Supplemental Materials for Nicholas G. Rupp and Curtis R. Taylor, "Who Initiates Recalls and Who Cares? Evidence from the Automobile Industry," *The Journal of Industrial Economics.*

1 Theory and Empirical Implications

We model the recall process as a two-period multi-stage game of incomplete information. The players in the game are a government agency, the manufacturers, and the car owners. The game begins at $t = 0$ when the manufacturers produce and sell cars. There are J manufacturers, n_m units of model m sold, and a total of M different models on the market. The probability that a unit of model m will have a safety failure in either period is p_m. A safety failure induces an expected utility loss of k dollars on a car owner, who is assumed to be risk neutral for convenience. The parameter p_m is the outcome of a random variable p with continuous distribution $F(p)$ on $[0, \overline{p}]$. At $t = 0$, p_m is unknown, which means that all car models are ex ante identical.

During the first period ($t = 0$ to $t = 1$), car owners experience safety failures which they report to the manufacturers and the government agency. Each time a unit of model m has a safety failure, the manufacturer of that model (if it is still in business) incurs an expected liability cost of k dollars; i.e., a car owner is fully compensated for his injuries. There is, however, some exogenous probability that a manufacturer will become bankrupt and, therefore, be unable to make liability payments. The probability that firm j becomes insolvent in the second period is $1 - q_j \in [0, 1]$.[1]

The recall process begins at $t = 1$. The government agency is the 'first-mover' in the game. In particular, there is assumed to be an objective safety standard, $\underline{p} < \overline{p}$ such that model m is deemed to be unsafe if $p_m > \underline{p}$. The agency's payoff is increasing in the number of unsafe cars that are repaired.[2] In order to recall model m, it is necessary first to conduct an engineering study at a cost of c. A study reveals the value of p_m and the cost r_m of repairing each unit of the model. The repair cost is $r_L \geq 0$ with probability $1 - \alpha$ or $r_H > r_L$ with probability α, and r is distributed independently of p.[3] A repaired car has zero probability of having a safety failure in the second period.

The government agency is endowed with a budget of b dollars with which it can conduct $b/c < M$ engineering studies. If the agency performs an engineering study on model m and discovers $p_m > \underline{p}$, then it 'orders' the manufacturer to recall the model. If it discovers $p_m \leq \underline{p}$, then it does not initiate a recall. In either case, the agency informs the manufacturer of the outcome of the study. Once the agency has exhausted its budget, each of the manufacturers decides whether it would like to perform any engineering studies. If a manufacturer performs an engineering study on model m, then it privately learns the values of p_m and r_m. It, then, recalls model m only if its expected benefit from doing so outweighs its expected cost; i.e., if the expected reduction in liability payments is greater than the cost of repairing the cars of owners who respond to the recall.

The final movers in the recall game are car owners. Specifically, if model m is recalled, then the manufacturer notifies the n_m car owners of the problem (i.e., tells them the value of p_m) and offers to perform the necessary repairs free-of-charge. There is, however, an implicit cost in time

[1] The probability of bankruptcy in the first period has no bearing on the recall process because an insolvent firm cannot recall its cars.

[2] We do not specify the agency's payoff function explicitly because of the long-standing debate over bureaucratic objectives. It seems reasonable, however, that the agency's payoff is increasing in the number of unsafe cars repaired.

[3] The analysis generalizes easily to more than two repair cost outcomes.

and effort to each owner from bringing his car in for repair. If an owner receives a recall notice and does not bring his car in for repair, then the manufacturer is absolved of any liability in the second period (from $t = 1$ to $t = 2$).[4] Hence, a car owner with time cost of z compares this with the expected cost of being injured when deciding whether to respond to a recall notice. The proportion of consumers with time cost less than z is given by the distribution function $G(z)$ with density $g(z)$.

It is notationally convenient to normalize k to one monetary unit. Also, the following assumptions, which are interpreted below, greatly simplify the analysis:

$$G(0) > 0, \tag{A1}$$

$$r_H G(\overline{p}) < q_j \overline{p}, \quad j = 1, \dots, J, \tag{A2}$$

$$r_H g(p) < q_j, \quad j = 1, \dots, J. \tag{A3}$$

As usual, this game is solved via backward induction. Therefore, first consider a car owner with time cost z who receives a recall notice informing him that his car will have a safety failure in the second period with probability p_m. This owner will respond to the recall notice if and only if z □ p_m. Hence, the total expected repair cost to firm j if model m is recalled is $r_m n_m G(p_m)$. On the other hand, its expected liability cost if it does not recall model m is $q_j n_m p_m$. Hence, firm j will 'voluntarily' recall model m if and only if an engineering study reveals

$$r_m G(p_m) \; □ \; q_j p_m.$$

In light of this expression, conditions (A1), (A2), and (A3) are now easily interpreted. Condition (A1) says that there is a group of car owners who will respond to any recall. Condition (A2) says that any firm would always voluntarily recall the most unsafe cars. Condition (A3), which is largely technical, says that $G(\cdot)$ does not rise too fast. Together, these imply the following.

LEMMA 1 *There exists a function $p^*(q_j, r_m)$ such that firm j will recall model m if and only if $p_m > p^*(q_j, r_m)$. Moreover, p^* is decreasing in q_j and increasing in r_m (see Figure 1).*

PROOF: Note that (A1) and (A2) imply that $r_m G(p)$ and $q_j p$ cross at least once, while (A3) implies that they cross at most once. Hence, p^* is the unique number satisfying

$$r_m G(p^*) = q_j p^*. \tag{1}$$

The comparative statics follow from implicitly differentiating this and invoking (A3). ||

The next step in solving the game is to determine the models (if any) on which a firm will perform engineering studies. First, note that a firm need not conduct a study on a model that the government has already investigated because p_m and r_m are already known. Now, consider a

[4] In fact, failure to respond to a recall does not mean that an owner cannot sue under products liability or tort law for subsequent injury. The manufacturer may, however, argue in its defense that the plaintiff's failure to respond to the recall constitutes assumption of the risk, contributory negligence, or comparative negligence. Our results are qualitatively unchanged so long as failure to respond to a recall weakens the plaintiff's case to some degree.

model for which n units were sold and x first-period safety failures were reported. Bayes' Rule gives

$$F(p|x,n) = \frac{\int_0^p \tilde{p}^x(1-\tilde{p})^{n-x}dF(\tilde{p})}{\int_0^{\overline{p}} y^x(1-y)^{n-x}dF(y)}. \tag{2}$$

LEMMA 2: *The CDF $F(p|x,n)$ is decreasing in x and increasing in n.*

PROOF: Suppose not. Then, differentiating (2) with respect to x gives

$$\begin{aligned}&\int_0^p [\ln(\tilde{p}) - \ln(1-\tilde{p})]\tilde{p}^x(1-\tilde{p})^{n-x}dF(\tilde{p}) \int_0^{\overline{p}} y^x(1-y)^{n-x}dF(y)\\ &> \int_0^p \tilde{p}^x(1-\tilde{p})^{n-x}dF(\tilde{p}) \int_0^{\overline{p}} [\ln(y) - \ln(1-y)]y^x(1-y)^{n-x}dF(y).\end{aligned}$$

Now, break the integrals running from 0 to $\overline{p}$ into the sum of two integrals running from 0 to p and from p to $\overline{p}$. Then subtract like terms from both sides to get

$$\begin{aligned}&\int_0^p [\ln(\tilde{p}) - \ln(1-\tilde{p})]\tilde{p}^x(1-\tilde{p})^{n-x}dF(\tilde{p}) \int_p^{\overline{p}} y^x(1-y)^{n-x}dF(y)\\ &> \int_0^p \tilde{p}^x(1-\tilde{p})^{n-x}dF(\tilde{p}) \int_p^{\overline{p}} [\ln(y) - \ln(1-y)]y^x(1-y)^{n-x}dF(y).\end{aligned}$$

or

$$\begin{aligned}&\int_0^p \int_p^{\overline{p}} [\ln(\tilde{p}) - \ln(1-\tilde{p})]\tilde{p}^x(1-\tilde{p})^{n-x}y^x(1-y)^{n-x}f(\tilde{p})f(y)dyd\tilde{p}\\ &> \int_0^p \int_p^{\overline{p}} [\ln(y) - \ln(1-y)]\tilde{p}^x(1-\tilde{p})^{n-x}y^x(1-y)^{n-x}f(\tilde{p})f(y)dyd\tilde{p}.\end{aligned}$$

At every point $(\tilde{p}, y) \in [0,p] \times [p,\overline{p}]$, however, $\tilde{p} < y$, which implies

$$\frac{\tilde{p}}{1-\tilde{p}} < \frac{y}{1-y},$$

which implies

$$\ln(\tilde{p}) - \ln(1-\tilde{p}) < \ln(y) - \ln(1-y),$$

which implies

$$[\ln(\tilde{p}) - \ln(1-\tilde{p})]\tilde{p}^x(1-\tilde{p})^{n-x}y^x(1-y)^{n-x}f(\tilde{p})f(y) < [\ln(y) - \ln(1-y)]\tilde{p}^x(1-\tilde{p})^{n-x}y^x(1-y)^{n-x}f(\tilde{p})f(y).$$

Integrating both sides over $[0,p] \times [p,\overline{p}]$ yields a contradiction. The same method of proof establishes the claim for n. ||

This lemma shows that the updated distribution, $F(p|x,n)$, is decreasing in the number of first-period safety failures, x, and increasing in the number of units on the road, n. In other words, if $x' > x$ then distribution $F(p|x', n)$ first-order stochastically dominates $F(p|x, n)$, and results in a higher expected value of p. On the other hand, if $n' > n$, then $F(p|x, n')$ is dominated by $F(p|x, n)$ and results in a lower expected value of p.

Combining Lemma 1 with (2), the expected total cost to firm j if it performs an engineering study on model m is

$$\begin{aligned} c \quad &+(1-\alpha)n_m \left[\int_0^{p^*(q_j, r_L)} q_j p dF(p|x_m, n_m) + \int_{p^*(q_j, r_L)}^{\overline{p}} r_L G(p) dF(p|x_m, n_m) \right] \\ &+\alpha n_m \left[\int_0^{p^*(q_j, r_H)} q_j p dF(p|x_m, n_m) + \int_{p^*(q_j, r_H)}^{\overline{p}} r_H G(p) dF(p|x_m, n_m) \right] \end{aligned}.$$

The first integral in each set of brackets represents a case in which the engineering study does not result in a recall while the second integral represents a case in which the firm recalls the model. If firm j does not perform an engineering study on model m, then it cannot recall it, which yields an expected liability cost of

$$n_m \int_0^{\overline{p}} q_j p dF(p|x_m, n_m).$$

Combining this expression with the previous one, firm j will perform a study on model m if and only if

(3)
$$c < n_m \left[(1-\alpha) \int_{p^*(q_j, r_L)}^{\overline{p}} (q_j p - r_L G(p))\, dF(p|x_m, n_m) + \alpha \int_{p^*(q_j, r_H)}^{\overline{p}} (q_j p - r_H G(p))\, dF(p|x_m, n_m) \right].$$

Both integrals on the right side of this expression are positive because the integrands are increasing functions (by (A3)) which equal zero at the lower limit of integration (by (1)). Indeed, $q_j p_m - r_m G(p_m)$ is the cost savings from conducting a recall. Hence, the right side of (3) is the expected benefit from performing an engineering study while the left side is the cost.

PROPOSITION 1: *The expected benefit for the firm from performing an engineering study is: (i) increasing in x_m; (ii) increasing in q_j; and (iii) decreasing in α.*

PROOF: Integrating the right side of (3) by parts yields

(4)
$$\begin{aligned} B = \quad & n_m(1-\alpha) \int_{p^*(q_j, r_L)}^{\overline{p}} (q_j - r_L g(p))\, (1 - F(p|x_m, n_m))\, dp \\ &+ n_m \alpha \int_{p^*(q_j, r_H)}^{\overline{p}} (q_j - r_H g(p))\, (1 - F(p|x_m, n_m))\, dp. \end{aligned}$$

(i) Expression (A3) ensures that both integrands on the right side of (4) are always positive. Now, since $F(p|x_m, n_m)$ is decreasing in x_m (by Lemma 2), B is evidently increasing in x_m. *(ii)* Use

Leibnitz Rule to obtain

$$\begin{aligned}\frac{\partial B}{\partial q_j} = \; & n_m(1-\alpha)\left(p^*(q_j,r_L)(1-F(p^*(q_j,r_L)|x_m,n_m)) + \int_{p^*(q_j,r_L)}^{\overline{p}} (1-F(p|x_m,n_m)dp\right) \\ & + n_m\alpha\left(p^*(q_j,r_H)(1-F(p^*(q_j,r_H)|x_m,n_m)) + \int_{p^*(q_j,r_H)}^{\overline{p}} (1-F(p|x_m,n_m))dp\right).\end{aligned}$$

The right side of this is evidently positive, which establishes the second claim. *(iii)*

$$\begin{aligned}\frac{\partial B}{\partial \alpha} = \; & -n_m \int_{p^*(q_j,r_L)}^{p^*(q_j,r_H)} (q_j - r_L g(p))\,(1-F(p|x_m,n_m))\,dp \\ & -n_m \int_{p^*(q_j,r_H)}^{\overline{p}} (r_H - r_L) g(p)\,(1-F(p|x_m,n_m))\,dp.\end{aligned}$$

Since both integrands on the right are positive, the expression is evidently negative, which establishes the claim. ‖

The intuition behind this result is easily grasped. A manufacturer is more likely to conduct an engineering study on a model if it had a large number of first-period safety failures, if the probability of becoming insolvent is small, and if repairs are likely to be low cost. In other words, a firm will investigate the models that it is most likely to recall. This makes sense because the informational value associated with an engineering study stems entirely from identifying those models that the firm would benefit from recalling. There is no point in investigating a model that it is unlikely to recall.

The final step in solving the recall game is to determine the models on which the government agency will perform engineering studies. At first blush, it seems like the agency would most effectively use its scarce budgetary resources by investigating the models with the worst first-period safety records. This, however, is not correct.

Although the agency benefits most when the models with the worst safety records are investigated, the preceding discussion indicates that these are exactly the models that the manufacturers will investigate on their own. While there may be a discrepancy between $\underline{p}$ and $p^*(q_j, r_m)$, manufacturer j would not bother to perform an engineering study on model m if it did not believe that it was fairly likely to recall it. Hence, the government agency need not expend precious budgetary resources performing engineering studies on models that the manufacturers have private incentives to investigate. Rather, the agency can 'stretch' its budget by conducting studies on models that would otherwise not be investigated. Proposition 1 provides some guidance about which models these are likely to be.

Corollary 1: *In the subgame perfect Nash equilibrium of the recall game, the government agency spends its budget investigating the models with the worst safety records that the manufacturers do not possess private incentives to investigate. These are models for which x_m/n_m is not especially high, q_j is relatively low, and α is relatively high.*

Given that we have very incomplete data on the safety records of recalled cars, what empirical implications can be drawn from Proposition 1 and Corollary 1? First, since the manufacturers investigate the models with the very worst safety records, the recalls they initiate will on average be associated with higher values of p_m than the recalls initiated by the government. In other words, recalls involving the most hazardous conditions will tend to be initiated by the manufacturers, while recalls involving marginally hazardous conditions will more often be initiated by the government. Second, a firm with a low value of q_j heavily discounts future liability payments as compared with the immediate costs associated with a recall. Hence, firms with relatively poor financial standing require more policing by the government regarding the initiation of recalls. Third, the manufacturers are more likely to initiate recalls for which the repair costs are low. Also, while the analysis performed above does not address vehicle age directly, it is relatively straightforward (though quite messy) to augment the theory to account for model vintage. The empirical implications of such an extension are, however, clear. Since old cars will be retired sooner than new ones and since the probability of a safety failure is roughly proportional to the remaining life of a car, manufacturers should be less likely than the government to initiate recalls involving older automobiles.[5] Finally, the implications regarding the number of vehicles, n_m, are somewhat more ambiguous. Specifically, both the manufacturer and the government benefit from large recalls because large recalls reduce the risk of injury to more owners than small ones. On the other hand, Lemma 2 indicates that a small value of x_m/n_m means that a recall is unlikely to be warranted and hence, that a manufacturer is less likely to investigate such a model. This suggests that manufacturers may be less likely than the government to initiate recalls involving a large number of units.

The empirical implications regarding the percentage of owners who respond to a recall are quite straightforward. In particular, the fraction of consumers who respond to a recall of model m is $G(p_m)$. Hence, recalls involving more hazardous conditions should give rise to higher response rates. Also, q_j, r_m, and who initiated the recall should not significantly impact response rates. There are also two indirect empirical implications of the theory that can be tested. First, firms with more dealerships should have higher response rates because owners will have to travel shorter distances on average to have their cars repaired. Second, recalls that receive significant publicity should have higher response rates because owners may have mistaken recall notices as junk mail or may, for some reason, not have received a notice at all. The remainder of the paper is devoted to estimating these effects and testing their significance.

[5] The fact that the probability of a safety failure is roughly proportional to the remaining life of a car comes from the Poisson approximation to the binomial for small values of p. See, for example, Casella and Berger (1990, pp. 94–5).