} - x) = px + \frac{x^2}{2},$$

which gives us the solution for the constant value of x which we shall denote by $\hat{x}$:

$$\begin{aligned} \hat{x} &= -(p+r) + \sqrt{(p+r)^2 + 2rv^{\mathrm{g}}} \\ &= -(p+r) + \sqrt{(p+r)^2 + 2\left(p\lambda - \frac{\lambda^2}{2}\right)} \\ &< \lambda \end{aligned}$$

Solving for v_0, we have

$$\begin{aligned} v_0 &= v^{\mathrm{g}} - x = \frac{1}{r}\left(p\lambda - \frac{\lambda^2}{2}\right) + (p+r) - \sqrt{(p+r)^2 + 2\left(p\lambda - \frac{\lambda^2}{2}\right)} \quad (27) \\ &> \frac{1}{r}\left(p\lambda - \frac{\lambda^2}{2}\right) - \lambda. \end{aligned}$$

Therefore, when $p \geq r + \frac{\lambda}{2}$, $v_0 > 0$.

Appendix 4A shows that

$$F(t \mid \theta = 1) = \frac{\lambda - \hat{x}}{\lambda}\left(1 - e^{-\lambda t}\right) \to \frac{\lambda - \hat{x}}{\lambda} = F(\infty \mid \theta = 1), \qquad (28)$$

and its hazard rate is

$$h(t) = \frac{(\lambda - \hat{x})\, e^{-\lambda t}}{1 - \frac{\lambda - \hat{x}}{\lambda}\left(1 - e^{-\lambda t}\right)} = \frac{\lambda(\lambda - \hat{x})\, e^{-\lambda t}}{\lambda - (\lambda - \hat{x})\left(1 - e^{-\lambda t}\right)} \to 0.$$

2.3.2 The case $\theta \in (0,1)$

This case is especially interesting because we shall estimate $\theta \approx 0.3$. Let x_1 and x_2 be the roots for x of the quadratic expression on the RHS of (26). We have

$$x_1 = \frac{\lambda + r - (1-\theta)p - \sqrt{(\lambda + r - (1-\theta)p)^2 + 2(1-\theta)\left[(\theta\lambda + r)v_0 - \theta(\lambda + r)v^g\right]}}{1-\theta} \qquad (29)$$

$$x_2 = \frac{\lambda + r - (1-\theta)p + \sqrt{(\lambda + r - (1-\theta)p)^2 + 2(1-\theta)\left[(\theta\lambda + r)v_0 - \theta(\lambda + r)v^g\right]}}{1-\theta} \qquad (30)$$

which (i) coincide with (10) and (11) when $\theta = 0$ and (ii) are continuous in θ around that point.

The initial condition for x is $x(0) = \theta\left(v^{\mathrm{g}} - v_0\right) \equiv x_0$. For the equilibrium with increasing effort and value to exist, we need $x_1 < x_0 < x_2$**.** As long as $v_0 > 0$, this condition holds for small θ, and then we have the solution

$$x_t = \frac{x_2 + \frac{x_2 - x_0}{x_0 - x_1} x_1 \exp\{-\frac{1-\theta}{2}(x_2 - x_1)t\})}{1 + \frac{x_2 - x_0}{x_0 - x_1}\exp\{-\frac{1-\theta}{2}(x_2 - x_1)t\}} \to x_2 \qquad (31)$$

as $t \to \infty$. This solution features a decreasing hazard for product recalls, such as the data show qualitatively.

The case where $\lambda < x_2$ is discussed in Appendix 4. No equilibrium appears to exists for this case even when $\theta > 0$, just as it does not exist when $\theta = 0$.

The probability of no recalls taking place, ever

In the mixed equilibrium a positive fraction of products never experiences a recall, we now introduce a random time at which expectations shift to a high grim-trigger equilibrium in which there are never any recalls. When $\theta = 0$ we are in a ratchet equilibrium forever. Let $\tau =$ the date when there is a regime switch to grim trigger, and let $t =$ the recall date, so that

$$\Pr\left(\tilde{t} \leq t \mid \tau\right) = F\left(\min\left(t, \tau\right)\right)$$

Unconditional is

$$F\left(t \mid \theta\right) = \int_0^{\infty} F\left(\min\left(t, \tau\right)\right) \theta x_\tau \exp\left(-\int_0^{\tau} \theta x_s ds\right) d\tau \qquad (32)$$

The probability that a recall will ever take place is

$$F\left(\infty \mid \theta\right) = \int_0^{\infty} F\left(\tau\right) \theta x_\tau \exp\left(-\int_0^{\tau} \theta x_s ds\right) d\tau$$

Fig. 4 plots $F\left(\infty \mid \theta\right)$ on θ while fixing the other parameter values at their estimates in Table 1 (which is in the next section):

At the estimate $\hat{\theta} = 0.015$ reported in Table 1A, as many as 99% of the products are eventually recalled. If course, a high enough θ could produce far fewer recalls in that

$$\lim_{\theta \to \infty} F\left(\infty \mid \theta\right) = 0.$$

3 Reputation and product recalls

A product recall is still a relatively rare event and it represents bad news for the producer, and the company's stock price falls.[6] In the ratchet equilibrium

[6] Barber and Darrough (1996) find that on average, over their 20-year sample period, the number of vehicles recalled in each calendar year by the Big Three automakers was

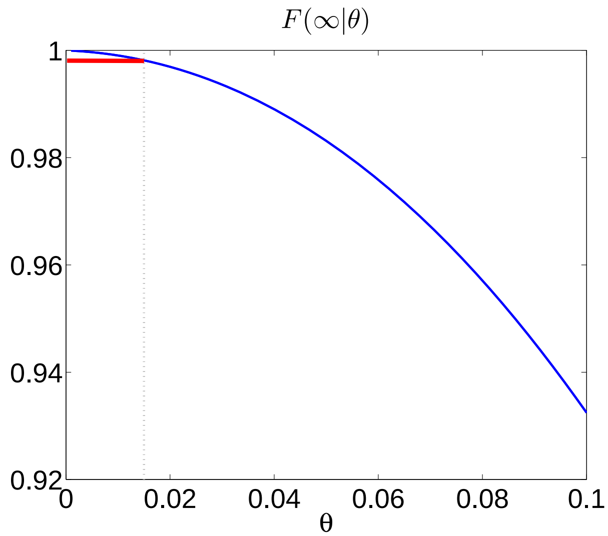

Figure 4: $F(\infty \mid \theta)$: The probability that a recall ever occurs

the drop is by the absolute amount $v - v_0$, and relative to pre-recall value it is $\frac{v-v_0}{v} = \frac{x}{x+v_0}$.

Product recalls, and Jarrell & Peltzman (1985) estimate the losses borne by owners of a firm that recalls a defective product from the market. The fall in value is too big to reflect only the cost of repairing the defective goods and compensating the owners. Instead, it is a reflection on the future dividends as they reflect the firm's quality. Jarrell and Peltzman write: "Our answer-for producers of drugs and autos that were recalled from the market-is that the shareholders bear large losses. They are substantially greater than the costs directly emanating from the recall-for example, costs of destroying or repairing defective products. In fact, they are plausibly larger than all the costs attributable specifically to the recalled product; the losses spill over to the firm's 'goodwill'." They speculate that this translates into reduced sales, increased quality costs, and so on. For automobile companies, Jarrell and Peltzman estimate the cumulative abnormal returns effects to be between -0.004 and -0.035, Rupp (2004) finds it to be -0.077, Barber and Darrough

equivalent to 50 percent (for Chrysler), 49 percent (for Ford), and 45 percent (for GM) of total vehicles sold in the United States in the preceding three years by each automaker. The recall rate by the three Japanese firms was equivalent to 37 percent (for Honda), 25 percent (for Nissan), and 21 percent (for Toyota).

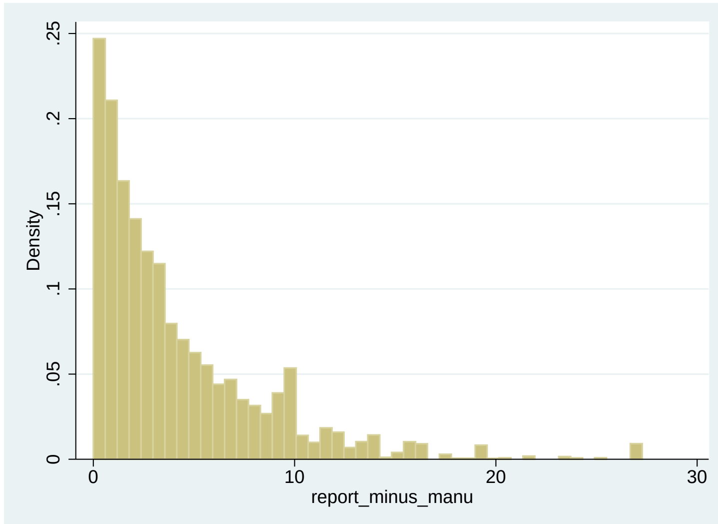

Figure 5: Recall frequency as a function of product age

(1996) estimate the returns declines to be between -0.001 and -0.013, and Hoffer, Pruitt and Reilly (1988). The mid point of these estimates is roughly -0.027, and we shall target this value to begin with and then experiment with a lower value.

Product-recall data.—We updated the auto recall data from the Department of Transportation, obtaining 48,000 observations covering the period 1978-2007. We measure "age at recall" as the difference between the product's recall date and the "start of manufacture" of the product to be 4.14 years. Details are in Appendix 5. Fig.5 shows the frequency distribution of the ages of the products at recall.At first look the data appear to be exponentially distributed, but in fact hazard is decreasing by more than six tenths of a percent per year. This is shown in Fig 6 which reports the annualized hazard rate

$$h(t) = \frac{f(t)}{1 - F(t-1)}$$

where F is the empirical CDF. On the face of it, this supports the "ratchet equilibrium" implication that x_t increases with t so that the hazard $h \equiv \lambda - x_t \searrow \lambda - x_2$ declines monotonically. The trigger equilibrium has a hazard that is either zero or constant.

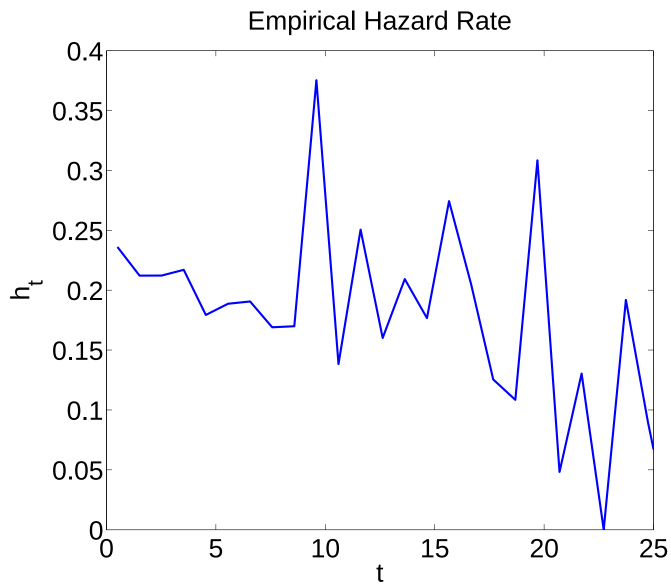

Figure 6: RECALL HAZARD AS A FUNCTION OF PRODUCT AGE

3.1 Estimates under homogeneity and $\theta = 0$

We shall first estimate the ratchet equilibrium, i.e., we assume that $\theta = 0$. The first round of estimates presume that all sellers share the same parameters (r, λ, p, v_0).

Interpretation.—We interpret F as the age distribution of the product at time of recall. We interpret G as the distribution of x at time of recall so that that the expected losses $E(z)$ in (19) is the average loss in value following a product recall, and we shall use the inequality (19) to constrain the ML estimates.

Estimation.—The time unit is a year and we set $r = 0.05$. Using the densities f and g shown in (40) and (41) leads to the constrained maximum likelihood estimation problem

$$\max_{(p,v_0,\lambda)} \prod_i f(t_i) \qquad \text{s.t.} \qquad E(z) \geq 0.027.$$

The constraint we impose is that the loss in firm value caused by a recall are

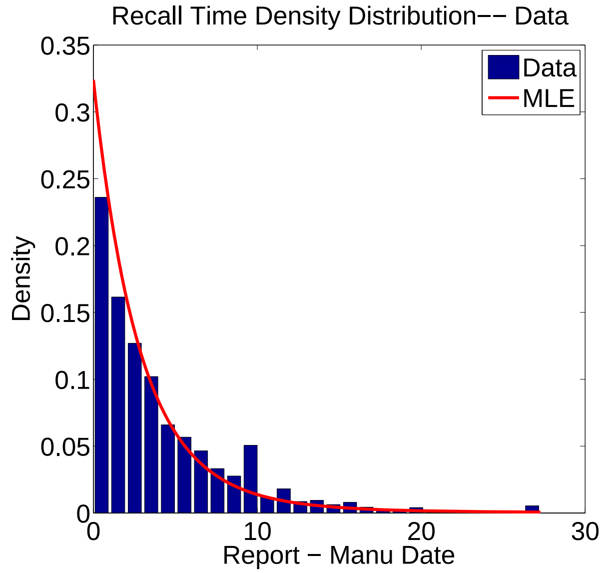

Figure 7: RATCHET EQUILIBRIUM ($\theta = 0$): FIT OF THE ML-ESTIMATED $f(t)$ UNDER HOMOGENEITY IN λ AND $E(z) \geq .027$.

at least 2.7 percent. The parameter estimates are reported in Table 1

Table 1: CONSTRAINED ML ESTIMATES

r	λ	p	v_0	x_1	x_2	$E(z)$
0.050	0.324	0.718	0.169	−0.712	0.024	0.027

First, the table shows that the constraint $E(z) \geq 0.027$ was binding. As a result we shall loosen this constraint in the next subsection. Second, the estimates do satisfy the constraint that $x_2 < \lambda$, which was sufficient for an interior solution for x in eq. (8). Note, finally, that at the estimated values in Table 1, (4) holds and the high grim-trigger equilibrium with $x^{\mathrm{H}} = \lambda$ therefore exists. The equilibrium $x^{\mathrm{L}} = 2(r + \lambda - p)$ does not. We shall compare their welfare levels later.

Maximal value loss.—Maximal loss in value is just over 12 percent because at the estimates in Table 1, the expression for the largest value yields

$$\frac{1}{1 + v_0/x_2} = 0.123.$$

Fig. 7 shows the model fit to the frequency distribution of the ages of the products at recall, The fit is very good, and the condition (13) is met so that Proposition 1 is valid at the estimated parameters.

Evidently the estimated density is *too steep*. The model appears to fit well but it overpredicts the hazard of the signal early on. The predicted hazard under homogeneity is the solid black line in Fig 11. It is well known that in the presence of heterogeneity in hazard rates, survivorship bias induces a downward slope in the estimated hazards. This deficiency caused us to introduce some unobserved heterogeneity. We shall estimate a multinomial distribution for λ in the hope that the density will flatten out a bid, especially early on.

The density f is, for the mixed case, also derived in Appendix 4, is needed for the ML estimation

The MLE of the mixed equilibrium is

$$\max_{(p,v_0,\lambda,\theta)} \prod_i f(t_i) \quad s.t. \quad E(z) \geq 0.027.$$

conditional on $x_0 = \theta(v^g - v_0)$ and

$$v^g = \begin{cases} \frac{1}{r}\left(p\lambda - \frac{1}{2}\lambda^2\right) \equiv x^{\mathrm{H}} \\ 2(r + \lambda - p) \equiv x^{\mathrm{L}} \end{cases}$$

The parameter estimates are reported in Table 1A:

Table 1A: CONSTRAINED ML ESTIMATES – MIXED

r	λ	p	v_0	θ	x_1	x_2	$E(z)$
0.050	0.328	0.389	0.207	0.015	-0.086	0.075	0.027

Fig 8) plots the fitted density and the data, andand Fig. 9 shows how the hazard rate compares to data.

3.2 Eliminating survivorship bias under the constraint that $\theta = 0$

We shall now suppose that λ differs over the firms and that it is captured by a discrete support, i.e., a multinomial distribution. In a related context, Heckman-Singer (1984) assume a continuous distribution and argue that convergence is rapid when there are even just a few points of support.

Formally, we adopt the Kiefer-Wolfowitz (1954) method of estimating the likelihood function along with a mixing distribution for the unobserved

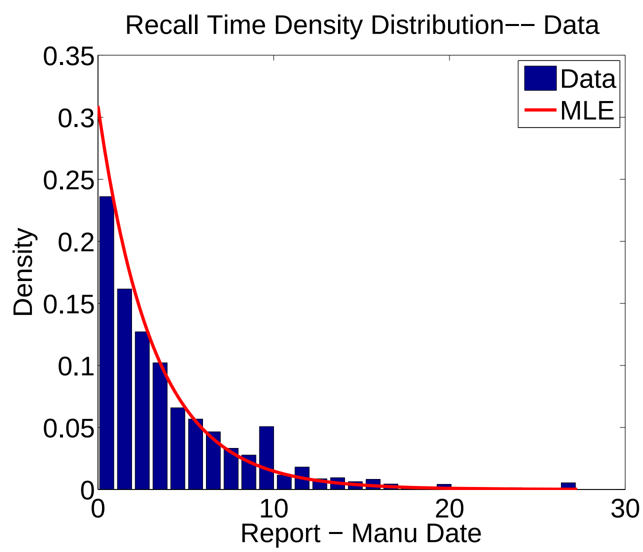

Figure 8: Mixed equilibrium fit under homogeneity of λ

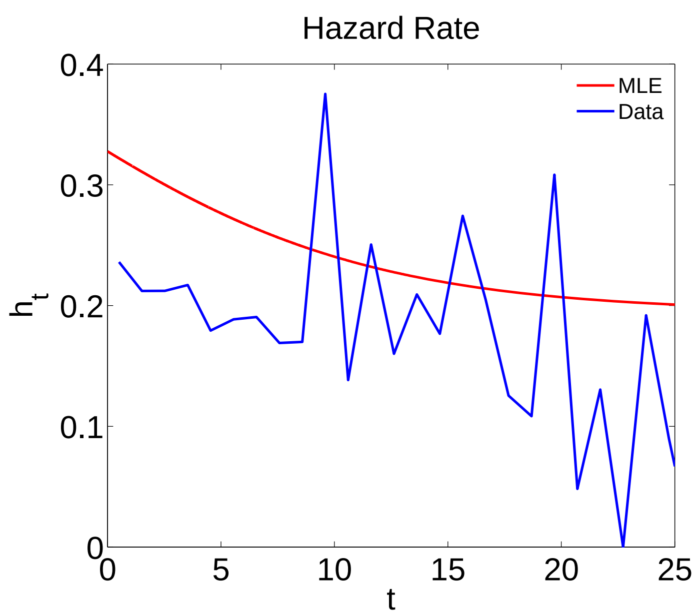

Figure 9: Mixed equilibrium fit to the hazard under homogeneity of λ

parameter that is assumed to vary. This will lead to a statistically significant estimate of the heterogeneity. Denote the CDF of λ to be and $\Phi(\lambda)$. This is the mixing distribution. The new likelihood of the observed duration t

$$\Lambda(t) = \int F(t \mid \lambda)\, d\Phi(\lambda)$$

where conditional on λ, is given in (17) where we now make explicit the dependence of x_1 and x_2 on λ

$$F(t \mid \lambda) = 1 - \left(\frac{1 - \frac{x_2(\lambda)}{x_1(\lambda)} \exp\{-\frac{1}{2}(x_2(\lambda) - x_1(\lambda))t\}}{1 - \frac{x_2(\lambda)}{x_1(\lambda)}} \right)^2 \exp\left(-(\lambda - x_2(\lambda))t\right). \tag{33}$$

Following Kiefer and Wolfowitz (1954) the likelihood becomes $\prod_i \Lambda(t_i)$, which is to be maximized with respect to the parameters of F and Φ.

We assume that $\Phi(\lambda)$ has a discrete support $\{\lambda_i\}_{i=1}^N$ with associated probability $\{\pi_i\}_{i=1}^N$ that satisfies $\sum \pi_i = 1$. We maintain the constraint (19), with the following change in the distribution of z:

$$\Psi(z) = \sum_{i=1}^{6} \pi_i \min\left\{1, G\left(v_0 \frac{z}{1-z} \mid \lambda_i\right)\right\} \qquad \text{for } z \geq 0. \tag{34}$$

So the objective function reads,

$$\max_{(p, v_0,, \lambda_1, \ldots, \lambda_{\mathrm{N}}, \pi_1,, \ldots \pi_{\mathrm{N}})} \prod_j \sum_{i=1}^{\mathrm{N}} \pi_i f(t_j \mid \lambda_i) \text{ s.t.} \quad E(z) \geq 0.027.$$

We have heterogeneity in hazard rates indexed by λ. But x_t also depends on λ. Write the generic hazard as $h_t = \lambda - x_t(\lambda)$. Heckman and Singer (1984) argued[7] that even moderately small values of N suffice to capture the heterogeneity in hazard rates, and we shall set $N = 6$. The results are in Table 2:

[7]Their hazard was Weibull, i.e., not of the same functional form as ours here. Recall that in addition to appearing additively, λ affects x_t directly through the roots (x_1, x_2) given in eqs (29) and (30).

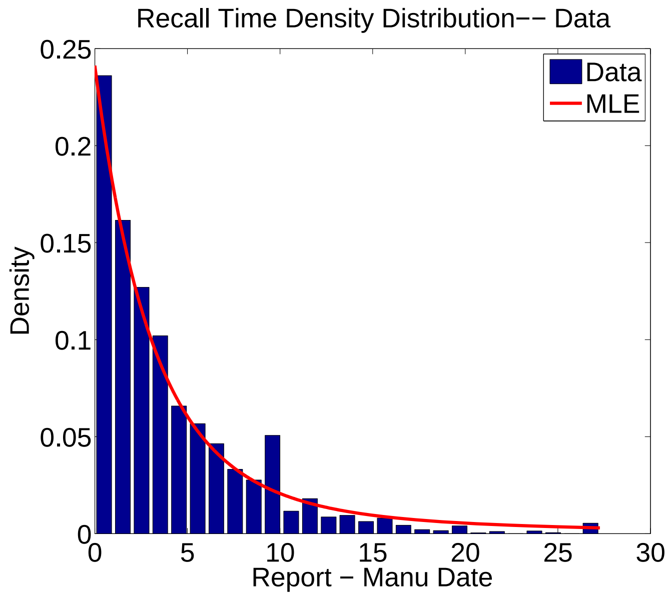

Figure 10: Ratchet equilibrium ($\theta = 0$): Fit of the ML-estimated $f(t)$ when $N = 6$ and $E(z) \geq .027$

Table 2 : ML estimates constrained by $E(z) \geq 0.027$ and $\theta = 0$

r	p	v_0	$E(z)$
0.050	0.306	0.117	0.027

Distribution $\Phi(\lambda)$ and the implied (x_1, x_2)

i	1	2	3	4	5	6
π	0.160	0.201	0.193	0.189	0.184	0.073
λ	0.241	0.289	0.316	0.331	0.349	0.400
x_1	-0.123	-0.080	-0.063	-0.056	-0.049	-0.036
x_2	0.094	0.147	0.184	0.206	0.236	0.324

Fig 10 shows the model fit to the frequency distribution of recalls. The model no longer overpredicts the density at young ages, rather the opposite, as the first four years are underpredicted..

Likelihood ratio statistic.—We calculate the statistic

$$D = -2\ln \frac{L\left(\text{data}|\lambda_1 = \lambda_2 = ... = \lambda_6 = 0.33\right)}{L\left(\text{data}|\hat{\lambda}_1, ..., \hat{\lambda}_6\right)} \sim \chi^2_{df=11}$$

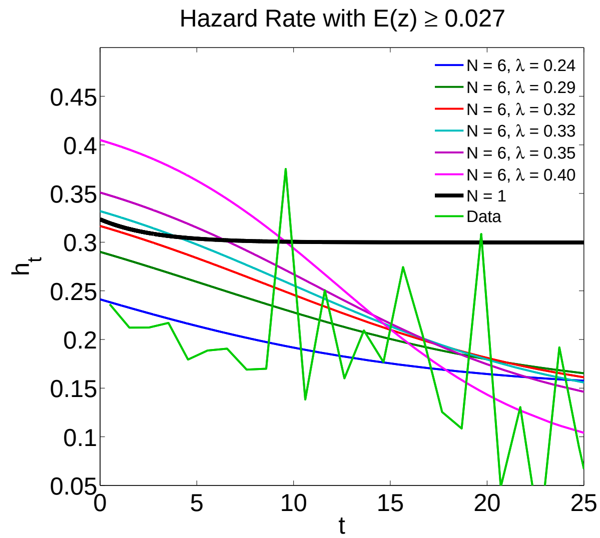

Figure 11: Hazard rates for each of the 6 λ values and for the baseline case $N = 1$.

That is, the statistic D follows a chi-square distribution with degree of freedom equal to the difference in the number of parameters for the two cases $N = 6$ and $N = 1$. When $N = 1$, we have 3 parameters (λ, p, v_0), whereas when $N = 6$, we have 2 $(p, v_0) + 6$ (for λ_i) $+ 6$ (for π_i) $= 14$ parameters. We have

$$\begin{aligned} \ln L\left(\text{data}|\hat{\lambda}_1, ..., \hat{\lambda}_6\right) &= -131.814, \quad \text{and} \\ \ln L\left(\text{data}|\lambda_1 = \lambda_2 = ... = \lambda_6 = 0.33\right) &= -144.267. \end{aligned}$$

Then

$$D = 2\left(-131.814 - (-144.267)\right) = 24.906,$$

whereas the 99% critical value for $\chi^2_{df=11}$ is 24.7250. Therefore the model with $N = 6$ does better than the $N = 1$ model at the 1% significance level.

More revealing than Fig 10 is the plot of the six estimated hazard rates $h_t(\lambda_i)$. They are shown in Fig. 11 together with the original estimate for which there was no heterogeneity. The range of t is the same as the range of of the data. Note that the six hazard rates do cross.

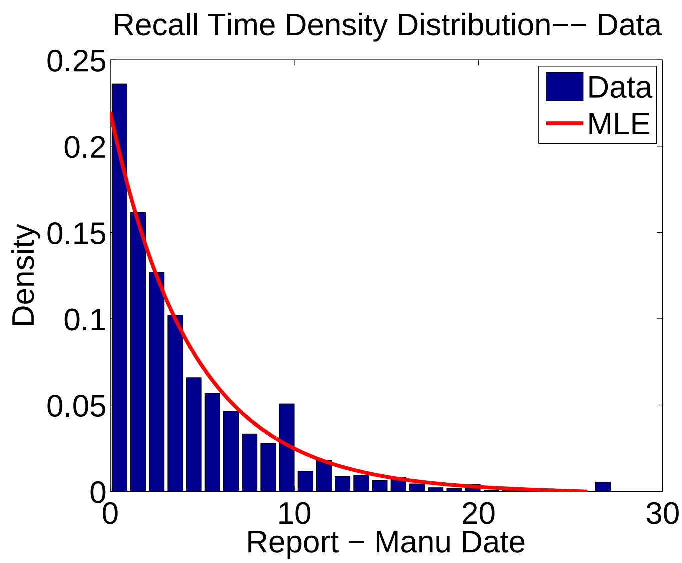

Figure 12: Mixed equilibrium: Fit of the ML estimate.

3.3 ML Estimation of the mixed equilibrium case

The ratchet equilibrium (i.e., the case $\theta = 0$) features every product being recalled eventually, because the existence of such an equilibrium requires that $x_2 < \lambda$

Table 2A: ML estimates constrained by $E(z) \geq 0.027$

	r	p	v_0	θ	$E(z)$
	0.050	0.255	0.398	0.302	0.027
Distribution $\Phi(\lambda)$ and the implied (x_1, x_2)						
i	1	2	3	4	5	6
π	0.017	0.017	0.017	0.017	0.017	0.016
λ	0.240	0.275	0.344	0.419	0.490	0.563
x_1	−0.109	−0.101	−0.088	−0.073	−0.060	−0.039
x_2	0.211	0.260	0.280	0.300	0.319	0.339

and the plot

Likelihood ratio statistic.—We use this statistic to check for whether relaxing the constraint $\theta = 0$ leads to significantly better fit. The likelihood ratio is

$$\begin{aligned} D &= -2\ln\frac{L\left(\text{data}|\lambda_1=\lambda_2=...=\lambda_6=0.33\right)}{L\left(\text{data}|\hat{\lambda}_1,...,\hat{\lambda}_6\right)} = 2\left(-128.3870-(-144.267)\right) \\ &= 31.76 \end{aligned}$$

which fits the data better than $\theta = 0$ above. Visual comparison of Fig 12 to Fig. 10 shows a better fit for the mixed equilibrium for ages 2-4, and 12+ but a worse fit for some of the ages in between.

3.3.1 Replacing constraint (19) by $E(z) = 0.005$

Hoffer, George E., Stephen W. Pruitt and Robert J. Reilly.(1988) argue that Jarrell and Peltzman (1984) overestimated the losses induced by product recalls and there may have been no significant effect on values. Indeed, Tables 1 and 2 show that the constraint 19) was binding. Thus we relax the constraint, replacing it by $E(z) = 0.005$, i.e., a product recall leads to half-percentage point loss in value. We re-estimate and report the results in Table 3. The fit is decidedly worse, in terms of the likelihood as well as the hazard-rate fit.

Table 3 : ML ESTIMATES CONSTRAINED BY $E(z) = 0.005$ AND $\theta = 0$

r	p	v_0	$E(z)$
0.050	0.324	0.131	0.005

Distribution of Heterogeneity Parameter λ

i	1	2	3	4	5	6
π	0.170	0.188	0.192	0.189	0.180	0.081
λ	0.240	0.300	0.325	0.351	0.380	0.418
x_1	−0.153	−0.091	−0.074	−0.061	−0.050	−0.040
x_2	0.086	0.144	0.177	0.216	0.263	0.328

Likelihood ratio statistic.—

$$\begin{aligned} D &= -2\ln\frac{L\left(\text{data}|\lambda_1=\lambda_2=...=\lambda_6=0.33\right)}{L\left(\text{data}|\hat{\lambda}_1,...,\hat{\lambda}_6\right)} = 2\left(-133.264-(-144.267)\right) \\ &= 22.006 \end{aligned}$$

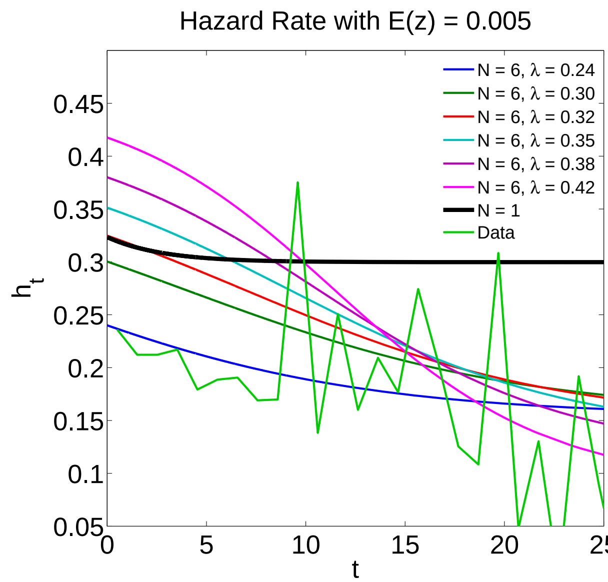

Figure 13: Hazard rates for each of the 6 λ values and for the baseline case $N = 1$.

which does not do as well as the previous version, but still better than the N = 1 version. Comparing Figures 11 and 13, we see that all 6 hazards are overpredicted when $E(z) = 0.005$.

3.3.2 Mixed equilibrium estimates with heterogeneous λ

The MLE mixed equilibrium estimates are reported in Table 3A

Table 3A : ML estimates constrained by $E(z) = 0.005$

		r	p	v_0	$E(z)$	θ
		0.050	0.298	0.134	0.005	0.025
Distribution of Heterogeneity Parameter λ						
i	1	2	3	4	5	6
π	0.213	0.215	0.217	0.218	0.110	0.027
λ	0.240	0.249	0.258	0.445	0.439	0.435
x_1	−0.007	−0.030	−0.030	−0.029	−0.017	−0.017
x_2	0.053	0.058	0.064	0.419	0.407	0.397

Likelihood ratio statistic.—

$$\begin{aligned} D &= -2\ln\frac{L\left(\text{data}|\lambda_1=\lambda_2=...=\lambda_6=0.33\right)}{L\left(\text{data}|\hat{\lambda}_1,...,\hat{\lambda}_6\right)} = 2\left(-129.9953-\left(-144.267\right)\right) \\ &= 33.28 \end{aligned}$$

and no surprise that the model with general θ fits the data better than $\theta = 0$ model with larger D.

3.3.3 Crossing of the hazard rates when $\theta = 0$

Figs 11 and 13 both show a tendency for the hazard rates to cross if they are associated with different values of λ. To understand these curves we note the following fact

Claim 2. *As λ varies in the region where*

$$\lambda > p - r, \tag{35}$$

the hazards must cross at least once.

Proof. The proof consists of showing that h_0 rises with λ whereas h_∞ moves in the opposite direction. Since $x_0 = 0$,

$$h_0\left(\lambda\right) = \lambda$$

is increasing in λ. On the other hand,

$$h_\infty \equiv \lim_{t\to\infty} h_t\left(\lambda\right) = \lambda - x_2 = p - r - \left(\left(p-r-\lambda\right)^2 + 2rv_0\right)^{1/2}.$$

Differentiating, (35) implies.

$$\frac{d}{d\lambda}h_\infty = \left(\left(p-r-\lambda\right)^2 + 2rv_0\right)^{-1/2}\left(p-r-\lambda\right) < 0.$$

■

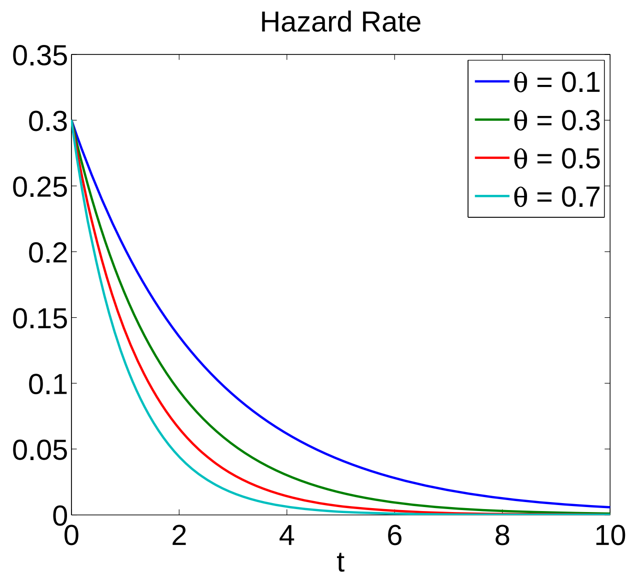

Figure 14: THE HAZARD RATE AS A FUNCTION OF θ

3.3.4 Hazards as θ varies

When $\theta > 0$, the recall density is scaled by $\Pr(\tau \geq t) = \exp\left(-\int_0^t \theta x_s ds\right)$. That is,

$$f(t \mid \theta) = (\lambda - x_t)\exp\left(-\int_0^t \theta x_s ds\right)$$

Then the hazard rate is

$$h(t \mid \theta) = \frac{(\lambda - x_t)\exp\left(-\int_0^t \theta x_s ds\right)}{1 - F(t \mid \theta)}$$

where $F(t \mid \theta)$ is defined in (32). The hazard rate thus decreases in θ, as Fig 14 shows

The other parameters used are given in Table 1. When $\theta \to \infty$, we have $\lim_{\theta\to\infty} F(\infty \mid \theta) = 0$ and hence

$$h(t|\infty) = \frac{(\lambda - x)\, e^{-[\theta+(\lambda-x)]t}}{1 - F(t \mid \theta)} \to 0$$

When $\theta \to 0$, we have

$$h(t|0) \to \lambda - x_t$$

4 Intangible capital and welfare

First best investment.—Since buyers are risk neutral, the first-best level of x is equal to

$$x^{\text{best}} \equiv \arg\max_{x \leq \lambda} \left\{ px - \frac{x^2}{2} \right\} = \min(p, \lambda) \tag{36}$$

Let us use the estimates from Table 1, from which we find that $\hat{\lambda} = 0.32 < \hat{p}$. From the second panel of Fig x we see the time path of x, and therefore much of the time x_t is well below $x_2 = .024$, and so

$$\frac{x_t}{x^{\text{best}}} \leq \frac{x_2}{x^{\text{best}}} = \frac{0.024}{0.32} = 0.075$$

Welfare.—All the rents go to the sellers, and their present value is $v_0 = 0.17$. Maximal rents are

$$v^{\text{best}} = \frac{1}{r} \max_x \left\{ px - \frac{x^2}{2} \right\} = \frac{1}{r} \left(p\lambda - \frac{\lambda^2}{2} \right) = 3.6 \tag{37}$$

Therefore society gets only a fraction

$$\frac{v_0}{v^{\text{best}}} = \frac{0.17}{3.6} = 0.047$$

of the potential rents, i.e., slightly under five percent.

If, instead of the Jarrell-Peltzman estimates of value loss, we were to use the estimates by Hoffer, Pruitt and Reilly (1988); they imply an even lower v_0, and an even lower fraction of potential rents that are realized in equilibrium.

How does welfare rank compared to what the grim-trigger equilibria would deliver and to what the mixed equilibrium with $\theta = 1$ would deliver? We noted that at the parameter estimates in Table 1, (4) holds so that if they chose to, players could coordinate on the high grim-trigger equilibrium. The low trigger equilibrium does not exist because the second inequality in (5) fails.

Table 4: Welfare		
ratchet equilibrium[8]	v_0 or $< v_0 + x_2$	0.169 or $<$ 0.193
grim trigger eq. low	$2(r+\lambda-p)$	does not exist
grim trigger eq. high	$\frac{1}{r}\left(p\lambda-\frac{\lambda^2}{2}\right)$	3.6
mixed eq. $\theta=1$	see eq. (27)	3.4
First best	see (37)	3.6

Thus the high grim-trigger equilibrium at $x^{\mathrm{H}} = \lambda$ delivers much higher – first best – welfare than the ratchet equilibrium. The low trigger equilibrium does not exist. The mixed equilibrium at $\theta = 1$ delivers almost as high a welfare level as the grim trigger equilibrium.

Admittedly, the parameter estimates were obtained on the assumption that the pure ratchet equilibrium was operative throughout the sample period. Had we assumed, instead, that one of the other equilibria had been in force, the parameter estimates would generally have been different.

5 Durable goods version

We now extend the model to a durable good – a “car” – the quality of which is interpreted as its breakdown rate, $\tilde{\rho}$ and mean lifetime $\tilde{\rho}^{-1}$. Flow utility of owning a car is u in units of the outside good which we take as the numeraire. Then given $\tilde{\rho}$, the probability that the car will still function τ periods from now is $e^{-\rho\tau}$, and so your willingness to pay for a new car at date t is

$$p_t = E\left(\int_t^\infty e^{-(r+\tilde{\rho})(s-t)}ds \mid \text{seller's public history}\right) = E_t\left(\frac{u}{r+\tilde{\rho}}\right).$$

As in Keller and Rady (2015), $\tilde{\rho}$ takes on just two values: $\tilde{\rho} \in \{0, \rho\}$, with

$$\tilde{\rho} = \begin{cases} 0 & \text{with prob. } x, \\ \rho & \text{with prob. } 1-x. \end{cases}$$

Then

$$p_t = x\frac{u}{r} + (1-x)\frac{u}{r+\rho} = \frac{u}{r+\rho} + \frac{\rho}{r(r+\rho)}ux.$$

[8] Welfare is most naturally measured at the starting date, $t = 0$, i.e., v_0. The upper bound for welfare is $\lim_{t\to\infty} v_t = v_0 + x_2$.

Let $p = Ax + B$ where

$$A = \frac{\rho}{r\,(r+\rho)} u \quad \text{and } B = \frac{u}{r+\rho}.$$

The public signal.—The signal is generated as the car rolls off the assembly line, which means that at t, the relevant effort variable is x_t. This is quite different from what would occur if the fault was discovered in a previously produced car, for which past levels of x_t would be relevant; we avoid this specification because it would lead to a second order differential equation for x.

Let μ be the rate at which a fault is detected conditional on the fact that a car is faulty. Since $1 - x$ is the probability of the car being faulty, the hazard for the public signal becomes $1 - x$. The Bellman equation then becomes

$$rv = B + Ax^{*} - \frac{1}{2}x^2 + \mu\,(1-x)\,(v_0 - v) + \frac{dv}{dt}$$

Then the FOC is

$$x = \mu\,(v - v_0)\,,$$

so that $dx/dt = \mu\,[dv/dt]$, and so that the ODE for x is

$$\begin{aligned} r\left(v_0 + \frac{x}{\mu}\right) &= B + Ax - \frac{1}{2}x^2 - \mu\,(1-x)\,\frac{x}{\mu} + \frac{1}{\mu}\frac{dx}{dt} \\ &= B + Ax + \frac{1}{2}x^2 - x + \frac{1}{\mu}\frac{dx}{dt}, \end{aligned}$$

i.e.,

$$\begin{aligned} \frac{1}{\mu}\frac{dx}{dt} &= r\,(v_0 + x) - B - Ax - \frac{1}{2}x^2 + x \\ &= rv_0 - B + \left(1 - A + \frac{r}{\mu}\right)x - \frac{1}{2}x^2 \end{aligned}$$

or

$$\frac{dx}{dt} = \mu\left[r\tilde{v}_0 - B + \left(1 - A + \frac{r}{\mu}\right)x - \frac{1}{2}x^2\right].$$

We can identify A, but not $\tilde{v}_0$ and B separately. Let us therefore set $\mu = 1$ (we cannot identify this parameter). The new ODE for x is equivalent to the old ODE $\frac{dx}{dt} = rv_0 + (r + \lambda - p)\,x - \frac{1}{2}x^2$ if the following restrictions hold:

$$\begin{aligned} 1 - A &= \lambda - p \\ r\tilde{v}_0 - B &= rv_0 \end{aligned}$$

Therefore we do not need to re-estimate, we simply calculate the new parameter set as a function of the old estimates. We calculate the implied value for A and for $r\tilde{v}_0 - B$ in terms of the old estimates of (p, v_0) given in Tables 1 and 1A. The resulting estimates are in Table 5:

Table 5: Constrained ML estimates: Homogeneity and $\theta = 0$

r	λ	v_0	x_1	x_2	θ	$E(z)$	A	$r\tilde{v}_0 - B$
0.050	0.324	0.169	-0.712	0.024	-	0.027	1.3940	0.0085
Constrained ML estimates – Homogeneity and $\theta \geq 0$								
0.050	0.328	0.207	-0.086	0.075	0.015	0.027	1.0610	0.0104

6 Conclusion

We modeled a seller that faces a continuum of risk-neutral, short-lived buyers, and whose effort is not observed and payment for its services is made up front. Punishment for low effort was periodic, arriving through a sequence of signals interpreted as product recalls.

Few papers have estimate a structural model of reputation building and – aside from the explicit solutions that appear in a few places – this is the paper's main contribution. Product recalls – when made by publicly traded firms – are typically accompanied by stock price reductions. Both the recall data and the stock-price data were used to estimate the model which fits quite well. The estimates imply that equilibrium delivers a very small fraction of maximal welfare.

References

[1] BARBER, B. & M. DARROUGH. "Product Reliability and Firm Value: The Experience of American and Japanese Automakers, 1973-1992." *Journal of Political Economy* 104, No. 5 (Oct., 1996), pp. 1084-1099.

[2] BASS, F. M. (1969). A new product growth for model consumer durables. Management science, 15(5): 215-227.

[3] Jess Benhabib, Feng Dong, Pengfei Wang. "Adverse Selection and Self-ful lling Business Cycles." May 2016

[4] HALL, R. "Struggling to Understand the Stock Market." *AEA Papers & Proc.* May 2001.

[5] HECKMAN, J. & B. SINGER. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica*, Vol. 52, No. 2 (Mar., 1984), pp. 271-320

[6] KELLER, G., & RADY, S. (2015). Breakdowns. *Theoretical Economics*, 10(1), 175-202.

[7] KIEFER, J., AND J. WOLFOWITZ. "Consistency of the Maximum Likelihood Estimator in the Presence of Infinitely Many Incidental Parameters." *Annals of Mathematical Statistics* 27(1956), 887-906.

[8] MCGRATTAN, E. R. & E. C. PRESCOTT. "Is the Stock Market Overvalued?" *Federal Reserve Bank of Minneapolis Quarterly Review*, Fall 2000, 24(4), pp. 20–40

[9] HOFFER, G. E., S. W. PRUITT, & R. J. REILLY. "The Impact of Product Recalls on the Wealth of Sellers: A Reexamination." *Journal of Political Economy* 96, No. 3 (Jun., 1988), pp. 663-670

[10] JARRELL, G. AND S. PELTZMAN. "The Impact of Product Recalls on the Wealth of Sellers." *Journal of Political Economy* 93 (June 1985): 512-36.

[11] ROB, R. & A. FISHMAN. "Is Bigger Better? Customer Base Expansion through Word-of-Mouth Reputation." *Journal of Political Economy* 113, No. 5 (2005): 1146-1162.

[12] Rupp, N. "The Attributes of a Costly Recall: Evidence from the Automotive Industry." (2004). The attributes of a costly recall: Evidence from the automotive industry. *Review of Industrial Organization* 25(1), 21-44.

[13] Sultan, F., Farley, J. U., & Lehmann, D. R. (1990). "A meta-analysis of applications of diffusion models." *Journal of marketing research*, 70-77.

Appendix

1. Proof of Proposition 1

PROOF: We solve (9) using separation of variables. We can write (9) as

$$\frac{1}{A(x-x_1)(x-x_2)}dx = dt,$$

where $A = -\frac{1}{2}$. This is equivalent to

$$\frac{1}{A(x_2-x_1)}\left(\frac{1}{x-x_2} - \frac{1}{x-x_1}\right) = dt.$$

Integrating both sides, we get

$$\frac{1}{A(x_2-x_1)}(\ln|x-x_2| - \ln|x-x_1|) = t + C$$

Noting that $x_1 < x < x_2$, we have

$$\frac{1}{A(x_2-x_1)}\ln\frac{x_2-x}{x-x_1} = t + C$$

Therefore, the general solution is

$$x_t = \frac{x_2 + x_1\exp\{A(x_2-x_1)(t+C)\}}{1+\exp\{A(x_2-x_1)(t+C)\}} \tag{38}$$

Using the initial condition $x(0) = 0$, we get

$$x_t = \frac{x_2(1-\exp\{-\frac{1}{2}(x_2-x_1)t\})}{1-\frac{x_2}{x_1}\exp\{-\frac{1}{2}(x_2-x_1)t\}} \tag{39}$$

Notice that x_t is strictly increasing in t, because the numerator is strictly increasing and the denominator is strictly decreasing (remember $x_1 < 0$). Also, x_t converges from below to x_2 as $t \to \infty$. Therefore, if $\lambda > x_2$, we have the solutions for (x, v) given by

$$x_t = \frac{x_2(1-\exp\{-\frac{1}{2}(x_2-x_1)t\})}{1-\frac{x_2}{x_1}\exp\{-\frac{1}{2}(x_2-x_1)t\}}$$
$$v_t = v_0 + x_t$$

with $x_t \to x_2$ and $v_t \to v_0 + x_2$.

With the solution to x_t, we can compute F explicitly as in (2). First,

$$\begin{aligned} x_t &= \frac{x_2(1-\exp\{-\frac{1}{2}(x_2-x_1)t\})}{1-\frac{x_2}{x_1}\exp\{-\frac{1}{2}(x_2-x_1)t\}} \\ &= x_2 + \frac{\frac{x_2}{x_1}(x_2-x_1)\exp\{-\frac{1}{2}(x_2-x_1)t\}}{1-\frac{x_2}{x_1}\exp\{-\frac{1}{2}(x_2-x_1)t\}}. \end{aligned}$$

Therefore,

$$\begin{aligned} \int_0^t (\lambda - x_s)ds &= (\lambda - x_2)t - \int_0^t \frac{\frac{x_2}{x_1}(x_2-x_1)\exp\{-\frac{1}{2}(x_2-x_1)s\}}{1-\frac{x_2}{x_1}\exp\{-\frac{1}{2}(x_2-x_1)s\}}ds \\ &= (\lambda - x_2)t - 2\ln\frac{1-\frac{x_2}{x_1}\exp\{-\frac{1}{2}(x_2-x_1)t\}}{1-\frac{x_2}{x_1}} \end{aligned}$$

Substituting into (2) and simplifying yields (17).

2. Derivation of f and g

Its derivative then yields the density

$$\begin{aligned} f(t) &= -\frac{\frac{x_2}{x_1}(x_2-x_1)\exp(-(\lambda-x_2)t)}{\left(1-\frac{x_2}{x_1}\right)^2}\left[\left(1-\frac{x_2}{x_1}\exp\{-\frac{1}{2}(x_2-x_1)t\}\right)\exp\{-\frac{1}{2}(x_2-x_1)t\}\right] \\ &\quad +(\lambda-x_2)\left(\frac{1-\frac{x_2}{x_1}\exp\{-\frac{1}{2}(x_2-x_1)t\}}{1-\frac{x_2}{x_1}}\right)^2\exp(-(\lambda-x_2)t) \end{aligned} \tag{40}$$

Now we derive the distribution of x using (17). Note that $x_t = x_1 + \frac{x_2-x_1}{1-\frac{x_2}{x_1}\exp\left\{-\frac{1}{2}(x_2-x_1)t\right\}}$. If $x_t = x$, the corresponding t satisfies $1-\frac{x_2}{x_1}\exp\left\{-\frac{1}{2}(x_2-x_1)t\right\} = \frac{x_2-x_1}{x-x_1}$ and $\exp\left\{-\frac{1}{2}(x_2-x_1)t\right\} = \frac{x_1}{x_2}\frac{x-x_2}{x-x_1}$. Therefore,

$$\begin{aligned} \Pr(x_t \le x) &= \Pr\left(t \le \phi^{-1}(x)\right) = F\left(x^{-1}(x)\right) \\ &= 1 - \left(\frac{\frac{x_2-x_1}{x-x_1}}{1-\frac{x_2}{x_1}}\right)^2\left(\frac{x_1}{x_2}\frac{x-x_2}{x-x_1}\right)^{\frac{2(\lambda-x_2)}{x_2-x_1}} \\ &= 1 - \left(\frac{x_1}{x-x_1}\right)^{\frac{2(\lambda-x_1)}{(x_2-x_1)}}\left(\frac{x-x_2}{x_2}\right)^{\frac{2(\lambda-x_2)}{(x_2-x_1)}} \end{aligned}$$

and

$$
\begin{aligned}
g(x) &= -\left[\begin{array}{c} -\frac{2(\lambda-x_1)}{(x_2-x_1)}\left(\frac{x_1}{x-x_1}\right)^{\frac{2(\lambda-x_1)}{(x_2-x_1)}-1}\left(\frac{x-x_2}{x_2}\right)^{\frac{2(\lambda-x_2)}{(x_2-x_1)}} x_1\left[x-x_1\right]^{-2}+ \\ +\frac{2(\lambda-x_2)}{(x_2-x_1)}\left(\frac{x_1}{x-x_1}\right)^{\frac{2(\lambda-x_1)}{(x_2-x_1)}}\left(\frac{x-x_2}{x_2}\right)^{\frac{2(\lambda-x_2)}{(x_2-x_1)}-1} \frac{1}{x_2} \end{array}\right] \\
&= \frac{2(\lambda-x_1)}{(x_2-x_1)}\left(\frac{x_1}{x-x_1}\right)^{\frac{2(\lambda-x_1)}{(x_2-x_1)}-1}\left(\frac{x-x_2}{x_2}\right)^{\frac{2(\lambda-x_2)}{(x_2-x_1)}} x_1\left[x-x_1\right]^{-2} \qquad (41) \\
&\quad -\frac{2(\lambda-x_2)}{(x_2-x_1)}\left(\frac{x_1}{x-x_1}\right)^{\frac{2(\lambda-x_1)}{(x_2-x_1)}}\left(\frac{x-x_2}{x_2}\right)^{\frac{2(\lambda-x_2)}{(x_2-x_1)}-1} \frac{1}{x_2}
\end{aligned}
$$

From (8) we then have $v_t = v_0 + x_t$.

3. Derivation of (14)

Eq. (13) implies

$$\lambda > r + \lambda - p + \sqrt{(r + \lambda - p)^2 + 2rv_0}$$

implies, i.e., $\sqrt{(r + \lambda - p)^2 + 2rv_0} < p - r$.

Since $(r + \lambda - p)^2 + 2rv_0 < (p - r)^2$, expanding the first squared term leads to the inequality

$$\lambda^2 + (r - p)^2 - 2\lambda (p - r) + 2rv_0 < (p - r)^2$$

i.e., $\lambda^2 - 2\lambda (p - r) + 2rv_0 < 0$, i.e.,

$$v_0 < \frac{p - r}{r}\lambda - \frac{\lambda^2}{2r},$$

and this is equivalent to eq. (14).

4. mixed equilibrium: Discussion of the case $\lambda < x_2$ and derivation of F and f.

The case where $\lambda < x_2$ is discussed in the Appendix – no equilibrium appears to exists for this case even when $\theta > 0$

With respect to the issue that $\lambda < x_2$, note first that continuity requires that $v^g = v_{t_\lambda} = \frac{1}{1-\theta}\left(\lambda + v_0 - \theta v^{\mathrm{g}}\right)$. This implies that

$$v_0 = v_g - \lambda$$

Therefore, an equilibrium where the effort reaches λ in finite time exists if at $v_0 = v_g - \lambda$, we have $\lambda < x_2$, i.e.

$$(1-\theta)\lambda < \lambda + r - (1-\theta)p + \sqrt{(\lambda + r - (1-\theta)p)^2 + 2(1-\theta)\left[(\theta\lambda + r)v_0 - \theta(\lambda + r)v^g\right]}$$

i.e. (assuming $p \geq \frac{\theta\lambda + r}{1-\theta}$)

$$(1-\theta)\lambda - [\lambda + r - (1-\theta)p] < \sqrt{(\lambda + r - (1-\theta)p)^2 + 2(1-\theta)\left[(\theta\lambda + r)v_0 - \theta(\lambda + r)v^g\right]}$$

i.e. (using $v_0 = v^g - \lambda$)

$$(1-\theta)^2\lambda^2 - 2(1-\theta)\lambda[\lambda + r - (1-\theta)p] < 2(1-\theta)\left[(1-\theta)rv^g - (\theta\lambda + r)\lambda\right]$$

i.e.

$$(1-\theta)\lambda^2 < 2\lambda[\lambda + r - (1-\theta)p] + 2\left[(1-\theta)rv^g - (\theta\lambda + r)\lambda\right]$$

i.e.

$$\frac{(1-\theta)\lambda^2}{2} < \lambda^2 + \lambda r - \lambda(1-\theta)p + (1-\theta)rv^g - \theta\lambda^2 - r\lambda$$

i.e.

$$\frac{(1-\theta)\lambda^2}{2} < (1-\theta)rv^g - \lambda(1-\theta)p + (1-\theta)\lambda^2$$

i.e. (using $v^g = \frac{1}{r}\left(p\lambda - \frac{\lambda^2}{2}\right)$)

$$\frac{\lambda^2}{2} < p\lambda - \frac{\lambda^2}{2} - \lambda p + \lambda^2$$

i.e.

$$0 < 0$$

which is impossible. Therefore, when $p \geq \frac{\theta\lambda + r}{1-\theta}$, there is no differentiable solution s.t. x reaches λ in finite time and then stays there forever.

On the other hand, when $p < \frac{\theta\lambda + r}{1-\theta}$, substituting $v_0 = v^g - \lambda$, we have $x_2 = \frac{1}{1-\theta} X_2$, where

$$\begin{aligned}
X_2 &= \lambda + r - (1-\theta)p + \sqrt{(\lambda + r - (1-\theta)p)^2 + 2(1-\theta)\left[(\theta\lambda + r)v_0 - \theta(\lambda + r)v^g\right]} \\
&= \lambda + r - (1-\theta)p + \sqrt{(\lambda + r - (1-\theta)p)^2 + 2(1-\theta)\left[(\theta\lambda + r)(v^g - \lambda) - \theta(\lambda + r)v^g\right]} \\
&= \lambda + r - (1-\theta)p + \sqrt{(\lambda + r - (1-\theta)p)^2 + 2(1-\theta)\left[(1-\theta)rv^g - (\theta\lambda + r)\lambda\right]} \\
&= \lambda + r - (1-\theta)p + \\
&\quad + \sqrt{\lambda^2 + r^2 - (1-\theta)^2 p^2 + 2\lambda r - 2(1-\theta)pr - 2(1-\theta)p\lambda + 2(1-\theta)\left[(1-\theta)\left(p\lambda - \frac{\lambda^2}{2}\right)\right.} \\
&= \lambda + r - (1-\theta)p + \sqrt{\theta^2\lambda^2 + r^2 - (1-\theta)^2 p^2 + 2\theta\lambda r - 2(1-\theta)pr - 2(1-\theta)\theta p\lambda} \\
&= \lambda + r - (1-\theta)p + \left|\theta\lambda + r - (1-\theta)p\right|,
\end{aligned}$$

so that

$$x_2 > \lambda$$

where the inequality follows from $p < \frac{\theta\lambda + r}{1-\theta}$. As a result, in such an equilibrium effort would increase until it reached λ, and then stays there forever. In such an equilibrium, for $t \in [0, t_\lambda]$,

$$\begin{aligned}
x_t &= \frac{x_2 + \frac{x_2 - x_0}{x_0 - x_1} x_1 \exp\{-\frac{1-\theta}{2}(x_2 - x_1)t\})}{1 + \frac{x_2 - x_0}{x_0 - x_1} \exp\{-\frac{1-\theta}{2}(x_2 - x_1)t\}} \\
v_t &= \frac{1}{1-\theta}\left(x_t + v_0 - \theta v^{\mathrm{g}}\right)
\end{aligned}$$

where $x_0 = \theta\left(v^{\mathrm{g}} - v\right) = \theta\lambda$ and $v_0 = \frac{1}{r}(p\lambda - \frac{\lambda^2}{2}) - \lambda$. The effort reaches λ at $t_\lambda = \frac{2}{(1-\theta)(x_2 - x_1)} \ln \frac{x_2 - x_0}{x_2 - \lambda} \frac{\lambda - x_1}{x_0 - x_1}$.

Let us show that such an equilibrium cannot be achieved when $\theta = 0$. In this case, when $p < r$, we have

$$\begin{aligned}
x_t &= \frac{x_2(1 - \exp\{-\frac{1}{2}(x_2 - x_1)t\})}{1 - \frac{x_2}{x_1} \exp\{-\frac{1}{2}(x_2 - x_1)t\}} \\
v_t &= x_t + v_0
\end{aligned}$$

with $x_0 = 0$ and $v_0 = \frac{1}{r}(p\lambda - \frac{\lambda^2}{2}) - \lambda$. But the latter requires that $v_0 < 0$.. When $\theta > 0$, however, there is an additional incentive to exert effort, namely in order to move to the trigger equilibrium, so such an equilibrium may exist.

Derivation of f and F in the mixed case.—$F(t)$ and $f(t)$ are needed for the MLE of the mixed equilibrium with $\theta \in (0,1)$.

From equation (32), we have that

$$F(t \mid \theta) = \int_0^\infty F(\min(t,\tau))\,\theta x_\tau \exp\left(-\int_0^\tau \theta x_s ds\right) d\tau$$

where x_t is defined in equation (31)

$$\begin{aligned} x_t &= \frac{x_2 + \frac{x_2-x_0}{x_0-x_1} x_1 \exp\{-\frac{1-\theta}{2}(x_2-x_1)t\})}{1+\frac{x_2-x_0}{x_0-x_1}\exp\{-\frac{1-\theta}{2}(x_2-x_1)t\}}. \\ &= x_2 + \frac{\frac{x_2-x_0}{x_0-x_1}(x_1-x_2)\exp\{-\frac{1-\theta}{2}(x_2-x_1)t\}}{1+\frac{x_2-x_0}{x_0-x_1}\exp\{-\frac{1-\theta}{2}(x_2-x_1)t\}} \end{aligned}$$

It's easy to show that when $\theta \to 0$, we get the same expression for $F(t)$ in the rachet equilibrium. Its derivative defines the density function explicitly

$$\begin{aligned} f(t) &= \frac{\frac{x_2-x_0}{x_0-x_1}(x_1-x_2)\exp(-(\lambda-x_2)t)}{\left(1+\frac{x_2-x_0}{x_0-x_1}\right)}\left[\frac{\left(1+\frac{x_2-x_0}{x_0-x_1}\exp\{-\frac{1-\theta}{2}(x_2-x_1)t\}\right)}{\left(1+\frac{x_2-x_0}{x_0-x_1}\right)}\right]^{\frac{2}{1-\theta}-1} \\ &\quad +(\lambda-x_2)\left(\frac{1+\frac{x_2-x_0}{x_0-x_1}\exp\{-\frac{1-\theta}{2}(x_2-x_1)t\}}{1+\frac{x_2-x_0}{x_0-x_1}}\right)^{\frac{2}{1-\theta}}\exp(-(\lambda-x_2)t) \end{aligned}$$

When $\theta \in (0,1)$, there is no closed-form solutions to $F(t)$ as x_t depends on t. The next section shows a closed-form solution is available when $\theta = 1$.

6.0.5 4A: Derivation of $F(t \mid \theta = 1)$

When $\theta = 1$, x is fixed and

$$\begin{aligned}
F(t \mid \theta = 1) &= \int_0^t \left(1 - e^{-(\lambda - x)\tau}\right) x e^{-x\tau} d\tau + \left(1 - e^{-(\lambda - x)t}\right) e^{-xt} \\
&= 1 - e^{-xt} - x \int_0^t e^{-x\tau - (\lambda - x)\tau} d\tau + \left(1 - e^{-(\lambda - x)t}\right) e^{-xt} \\
&= 1 - e^{-xt} - \frac{x}{x + \lambda - x} \left(1 - e^{-[x + (\lambda - x)]t}\right) + \left(1 - e^{-(\lambda - x)t}\right) e^{-xt} \\
&= 1 - \frac{x}{x + \lambda - x} \left(1 - e^{-[x + (\lambda - x)]t}\right) - e^{-[x + (\lambda - x)]t} \\
&= 1 - \frac{x}{x + \lambda - x} - e^{-[x + (\lambda - x)]t} \left(1 - \frac{x}{x + \lambda - x}\right) \\
&= \left(1 - \frac{x}{x + \lambda - x}\right) \left(1 - e^{-[x + (\lambda - x)]t}\right) \\
&= \frac{\lambda - x}{x + \lambda - x} \left(1 - e^{-[x + (\lambda - x)]t}\right) \\
&= \frac{\lambda - x}{\lambda} \left(1 - e^{-\lambda t}\right) \to \frac{\lambda - x}{\lambda}
\end{aligned}$$

as claimed in eq. (28).

5. Product recall data and calculations

The data are taken from the National Highway Traffic Safety Administration of the Department of Transportation. The data contain all NHTSA safety-related defect and compliance from late 1960s, including detailed report receive date, record creation date, model of the car, and date of manufacture. We construct the quarterly recall data as follows. We

1. Removed the observations with missing recall report date, and/or start of manufacture date, and/or end of manufacture date. After such removal, we end up with 48014 total cases,

2. Sorted the cases by the report date, and create a quarterly bin from 1966Q4 to 2012Q3,

3. Calculated the number of total recalls in each bin,

4. Further removed the bins with consecutive zero observations and end up with sample spanning 1978Q1 to 2007Q3,

5. Took logs, de-trended the observations in each remaining bin, and compared with the log and de-trended real PCE.

6. Relation to the Bass model

The Bass model features an ODE for the population fraction of informed that is also quadratic and that for that reason has the same general solution as does this model.

https://en.wikipedia.org/wiki/Bass_diffusion_model

Let $\xi \equiv$ the fraction of informed people in the population, and

$$\frac{d\xi}{dt} = b\left(1-\xi\right) + q\xi\left(1-\xi\right)$$

where

$q =$ the copying parameter – those who don't know learn from those who do

$b =$ information arrival from the outside

The ODE is also a Ricatti equation, i.e., quadratic. The quadratic term also comes in negatively.

Denote $n\left(t\right)$ to be the cumulative proportion of adopters in the population. Parameter b captures the **innovation** and parameter q captures the **imitation**, the Bass diffusion model is

$$\frac{dn}{dt} = \left(b + qn\right)\left(1-n\right) = b + \left(q-b\right)n - qn^2 \tag{42}$$

The present model has the ODE

$$\frac{dx}{dt} = rv_0 + \left(r+\lambda-p\right)x - \frac{1}{2}x^2, \tag{43}$$

To make (43) comparable to (42), divide both sides by x_2. Let $z = x/x_2$. Then the admissible values are $z \in [0,1]$, and its ODE is $\frac{1}{x_2}\frac{dz}{dt} = \frac{rv_0}{x_2} + \left(r+\lambda-b\right)\frac{x}{x_2} - \frac{1}{2x_2}x^2$, i.e.,

$$\frac{dz}{dt} = \frac{rv_0}{x_2} + \left(r+\lambda-p\right)z - \frac{x_2}{2}z^2. \tag{44}$$

We can match coefficients in (2) and (44) to make the two identical, but this is artificial, there is no reason why that should arise. Otherwise they both just have the quadratic form, but the coefficients are not related to one another in the same way. The Bass model has $\lim_{t\to\infty} n(t) = 1$ which the present model attains only if $\theta = 0$. It is instructive to compare the two models on a level playing field, i.e., under the $\theta = 0$ constraint.

Estimates of innovation and imitation.—In Sultan, Farley and Lehmann (1990), they estimated that on average over various products and technologies, copying was far more important than inventing. That is,

$$\begin{aligned} b &= 0.03 \\ q &= 0.38 \end{aligned}$$

so that once you had enough informed, learning would take off. This is what produces the **S shape** in the $n(t)$. The S shape appears if an only if $q > b$, and that is indeed the case empirically. To prove this, think back to how we proved the S shape in the Fishman Rob case. In the FR case, we found that an S shape emerged iff $p < r + \lambda$, because then initially dx/dt was increasing in t. We can use the same method here.

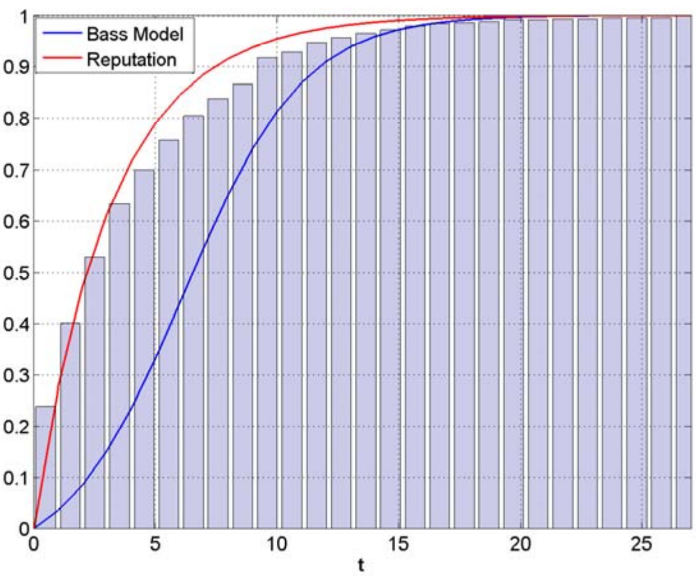

Figure 15: The bass model with $p = 0.03$ and $q = 0.38$ compared to te fit of the ML estimate constrained by $\theta = 0$)