Product Recalls and Firm Reputation:

Boyan Jovanovic[*]

October 11, 2017

Abstract

We model reputation capital as a reward for good behavior of sellers of a product the quality of which is not contractible. The market reacts unfavorably to product recalls which are the result of product defects. A recall triggers a reduction in the firm's value which then rises steadily until its next defect occurs. We fit the model to data on product recalls in the transportation-equipment sector, and on stock-price reductions following such recalls and find that reputation accounts for about 11.2 percent of firm value. Contract incompleteness leads to a welfare level of 49 percent of first best. A simple policy intervention attains first best, namely a recall tax and a production subsidy.

1 Introduction

This paper estimates the value of a reputation using product-recall data from the transportation-equipment sector and information about stock price movements in response to product-recall announcements. Reputation causes the values of firms to rise above the reproduction value of their physical assets.

Reputation, by our estimate, accounted for about 11.2 percent of firm value in the transportation-equipment sector over the 1978-2007 period. This

[*]NYU Economics. I thank M. Cerda, Z. Haque, S. Ma, H. Singh and D. Wei for help with the research, the NSF and C.V. Starr Center for financial assistance, and K. Borovickova, P. Chintagunta, M. Darrough, C. Flinn, E. Kamenica, R. Lagos, C. Phelan, R. Radner, D. Siegel, and E. Stacchetti for comments.

estimate derives from how much a firm's value falls when it recalls one of its products. Stock prices of publicly traded firms show this mechanism at work. Product recalls are common, but still rare enough that a recall represents significant news that produces a negative stock-price impact, often far larger than the direct costs associated with a recall, and this excess we interpret as reputational loss.

Welfare is estimated to be 49 percent of its first-best level. Although the estimates by Jarrell and Peltzman and others indicate that the reputational loss is up to twelve times higher than their direct recall cost, it is not large enough to correct for the inefficiency. The welfare analysis indicates that recalls should be taxed by an even larger multiple of their direct cost: The recall tax that is accompanied by a production subsidy, and it attains first best. The policy is in the broad class of policies under which one taxes or subsidizes marginal cost and then corrects the distributional implications by a lump-sum tax or subsidy.

The model is that of a market for a homogenous, non-durable good with a continuum of buyers and sellers. Firms share the same production function but are hit by occasional shocks to product quality or "defects." Defects *are* contractible and the firm must fully compensate its customers for them. Otherwise the quality of a seller's output is not observed before purchase and is not contractible; payment is up front. Interactions between buyers and sellers are short-term – there are no repeat meetings or long-term contacts.

Costly effort reduces the probability that a defect will occur but if it does occur, the firm's revenue drops and then recovers gradually until the episode repeats itself. A defect signals that effort was low, and the signal is public. A firm's "reputation" is the history of its public signals, i.e., of its past defects or recalls.

We study a steady state in which the product price is constant. Free entry of firms means that all the welfare gains go to the consumers. The model has multiple steady states. Consumer welfare and average reputation are positively correlated across these states, and welfare is negatively related to the price of the product. In other words, equilibria that feature higher reward to good performance yield higher social welfare – not surprisingly.

Within the class of equilibria in which only the history of recalls matters there is a continuum of equilibria indexed by the size of the punishment induced by the recall. We let the data pick one equilibrium from this class, and that equilibrium fits the data on product recalls and stock-price reductions fairly well – the maximum likelihood estimate is compared to data and the two are close in several dimensions.

One dimension is the endogenous recall hazard. An implication of the model is that a declining hazard of product recalls reflects the build-up of reputation. A bias may arise if firms have different hazard parameters and if the low-hazard firms survive longer without experiencing a recall. Surprisingly, however, allowing for heterogeneity raises the estimated reputation building because the individual hazards are steeper than the combined hazard. The reason seems to be that when the heterogeneity is in the products' hazard-proneness itself, firms make greater efforts to prevent it, causing the reduced-form hazards to cross. This is explained in the last proposition of the paper.

A second dimension is the stock-price drop following a recall; that too is something that the model fits. Value drops exceed direct recall costs, which suggests that a recall damages the firm's reputation more broadly, perhaps affecting consumer expectations of the quality of the modified or upgraded versions of the recalled product itself, or even of the quality of its other products.

A third dimension is the direct recall cost. Moreover, product recalls have been found to reduce the value of the firm over and above their direct effects on the firms' profits or dividends. It seems a clear indication of a loss of reputation, of a decline in a firm's goodwill as Jarrell & Peltzman (1985) point out, and it seems that the value of the goodwill lost is as much as twelve times the direct costs surrounding the recall – this too is something that the model fits. Other papers that estimate the effect of recalls on firm values are Hoffer, Pruitt and Reilly (1988), Barber and Darrough (1996) and Rupp (2004). This paper adds to the literature by structurally estimating a model of reputation building, and the policy conclusion is that recalls should be taxed because an even larger punishment is warranted than any of the equilibria provide.

Fig. 1 plots the steady state distribution of estimated values of reputation among sellers relative to their capital stocks. In other words it is the excess of market value over book value. The mean is 11.2 percent of book value. This estimate is for the 1978-2007 period and for the transportation-equipment sector. Now, Hall (2001) reports a value of 90 percent for that sector in 1998, but that was a year when market valuations in all sectors were unusually high. Broad measures of intangibles should include the value of a customer base as stressed by Gourio and Rudanko (2014) and advertising or information as modeled by Butters (1977) and Milgrom and Roberts (1984), both of which

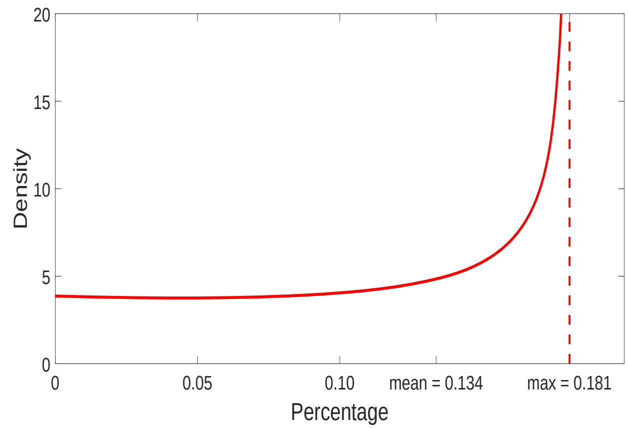

Figure 1: Market value of reputation relative to book value of capital

Hall's estimate includes but which my model leaves out. The model also leaves out convex capital adjustment costs.[1]

Azoulay, Bonatti and Krieger (2017) show that citations data can be explained by a model where learning about types occurs via periodic bad news signals similarly to what I assume here. Pakes (1985) has studied how a firm's stock price reacts to changes in its patents stock which are good news for the firm.

The model resembles Hopenhayn (1992) with a continuum of firms in facing periodic shocks, with free entry, and with the industry being in a long run steady state. Other models have focused on an agent facing a stream of short run buyers. Rob and Fishman (2005) also focus on equilibria in which seller reputation consists of the time elapsed since the seller's last public bad-news signal. Holmstrom (1999) has not only a hidden action but exogenous types as do Board and Meyer-ter-Vehn (2013) where an agent has hidden investment in an evolving state and periodically generates a public signal and Cisternas (2016) where the outcome has similarities to some of the ones here. Watson (1999, 2002) studies repeated interactions in which the level of trust starts at a low level and gradually rises as in the equilibria I focus on. Related are models of governments that wish to borrow money or to refrain

[1] Hall argues that movements in the value of intangibles and not capital adjustment costs are responsible for the large swings in aggregate Tobin's Q values since 1945.

from confiscatory taxation, but have no ability to commit – Kydland and Prescott (1977) and Phelan and Stacchetti (2001).

Section 2 lays out the model and its implications for the value of reputation and for optimal policy, Section 3 describes the data, the estimation, and the implications of the estimates for the magnitude of reputation capital and for policy. Section 4 discusses some extensions of the model and Section 5 concludes. Some proofs and data details are in the Appendix.

2 Model

There is a continuum of buyers and sellers.

Buyers.—A buyer has a utility function $U(q, z)$ defined over consumption of quality units q of a "reputation good," and on the number of physical units z of an outside good. A buyer's income each period is m, he takes as given the price p per unit of the reputation good q, i.e., the price of quality. Both m and p are measured units of the numeraire good z. The buyer faces the period budget constraint $z + pq \leq m$. A single customer's demand for quality then is

$$D(p) = \arg \max_{q \in [0,(m-z)/p]} \{U(q, m - pq)\}. \tag{1}$$

There is a continuum of buyers of measure one so that market demand also is $D(p)$. We assume that $\lim_{\mathrm{q} \to 0} \partial U / \partial q = +\infty$. The buyers' parameters are (m, U).

Sellers.—Sellers are risk neutral with discount rate r. Each can sell up to one divisible physical unit per period[2] that is bought by many customers. Payment is in advance and between the continuum of customers and firms there are no repeat interactions. The quality of the output is equal to a seller's effort, x. Effort is unobserved, and the price per physical unit sold is px^*, where x^* is effort that buyers expect the seller to exert. The seller's cost of effort is $x^2/2$. All firms get the same price per unit of quality supplied, and the firm's "normal" payoff is

$$px^* - \frac{1}{2}x^2.$$

[2] Sec. 5.1 shows that the results extend to the case where firms differ in their physical scale as long as returns to scale are constant.

Defects.—A customer can buy from various sellers. Periodically, the seller's output has a defect. A defective product reduces a customer's utility in proportion to the number of units he bought from that seller. If defects were not compensated, a buyer i's utility would be

$$U\left(q, z - c\sum_{j=1}^{M_i} \kappa_{i,j}\right)$$

where $\kappa_{i,j}$ is the number of defective physical units customer i bought from seller j, where $j \in \{1, 2, ..., M_i\}$ is the set of sellers to customer i that had a recall, and where c is a parameter.

Compensation for defects.—A seller must, by law, compensate each customer by the full amount of the loss; if seller j sells quantities $(\kappa_{i,j})_{i=1}^{N_j}$ to N_j customers and if he has a recall, his total recall cost is

$$c\sum_{i}^{N_j} \kappa_{i,j} = c, \tag{2}$$

because $\sum_i^{N_j} \kappa_{i,j} = 1$, i.e., because his physical quantity is unity. This payment restores each of seller j's customers' utilities to their no-defect level of $U(q, z)$. There is, in other words, a full warranty that a seller has to honor in the event that the product turns out to be defective, and the customer is fully insured.[3]

Public histories.—Buyers cannot share their purchasing experience. The only way buyers learn about a seller's performance is through public signals. A seller can try to avoid a recall by exerting effort x. Conditional on an effort path $(x_\tau)_{\tau=0}^{\infty}$, the waiting time τ until the next defect has CDF

$$\Pr(\tau \leq t) \equiv F(t) = 1 - \exp\left(-\int_0^t (\lambda - x_\tau)_+ \, d\tau\right). \tag{3}$$

Thus, the defect hazard is $(\lambda - x)_+$, where $\lambda > 0$ is a parameter. There are no inherent differences among sellers, but their public histories will generally differ, and may influence buyers' expectations.

[3] The warranty is treated as exogenous. In the present model a warranty could be used as a commitment device – promising to pay a multiple of the damages would induce higher effort and draw a higher price. We do not pursue this idea here.

Public signals and reputation.—A defect is the only public signal about the seller's effort. Let t denote time elapsed since the last defect or, if the seller has had no defects yet, since the date of entry. Reputation matters if x_t depends on t. When its next defect occurs, the price of the seller's output will fall and its market value will drop to k.[4]

The HJB equation.—Suppose that t is the only variable determining quality and not, for instance, the number of past recalls. Let r be the rate of discount. Conditional on $(x_t^*)_0^\infty$, the Bellman equation for the lifetime value v is

$$rv_t = \max_{x \leq \lambda} \left(px_t^* - \frac{x^2}{2} - (\lambda - x)(v_t - k + c) + \frac{dv}{dt} \right). \tag{4}$$

The problem is concave in x and the seller's first-order condition is

$$x_t = v_t - k + c. \tag{5}$$

Free entry of sellers.—The supply of entrants is infinitely elastic at the value k. The initial condition in eq. (4) reads

$$v_0 = k, \tag{6}$$

and it embodies two assumptions: A free entry condition at the cost k, and the assumption that the seller can sell the business to a new entrant for the price of k. If an incumbent's value ever dropped below k, it would be taken over by an entrant, and therefore

$$v_t \geq k. \tag{7}$$

The seller cannot escape the cost of recall, however, regardless of whether or not it is taken over.[5] The inequality (7) limits the punishment that the

[4] If the firm can continue to operate under a new name without having to pay k again, its value cannot fall below k. Smaller punishments and more complicated history dependence are both possible equilibria. We shall assume, however, reputation works exclusively via the pair (t, k). The drop in value at recall is an on-path punishment, roughly as in Green and Porter (1984).

[5] A product recall often results a takeover of the firm in question. Some examples are listed in Appendix 5. Jovanovic and Rousseau (2002) introduce takeovers into a Hopenhayn (1992) type of model and treat the value of acquired capital as k net of a "salvage cost." The salvage cost represents the costs of transferring the capital to new owners, a process that often involves a private equity firm. See Appendix 5 and particular Table A2 for examples where a takeover occurred following a product recall.

market can inflict on a firm should it have a recall. In particular, it rules out a grim-trigger type of equilibrium. Sec. 5.3 revisits the issue of other equilibria.

Market clearing.—Let n be the number of sellers. Let $\bar{x}$ denote the average quality per seller (defined precisely in eq. (16) below) the market-clearing price p satisfies

$$D(p) = \bar{x}n. \tag{8}$$

When t does affect x, quality varies over sellers. The parameters are (r, m, U, c, k, λ). We consider a steady state in which the aggregates (q, p, n) are fixed.

Equilibrium.— Equilibrium consists of the triple (q, p, n) and the pair of sequences of real numbers $(x_t, v_t)_{t=0}^{\infty}$ such that for all $t \geq 0$, (i) $v_t \geq k$ satisfies (4), (ii) $x_t \geq 0$ satisfies (5), (iii) buyers' expectations are correct in that $x_t = x_t^*$, (iv) the market clears so that (8) holds.

We observe first that if c exceeds λ, equilibrium is unique and it has zero recalls:

Proposition 1. *If*

$$c \geq \lambda,$$

the only equilibrium is at $x_t = \lambda$.

Proof. The stage payoff is $px_t^* - \frac{x^2}{2} - c(\lambda - x_t)_+$, and the marginal payoff is $c - x > 0$ for all $x \in [0, \lambda)$. Payoff is maximized at $x = \lambda$. ■

From now on we shall assume that $c < \lambda$.

Solution for x_t.—We now use (5) to eliminate v_t from (4), and noting that $dx/dt = dv/dt$, we get

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

Since $v_0 = k$, the initial condition is $x_0 = c$. Let x_1 and x_2 be the two roots of x at which the RHS of (9) is zero:

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

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

We shall assume that the entry cost exceeds the recall cost,

$$k - c > 0, \tag{12}$$

so that the roots are real. Then Appendix 1 proves that x_t is a monotone increasing function of t. In particular, we have the following claim:

Proposition 2. *If $x_2 < \lambda$, the ODE (9) has the solution*

$$x_t = x_1 + \frac{x_2 - x_1}{1 + \frac{x_2 - c}{c - x_1} \exp\{-\frac{1}{2}(x_2 - x_1)t\}} \tag{13}$$

for all $t \geq 0$, with $x_1 < x_2$ as shown in (10) and (11).

The following corollary guarantees that the inter-arrival hazard of recalls declines monotonically as a function of time elapsed since the last recall:

Corollary 1. *Equilibrium p must be such that when x_2 is given in (11),*

$$x_2 \geq c.$$

and x_t is monotonically increasing.

Proof. Since x_t is monotone, (13) and $x_2 < c$ would imply that $dx/dt < 0$ for all t. But $v_0 = k$ and (5) then implies that $v_t < k$ which contradicts (7). ■

Since $x_2 > c > 0 > x_1$, x_t increases with t, from $x_0 = c$ to $\lim_{t\to\infty} x_t = x_2$. However at $p = p_{\max}$ (given in eq. (18)), x_t is constant at $x_t = x_2 = c$.

Typical equilibrium play is illustrated in Fig. 2. As x varies on the horizontal x axis, dx/dt in (9) is inverted-U-shaped and it peaks at $x = r + \lambda - p$. Since $k > c$, we have $x_1 < 0 < x_2$, so that the curve crosses the positive half line exactly once. A larger view of (9) is again shown in Panel 1 of Fig. 9 in Sec. 3.4.

Whenever $x_2 > x_1$, the function x_t is strictly increasing. Let $t = \phi(x)$ denote the inverse of x_t. Inverting the function in eq. (13) yields

$$\phi(x) = -\frac{2}{x_2 - x_1} \ln\left(\frac{c - x_1}{x - x_1}\frac{x_2 - x}{x_2 - c}\right) \geq 0, \tag{14}$$

because $x_2 \geq c$.

The stationary "age" distribution μ.—To calculate long run industry supply we shall need the long-run distribution of x which in turn depends on t.

Figure 2: THE RHS OF (9) AS A FUNCTION OF x.

Let $\mu(t)$ be the pdf of sellers for whom the time elapsed since the last recall is t. Then[6]

$$\mu(t) = \frac{1 - F(t)}{\int_0^\infty [1 - F(s)]\, ds}. \tag{15}$$

The estimated μ is plotted in panel 3 of Fig 9 in Sec. 3.4 and it is not very different from f. Then, the average quality in the market, $\bar{x}$, is:

$$\bar{x} = \int x_t \mu(t)\, dt. \tag{16}$$

Appendix 2 shows that the numerator of (15) evaluated at $\phi(x)$ is:

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

using (13) for the second step.

2.1 Equilibrium and Welfare

Propositions 1 and 2 showed that (a) if $\lambda \leq c$, the only equilibrium is one with no reputation building and (b) if $\lambda > c$, there is a continuum of equilibria indexed by $p \in [p_{\min}, p_{\max}]$; we shall solve for these prices presently.

[6] For proof see "Age Processes." and Theorem 18 in http://www.math.uah.edu/stat/renewal/LimitTheorems.html.
Analytically solvable distributions include the uniform $[0,1]$: $F(t) = t \Rightarrow \mu(t) = 2(1-t)$, and the exponental: $F(t) = e^{-\alpha t} \Rightarrow \mu(t) = \alpha e^{-\alpha t}$.

Since sellers get zero rents, all rents go to the buyers. Welfare therefore declines with p – the equilibria are Pareto ranked. Even in the most efficient equilibrium, $p_{\min}$, welfare is below its maximum level.

The no-reputation equilibrium at $p_{\max}$

The firm must compensate its customers for recall disutilities and this induces a minimal level of effort at which (x, v) are constant. It entails $x_t = c$ and $v_t = k$ for all t. Eqs. (6), (4), and the fact that $dv/dt = 0$ imply that $p_{\max}$ solves $rk = p_{\max}c - \frac{c^2}{2} - (\lambda - c)\, c = p_{\max}c + \frac{c^2}{2} - \lambda c$, so that the firm's discounted profits equal its entry cost plus discounted expected recall costs. This gives us the no-reputation equilibrium price

$$p_{\max} = \frac{rk}{c} + \lambda - \frac{c}{2}, \tag{18}$$

which is increasing in λ. And $p_{\max}$ is decreasing in c because c raises x, and therefore the firm's sales, px^*, rise by more than the production and recall costs do.

The no-reputation equilibrium entails the smallest welfare and the highest p. This equilibrium exists if $p_{\max} \geq 0$. If the RHS of (18) is negative, there is no positive price at which the firm's rents can cover its entry cost and the market shuts down if there are no reputations.

The highest-welfare equilibrium at $p_{\min}$

The maximal effect that reputation concerns have on effort is $x_2 = \lim_{t\to\infty} x_t$. By eq. (5), x_2 is also the maximal effect of a recall on the value of the firm. Fig. 2 shows that x_2 is the maximal value of x_t that is reached only as $t \to \infty$ if no defect occurs.

Lemma 1. *When $x_2 > \lambda$ equilibrium does not exist.*

Proof. Assume, on the contrary, that an equilibrium entails $x_2 > \lambda$. Using (11), $x_2 > \lambda$ is equivalent to $\sqrt{(r + \lambda - p)^2 + 2r(k - c)} > (p - r)$. Squaring both sides, this is equivalent to $(r-p)^2 + 2\lambda(r-p) + \lambda^2 + 2r(k-c) > (p-r)^2$. Canceling, this leaves

$$\begin{aligned} 2r(k-c) > 2\lambda(p-r) - \lambda^2 &\Leftrightarrow (k-c) > \frac{\lambda p}{r} - \lambda - \frac{\lambda^2}{2r} \\ &\Leftrightarrow (k - c + \lambda) > \frac{1}{r}\left(p\lambda - \frac{\lambda^2}{2}\right). \end{aligned} \tag{19}$$

Now, if $x_2 > \lambda$ and since $dx/dt > 0$, there is a date t_λ such that $x_t = \lambda$ for all $t \geq t_\lambda$. Evaluating (5) at $t = t_\lambda$ and $x_{t_\lambda} = \lambda$,

$$v_{t_\lambda} = k - c + \lambda. \tag{20}$$

But for $t \geq t_\lambda$, we have a zero probability of a defect, and the flow profit at $x = \lambda$ is $p\lambda - \lambda^2/2$. That implies

$$v_{t_\lambda} = \frac{1}{r}\left(p\lambda - \frac{\lambda^2}{2}\right). \tag{21}$$

Then (20) and (21) imply $(k - c + \lambda) = \frac{1}{r}\left(p\lambda - \frac{\lambda^2}{2}\right)$, which contradicts (19).
■

Lemma 1 implies that $x_t < \lambda$. In words, if $x_2 > \lambda$, the defect hazard would become zero in finite time at which point v can no longer grow. This would cause a discontinuity in dv/dt which causes the non-existence. The restriction that $x_2 < \lambda$ implies (see Appendix 3) that

$$0 < k < \frac{p\lambda - \lambda^2/2}{r} - \lambda + c. \tag{22}$$

Next, x_2 is strictly decreasing in p. In fact, using (12) we have that:

$$\frac{dx_2}{dp} = -1 - \frac{r + \lambda - p}{\sqrt{(r + \lambda - p)^2 + 2r(k - c)}} < 0.$$

Therefore, the lowest admissible p is one at which $x_2 = \lambda$. So, using (20) and (21),

$$p_{\min} = r\left(1 + \frac{k - c}{\lambda}\right) + \frac{\lambda}{2}. \tag{23}$$

The first-best, contractible-x equilibrium at $\hat{p}$

The contractible-x equilibrium is unique and it coincides with the first-best outcome. This equilibrium consists of the triple (n, x, p) that solves (24), (26), and (27). Price of quality equals its marginal utility:

$$p = U_q(xn, m - pnx). \tag{24}$$

The social opportunity cost of an additional seller is k. Equating k to the discounted net benefit of the seller's output yields

$$k = \max_x \left(\frac{1}{r} \left(px - \frac{x^2}{2} \right) - C(x) \right), \tag{25}$$

where C is the present value of compensation costs:

$$C(x) = \int_0^\infty e^{-rt} (c + w(x)) (\lambda - x) e^{-(\lambda - x)t} dt = (c + w) \frac{\lambda - x}{r + \lambda - x} = \frac{\lambda - x}{r} c.$$

Substituting for C into (25) implies

$$rk = \max_x \left(px - \frac{x^2}{2} - (\lambda - x) c \right). \tag{26}$$

Denote the first-best values by hats. From (26),

$$\hat{x} = \hat{p} + c. \tag{27}$$

Substituting into (26), $rk = \frac{1}{2}(\hat{p} + c)^2 - \lambda c$, which implies

$$\hat{p} = \sqrt{2(rk + \lambda c)} - c, \tag{28}$$

which is decreasing in c because recalls are the only direct discipline on effort – the higher is c the higher is the effort that sellers provide as shown by (5) and (28). Since all rents go to buyers, first best implies that

$$p \geq \hat{p}, \tag{29}$$

and total quality consumed is at its highest.

2.2 Optimal policy

A simple tax-subsidy scheme attains first best. It consists of

a) A tax T per recall
b) A subsidy S paid to each firm each period
c) Full compensation by firms to consumers for defects .

The tax raises firms' effort to its first-best level $\hat{x}$, but at the first-best price $\hat{p}$ firms would then be making a loss, because in addition to compensating customers for recall they would be facing the additional expense T, adding up in expectation to $(\lambda - x) T$. We have the following characterization:

Proposition 3. *The tax-subsidy scheme that attains first best is*

$$\begin{aligned} T &= \hat{p}, & (30) \\ S &= (\lambda - c - \hat{p})\, T > 0. & (31) \end{aligned}$$

Proof. Note first that since $c < \lambda$, $\hat{p} > c$ and so $T > 0$. The HJB eq. now is

$$rk = S + \max_{x \leq \lambda} \left(px_t^* - \frac{x^2}{2} - (\lambda - x)(c + T) \right) \tag{32}$$

and the FOC is

$$x_t = c + T \tag{33}$$

Eqs. (27) and (30) imply that $x = \hat{x}$. Substituting for x and for S into (32), the latter reads

$$rk = p\hat{x} - \frac{\hat{x}^2}{2} - (\lambda - \hat{x})c,$$

which is identical identical to (26) evaluated at $x = \hat{x}$. Finally, $x = \hat{p} + c$, and therefore $S = (\lambda - x)\, T$. But from Lemma 1 existence of equilibrium requires that $\lambda > x$, and so $S > 0$. ■

2.3 The private value of a reputation

The private value of the firm's reputation is $v_t - k$. Using (5) and the fact that $x_t \to x_2$,

$$w_t \equiv \frac{v_t - k}{k} = \frac{x_t - c}{k} \leq \frac{x_2 - c}{k} \equiv w_{\max}, \tag{34}$$

with x_2 given in (11). If the firm is publicly traded and if k is the firm's book value[7], the firm's market to book value is v/k.

The CDF of w is $\int_0^{\phi(c+wk)} \mu(t)\, dt$, where μ and ϕ are defined in (15) and in (14) so that the pdf of w is

$$\zeta(w) = k\phi'(c + wk)\, \mu(\phi(c + wk)) = \frac{2k}{(c + wk - x_1)(x_2 - c - wk)} \mu(\phi(c + wk)) \tag{35}$$

[7]Relatedly, Hopenhayn (1992b) assumes that the cost of entry is an irreversible initial investment.

because $\phi'(x) = \frac{2}{(x_2 - x)(x - x_1)}$, as shown in Appendix 2. By L'Hôpital's rule in (35) and using (15), we find in Appendix 2 that

$$\lim_{w \to w_{max}} \zeta(w) = \begin{cases} 0 & 2\lambda + x_1 > 3x_2 \\ \infty & 2\lambda + x_1 < 3x_2 \end{cases} . \tag{36}$$

At the estimated parameter values, $2\lambda + x_1 < 3x_2$ so that $\lim_{w \to w_{max}} \zeta(w) = \infty$.

2.4 Two distributions needed for the estimation procedure

Percentage in value .—Define the percentage loss at recall as z:

$$z = \frac{v - (k - c)}{v} = \frac{x}{x + k - c} \tag{37}$$

which is an increasing function of x. The larger is z, the larger is the loss in value. Since $x = (k - c)\frac{z}{1-z}$, the CDF of z is

$$\Psi(z) = B\left(\phi\left[(k - c)\frac{z}{1 - z}\right]\right) \qquad \text{for } z \in \left[0, \frac{x_2}{x_2 + k - c}\right] \subset [0, 1) \tag{38}$$

in light of (12).

Distribution of t conditional on recall occurring at t.—Call this density $b(t)$, and CDF $B(t)$. The likelihood of a recall at t is f, and the log-run frequency of a recall by a seller at t is $\mu(t)$ in eq. (15) which, being the unconditional age distribution, acts as the prior on t. Using Bayes rule,

$$b(t) = \frac{f(t)\mu(t)}{\int_0^\infty f(s)\mu(s)\,ds}, . \tag{39}$$

where f is derived analytically in Appendix 2 eq. (51). Then the distribution of x conditional on recall is

$$G(x) = B(\phi(x))$$

Denote the density by $g(x)$; it is plotted in panel 4 of Fig. 9 in Sec. 3.4. This is the distribution we shall use when we constrain the average drop in value at recall, $E(x) \equiv \int x dG(x)$.

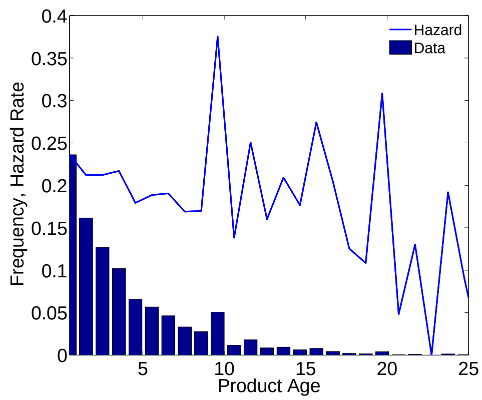

Figure 3: The data: the frequency distribution and its hazard.

3 Estimation

Equilibrium implies that reputation is related to the slope of the recall hazard; a recall is judged to be less likely if it has not occurred for a long time. In comparing the large recall-related stock-price drop of Toyota in 2008 to smaller drops in GM's stock following its recalls, Brauer (2014) writes:

"While no recall is good for an automaker, Toyota's history of uninterrupted success in the U.S. market had established very high expectations for the brand. By comparison, GM doesn't have the same infallible reputation Toyota possessed in 2008, meaning a recall (even a massive one) doesn't impact GM to the same degree."

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. The average age at recall is 4.14 years. Details are in Appendix 4. The resulting data are portrayed in two figures: Fig.3 shows the frequency distribution of the ages of the products at recall, and the annualized hazard rate $h(t) = \frac{f(t)}{1-F(t-1)}$.

Recall costs.—Recall costs are included in warranty expenses; they are not a separate line item. The 10Ks and 10Qs (the official financial reports filed with the SEC) sometimes include information on large recalls, but recall costs are not mandated disclosure items. We therefore rely on the estimates of several studies of the stock-price impact of recalls. The first to estimate this were Jarrell and Peltzman who report (1985, p. 521) that for publicly traded firms value loss is twelve times recall costs. Because others Hoffer, Pruitt and Reilly (1988), Barber and Darrough (1996) and Rupp (2004) have claimed that the multiple is smaller, and we shall constrain the estimates by a loss-to-costs ratio of six; we also report some of the results with the alternative ratio equaling twelve.

Stock-price impact of recalls.—We shall rely on the results reported in the papers listed above. Some recent examples are portrayed graphically in Appendix 7. Only a small fraction of the 1636 firms involved were publicly traded, but under the constant returns to scale assumption detailed in Sec. 5.1 this does not bias the results. Assuming that c is financed by debt or from future profits, we may interpret $v_t + c - k$ as the stock-price reduction at the time of the recall, and we now calculate its distribution in the population of all recalls. From (5), the loss is equal to x_t, and so we shall need the distribution of x_t conditional on recall, i.e., G.

We proceed in two steps. We first solve for the distribution of age at recall, $b(t)$. This is generally different from F unless x is constant so that F is exponential. We then invert the solution for x to figure out the distribution $G(x)$ of the size of the drop in value.

Percentage declines in value at recall.—A product recall is still a relatively rare event and it represents bad news for the producer, and the company's stock price falls. Jarrell and Peltzman (1985) were the first to estimate the losses borne by owners of a firm that recalls a defective product from the market. Table 1 of their paper shows that 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, Hoffer, Pruitt and Reilly (1988) find them about half as large in absolute value. Rupp

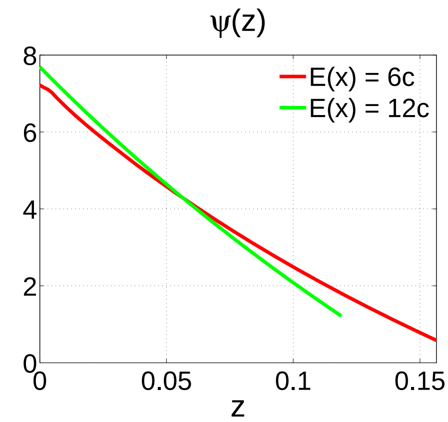

Figure 4: The density of percentage loss of value

(2004) finds it to be -0.077, Barber and Darrough (1996) estimate the returns declines to be between -0.001 and -0.013. In the model the drop is by the absolute amount $v - k + c$, and relative to pre-recall value is the value z in (37). We shall check ex post that the mean percentage loss value defined in (37) is roughly within the range of the estimates summarized in this paragraph, i.e.,

$$E(z) = \int_0^1 z\psi(z)\,dz \in [0.004, 0.077]\,. \tag{40}$$

3.1 Estimates under homogeneity

The first round of estimates presumes that all sellers share the same parameters (r, λ, p, k, c). The estimates are constrained by the requirement that the roots (10) and (11) be real, i.e., that $(r + \lambda - p)^2 + 2r(k - c) > 0$.

Using eq. (38) we plot the density $\psi(z) \equiv \Psi'(z)$ in Fig. 4. From (38) and the estimates, the support of ψ is $[0, 0.16]$ if $E(x)/c = 6$ and $[0, 0.12]$ if $E(x)/c = 12$.

Interpretation.—We interpret $B()$ as the age distribution of the product at time of recall. We interpret $G()$ as the distribution of x at time of recall. To constrain the ML estimates we shall use

$$E(x) = \int xg(x)\,dx = 6c \text{ or } 12c. \tag{41}$$

Estimation.—The time unit is a year and we set $r = 0.05$. Using the densities f and g shown in (51) and (54) leads to the constrained maximum likelihood estimation problem

$$\max_{(p,k,\lambda)} \prod_i f(t_i) \quad \text{s.t.} \quad (41). \tag{42}$$

This likelihood does not depend on n directly. Rather, it depends only on c which we will estimate as a parameter. Similarly, the estimates do not depend on the demand curve $D(\cdot)$; only p matters, and it too will be treated as a parameter. Therefore we cannot infer n from these estimates because we have no direct measure of c.

We pre-set r at 0.05. The parameter estimates are reported in Table 1 which also reports x_1 and $x_2, E(z)$ and $E(x)/c$ even though they are fully determined by (λ, p, k, c).[8]

Table 1: CONSTRAINED ML ESTIMATES

λ	p	k	c	x_1	x_2	$E(z)$	$E(x)/c$	Ln ML
0.272	0.527	0.696	0.002	−0.540	0.129	0.016	6	−2.396
0.384	0.750	0.802	0.001	−0.740	0.108	0.015	12	−2.230

There are two sets of estimates. Line 1 reports estimates constrained by a value-drop-recall ratio of 6, so as to reflect the estimate averaged over the various papers cited above. Line two of the table reports the estimates constrained by value-drop-to-recall-cost ratio 12 as reported by Jarell and Peltzman (1985 p. 521). The estimates are not too different. Visually, the plots in Fig. 5 favor the estimates in Line 1 which will be used in subsequent calculations, but the log likelihood is slightly higher for the second set.

[8] To speed up the computation of $E(x)$ we use the following approximation for $G(x)$. The distribution of x at *first* recall in a cohort of sellers of the same vintage, 0, is solvable in closed form, and is derived analytically in Appendix 2 eq. (53):

$$G_0(x) = F(\phi(x)) \tag{43}$$

$$= 1 - \left(\frac{c - x_1}{x - x_1}\right)^{\frac{2(\lambda - x_1)}{(x_2 - x_1)}} \left(\frac{x_2 - x}{x_2 - c}\right)^{\frac{2(\lambda - x_2)}{(x_2 - x_1)}} \tag{44}$$

At $p_{\max}$ in (18) the two distributions are identical because they approach each another as the x_t profile becomes flat and $\mu(t)$ in (15) converges to the exponential distribution – see note 6.

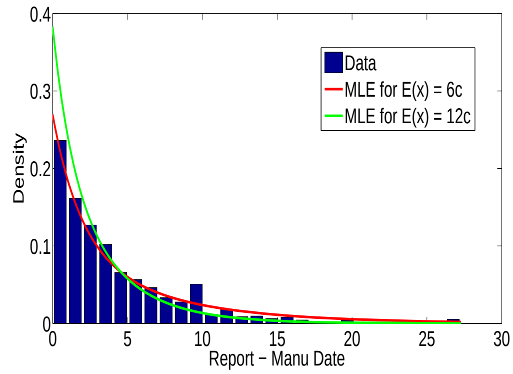

Figure 5: Model fit

As one reality check on our estimate of c, note that that warranty expenses are an upper bound on recall costs. Cohen *et al.* (2011) report that warranty expenses are from 1.45 and 1.82 percent of sales revenue in the industries they studied. Our estimates say that relative to sales recall costs, c/p, are between 0.13 (using row 2) and 0.38 (row 1) percent of sales revenue, i.e., at most about one fifth of all warranty expenses. A second check is the magnitude of $E(z)$ which is well within the range given in (40).

Fig. 5 shows the model fit to the frequency distribution of the ages of the products at recall under the two respective constraints on the average price decline. The fit is fairly good, and slightly better under the first constraint, namely $E(x) = 6c$, i.e., with the parameter values in reported in the first line of Table 1.

3.2 Reducing potential heterogeneity bias: Two markets

Fig 7 shows that the fit is good, but the hazard is overpredicted and is too flat. Yet the model leads us to estimate the value of reputations from the decline in the recall hazard and if this decline is underestimated the effect on reputation capital is biased down.

This is opposite to the bias that we would have if we had targeted the hazard because if there is heterogeneity in hazard rates, survivorship bias induces a downward slope in the estimated hazards. This is the classic "moverstayer" problem in the analysis of durations data, and it may generate an upward bias in our estimates of the value of reputation.

As a first pass at this issue, and in order to see if we can flatten the estimated hazard, we consider two groups and estimate p, k, c, λ separately across groups, treating them as separate markets. We also estimate the fraction π_1 of the observations that fall in group 1, with the remaining fraction π_2 falling in group 2. Figure 7 plots the two hazards and compares them to the single market estimated hazard.

Table 2: ML ESTIMATES; $E(x|\lambda_i) = 6c$ for both i

Parameter estimates and the implied (x_1, x_2)

i	λ	p	k	c	x_1	x_2	$E(z\|\lambda_i)$
1	0.243	0.311	0.166	0.0438	−0.280	0.044	0.0196
2	0.380	0.348	0.211	0.0244	−0.158	0.119	0.011

The group 1 density overestimates the tail and underestimates the early ages, whereas the group 2 density does the opposite, but otherwise the two estimated densities do not differ much. The hazard rates differ more starkly with group 2 having the steeped hazard. Within each group, the estimates of $(c, E(z))$ pass the reality check described in the context of Table 1.

Crossing of the hazard rates.—Crossing as shown in Figs 19 will generally occur for different values of λ. We have

Claim 1. *Let $c = 0$ and fix $p \in (p_{min}, p_{max})$. As λ varies in the region where*

$$\lambda > p - r, \tag{45}$$

the hazards must cross at least once. In particular, it holds at p_{min} in eq. (23) if $r(k-c) < \frac{\lambda^2}{2}$.

Proof. We show that h_0 rises with λ whereas h_∞ moves in the opposite direction. Since $x_0 = 0$, $h_0(\lambda) = \lambda$ is increasing in λ. On the other hand,

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

Differentiating,

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

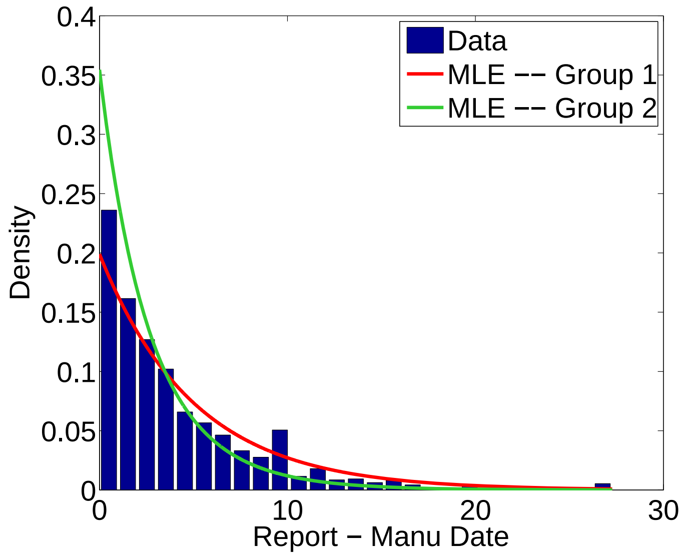

Figure 6: Two-market density estimates

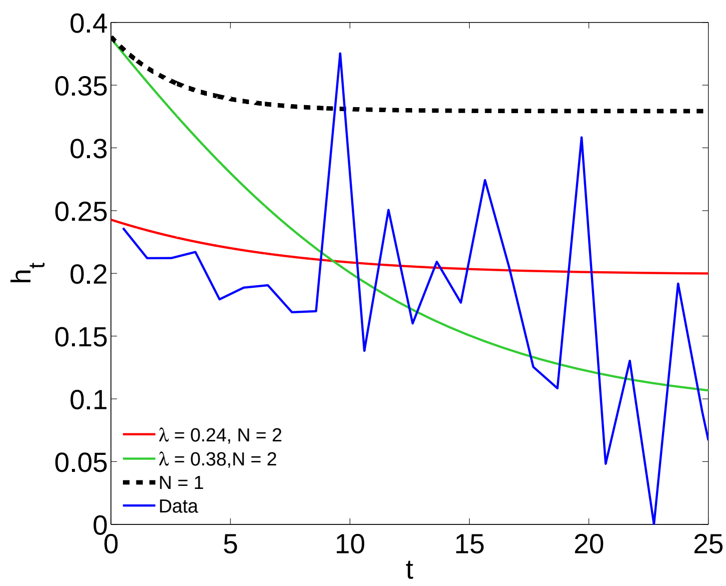

Figure 7: Comparison of the one-market hazard and the 2 two-market hazards

because of (45). ■

Holding p constant, then, as λ rises the hazard turns clockwise. At the estimates of Table 1 (45) does not hold and at the estimates of Table 2 it holds for the estimates for market 2.

Crossing hazards are known to often arise in analyses of survivorship following distinct medical procedures (Liu, Qiu and Sheng 2007). If two procedures have hazards that cross at t^*, and if such data are aggregated into a single waiting time distribution,the slope of the hazard will be biased down for $t \in [0, t^*)$, biased up on a neighborhood just to the right of t^* and so it appears that the overall impact would be to convexify the estimated hazard.

3.3 Implications of the estimates

We now refer back to the model and the issues it raises in light of these estimates. We shall focus on the estimates in Table 1 except for the welfare results where we also include the two-market results in Table 2.

3.3.1 Welfare

We use the estimates of Tables 1 and 2. Thus in Table 3 for the two market case, $p_{\min}$ and $p_{\max}$ are calculated in each market using its estimated parameters (λ, k, c) reported in Table 2

Table 3: WELFARE

	One Market		Two Markets	
	source of info	$1/p$	source of info	$\frac{\pi_1}{p_1} + \frac{\pi_2}{p_2}$
First best (contractible)	eq. (28)	3.79	eq. (28)	$\frac{.53}{.15} + \frac{.47}{.17} = 6.30$
highest-welfare equilib. $p_{\min}$	eq. (23)	3.19	eq. (23)	$\frac{.53}{.20} + \frac{.47}{.26} = 4.46$
estimated equilib.	Table 1	**1.90**	Table 2	$\frac{.53}{.31} + \frac{.47}{.35} = \mathbf{3.09}$
worst equilib. $p_{\max}$	eq. (18)	0.06	eq. (18)	$\frac{.53}{.41} + \frac{.47}{.80} = 1.88$

Therefore two markets have higher welfare than the one-market estimates, but the estimated first-best level is also higher. In both cases welfare is estimated to be 49 percent of its first-best level.

3.3.2 The optimal policy intervention

Evaluating (30) and (31) at the parameter estimates in Table 1, we derive the estimates of the optimal policy in Table 4:

Table 4: OPTIMAL POLICY

Assumed average price drop	$E(x)/c$	6	12
Recall tax	$T = \hat{p}$	0.264	0.824
Period subsidy	$S = (\lambda - c - \hat{p})\, T$	0.005	0.010
Tax relative to firm value	$T/E(v)$	0.340	0.329
Subsidy relative to firm value	$S/E(v)$	0.007	0.115
equilibrium recall rate (est.)	$\lambda - \bar{x}$	0.1897	0.3229
optimal recall rate	$\lambda - \hat{x}$	0.0081	0.1005
recall rate difference (equilib. minus opt.)		0.1815	0.2224
recall rate % reduction		0.9571	0.6889

where $E(v) = k - c + \int x_t \mu(t)\, dt$ and μ is given in (15).

A couple of things to note here: First, T is about one third of the average market value, and an order of magnitude higher than $E(z)$, the equilibrium market punishment of 1.5 or 1.6 percent. It is high because recalls are relatively infrequent and the policy has to leverage them to induce higher effort in all periods. Second, compared to equilibrium, recalls are much smaller under the optimal policy, again by an order of magnitude.

3.3.3 Private value of a reputation

Reputation is only one component of the type of intangible leading to the difference between a firm's market value and the book value of its capital. Fig. 8 shows the value of reputation in four equilibria. Fig 8 and Table 5 summarize the findings. The steady state distribution of $w \equiv \frac{v-k}{k}$ defined in eq. (34).it plotted in Fig. 8

The larger is x_2, the greater the tendency for firms to be bunched near the maximum where the recall hazard is the lower.

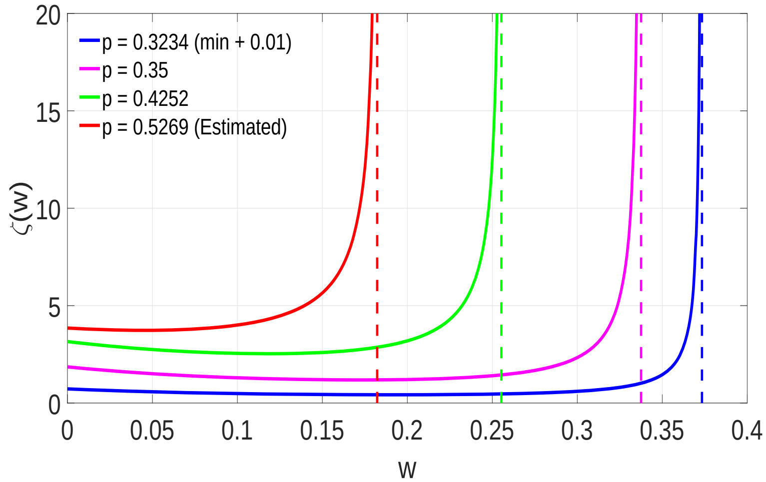

Figure 8: THE DENSITY $\zeta(w)$ OF THE VALUE OF REPUTATION, w, IN FOUR EQUILIBRIA.

Table 5: AVERAGE VALUE OF REPUTATION, $E(w)$

equilibrium p	$E(w)$
0.3234 (p_{min} + .01)	0.254
0.35	0.252
0.41	0.163
0.53 (p_{est})	**0.115**

3.3.4 Consistency of the estimates with the assumptions made on the parameter values

Several inequality restrictions were imposed and now we check that they are satisfied at the parameter estimates in Table 1.

Regarding Fig. 2.— Since $p > r + \lambda$, in the positive orthant of the figure dx/dt is positive and decreasing in x as shown in Fig. 2. A larger version of the equation system is in Fig. 9 of Sec. 3.4 below.

On the existence of the equilibrium at $p_{\max}$.—The RHS of (18) is estimated to be positive which means that the $p_{\max}$ equilibrium exists.

Checking that equilibrium p *exceeds first best* p.—Since $p > \hat{p}$ and since (28) implies that $\lambda > \sqrt{2rk} + \frac{c}{2}$, we should have $p > \sqrt{2rk}$, and this is indeed

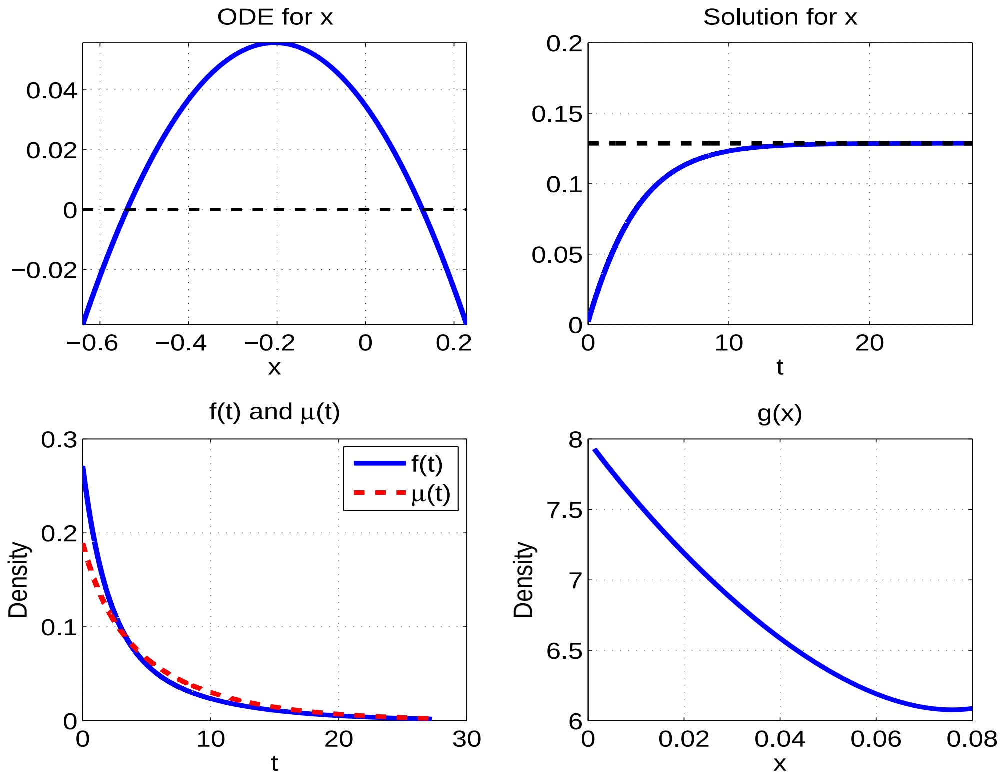

Figure 9: Plot of eqs (9) in Panel 1, (13) in Panel 2, (15) and (51) in Panel 3, and (54) in Panel 4.

so at the estimates.

Checking condition (12).—and the condition (12) is met so that Proposition 1 and the solution for x in eq. (5) is valid.

3.4 Simulations

The solutions are depicted in the 4-quadrant Fig. 9, evaluated at the parameter estimates in the top row of Table 1. Panel 1 is an expanded version of Fig. 2, where we see the RHS of (9) crossing the zero axis at $x_1 = -0.54$ and at $x_2 = 0.13$. These are the values given in (10) and (11).

Panel 2 shows that x approaches x_2 and as x rises, v also rises. Additionally, v rises because the hazard rate $\lambda - x$ declines, reducing the likelihood of a reversion to $v_0 = k$. Panel 3 of Fig. 9 shows that f is quite close to μ, and that as $f(t) \gtrless \mu(t)$ as $t \gtrless 2.30$.[9] Panel 4 shows the distribution of x conditional on a recall which, by eq. (5), is equal to c plus the loss in value.

The simulation assumes the utility function

$$U(q,z) = q^{\alpha} z^{\beta} \text{ and } \Rightarrow D(p) = \frac{A}{p} \quad \text{where } A = \frac{\alpha}{\alpha+\beta} m \tag{46}$$

where m is income per head. We simulate the equilibrium using (15), (13) and (8) and the solution for x_t in (13). A simulation based on parameters in the top row of Table 1 is in Fig 10. Once p is specified, equations (15), (13) and (8) and our x_t do not depend on the demand parameters. Because the entry cost is constant at k, n is proportional to A and so we set $A = 1$. The other parameters used in Fig 10 are listed in Table 1.

Panel 1 shows that the planner wants fewer firms and more effort per firm. First-order dominance of the waiting times distributions is apparent in the fourth quadrant.

4 Discussion and extensions

This section discusses three theoretical modifications.

4.1 Variable capacity

The data cover 1600 firms that vary in size, i.e., in the capacity to produce physical output. For the model to apply we need a constant returns assumption. From (2) we already have recall costs so that costs of entry and of effort x are proportional to capacity. Let A denote an (exogenous) capacity so that total efficiency units of quality produced are $q = Ax$. And in eq. (2) we have

[9] This is a manifestation of the "inspection paradox" in renewal theory. The paradox is that if we wait some predetermined time t and then observe how large the renewal interval containing t is, we should expect it to be typically larger than a renewal interval of average size. https://en.wikipedia.org/wiki/Renewal_theory

To illustrate the point for the uniform case, $F(1) = 1$, $a(t) = 2(1-t)$ and it crosses $f(t) = 1$ from above at $t = 1/2$.

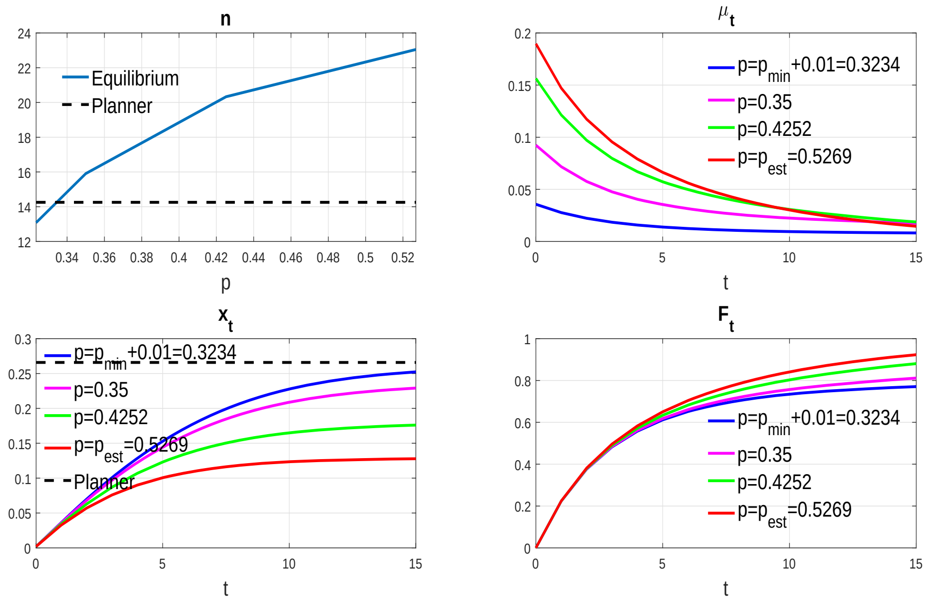

Figure 10: $A = 1$. Param. values from the top row of Table 1

$\sum_i^{N_j} \kappa_{i,j} = A$ with the total recall cost being cA. The Bellman equation then reads

$$V = \max_x \left(pAx^* - A\frac{x^2}{2} - (\lambda - x)(V - Ak + Ac) + \frac{dV}{dt} \right)$$

and the FOC is

$$x = \frac{V}{A} - k + c \tag{47}$$

Then the solution for V is $V_t = Av_t$ so that (47) is the same as (5); A drops out of the solution and the results go through. In particular, the estimates reported in the tables and in Fig. 1 still apply.

4.2 Durable goods

We now interpret q as durable-good services. Assume as before that the manufacturer must replace a failed product at the cost c in units of z, but that recalls affect only the latest generation of products.

Consumers.—Then (1) still applies with the following modification: Investment by consumers in durables is I, measured in quality units and services of durables consumed in quality units are

$$q_t = \int_0^\infty e^{-\delta s} I_{t-s} ds.$$

Since bad products are replaced, recalls do not appear in the above equation. The budget constraint is $m = pI + z$. In steady state $q = \delta^{-1} I$ is a constant and assuming all income m is spent each period, the consumer's FOC is

$$p = \int_t^\infty e^{-(r+\delta)(s-t)} U_q \left(\delta^{-1} I, m - pI\right) ds = \frac{1}{r+\delta} U_q \left(\delta^{-1} I, m - pI\right). \quad (48)$$

Producers.—Conditional on p, (4) is unchanged and the main propositions and implications again go through subject to a reinterpretation of p as given in (48).

4.3 Other equilibria

If we disallow takeovers so that the firm and its assets cannot re-enter under a new name without paying k, and so that (7).fails, then other equilibria exist. In particular, if the punishment following a recall can force its value below k; value can drop to any $v_{\mathrm{R}} \in (0, k)$, and then behavior would be as described in the previous sections, but with a lower post-recall punishment value $v_{\mathrm{R}} < k$. The Bellman equation would read

$$rv = \max_{x \leq \lambda} \left(px^* - \frac{x^2}{2} - (\lambda - x)(v - v_{\mathrm{R}} + c) \right),$$

with a post-entry initial condition $v_0 = k$. These equilibria elicit more effort, the FOC being $x = v - v_{\mathrm{R}} + c$. They are indexed by the additional parameter v_{R}, and, like the equilibrium discussed in the previous sections, they feature rising effort in between recalls.

Further equilibria entail punishments that depend on recall order. Let v_{R}^j denote the value following the jth recall. Starting with $v_0 = k$, any decreasing sequence $\left(v_{\mathrm{R}}^j\right)_{j \geq 1}$is an equilibrium. And when the second punishment entails $v_{\mathrm{R}}^2 = 0$, this would resemble one of the the outcomes in Board and Meyer-ter-Vehn (2013 p. 2412) in which the second piece of bad news finishes the firm off.

5 Conclusion

The paper has structurally estimated model of reputation building in a market in which firm reputations consist of the public histories of their product recalls. On-path punishments were periodic, arriving through a sequence of defects interpreted as product recalls. Product recalls – when made by publicly traded firms – are typically accompanied by stock price reductions. The recall data and information from the stock-prices were used to estimate the model. The model fits the recall hazards pretty well especially when one accounts for heterogeneity. Reputation appears to accounts for about 11.2 percent of firm value in the transportation-equipment sector, and the estimates were then used to draw policy conclusions.

Welfare was estimated at 49 percent of first best, but an easily implementable policy can attain first best. First best is attained by a recall tax that is substantially larger than the direct recall cost; to maintain the right incentives for firms to enter, the tax has to be accompanied by a subsidy that can be paid every period. That would bring the recall rate down by an order of magnitude and raise quality of goods produced to its first best level.

References

[1] Azoulay, P., A. Bonatti, and J.L. Krieger. "The career effects of scandal: Evidence from scientific retractions." *Research Policy* (2017).

[2] Asquith, Paul. "Merger bids, uncertainty, and stockholder returns." *Journal of Financial Economics* 11, no. 1-4 (1983): 51-83.

[3] 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.

[4] Brauer, Karl. "Why A Massive Safety Recall Hurt Toyota More Than GM." *Forbes* 2014, Jul. 1.[10]

[5] Board, Simon, and Moritz Meyer-ter-Vehn. "Reputation for quality." *Econometrica* 81, no. 6 (2013): 2381-2462.

[10] https://www.forbes.com/sites/kbrauer/2014/07/01/why-massive-safety-recall-hurt-toyota-more-than-gm/#67948153328c.

[6] Butters, G. R. (1977). "Equilibrium distributions of sales and advertising prices." *Review of Economic Studies*, 465-491.

[7] Cohen, D., Darrough, M. N., Huang, R., & Zach, T. (2011). "Warranty reserve: Contingent liability, information signal, or earnings management tool?" *Accounting Review* 86(2), 569-604.

[8] Gourio, F., & Rudanko, L. (2014). "Customer capital." *Review of Economic Studies*, 81(3), 1102-1136.

[9] Green, E. J., & Porter, R. H. (1984). "Noncooperative collusion under imperfect price information." *Econometrica* 87-100.

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

[11] 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

[12] Holmström, B. (1999). "Managerial incentive problems: A dynamic perspective." *Review of Economic studies*, 66(1), 169-182.

[13] Hopenhayn, H. (1992). "Entry, exit, and firm dynamics in long run equilibrium." *Econometrica* 1127-1150.

[14] Hopenhayn, H. "Exit, selection, and the value of firms." *J. Economic Dynamics and Control*, 1992.

[15] Jarrell, G. and S. Peltzman. "The Impact of Product Recalls on the Wealth of Sellers." *Journal of Political Economy* 93 (June 1985): 512-36.

[16] Jovanovic, B., & Rousseau, P. L. (2002). The Q-Theory of Mergers. *American Economic Review*, 92(2), 198-204.

[17] Keller, G., & Rady, S. (2015). "Breakdowns." *Theoretical Economics*, 10(1), 175-202.

[18] Liu, K., Qiu, P., & Sheng, J. (2007). "Comparing two crossing hazard rates by Cox proportional hazards modelling." *Statistics in medicine*, 26(2), 375-391.

[19] 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.

[20] Milgrom, P., & Roberts, J. (1986). "Price and advertising signals of product quality." *Journal of political economy*, 94(4), 796-821.

[21] Pakes, A. (1985). On patents, R & D, and the stock market rate of return. *Journal of political economy* 93(2), 390-409.

[22] Phelan, C., and E. Stacchetti. "Sequential equilibria in a Ramsey tax model." Econometrica 69.6 (2001): 1491-1518.

[23] 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.

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

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

[26] Watson, J. (1999), "Starting Small and Renegotiation," *Journal of Economic Theory* 85, 52–90.

[27] Watson, J. (2002), "Starting Small and Commitment," *Games and Economic Behavior* 38, 176–199.

Appendix

1. Proof of Proposition 1

Proof of Proposition 1: We solve (9) using separation of variables. We can write (9) as

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

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

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

Integrating both sides, we get

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

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

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

Therefore, the general solution is

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

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

$$c = \frac{x_2 + x_1\exp\{\rho(x_2-x_1)(C)\}}{1+\exp\{\rho(x_2-x_1)(C)\}}$$

or

$$\exp\{\rho(x_2-x_1)(C) = \frac{x_2-c}{c-x_1}$$

This implies that

$$x_t = \frac{x_2 + x_1\frac{x_2-c}{c-x_1}\exp\{\rho(x_2-x_1)t\}}{1+\frac{x_2-c}{c-x_1}\exp\{\rho(x_2-x_1)t\}} \tag{49}$$

or, alternatively

$$x_t = x_1 + \frac{x_2 - x_1}{1 + \frac{x_2 - c}{c - x_1}\exp\{\rho(x_2 - x_1)t\}}$$

or,

$$x_t = x_2 - \frac{(x_2 - x_1)\frac{x_2 - c}{c - x_1}\exp\{\rho(x_2 - x_1)t\}}{1 + \frac{x_2 - c}{c - x_1}\exp\{\rho(x_2 - x_1)t\}}$$

which implies (13) since $\rho = -1/2$. Notice that x_t is strictly increasing in t, because the denominator is strictly decreasing (using the first definition for x_t and remembering $x_1 < 0 < x_2$). Also, x_t converges from below to x_2 as $t \to \infty$. Finally $x_t \to x_2$ implies $v_t \to k - c + x_2$ (since the solution for v_t is $v_t = k - c + x_t$).

2. Derivation of $F(t)$, $f(t)$, $g(x)$, $\phi(x)$ and Eqs. (43), (35) and (36)

Note that by definition we have that $F(t) = 1 - exp\left(- \int_0^t (\lambda - x_\tau) d\tau \right)$. Then, since (13) can be written as

$$x_t = x_2 - \frac{(x_2 - x_1)\frac{x_2 - c}{c - x_1}\exp\{-\frac{1}{2}(x_2 - x_1)t\}}{1 + \frac{x_2 - c}{c - x_1}\exp\{-\frac{1}{2}(x_2 - x_1)t\}}, \tag{50}$$

we have

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

Therefore,

$$F(\phi(x)) = 1 - \left[\frac{1 + \frac{x_2 - c}{c - x_1}\exp\{-\frac{1}{2}(x_2 - x_1)t\}}{1 + \frac{x_2 - c}{c - x_1}} \right]^2 \exp(-(\lambda - x_2)t).$$

Using (13) we have

$$1 - F(\phi(x)) = \left(\frac{1 + \frac{x_2 - x}{x - x_1}}{1 + \frac{x_2 - c}{c - x_1}} \right)^2 \exp\left(-(\lambda - x_2)\phi(x)\right),$$

But

$$\frac{1+\frac{x_2-x}{x-x_1}}{1+\frac{x_2-c}{c-x_1}}=\frac{c-x_1}{x-x_1},$$

and this yields (17). Its derivative then yields the density $f(t)$ which is used in (39) and in the likelihood function (42).

$$f\left(t\right)=\left(\frac{c-x_1}{x_2-x_1}\right)^2\exp\left(-(\lambda-x_2)t\right)\left(1+\frac{x_2-c}{c-x_1}\exp\left(-\frac{1}{2}(x_2-x_1)t\right)\right) \times\left[(\lambda-x_2)+(\lambda-x_1)\frac{x_2-c}{c-x_1}\exp\left(-\frac{1}{2}(x_2-x_1)t\right)\right] \quad (51)$$

Showing that at $x_2=\lambda$, $\int_0^\infty\left[1-F\left(s\right)\right]ds=+\infty$

First, note that at $x_2=\lambda$, (17) implies that

$$1-F\left(t\right)=\left(\frac{c-x_1}{x_2-x_1}\right)^2\left(1+\frac{x_2-c}{c-x_1}\exp\left(-\frac{1}{2}(x_2-x_1)t\right)\right) \quad (52)$$

Therefore:

$$\begin{aligned}\int_0^\infty\left[1-F\left(s\right)\right]ds &= \left(\frac{c-x_1}{x_2-x_1}\right)^2\int_0^\infty\left(1+\frac{x_2-c}{c-x_1}\exp\left(-\frac{1}{2}(x_2-x_1)s\right)\right)ds\\ &= \left(\frac{c-x_1}{x_2-x_1}\right)^2\times\\ &\times\left[t-\frac{4}{x_2-x_1}\frac{x_2-c}{c-x_1}e^{\left(-\frac{1}{2}(x_2-x_1)s\right)}-\frac{1}{x_2-x_1}\left(\frac{x_2-c}{c-x_1}\right)^2e^{\left(-(x_2-x_1)s\right)}\right]\Bigg|_0^\infty\\ &= \left(\frac{c-x_1}{x_2-x_1}\right)^2\left[\infty+\frac{4}{x_2-x_1}\frac{x_2-c}{c-x_1}+\frac{1}{x_2-x_1}\left(\frac{x_2-c}{c-x_1}\right)^2\right]=\infty\end{aligned}$$

Therefore as $p\rightarrow p_{\min}$,

$$\mu\left(t\right)=\frac{1-F\left(t\right)}{\int_0^\infty\left[1-F\left(s\right)\right]ds}.\rightarrow 0$$

and thus $\xi\left(P\right)\rightarrow 0$

Derivation of (43).—Now we derive the distribution of x using (56). Note that

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

If $x_t = x$, the corresponding t satisfies $1 + \frac{x_2-c}{c-x_1}\exp\left\{-\frac{1}{2}(x_2-x_1)t\right\} = \frac{x_2-x_1}{x-x_1}$ and $\exp\left\{-\frac{1}{2}(x_2-x_1)t\right\} = \frac{c-x_1}{x_2-c}\frac{x_2-x}{x-x_1}$. Therefore, with the definition in (14),

$$\begin{aligned}
\Pr(x_t \le x) = \Pr(t \le \phi(x)) &= F\left(x^{-1}(x)\right) \\
&= 1 - \left(\frac{1 - \frac{x_2-c}{x_1-c}\frac{x_1-c}{x_2-c}\frac{x-x_2}{x-x_1}}{1-\frac{x_2-c}{x_1-c}}\right)^2 \exp\left(-(\lambda-x_2)t\right) \\
&= 1 - \left(\frac{c-x_1}{x-x_1}\right)^{\frac{2(\lambda-x_2)}{x_2-x_1}+2}\left(\frac{x_2-x}{x_2-c}\right)^{\frac{2(\lambda-x_2)}{(x_2-x_1)}} \\
&= 1 - \left(\frac{c-x_1}{x-x_1}\right)^{\frac{2(\lambda-x_1)}{(x_2-x_1)}}\left(\frac{x_2-x}{x_2-c}\right)^{\frac{2(\lambda-x_2)}{(x_2-x_1)}} \qquad (53)
\end{aligned}$$

and

$$g(x) = \frac{2}{x_2-x_1}\left(\frac{c-x_1}{x-x_1}\right)^{\frac{2(\lambda-x_1)}{x_2-x_1}}\left(\frac{x_2-x}{x_2-c}\right)^{\frac{2(\lambda-x_2)}{x_2-x_1}}\left(\frac{\lambda-x_1}{x-x_1}+\frac{\lambda-x_2}{x_2-x}\right). \qquad (54)$$

From (5) we then have $v_t = k - c + x_t$.

Derivation of (35).—Differentiating in eq. Eq. (14) we have

$$\begin{aligned}
\phi'(x) &= -\frac{2}{x_2-x_1}\left(\frac{c-x_1}{x-x_1}\frac{x_2-x}{x_2-c}\right)^{-1}\left(\frac{c-x_1}{x_2-c}\right)\left\{\frac{d}{dx}\left(\frac{x_2-x}{x-x_1}\right)\right\} \\
&= -\frac{2}{x_2-x_1}\left(\frac{c-x_1}{x-x_1}\frac{x_2-x}{x_2-c}\right)^{-1}\left(\frac{c-x_1}{x_2-c}\right)\left\{\frac{(x-x_1)(-1)-(x_2-x)}{(x-x_1)^2}\right\} \\
&= -\frac{2}{x_2-x_1}\left(\frac{c-x_1}{x-x_1}\frac{x_2-x}{x_2-c}\right)^{-1}\left(\frac{c-x_1}{x_2-c}\right)\left\{\frac{x_1-x_2}{(x-x_1)^2}\right\} \\
&= \frac{2}{1}\frac{x-x_1}{x_2-x}\left\{\frac{1}{(x-x_1)^2}\right\} \\
&= \frac{2}{(x_2-x)(x-x_1)},
\end{aligned}$$

and since $dx/dw = k$, this implies (35).

Finally we show that $\lim_{w\to w_{\max}}\zeta(w) = +\infty$. By L'Hôpital's rule in (35) and using (15), and letting $L := \frac{2k(c-x_1)^{\frac{2\lambda-2x_1}{x_2-x_1}}(x_2-c)^{\frac{2x_2-2\lambda}{x_2-x_1}}}{E(t)}$

$$\begin{aligned}
&\lim_{w\to w_{\max}} \zeta(w) \\
=& \lim_{x\to x_2} \frac{2k\mu(\phi(x))}{(x-x_1)(x_2-x)} \\
=& \lim_{x\to x_2} \frac{L(x-x_1)^{\frac{2x_1-2\lambda}{x_2-x_1}}(x_2-x)^{\frac{2\lambda-2x_2}{x_2-x_1}}}{(x-x_1)(x_2-x)} \\
=& \lim_{x\to x_2} \frac{L(\frac{2x_1-2\lambda}{x_2-x_1})(x-x_1)^{\frac{3x_1-2\lambda-x_2}{x_2-x_1}}(x_2-x)^{\frac{2\lambda-2x_2}{x_2-x_1}} - L(\frac{2\lambda-2x_2}{x_2-x_1})(x-x_1)^{\frac{2x_1-2\lambda}{x_2-x_1}}(x_2-x)^{\frac{2\lambda+x_1-3x_2}{x_2-x_1}}}{x_2-2x+x_1}
\end{aligned}$$

where the second line follows from substitution and the third line follows from L'Hôpital's rule. Hence,

$$\lim_{w\to w_{max}} \zeta(w) = \begin{cases} 0 & 2\lambda + x_1 > 3x_2 \\ \infty & 2\lambda + x_1 < 3x_2 \\ L(\frac{2\lambda-2x_2}{(x_2-x_1)^2})(x_2-x_1)^{\frac{2x_1-2\lambda}{x_2-x_1}} & 2\lambda + x_1 = 3x_2 \end{cases}$$

which proves (36)

3. Derivation of (22)

Derivation of (22).—Since $\lambda > x_2$, eq. (11) implies

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

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

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

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

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

$$k < \frac{p-r}{r}\lambda - \frac{\lambda^2}{2r} + c,$$

and this is equivalent to eq. (22).

4. data and calculations

The data recall are taken from the National Highway Traffic Safety Administration of the Department of Transportation. The website is

https://www-odi.nhtsa.dot.gov/downloads/

The data contain all NHTSA safety-related defects and compliance from late 1960s, and involve 1636 firms. For each recall they include the report-received date, record-creation date, a description of the recalled item such as model of the car, the name of the manufacture and date of manufacture. We construct the quarterly recall data as follows.

1. Removed the observations with missing recall report date, and/or start of manufacture date, and/or end of manufacture date, leaving a total of 48014 cases;

2. Sorted the cases by the report date, and created quarterly bins from 1966Q4 to 2012Q3;

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

4. Further removed bins with consecutive zero observations and ended up with a sample spanning 1978Q1 to 2007Q3;

5. Took logs of the observations in each remaining bin and de-trended the series.

These are the data portrayed in the Figures in Sec. 3.

5. Recalls and takeovers

In many cases the sale is to a private equity firm, and it is not known who will manage the company's assets in the future. VC and buyout funds' evidence is relevant to takeovers by private equity groups. Table A2 reports some examples from various sectors where recall was soon followed by an acquisition. Following the table are Figures (11)-(15) which show the companies' stock-price series around the time of the recall and subsequent takeover, with the exception of Bausch and Lomb which was privately owned.

Table A1: Major recalls and Subsequent Takeovers

Firm	Recall date	Acquisition date	Acquirer
Takata[a]	Nov '14	Jun '17	Key Safety System[g]
Patties Foods[b]	Feb '15	Jun '16	Pacific Equity Partners
Bausch & Lomb[c]	Dec'10/Nov'12	Aug '13	Valeant Pharmaceut.
Cadbury/Schweppes[d]	Jun '06	Feb '10	Kraft
Merck[e]	Sep '04	Nov '09	Schering-Plough[h]
Keurig Green Mntn[f]	Dec '14	Dec '15	JAB Holding Co.

a) 65-70 million airbags (> 42 million vehicles) recalled for potential to deploy explosively, causing life-threatening injuries.

b) Thousands of frozen berries packets recalled after being linked to the Hepatitis A outbreak in Australia.

c) Approximately 150,000 contact lenses cleaner bottles recalled for not meeting sterility requirements. 2.5 million ophthalmic cannulas recalled for the potential to leak visoelastic material or detach during injection, creating the potential for serious injury.

d) Over 1 million chocolate bars recalled due to a Salmonella outbreak.

e) Vioxx, a prescription arthritis drug, was taken off shelves after being linked to heart problems and causing thousands of deaths.

f) Approximately 7 million coffee makers recalled due to over-heating, causing burn-related injuries.

g) The Department of Transportation's recall schedule for vehicles containing Takata airbags grew increasingly aggressive from Nov. 2014 to Dec. 2016, as the product was continually linked to more and more fatalities. After facing an expected $1 billion in fines by 2017, Takata declared bankruptcy in June 2017, and was then acquired by Key Safety Systems.

h) Merck acquired Schering-Plough After a joint venture in 2000 to develop a cholesterol-lowering drug, many believed Merck and Schering-Plough would merge in the future. After the Sept. 2004 Vioxx recall, Merrill Lynch assessed an increase in the likelihood of Merck acquiring Schering-Plough, as it would be considered a "strategic action" taken to improve investor per-

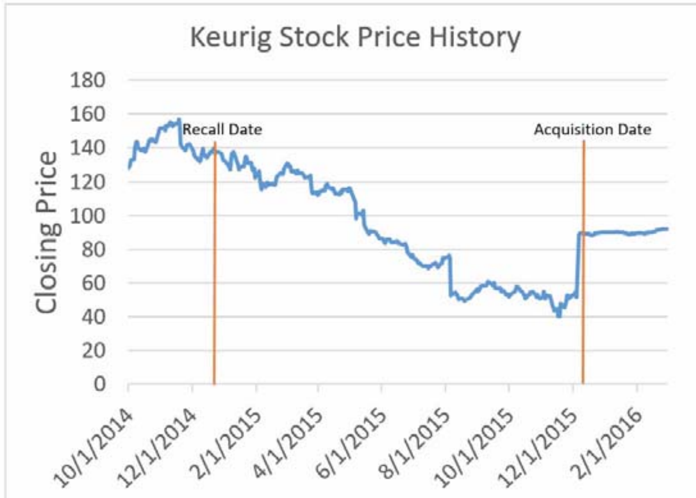

Figure 11: KEURIG

ception. [11] Nevertheless, no negotiations happened until Mar. 2009, when Merck decided to expand its laboratories via acquisition, due to increasing competition from generic drugs and years of declining sales. The merger was finalized in Nov. 2009, in which Merck bought Schering-Plough for $41 billion. The deal was a reverse merger, in which Schering-Plough was the surviving company, but was renamed Merck.[12]

[11] https://www.forbes.com/2004/09/30/0930automarketscan09.html

[12] The OECD has a website on recalls, at
https://globalrecalls.oecd.org/front/index.html#/recalls?scrollTop=659
but the only recalls the site has that are also listed in the above table are the Keurig recalls and the Takata recalls. However, for Takata, they only list a few of the vehicles recalled in Australia, and none from the U.S. The information is roughly the same as the one already in the table.

Figure 12: MERCK

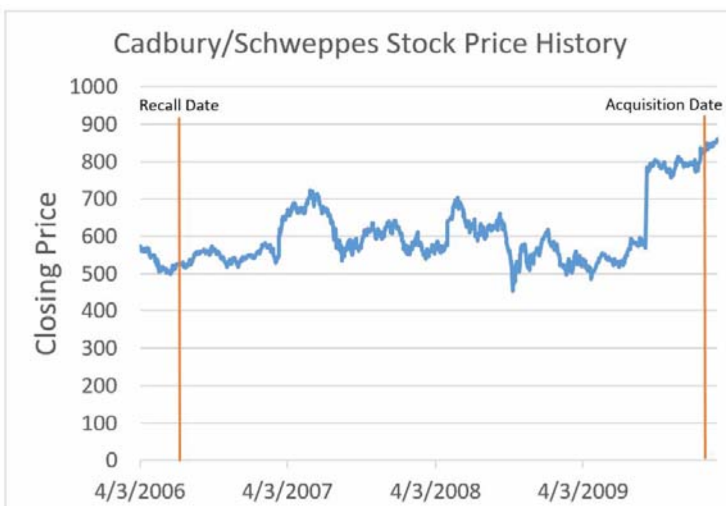

Figure 13: CADBURY/SCHWEPPES

Figure 14: TAKATA

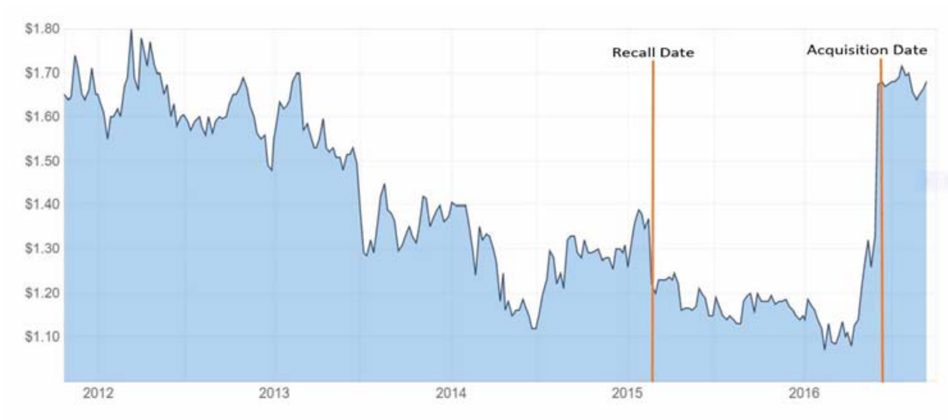

Figure 15: PATTIES FOODS PRICE HISTORY