LE 10:30 AM 8/1/01
VISCUSI/ZECKHAUSER

The Denominator Blindness Effect: Accident Frequencies and the Misjudgment of Recklessness*

by

W. Kip Viscusi
Cogan Professor of Law and Economics
Harvard Law School
Cambridge, MA 02138

and

Richard J. Zeckhauser
Ramsey Professor of Political Economy
John F. Kennedy School of Government
Harvard University
Cambridge, MA 02138

December 20, 2000

*This research was supported in part by the Sheldon Seevak Research Fund, the Olin Center for Law, Economics and Business, and a grant to the authors from the Exxon Corporation. DeYett Law provided superb programming assistance and expositional suggestions. Miriam Avins provided helpful comments.

Abstract

People's judgments of accident risk may be seriously flawed because they routinely neglect relevant information about exposure. Such risk judgments affect personal and public policy decisions, e.g., choice of a transport mode, but also play a vital role in legal determinations, such as assessments of recklessness. Experimental evidence for a sample of 420 jury-eligible adults indicates that people incorporate information on the number of accidents, which is the numerator of the risk frequency calculation. However, they appear blind to information on exposure, such as the scale of a firm's operations, which is the risk frequency denominator. Hence, the actual observed accident frequency of accidents/exposure is not influential. Assessments of the number of locations of firms are also biased in a manner that will tend to disadvantage firms with a large scale of operations.

I. Introduction

Juries faced with the task of assessing whether punitive damages should be awarded often must confront the issue of whether the defendant has engaged in reckless behavior. When the defendant is an individual, malice is often an additional concern but malice generally is not a factor for actions of corporate entities.[1] Standard jury instructions with respect to punitive damages often call for the jury to determine whether the behavior of the company was reckless. Such instructions, such as the following example from an actual court case, are typical:

> The purposes of punitive damages are to punish a defendant and to deter a defendant and others from committing similar acts in the future.
>
> Plaintiff has the burden of proving that punitive damages should be awarded by a preponderance of the evidence. You may award punitive damages only if you find that the defendant's conduct
>
> (1) was malicious; or
> (2) manifested reckless or callous disregard for the rights of others.
>
> Conduct is malicious if it is accompanied by ill will, or spite, or if it is for the purpose of injuring another.
>
> In order for conduct to be in reckless or callous disregard of the rights of others, four factors must be present. First, a defendant must be subjectively conscious of a particular grave danger or risk of harm, and the danger or risk must be a foreseeable and probable effect of the conduct. Second, the particular danger or risk of which the defendant was subjectively conscious must in fact have eventuated. Third, a defendant must have disregarded the risk in deciding how to act. Fourth, a defendant's conduct in ignoring the danger or risk must have involved a gross deviation from the level of care which an ordinary person would use, having due regard to all circumstances.
>
> Reckless conduct is not the same as negligence. Negligence is the failure to use such care as a reasonable, prudent, and careful person would use under similar circumstances. Reckless conduct differs from negligence in that it requires a conscious choice of action, either with knowledge of

[1] Polinsky and Shavell (1998) provide a detailed overview of the role of malice and other possible factors pertinent to assessing punitive damages.

serious danger to others or with knowledge of facts which would disclose the danger to any reasonable person.[2]

Despite the detail of these instructions, what we mean by recklessness is not in fact well defined. Because a risk-free society is not feasible, presumably recklessness is a failure to strike an appropriate balance between risk and cost in situations where additional expenditures would have reduced the risk. Experimental evidence suggests that people are not able to make such judgments reliably for a wide variety of legal contexts.[3] Hindsight bias often intrudes, as people view a risk situation after an accident as having been more preventable than it was.[4] In addition, if corporations undertake explicit efforts to balance risk and cost through a risk analysis, the very act of making such explicit tradeoffs in contexts where people are at risk may be viewed as a form of a reckless disregard for individual life.[5] Ideally, however, companies should be encouraged to balance these concerns, thereby ensuring that whatever risks remain were not addressed because the costs were too high.

Any judgments of whether a company was reckless in its risk-cost balancing entails some judgment of the resulting risk level. Can people think systematically about risks and accurately assess how hazardous were various activities? Various generic biases in risk belief are well documented, and in some cases may intrude on a jury's decision making.[6] For example, the observed pattern in which people overestimate small

---

[2] Jardel Co. Inc. et al. v. K. Hughes.

[3] See, among others, Hastie, Schkade, and Payne (1998, 1999a, 1999b); Kahneman, Schkade, and Sunstein (1998); Schkade, Sunstein, and Kahneman (2000); Sunstein, Kahneman, and Schkade (1998); and Sunstein, Schkade, and Kahneman (2000).

[4] Assessments of the role of hindsight for juror and judge decisions appear in Rachlinski (1998); Hastie and Viscusi (1998); Hastie, Schkade, and Payne (1999b); and Viscusi (1999).

[5] Viscusi (2000) presents evidence on the treatment of corporate risk analyses.

[6] Viscusi (1999, 2001) links many of the established pattern of bias to legal contexts.

risks may lead to an exaggerated response to the hazard.[7] In this paper, we examine whether people have systematic biases in how they process risk information pertaining to legal cases involving accidents. We present jury-eligible individuals with the kinds of information they are likely to receive in courtroom settings, and see if they process this information in a way that enables them to form sensible judgments pertaining to the risk.

More specifically, their task is to assess the risk level based on the observed accident history, where

$$\text{Frequency of Accidents} = \frac{\text{Number of Adverse Outcomes}}{\text{Level of Exposure}}. \tag{1}$$

Jurors will typically receive information about a particular accident, past accidents of that type that have occurred because of a firm's behavior, and information on the scale of the firm's operations that will generate the risk exposure. As equation (1) indicates, at least in theory the task of combining the number of adverse outcomes and the level of exposure to calculate the accident frequency is a straightforward arithmetic exercise.

How well do people process information pertaining to the number of accidents and the level of risk exposure? For example, a juror might be told how many accidents occurred in a given number of product deliveries. The number of accidents is the numerator and the number of deliveries is the denominator when determining the rate of accidents per delivery. Do people combine this information in a reliable manner in making judgments about the recklessness of particular activities? Our hypothesis is that people are much more responsive to information about the numerator, or the total number of accidents, than they are to the denominator, which is the measure of the total level of the particular economic activity. If so, then large-scale risky operations would be at a

[7] See Fischhoff et al. (1981) for discussion of the pertinent psychological literature.

disadvantage in terms of the public perception of the riskiness of the activity when compared to smaller-scale efforts. Such biases are quite apart from a range of "deep pocket" biases, which also tend to disadvantage large firms.

This type of bias does not appear to have been fully explored previously in the literature. Discussion of public risk perception efforts often focus on the risk numerator, such as the total number of people killed by a certain cause of death, which may account for the observed overestimation of such risks.[8] Related evidence suggests that for any given probability of winning a prize, people prefer the lottery with the greatest number of prizes. For example, people would prefer 10 chances out of 190 to win a prize to 1 chance out of 19. This is consistent with the hypothesis that people process the numerator more reliably than the denominator.[9] However, such a bias could also be due to distrust of experimental lotteries and a belief that they are more likely to be legitimate if there are many prizes.

This paper reports upon an experiment in which 420 jury-eligible respondents engaged in a series of analyses pertaining to cases in legal settings. Section II describes the sample and the general information given to subjects. It also outlines our model of how judgments of the risk probability affect assessments of recklessness. The principal case studies involve a pizza delivery risk example (Section III) and a hazardous chemical

---

[8] In particular, Viscusi (1992), p. 7 notes: "This pattern of overestimation may surprise many participants in the smoking debate, but it is quite consistent with other evidence on highly publicized hazards. People frequently overassess widely publicized risks, whether the risks are those of smoking or the chance of being killed by lightning or a tornado. One contributor to this overassessment of the risk is that these public accounts call individuals' attention to the adverse outcome but do not indicate the probability that the event will occur. Media accounts provide frequent and selective coverage of the numerator of the risk (e.g., the number of tornado deaths) without information on the denominator (e.g., the size of the reference population), making incorporation of public information into risk judgments difficult. The annual reports of the Surgeon General have a similar emphasis on tallies of the adverse health outcome without indicating the number of smokers or the intensity of the product's use."

[9] See Denes-Raj and Epstein (1994) for discussion of experimental evidence on this issue.

transportation risk (Section IV). Our experimental methodology presents scenarios that vary different parameters that determine the accident rate. We then assess the responsiveness of individuals to these parameters in their judgments of the company's recklessness. In each instance, we find that the scale of operations is not influential, but the number of accidents is. In Section V we focus on the underlying hypothesis that people fail to fully grasp the extent of the scale of operations by investigating how accurately respondents assess the number of locations of various firms in the area. These findings, in conjunction with those pertaining to the judgments of recklessness, all suggest that given the same accident frequencies people are more likely to judge large-scale operations as reckless than smaller companies.

II. Conceptual Framework and Sample Characteristics

Judging Recklessness

We presented people with information about the number of accidents and level of exposure as indicated by the total amount of economic activity. Based on this information, they were asked to assess whether the defendant was reckless. From a conceptual standpoint, for any given type of activity, a firm is reckless if, in the judgment of the juror,

$$p^* > s, \qquad (2)$$

where p* is the assessed probability of an accident associated with the company's activity and s is some critical value for that activity above which the juror will believe that the company has been reckless. The critical level s will vary depending on the character of the economic activity involved. For example, the construction industry has a fatal accident rate an order of magnitude larger than manufacturing industries, which in turn have fatal accident rates that dwarf those facing college professors. If a university was operated in a manner that college professors had the same fatality rate as an average construction firm, then one might well conclude that the school was being operated in a reckless manner. Thus, implicit in any judgment of recklessness is some sense of how expensive it is to achieve and enhance the safety level for that particular activity, and an assessment of what the resulting risk level implies about the balance the defendant has struck between risk and cost.

More specifically, suppose one views negligence in terms of a Learned Hand formulation of the negligence doctrine.[10] A firm is negligent if the assessed risk level that leads to a loss L creates expected losses in excess of the cost C of the safety measure needed to eliminate the risk, or

$$p^* L > C, \tag{3}$$

or

$$p^* > \frac{C}{L}. \tag{4}$$

A firm is negligent if inequality (4) holds. If p^* is greatly in excess of C/L, then this gap can be viewed as a measure of the extent of the firm's recklessness. For any particular context there may be some critical value s for which a firm is judged to be reckless and not simply negligent, if

$$p^* > s >> \frac{C}{L}. \tag{5}$$

[10] See Posner (1986) for a description of the Learned Hand formula.

Such notions of balancing costs and benefits are quite explicit in the Learned Hand formulation of negligence rules. When companies fail to undertake safety improvements for which the benefits exceed the costs, then they should be found guilty of negligence. If a company could achieve substantial safety benefits at very modest cost, then one might conclude that the behavior is more than negligent; it shows a reckless disregard for safety. Indeed, a previous experimental analysis of a railroad accident scenario, in which the respondents were state judges, found that when the judges assessed the benefits of safety measures as being at least six times larger than the cost of the safety improvements, this spread served, in effect, as the trigger for whether they thought that punitive damages would be pertinent (Viscusi, 1999).

Bayesian Learning

As a benchmark, we first consider individuals who learn in a rational Bayesian manner as prescribed by the discipline of decision analysis. We assume that individuals have prior beliefs about accident frequencies that are characterized by a beta distribution. This distribution is extremely flexible; it can assume a wide variety of shapes, both skewed and symmetric, and is commonly employed. Suppose that such individuals' prior risk beliefs are tantamount to having observed e accidents out of f trials. Then from their standpoint the risk of an accident is simply e/f as their prior belief pertaining to the accident frequency. Suppose then as part of the legal case the individuals receive information that the firm's activity led to g accidents out of h trials. Based on this information, the individual's posterior beliefs are governed by

$$p^* = \frac{e + g}{f + h}; \qquad (6)$$

that is, their initial beliefs were updated to e+g accidents out of f+h trials. For very high values of h relative to f, the information on the risk levels conveyed at trial will tend to be much more influential role than people's prior beliefs.

To see how such a learning process might work, suppose people begin with prior beliefs characterized by a value of e equal to 1 and f equal to 1,000. Thus, these prior beliefs would correspond to a prior probability of an accident of 0.001. People will, however, alter their risk beliefs based on experience. Suppose that at the trial the individual learns that there have been 3 accidents out of 1,000 situations in which an accident might have occurred. Thus, the actual accident frequency is 0.003. Combining this information with the individual's prior beliefs leads to a perceived probability of

$$p^* = \frac{1+3}{1{,}000+1{,}000} = 0.002\,. \qquad (7)$$

This value is closer to the actual accident frequency because of the large number of trials involved relative to the precision of the individual's prior beliefs.

If, however, the individual had observed 10 accidents out of 10,000 trials, then the posterior risk belief would be

$$p^* = \frac{1+10}{1{,}000+10{,}000} = 0.001\,, \qquad (8)$$

a value that is half as high despite representing many more accidents.

Characteristics of the Sample and General Instructions

Our experimental design presented subjects with different case scenarios. By comparing their responses to the different cases, we were able to assess the effects of different case characteristics. The only characteristics that were varied pertained to the number of accidents or to the scale of the firm's operations.

Our sample consisted of 420 jury-eligible adults. In July 2000 a marketing research firm in Austin, Texas, recruited this sample by phone. Subjects came to a central location where they participated in the study. Each individual received $40 as compensation for taking part in the survey, which lasted approximately half an hour.

The sample, which is summarized in Table 1, included a broad population cross section. One-third of the sample was black or Hispanic. The educational levels were quite diverse, as just under half of the sample had either completed high school or some college education, with the remainder being either college or post-college graduates. The mean age was 41, and women were somewhat overrepresented in the sample. Subsequent analysis will, in many cases, control for demographic characteristics that might have influenced answers.

Before beginning the survey, each respondent received general instructions indicating that their task would involve analysis of legal contexts:

> You will consider a series of legal case situations. You will be allowed as much time as you need to review the information. Please indicate your best judgment with respect to each question. In almost all instances there are no right or wrong answers. We are interested in your assessments, and people can feel differently about the cases.

In addition, respondents also received general guidance about punitive damages similar to the information often received as part of a jury's punitive damages instructions:

> Below you will consider a series of legal cases. In every instance, the trial jury has already ordered each defendant to pay compensatory damages as full compensation for the harm suffered by the plaintiff. We would like you to imagine that you are a member of the punishment jury. Your job is to decide whether and how much each defendant should be punished, in addition to paying compensatory damages.

As a jury member, you are instructed to award punitive damages if a preponderance of the evidence shows that the defendants acted either maliciously or with reckless disregard for the welfare of others. Defendants are considered to have acted maliciously if they intended to injure or harm someone or their property. Defendants are considered to have acted with reckless disregard for the welfare of others if they were aware of the probable harm to others or their property but disregarded it, and their actions were a gross deviation from the standard of care that a normal person would use.

Each respondent received one scenario for each type of case. Scenarios were assigned randomly to respondents, so there should be no systematic differences across the different samples of respondents who considered the different case scenarios. The different cases considered involved chemical spill accidents and pizza delivery accidents.

III. Chemical Spill Accidents: The Role of Numerators and Denominators

The first set of scenarios involved a company's delivery of hazardous chemicals by truck. Respondents were given information about the number of chemical spills and the number of deliveries. In this context, the number of deliveries measures the risk exposure. Subsequently, we shall use the terminology numerator and denominator to identify the number of accidents and the level of exposure.

The respondents were also told that chemical spills endangered fish and other wildlife, and were potentially hazardous to people if they contaminated the groundwater. The appendix includes the complete text for one version of this scenario. After reading the scenario, respondents assessed the probability that the company was reckless and should therefore be subject to punitive damages. Each respondent received a series of five possible probabilities ranging from 0 to 1 on a linear risk scale with intervals of 0.25

and verbal characterizations of the probabilities. For example, ½ was characterized as "possibly, 50-50."

We used four versions of this scenario (see Table 2 for an overview). In Scenario A, the chemical company had had two chemical spills out of 10,000 deliveries, or an accident frequency rate of 0.0002. In Scenario B there were 5 accidents out of 10,000 deliveries. Because the accident frequency is 2.5 times as great for Scenario B as for Scenario A, one would expect respondents to be more likely to judge the company as being reckless in that instance.[11]

Whereas Scenario B changes the numerator in Scenario A, Scenario C alters the denominator. Thus, rather than there being 10,000 deliveries, Scenario C assumes that there are 50,000 deliveries. The accident frequency rate is consequently one-fifth as great as in Scenario A, which presumably should decrease the assessed likelihood of punitive damages. The final scenario, Scenario D, involves five chemical spills out of 50,000 deliveries. Compared to Scenario C it represents an increase in the number of spills, and compared to Scenario B it has a greater number of deliveries but the same number of spills.

We tend to judge things as reckless when their risk level is significantly above what we expect in that context, which is the prior. This implies that a rational Bayesian decision maker might judge 10 accidents out of 10,000 trials as being reckless, but not 1 out of 1,000. If the initial prior were equivalent to 1 in 10,000, the first would give a

[11] There is no reason why if the risk posterior goes up from .001 to .002 that we should double the number of people who assign recklessness: (a) the prior plays and role, and (b) thresholds for recklessness need not have any particular distribution. Let's say that half the people had a threshold of .0009 and the other half were at .0025. Then this doubling would not affect the percentage assigning recklessness.

posterior judgment of 0.00055 , whereas the second would give 0.00018, or a third as great an estimated risk.

Notwithstanding these qualifications, the accident frequency should play a major role in determinations of recklessness alongside prior beliefs and the juror's recklessness threshold. Substantial evidence should swamp the influence of prior beliefs in the jurors' assessments. Because juror beliefs and the recklessness threshold are not observed, our empirical analysis focuses on the effect of accident frequency on the assessment of whether the firm is reckless.

We used the changes in the number of accidents (the numerator) and in the number of spills (the denominator) to assess whether people incorporated the information about the scale of the operation and the number of accidents in judging whether the company was reckless. The mean assessed probability that the company was reckless ranged from 0.25 for Scenario A to 0.36 for Scenario B.

Analysis of the results for pairs of scenarios indicate that our respondents used the information about the number of spills but completely ignored the information about the number of deliveries. Scenario B, in which there are five spills, yielded an assessed probability of recklessness of 0.36 as compared to 0.25 for Scenario A; this difference is statistically significant.[12] Similarly, the change in the number of spills from two in Scenario C to five in Scenario D yields a somewhat smaller increase in recklessness risk, from 0.26 to 0.33, which also proves statistically significant.[13]

---

[12] In particular, the t-value is 2.96, which is statistically significant at the 95% confidence level, two-tailed test.

[13] The calculated t-value for this comparison was 1.74, which is statistically significant based on a one-sided test at the 95% confidence level, which seems appropriate given that an increased number of spills should boost the assessed probability of recklessness rather than decrease it.

In contrast, changes in the number of deliveries (the denominator of the accident frequency calculation) failed to produce any significant differences in the likelihood that the company would be considered reckless. In Scenarios A and C, the number of spills is two, but increasing the number of deliveries from 10,000 to 50,000 altered the probability of being assumed reckless from 0.25 to 0.26, which is an insignificant difference.[14] Likewise, the shift in the number of deliveries from 10,000 to 50,000 in Scenarios B and D, which both involve five spills, altered the recklessness estimate from 0.36 in Scenario B to 0.33 in Scenario D; this difference is also not statistically significant.[15] Across all scenarios, shifts in the number of spills increase the assessed probability of recklessness, but changes in the number of deliveries do not.

The relationship between the frequency of accidents and the respondents' assessments of the probability that the company was reckless is intriguing. The recklessness assessment increased much less than proportionally. For example, Scenario B has an accident frequency that is 2.5 times as great as Scenario A, and respondents are 1.4 times as likely to believe that the company was reckless. This less than proportional response might be rational if people have strong prior beliefs -- i.e., large values for e and f -- about the likelihood that a company is reckless. They might have such beliefs if they were well informed about a risk.

As observed above, Scenario C has an accident frequency that is one-fifth as great as Scenario A but a mean assessed probability of recklessness that is almost identical; the frequency difference is due to the number of deliveries. Similarly, Scenario D has an accident frequency that is half as great as Scenario A, yet respondents are more likely to

[14] In particular, the calculated t-statistic is 0.41.
[15] The calculated t-statistic is 0.77.

view the company as reckless in Scenario D than they are in Scenario A. Again, the difference is in the denominator. What gets processed is the number of spills, not the risk level implied by the number of spills divided by the scale of the company's operations.

The personal characteristics of our respondents turned out to affect their propensity to find recklessness, as is shown in Table 3 and in a subsequent experiment. For example, Hispanics are 0.1 more likely to award punitive damages than are whites, who are the omitted group. Respondents with some college or who are college graduates are less likely to award punitive damages. Cigarette smokers, who have revealed through their decision to smoke a greater willingness to incur risks, are less likely to award punitive damages, by a probability of 0.09. Accepting more risky behavior by the company is consistent with smokers' own risk-taking patterns. Seatbelt use, however, does not predict any statistically significant difference.

That personal characteristics affect risk judgments, however, in no way diminishes our central results about the number of accidents being influential and the level of exposure being ignored when considering recklessness. Consider the coefficients for the three scenarios in Table 3, which show the impact of the scenario relative to Scenario A (the base or omitted case). In Scenario B, in which there are five spills out of 10,000 deliveries, respondents had a 0.12 higher probability of awarding punitive damages than in Scenario A, with 2 spills out of 10,000 deliveries. Similarly, respondents had a 0.08 higher probability of awarding punitive damages in Scenario D with five accidents out of 50,000 deliveries than in Scenario A. Both B and D increase the numerator of the risk calculation, increasing the likelihood that punitive damages will be awarded. In contrast, there is no statistically significant effect for Scenario C, in which the denominator is changed by a factor of five relative to Scenario A.

Table 4 presents the key coefficients from two different regression results for which the estimates of the personal characteristic variables are not reported because they closely parallel the findings in Table 3. The first specification reports the simple regression of the probability of awarding punitive damages on the number of deliveries and the number of spills. The number of deliveries has no statistically significant effect and has a negligible influence on the assessed probability of recklessness. In contrast, the increased number of spills boosts the assessed probability of recklessness by 0.03 per spill. The second specification in Table 4 regressed the natural logarithm of the probability that the company was reckless against the log of the number of deliveries and the log value of the number of spills.[16] For this formulation, which is commonly used in empirical analysis, the logarithm of the assessed risk should be positively related to the logarithm of the number of accidents and negatively related to the logarithm of the number of deliveries. In our calculations, however, we find that the number of deliveries does not play a statistically significant role, but the number of spills does and is statistically significant.

The consistent pattern that emerges is that it is not the level of risk that is influential but rather the absolute number of accidents. This result holds controlling for personal characteristics and is true for both specifications in Table 4. The level of

---

[16] This formulation would be appropriate if people based their risk assessments solely on the information provided in the survey, using the formula

$$p^* = \frac{\text{Number of Accidents}}{\text{Number of Deliveries}}.$$

Then

$$\ln p^* = \ln(\text{Number of Accidents}) - \ln(\text{Number of Deliveries}).$$

economic activity generating a series of accidents plays an insignificant role in respondents' assessments of recklessness.

IV. Pizza Delivery Accidents: Does the Scale of Operations Matter?

We conducted a second set of experiments involving a quite different setting, namely automobile accidents arising out of pizza deliveries. This enables us to determine whether our earlier results generalize, or whether they reflect the sensitive issue of chemical spills.

The experimental design held the number of accidents constant at three accidents per scenario, but varied the number of pizza locations to assess whether respondents would be sensitive to this manipulation of the scale of operations.

The appendix includes a copy of a representative pizza delivery operation scenario. The risk was that of automobile accidents that arose while a driver for the pizza chain was delivering pizzas. In each case there was property damage to vehicles but no personal injury. The scenarios asked respondents to assess the probability that the company called Best Pizza was reckless. A separate question asked respondents to rate the importance of different kinds of information, which helps us determine whether the scale of operations is influential in their thinking.

Table 5 summarizes the experimental design. In every instance there were three accidents. In Scenario A the firm was a local firm with an unspecified number of locations. Scenario B indicates that the firm is local but has 15 locations, whereas in Scenario C the local firm has two locations. Presumably, the decrease in the number of locations make liability judgments more likely, as the accident rate is 7.5 times as great for Scenario C as for Scenario B. In Scenario D there are 15 locations, as in Scenario B, but the company is a national chain, which may be a less sympathetic defendant. Respondents may also view the national chain as being a large-scale enterprise no matter how many locations it has in the area.

The assessed probabilities of recklessness in this example ranged from 0.41 in Scenario B to 0.48 in Scenario A. These assessed values of the probability of reckless behavior are higher than for the hazardous chemical delivery scenario.

We consider first the results for the scenarios in which the number of locations is specified. The dramatic increase in the number of accidents per location from Scenario B to Scenario C increases the mean assessed probability of recklessness modestly, from 0.41 to 0.46, but this difference is not statistically significant.[17] Similarly, the assessed probability of recklessness in Scenario C is almost identical to that in Scenario D even though the risk levels differ by a factor of 7.5.[18] That comparison involved not only a change in the number of locations but also a shift in the identity of the firm from a local to a national firm. We isolate the role of a national firm by comparing Scenarios B and D, for which the number of accidents and number of locations is identical. The shift to a national firm increases the assessed probability of recklessness from 0.41 to 0.47, a statistically significant difference.[19]

If the number of locations is unspecified, as in Scenario A, then the mean assessed probability of recklessness reaches its highest value of 0.48. This estimate is statistically

---

[17] In particular, the calculated t-statistic is 1.45, which falls short of statistical significance based on a one-tailed t-test at the 95% confidence level.

[18] The calculated t-statistic for this comparison is 0.37.

[19] The calculated t-statistic is 1.79, which is statistically significant at the 95% confidence level, one-tailed test. This result is plausible if one acts with the working hypothesis the jurors will be more likely to assess recklessness if the firm is not local.

different only from that in Scenario B, which has the lowest accident frequency rate, 0.2 per location, for a local firm.[20] In short, and parallel to our earlier results about chemical spills, the number of locations did not influence assessments of recklessness, despite its immediate link to level of exposure, the denominator of frequency of accidents.

We wished to determine whether personal characteristics affected recklessness assessments in the pizza case as they did with chemical spills. Table 6 reports a regression analysis that parallels Table 3. The results in Table 6 examine the determinants of the assessed probability of recklessness controlling for various personal characteristics. Female respondents are more likely to assess a greater degree of recklessness, as are Hispanic respondents and respondents who are in the other nonwhite group. The omitted education group variable consists of those with no more than a high school education, and this group assesses a greater degree of recklessness than do the three included education group variables for different levels of college education.

The omitted scenario indicator variable is that for Scenario A. Only Scenario B has a statistically significant influence, which implies a negative effect on the assessed probability of recklessness of 0.08. Being a local firm with a large number of locations proves to have some influence, though not perhaps as stark as one might expect based on the change in the number of accidents per location. Moreover, the comparable risk performance of the national firm in Scenario D does not play a significant role.

We gave respondents the additional task to rate from 1 to 5 the relative importance of five different accident-related factors that they might want to learn more about in determining whether punitive damages were warranted. Three of these factors

[20] The pertinent t-test for Scenario A in comparison to the other scenarios are 1.94 for Scenario B, 0.46 for Scenario C, and 0.08 for Scenario D.

were not related to the scale of operations: driver training experience, incentives for fast delivery, and car maintenance practices. The other two factors were related to the scale of activity: number of deliveries and average length of delivery trip. Consistent with our findings about level of exposure, the two factors that are most closely linked to the scale of operations received the lowest rank (see Table 7). Respondents' assessments of which factors should be considered when deciding whether to award punitive damages downplay influences that would affect the denominator of any risk frequency calculation.

V. Judging the Scale of Operations

Suppose that jurors are not provided with information concerning the number of locations or the scale of economic activity associated with the defendant's operations. If the defendant is a local business, they might have a reasonable assessment of what this denominator is. Or perhaps they misjudge the company's size. This would be particularly important if jurors consistently underestimated or overestimated the number of locations or other indicators of exposure.

To test for this hypothesis, the survey also included a list of local businesses, and respondents had to assess the number of locations in the Austin area (the site of this survey). The exact text of this question is included in the appendix. The types of businesses on the list included muffler repair shops, truck rental agencies, rental car agencies, pizza restaurants and delivery, fast food restaurants, banks, and pharmacies. For each type of business, the survey listed several representative national chains as well as some local enterprises when that seemed appropriate. In all, people assessed the number of locations for 25 businesses.

Table 8 shows the businesses' actual number of locations in the Austin area as well as the median value indicated by the survey respondents.[21] We use the median respondent to avoid the undue influence of outliers, such as a respondent who indicated 100 locations for some of these enterprises. Overall, the responses appear to be generally reasonable. People are aware that there are comparatively few locations for muffler repair and rental cars, and seem to be aware that there are many more fast food and pizza restaurants in the area. However, there is a substantial underestimation of the number of truck rental companies, banks, and pharmacies. These errors are not random, but in fact are quite systematic in terms of their relationship to the number of locations in the Austin area.

Figure 1 indicates the assessed number of locations for the 25 businesses on the vertical axis and the actual values on the horizontal axis. It also shows the regression line linking the median surveyed value of store locations to the actual values in the Austin area. There is the expected positive correlation; that is, respondents tend to assess a greater number of locations for larger enterprises than for the firms with a smaller number of locations. However, people tend to overestimate the number of locations of stores with a small number of locations and underestimate the number of locations of enterprises with a large number of locations. Thus, there is a tendency to homogenize these differences. This is a fairly conventional pattern of bias, not unlike that observed for biases of assessment of mortality risk values. Consequently, individuals will tend to underestimate the risk of an accident for firms with a small number of locations and overestimate the risk of an accident for firms with a large number of locations. Large

[21] The actual number of locations was determined by counting appropriate business listings in the Southwestern Bell Yellow Pages directory for Greater Austin. Alternative measures, such as counts using

firms consequently will be disadvantaged, and smaller firms advantaged, assuming that the frequency of accidents is a relevant factor in determining recklessness.

Do people in fact incorporate the local information about the number of business locations in their judgments? Or do they rely on their broader knowledge of national or regional representation? To address this question, we used the number of locations for these companies in the Houston area as a proxy for regional prominence. Presumably, the Houston figures would be more representative of the regional scale of operation because it is a much larger city than Austin. The second regression line in Figure 1 shows the perceived number of store locations in the Austin area as a function of the number of locations for these enterprises in Houston.[22] The goodness of fit in terms of the percentage of variation explained is comparable, but somewhat smaller for the Houston-based estimates.

A second test of the relative influence of the number of stores in Austin as compared to the number in Houston is to include both of these variables in one equation to explain the perception of the number of locations in Austin. When we included these variables in such an equation, both were statistically significant at the 95% level based on a one-tailed t-test, though only the Austin store locations variable is significant using a two-tailed t-test at the 95% confidence level.[23] The number of local store locations is influential, but respondents seem to be taking into account the regional scale as well, perhaps as a proxy for national operations.

---

internet directories, yielded similar results.

[22] The number of Houston locations is based on listings in the Southwestern Bell Yellow Pages directory for Greater Houston.

VI. Conclusion

Judging the magnitude of a risk -- how often an accident occurs per unit of exposure -- is essential to making any judgments whether the party responsible for the accident was reckless. Until one knows whether a risk is consequential or trivial, then it is impossible to assess whether efforts to address the risk were adequate. This kind of concern arises not only with respect to liability judgments but also with respect to regulatory policy. For example, the U.S. Supreme Court has ruled that the Occupational Safety and Health Administration can only regulate risks that are judged to be "significant"; any judgment of significance necessarily must entail some explanation of whether the risk is large.

To obtain a proper assessment of the risk, one must investigate the probability of various adverse consequences. The risk of an accident consists of two components, the number of adverse accidental outcomes divided by some measure of the economic activity that generates the accident. Thus, a primary task is to construct a measure of the accident frequency, such as the risk of automobile crashes per 100,000 miles driven or the probability that any given launching of a space shuttle will lead to a fatality.

The experimental evidence presented here indicates that people often do quite badly in making such judgments even when presented with all the information they need to assess accident frequency. The number of accidents influences assessments of recklessness, but people tend to ignore or give slight attention to information pertaining to the scale of the economic activity, the denominator of risk frequency. Moreover, their ranking of the relative importance of different kinds of information suggests that the scale

---

[23] In particular, the perceived number of store locations in Austin = 3.730 (1.003) + 0.178 (0.060) Austin Locations + 0.029 (0.015) Houston Locations, $R^2 = 0.54$, where numbers in parentheses are standard errors.

of operations is low on the list of things they view as important to consider when assessing liability.

In our experiments, we gave subjects information about both the number of accidents and the scale of economic activity. However, in many actual trials people only learn that a particular enterprise has been involved in an accident. Do they then impute an appropriate scale to the firm's economic activity based on its level of operations? Separate tests indicate that people tend to overestimate the level of economic activity for smaller firms and underestimate that for larger firms. If such beliefs were then incorporated into judgments of the risk level, which they well may not be, then this would further increase the tendency to penalize entities that will tend to have a large number of adverse events simply because of their large scale of operations.

That we detected these biases does not mean that jurors cannot be educated to think more analytically about risk frequency issues. However, our results suggest that eliminating such biases in risk belief is an important task that should be addressed in order to promote sounder judgments of liability.

References

Denes-Raj, Veronika, and Seymour Epstein, "Conflict Between Intuitive and Rational Processing: When People Behave Against Their Better Judgment," Journal of Personality and Social Psychology, Vol. 66, No. 5 (1994), pp. 819-829.

Fischhoff, Baruch, Sarah Lichtenstein, Paul Slovic, Stephen L. Derby, and Ralph L. Keeney, Acceptable Risk (Cambridge: Cambridge University Press, 1981).

Hastie, Reid, David Schkade, and John Payne, "A Study of Juror and Jury Judgments in Civil Cases: Deciding Liability for Punitive Damages," Law and Human Behavior, Vol. 22 (1998), pp. 287-314.

Hastie, Reid, David Schkade, and John Payne, "Juror Judgments in Civil Cases: Effects of Plaintiff's Requests and Plaintiff's Identity on Punitive Damage Awards," Law and Human Behavior, Vol. 23 (1999a), pp. 445-470.

Hastie, Reid, David Schkade, and John Payne, "Juror Judgments in Civil Cases: Hindsight Effects on Judgments of Liability for Punitive Damages," Law and Human Behavior, Vol. 23 (1999b), pp. 597-614.

Hastie, Reid, and W. Kip Viscusi, "What Juries Can't Do Well: The Jury's Performance as a Risk Manager," Arizona Law Review, Vol. 40 (1998), pp. 901-921.

Kahneman, Daniel, David Schkade, and Cass Sunstein, "Shared Outrage and Erratic Awards: The Psychology of Punitive Damages," Journal of Risk and Uncertainty, Vol. 16 (1998), pp. 49-86.

Kunreuther, Howard, et al., Disaster Insurance Protection: Public Policy Lessons (New York: John Wiley, 1978).

Lichtenstein, Sarah, et al., "Judged Frequency of Lethal Events," Journal of Experimental Psychology, Vol. 4 (1978), pp. 551-578.

Morgan, M. Granger, "On Judging the Frequency of Lethal Events: A Replication," Risk Analysis, Vol. 3 (1983), pp. 11-16.

Posner, Richard, Economic Analysis of Law, 3rd Edition (Boston: Little Brown, 1986).

Rachlinski, Jeffrey J, "A Positive Psychological Theory of Judging in Hindsight," University of Chicago Law Review, Vol. 65 (1998), pp. 571-625.

Schkade, David A., Cass R. Sunstein, and Daniel Kahneman, "Deliberating about Dollars: The Severity Shift," Columbia Law Review, Vol. 1000 (2000), pp. 1139-1175.

Sunstein, Cass, Daniel Kahneman, and David Schkade, "Assessing Punitive Damages (with Notes on Cognition and Valuation in Law)," Yale Law Journal, Vol. 107 (1998), pp. 2071-2153.

Sunstein, Cass, David Schkade, and Daniel Kahneman, "Do People Want Optimal Deterrence?" Journal of Legal Studies, Vol. 29 (2000), pp. 237-253.

Viscusi, W. Kip, Smoking: Making the Risky Decision (New York: Oxford University Press, 1992).

Viscusi, W. Kip, Rational Risk Policy (Oxford: Oxford University Press, 1998).

Viscusi, W. Kip, "How Do Judges Think about Risk?" American Law and Economics Review, Vol. 1/2 (1999), pp. 26-62.

Viscusi, W. Kip, "Corporate Risk Analysis: A Reckless Act?" Stanford Law Review, Vol. 52 (2000), pp. 547-597.

Viscusi, W. Kip, "Jurors, Judges, and the Mistreatment of Risk by the Courts," Journal of Legal Studies, (2001), forthcoming.

Table 1
Sample Characteristics

	Mean (Standard Deviation)
Age	41.31 (12.34)
Female	0.59 (0.49)
White	0.63 (0.48)
Black	0.12 (0.33)
Hispanic	0.20 (0.40)
Other nonwhite races	0.05 (0.21)
High school	0.14 (0.34)
Some college	0.32 (0.47)
College grad	0.36 (0.48)
Professional degree	0.17 (0.38)
Smoker	0.15 (0.36)
Seatbelt user	0.89 (0.32)

Note: Sample size is 422.

Table 2
Likelihood of Punitive Damages Being Awarded for Chemical Spills, by Scenario[a]

Case Scenario	Number of Spills	Number of Deliveries	Accident Frequency	Recklessness Estimate	
				Mean	Std. Error of Mean
A	2	10,000	0.0002	0.25	0.02
B	5	10,000	0.0005	0.36	0.03
C	2	50,000	0.00004	0.26	0.02
D	5	50,000	0.0001	0.33	0.03

[a]The question asked of respondents was: "How likely do you think it is that Apex [Chemical Company] was reckless in its delivery operations and hence should be subjected to punitive damages?"

Table 3
Regression of Probability of Punitive Damages Being Awarded for Chemical Spills on Personal Characteristics

Variable	Coefficient (Standard Error)
Constant	0.323* (0.073)
Age	$3.22E^{-4}$ (0.001)
Female	0.040 (0.027)
Black	0.024 (0.042)
Hispanic	0.102* (0.035)
Other nonwhite races	-0.005 (0.065)
Some college	-0.101* (0.042)
College graduate	-0.089* (0.042)
Professional degree	-0.027 (0.049)
Smoker	-0.091* (0.038)
Seatbelt user	-0.034 (0.043)
Scenario B (5 spills; 10,000 deliveries)	0.119* (0.037)
Scenario C (2 spills; 50,000 deliveries)	0.027 (0.038)
Scenario D (5 spills; 50,000 deliveries)	0.076* (0.037)

*Coefficient is significant at the 95% confidence level, two-tailed test.

Table 4
Probit Regression of the Probability of a Recklessness Finding as a Function of Spills and Deliveries

	Coefficient (Standard Error)	
Variable	Probability	Ln (Probability)
Spills	0.028** (0.009)	
Deliveries	$-2.14E^{-7}$ ($6.61E^{-7}$)	
Ln (Spills)		0.066** (0.021)
Ln (Deliveries)		-0.004 (0.012)

**Coefficients are significant at the 99% confidence level, two-tailed test.
Note: Each equation also includes the demographic variables listed in Table 3 and a constant term.

Table 5

Likelihood of Company Recklessness in Pizza Delivery Operations, by Scenario[a]

Case Scenario	Number of Accidents	Number of Locations	Accident Frequency	Firm	Recklessness Estimate: Mean	Recklessness Estimate: Std. Error of Mean
A	3	Unspecified	Unspecified	Local	0.48	0.03
B	3	15	0.2	Local	0.41	0.02
C	3	2	1.5	Local	0.46	0.03
D	3	15	0.2	National	0.47	0.03

[a]Respondents were asked to assess whether the court should award punitive damages against Best Pizza because they believe its delivery operations were reckless.

Table 6

Regression of Probability of Company Recklessness in Pizza Delivery Operations on Personal Characteristics and Scenarios

Variable	Coefficient (Standard Error)
Constant	0.492* (0.070)
Age	0.001 (0.001)
Female	0.053* (0.026)
Black	-0.043 (0.040)
Hispanic	0.086* (0.034)
Other nonwhite races	0.134* (0.062)
Some college	-0.081* (0.040)
College graduate	-0.074** (0.040)
Professional degree	-0.132* (0.047)
Smoker	-0.042 (0.037)
Seatbelt user	-0.016 (0.041)
Scenario B (3, 15, Local)	-0.077* (0.036)
Scenario C (3, 2, Local)	-0.017 (0.036)
Scenario D (3, 15, National)	-0.006 (0.036)

*Coefficient is significant at the 95% confidence level, two-tailed test.

**Coefficient is significant at the 95% confidence level, one-tailed test.

Table 7
Ranking of Additional Information in Determining Punitive Damages for Pizza Delivery Case[a]

	Mean Score
Driver training experience	2.12
Incentives for fast delivery	2.26
Car maintenance practices	3.13
Number of deliveries	3.25
Average length of delivery trip	3.98

[a]1 is most important, 5 is least important

Table 8

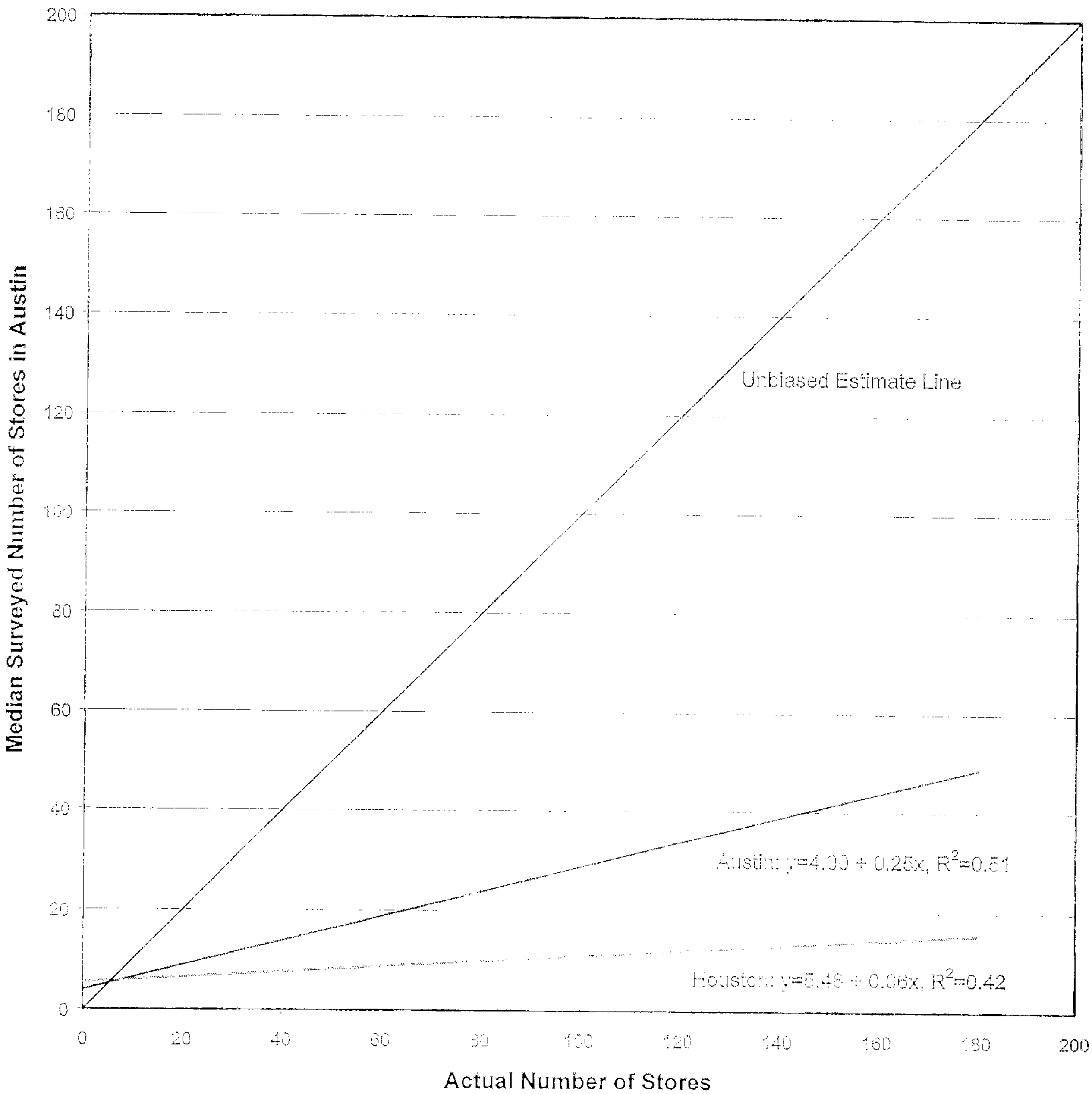
Number of Store Locations in Greater Austin

Business Name	Actual	Median Surveyed
Muffler Repair		
Brake Specialists	8	7
Meineke	5	5
Midas	4	6
Truck Rental		
Penske	19	3
Ryder	19	6
Uhaul	32	10
Rental Cars		
Avis	1	3
Budget	2	4
Hertz	1	4
Enterprise	28	5
Pizza Restaurants and Delivery		
Dominos	16	10
Doubledave's	10	5
Mr. Gatti's	29	10
Papa John's	16	8
Pizza Hut	27	12
Fast Food Restaurants		
Burger King	13	10
KFC	10	10
McDonald's	33	18
Taco Bell	29	15
Banks		
Bank of America[a]	30	13
Bank One	17	6
Wells Fargo[b]	38	12
Pharmacies		
Randall's	12	8
Eckerd	23	10
Walgreens	21	10

[a]Nationsbank merged with Bank of America, and is counted in the Bank of America category.

[b]First State Bank merged with Norwest, followed by a merger between Wells Fargo and Norwest. All three are counted in the Wells Fargo category.

Figure 1: Relationship between Perceived Number of Austin Stores and Actual Number of Austin and Houston Stores



Appendix

Chemical Spill Accident Scenario

The *Apex Chemical Company* transports hazardous chemicals for important industrial uses. These chemicals are toxic to fish and wildlife. Moreover, if the chemicals get into the water supply or the groundwater, they can create significant health hazards for people as well. Because these chemicals are transported by truck, there is some risk of a traffic accident, which in turn can cause a chemical spill. Last year, *Apex* had 2 chemical spills out of 10,000 deliveries.

How likely do you think it is that *Apex* was reckless in its delivery operations and hence should be subjected to punitive damages? Your best estimate will do.

0	¼	½	¾	1
9 9	9 9	9		
Not at All Likely	Somewhat Likely	Possibly, 50-50	Very Likely	Definitely

Pizza Delivery Accident Scenario

In calendar 1998, *Best Pizza*, a local pizza chain with 15 locations, had 3 of its employees involved in separate automobile accidents while delivering pizzas in the Austin, Texas area. Each of these accidents caused property damage to other vehicles, but no personal injury. You have been asked to assess whether the court should award punitive damages against *Best Pizza* because they believe its delivery operations were reckless. Improperly maintained vehicles, poor worker training, or emphasis on rapid delivery schedules that compromise safety all could be classified as reckless if they led to accidents.

How likely do you think it is that *Best Pizza* was reckless in at least one of these different safety dimensions? Use the scale below to indicate the probability that *Best Pizza* was reckless and did not exercise appropriate care, based on your best guess given the information you have been given above.

Probability That *Best Pizza* Was Reckless

0		¼		½	¾	1
9	9	9	9	9		
Not at All Likely		Somewhat Likely		Possibly, 50-50	Very Likely	Definitely

Rank the following different types of additional information that you would like to assist in your determination of whether punitive damages are warranted. Rate these factors from 1 to 5 with 1 being most important.

_____ Car maintenance practices
_____ Driver training and experience
_____ Incentives given to driver for fast delivery
_____ Number of deliveries
_____ Average length of delivery trip

Number of Store Locations Assessment Task

Listed below are different well known companies marketing various products in the Austin, Texas area. For each of the products, please indicate the number of stores or locations that you estimate there are of each type in this area. Please give your best estimate.

a) **Muffler repair**

Brake Specialists ________

Meineke ________

Midas ________

b) **Truck rental**

Penske ________

Ryder ________

U-Haul ________

c) **Rental Cars**

Avis ________

Budget ________

Hertz ________

Enterprise ________

d) **Pizza restaurants and delivery**

Domino's ________

Doubledave's Pizzaworks ________

Mr. Gatti's ________

Papa John's ________

Pizza Hut ________

e) **Fast food restaurants**

Burger King ________

KFC (Kentucky Fried Chicken) ________

McDonald's ________

Taco Bell ________

f) **Banks**

Bank of America ________

Bank One ________
First State Bank ________
Nationsbank ________
Wells Fargo (merged with *Norwest*) ________

g) **Pharmacies**

Randall's Food and Pharmacy ________
Eckerd ________
Walgreen's ________