TRESPASSING ON THE RAILROAD

IAN SAVAGE
Department of Economics
Northwestern University
2001 Sheridan Road
Evanston, Illinois 60208, USA
Ph: +1-847-491-8241
Fax: +1-847-491-7001
E-mail: ipsavage@northwestern.edu

ABSTRACT

More than half of the fatal injuries on the railroads in the United States are sustained by trespassers. The paper provides a statistical analysis of the demographics of trespassers, the activities the trespassers were engaged in, and the causes of injury. It also analyzes trends over time and between different parts of the country. The paper finds that the risks of injury and death are particularly acute for males in their 20s and 30s. The total number of annual casualties has remained relatively stable in recent decades because growing affluence, which tends to reduce risk-taking behavior, has been balanced by increases in railroad activity and the size of the population.

INTRODUCTION

In 2004, 480 people died while trespassing on the railroads in the United States. Since 1970, the annual fatality count has fluctuated between 400 and 500. The lack of a sustained improvement is in stark contrast to the considerable reduction in the risks faced by railroad employees and users of rail-highway grade crossings. In 2004, trespassers represented 54% of the 896 fatalities in railroad operations, whereas as recently as the late 1970s the proportion was only 25%.

A landmark year was 1997 when the number of trespasser fatalities exceeded the number killed in collisions at highway-rail grade crossings for the first time since 1941. In the late 1960s, crossing fatalities exceeded trespasser deaths by a ratio of three to one. Now there are 25% fewer crossing fatalities than trespassing deaths. A public outcry in the late 1960s led to a series of programs to improve crossing safety: federal funds were made available to install gates and/or warning lights at a greater proportion of crossings; lighting on the front of trains was upgraded to improve conspicuity; little-used crossings were closed; and a public information campaign called *Operation Lifesaver* was initiated. Taken together, these initiatives were and are remarkable successful in saving lives, and doing so in a cost-effective way (Mok and Savage, 2005; Savage, 2006). In contrast, solutions to the trespassing problem have been far more elusive. With the public policy spotlight shifting in the past decade from grade crossings to trespassing, there is a greater need for the professional community to understand the causes of trespassing and what can be done to reduce the annual casualty count.

DATA

The analysis in this paper concerns mainline railroads. It does not deal with urban mass transit or streetcars, but it does include commuter railroads. Railroads are required to report any death or injury (of any severity) to the Federal Railroad Administration (FRA) on Form FRA F 6180.55a. They have been required to do so since 1910. However, casualty data can be found for as far back as 1890. The data were published from 1901 to 1965 by the Interstate Commerce Commission (ICC), and since 1966 by the FRA. Throughout this paper "trespassers" will be defined as those people trespassing at locations other than grade crossings. (The term is also used to describe persons at highway-rail grade crossings who pass through or around closed crossing gates, but the data are reported separately.)

It is important to recognize that the FRA data exclude fatalities that are judged suicides by a coroner. We will refer to these as "documented suicides." Data on documented suicides, while not reported to the FRA, are reported by coroners to the federal Centers for Disease Control and Prevention (CDC). The CDC reports that there were 91 suicidal deaths in non-motor-vehicle land-based transportation in 2001 and 112 in 2002. In addition, 22 deaths in non-motor-vehicle land-based transportation in 2001, and 27 in 2002, were undetermined as to cause (Arias, 2003; Kochanek, 2004). The Federal Transit Administration (annual) reports that about 35 suicide deaths occurred on urban mass transit rail systems in each of 2001 and 2002.

If the other suicides and those due to undetermined causes occurred on mainline railroads, total trespasser deaths would be about 14% to 19% higher than that reported by the FRA. Documented suicides on railroads are less prevalent in the United States than in Europe. In Britain, documented suicides on the mainline railways are equal in number to the number of trespassing fatalities not deemed a suicide by a coroner (Rail Safety and Standards Board, 2005). Consistent with this, 2.6% of all documented suicides in Britain are by means of mainline trains, while the proportion in the United States is in the order of 0.2%. The explanation for the difference is that Europeans do not have ready access to firearms, which are used in 54% of suicides in the United States (Kochanek et al, 2004). A CDC study (CDC, 1999) of 132 trespasser deaths in the State of Georgia between 1990 and 1996 found that only nine (6.8%) were judged suicides by a coroner, and one was judged a homicide.

HISTORICAL PERSPECTIVE

While the annual fatality count has not changed much in recent decades, the situation is much improved compared with a century ago. The total number of casualties (deaths plus injuries) for each year from 1890 to the present is shown in Figure 1. The graph also distinguishes between deaths and injuries. Of course, some caution is required in comparing data over such a long period. In particular, suicides were included in the data at one time. Comparing 1904 and 2004, one is struck by the fact that fatalities were ten times more numerous (4,908 versus 480) at a time when the country was only a fourth as populous (82 million versus 294 million). The fatality risk per head of population in 1904 was forty times larger than it is today. It is difficult for the modern reader to comprehend the kinds of risks that our forebears took around railroad tracks.

Figure 2 shows the combined number of fatalities and injuries relative to two measures of exposure: population in millions, and line-haul train miles in tens of millions. (The latter measure excludes train miles in yard and switching operations, which have not been reported in a consistent manner over time. It should be noted that the majority of trespassing casualties occur on the main line.) Immediately noticeable from both Figures 1 and 2 is the very substantial decline in risk between 1915 and 1919, and again between 1939 and 1945. While some of this decline is understandable in that the segment of the population most likely to trespass (men in their 20s and 30s) was away in military service, the improvement persisted even after cessation of hostilities.

There is also evidence of a spike in trespassing during the Great Depression of the early 1930s. The image in popular culture of people riding freight trains while looking for work during the dust bowl years is not that inaccurate. In the 1930s, a quarter of the trespasser casualties were described as “hoboes or tramps.” The proportion fell to about 20% in the 1950s and was less than 10% in the early 1970s when the published reports ceased using this description. Similarly, about a quarter of the casualties in the 1930s occurred onboard trains (albeit that this would include injuries sustained aboard passenger trains by people who had not purchased a ticket). This had fallen to about 12% in the 1950s, and today it is relatively rare despite the use of freight trains by illegal immigrants in the Southwest.

Figure 3 is an enlarged version of Figure 2 showing the post-Second World War years, with the rates per head of population and per line-haul train mile shown as indices with 1947 set equal to 100. The rate per head of population declined quite steadily until 1960, when it leveled out at about 35% of the rate in 1947. There was then a substantial decrease of about 25% between 1967 and 1975, followed by more than 20 years of stagnation. There is evidence of a reduced risk in the past five years, but it is too early to tell whether this will be a sustained improvement. The rate per line-haul train mile improved between 1947 and 1955, but then fluctuated around a level that was 20% lower than in 1947 until the mid 1990s. Since the mid 1990s, there has been a considerable improvement of about 30% primarily because the number of train miles has increased but the number of trespasser casualties has not.

The improvement in casualty rates relative to both train miles and population size in recent years has been counterbalanced by the growth in both population and train miles (up by 25% and 60% respectively since the early 1980s). This explains why the absolute count of fatalities has remained relatively constant.

WHO, WHERE, AND WHAT

In Table 1, the 3,628 trespassing deaths and injuries that occurred between 2001 and 2004 are categorized by the event that caused the injury and the activity the trespasser was engaged in at the time. Three-quarters of the casualties occur when the trespasser is struck by a train. A further 10% are associated with slipping, falling or striking a fixed object while on railroad land or on trains. The remainder are a mixture of other events that include assault, exposure to the environment, and cases where the event and activity cannot be determined.

Especially notable is that 30% of all casualties occur when a train struck a trespasser who was sitting or lying down. These data are often cited to support the contention that some of the reported trespasser fatalities are by suicidal people who do not leave notes or other evidence of their intentions. Consequently, coroners are unable to determine whether or not that the fatality was a suicide. Even in the unlikely event that all 30% are "undocumented suicides," the proportion would be far lower than in Europe. Railroad management in Britain investigated the circumstances of all trespasser deaths not deemed a suicide by a coroner, and found a strong suspicion of suicidal intent in 60% of cases (Rail Safety and Standards Board, 2005). Again, the ready access to firearms in the United States seems to reduce the popularity of trains as a method of suicide.

An analysis of casualty age is presented in Table 2. The age distribution of all the casualties occurring between 1999 and 2001 is shown in the middle column. A risk rate is shown in the final column. This is calculated by dividing the average annual casualty count in an age group by total population data from the 2000 Census. There is a popular image in the press that children under the age of 10 are at particular risk. In reality, children in this age group represent only 2.2% of casualties and have casualty rate that is smaller than that for senior citizens. The majority of the risk is faced by people between the ages of 16 and 45 years old. This age group represents 45% of the general population but 75% of the trespasser casualties whose age is known. People in their early twenties are particularly at risk, and face an annual casualty risk of 1 in 150,000.

There have been three major studies that have attempted to shine more light on demographic characteristics. A National Transportation Safety Board (NTSB, 1978) study looked at 280 fatalities that occurred between March 1976 and October 1977. A CDC study (Pelletier, 1997) examined coroners' reports for all of the 138 trespasser deaths in North Carolina for the years 1990-94. Finally, another CDC (1999) study analyzed 132 fatalities and 156 injuries in Georgia between 1990 and 1996. The results of the three studies are almost identical.

More than 90% of victims were found to be adult males, with the vast majority between the ages of 20 and 49. Consequently, the risk rates shown in the final column of Table 2 considerably understate the risk to males (FRA data does not include a gender designation). Eighty percent of the adult victims were unmarried. Pelletier's study found that for those adults whose education was known, only 45% had graduated from high school. Pelletier found that blacks were overrepresented at 38% of the victims whereas they formed only 22% of the population of North Carolina.

Contrary to the popular image of trespassers as "hoboes or tramps," Pelletier found that only 10% of victims were transients, and 80% of deaths occurred within the victim's county of residence. The Georgia study found that 60% were killed in the city in which they resided, suggesting that trespassers die close to home. The trespasser problem appears to be an urban one with less than a quarter of fatalities occurring outside of city or town limits. The NTSB found that nearly all of the fatalities occurred on multiple-track mainlines. In 80% of the cases there was no fence erected to protect the right of way.

Alcohol would appear to be involved in most cases. A disproportionate number of fatalities occur at night on the weekends. Sixty percent of the victims in the NTSB study and 80% in Pelletier's study had been drinking heavily. The Georgia study found that 65% of fatalities tested positive for alcohol or drugs. The average blood-alcohol content was 0.23mg/100mL in the NTSB study, and the median was 0.26 in Pelletier's study and 0.22 in the Georgia study. These are about three times the legal limit for driving, and according to the National Safety Council puts a person in a state of "confusion." Twenty-eight percent of victims in Pelletier's study had previously received medical treatment for alcoholism.

At the risk of generalizing, at least two-thirds of the trespassers can be characterized as single adult males in their 20s and 30s under the influence of alcohol. It would appear that the railroad right of way is an attractive place for people to socialize and imbibe, or to sleep off the effects of alcohol or drugs. Almost a third of the trespasser casualties are sitting or lying in the right of way at the time of impact which clearly indicates considerable negligence on the trespasser's part or suicidal intentions. This would suggest that solutions to this segment of the trespassing problem fall more into the realm of public health rather than changes in railroad operations.

That said, the other third represent people on railroad property for the purposes of vandalism, thrill seeking, catching a ride on freight rain, or taking a short cut over or along the right of way. The railroad is generally unfenced, and it bisects many small towns. Pedestrian are tempted to take a short cut rather than walk to the nearest grade crossing or bridge. In urban areas the temptation to take a short cut leads to the destruction of existing fencing. In rural areas, there is evidence that the railroad right of way is used by hunters, fishermen, and the operators of snowmobiles and all-terrain vehicles. There is also evidence that residents of homes for senior citizens and those with developmental disabilities can become disoriented and wander onto neighboring railroad tracks.

TIME SERIES ANALYSIS 1947-2003

A time series of annual data from 1947 to 2003 is analyzed using two different regression techniques. Both techniques produce similar results. The first type of regression is a log-linear regression on the rate of trespasser casualties per head of population. The Prais and Winston (1954) AR(1) estimator is used to reduce the problems caused by serial correlation. It does so by transforming both the dependent and explanatory variables in the regression by subtracting a proportion, ρ, of the variable's value in the previous period. Hence variables take the form:

$$X_t - \rho X_{t-1}$$

The Prais-Winsten method also ensures that the regression does not "lose" one observation in making this transformation.

The second is a negative binomial regression with the count of casualties as the dependent variable, population as the exposure variable, and the other explanatory variables expressed in logarithms. The negative binomial regression is a more generalized version of the Poisson regression. It assumes that the mean, E(Y), and variance, Var(Y), of the count of casualties for a group of years with identical values of the explanatory variables have the following relationship:

$$Var(Y) = E(Y) + \alpha E(Y)^2$$

While the estimation algorithms of the two equations are very different, the functional form is very similar. The negative binomial equation can be usefully visualized as having the form:

$$\text{count of casualties} = \text{population} * e^{(\alpha + \sum \beta i \ln(Xi))} + \epsilon$$

while the log-linear function is:

$$\ln(\text{count of casualties/population}) = \alpha + \sum \beta_i \ln(X_i) + \epsilon$$

or

$$\text{count of casualties/population} = e^{(\alpha + \sum \beta i \ln(Xi))} + \epsilon$$

Consequently the magnitudes of the estimated coefficients of the explanatory variables in the two equations can be directly compared. Moreover, as all but one of the explanatory variables are expressed in logarithms, the coefficients can be interpreted as elasticities. In both regressions, the dependent variable measures trespassing casualties which is the combination of both fatal and non-fatal injuries. The use of this broader measure of victims is designed to reduce some of the random variations from year-to-year that occurs when the analysis is confined to just fatal injuries.

The regression results are shown in Table 3. The negative binomial regression has an alpha value significantly larger than zero, thereby rejecting the Poisson model. As the estimated values of α are positive, the data are referred to as overdispersed. This model has a pseudo R^2 of 0.25. In the log-linear model, a Durbin-Watson test finds that a Prais-Winsten AR(1) estimator, with a value of ρ of 0.55, removes serial correlation. The adjusted R^2 of the equation is very high. The overall goodness of fit of the equations can be seen in Figure 4 which shows the actual casualty rates per million population (shown as the dots) versus the predictions of the negative binomial (represented by the solid line) and log-linear AR(1) (dashed line) regressions. The two predicted regression lines are very similar, and track the actual data quite well, with the exception of the period between 1960 and 1967 when the actual rate reached a temporary plateau.

In interpreting the results, one should not forget that the size of the population is treated by both regressions as having a direct 1:1 effect on the number of casualties. The population has more than doubled from 143 million in 1947 to 291 million in 2003. Consequently, had nothing else changed, we would expect to have twice as many casualties in 2003 as there were in 1947.

The first explanatory variable is the national railroad route length, known in the industry as "road miles." This was obtained from the ICC's annual statistical publication and later from the Association of American Railroad's *Railroad Facts.* The national network shrunk by 35% between 1947 and 2003, with most of the reduction occurring in the decade between 1974 and 1984. The 1976 *Railroad Revitalization and Regulatory Reform Act*, and the 1980 *Staggers Act* gave railroads more freedom to abandon unremunerative lines. Reduction in the networks should reduce casualties as fewer people live in close proximity to the tracks. The estimated coefficient is very close to unity in both regressions. Indeed one cannot reject a null hypothesis that it is unity. Casualties have been reduced proportionately with the size of the network, with the implication that the parts of the rail network that have been abandoned had a similar trespassing experience to those that remain open.

The second variable is the average number of daily line-haul trains per mile of network. The number of national train miles was obtained from the ICC's annual statistical publication and later from the FRA's annual safety publication. The average number of daily trains is calculated as:

$$\text{Daily Trains} = \frac{\text{Annual Train Miles}}{\text{Route Miles} * \text{Days in Year}}$$

The average daily number of trains fell from 13 a day in 1947 to 8 in 1960. It then fluctuated around this number until 1991. Rail traffic density then started to increase, and by 2003 the number of daily trains had almost returned to its 1947 level. The estimated coefficient is in the range of 0.85 to 0.9, implying that casualties change slightly less than proportionately with rail activity. This is to be expected given that a quarter of all casualties occur due to slips and falls rather than as a result of being struck by a moving train.

The third variable measures the proportion of the population that is between the ages of 15 and 44 (U.S. Census Bureau, 2002). Table 2 indicates that persons in this age range have a disproportionate involvement in trespassing incidents. As the post-World War II "baby boom" generation has aged, the proportion of 15 to 44 year olds has followed a wave pattern. It decreased from 0.463 in 1947 to a low of 0.394 in 1961, increasing to a high of 0.479 in 1986/7, and has subsequently fallen to 0.431 in 2003. This variable is found to have a very strong effect on the number and rate of casualties.

The fourth variable measures the real per capita Gross Domestic Product. GDP data from the Bureau of Economic Analysis were converted to 2003 dollars using the Consumer Price Index, and expressed as a per capita rate. Real GDP per capita has increased from $14,000 in 1947 to $38,000 in 2003. Standard economic theory suggests that citizens demand more lifesaving activities as a country becomes richer. This manifests itself in increased health care expenditures, a demand for more product safety features, and perhaps a reduction in undertaking risky activities such as trespassing on the railroad. Consistent with this theory, the National Safety Council (annual) reports that the rate of non-work-related unintentional deaths in the United States fell from about 55 per 100,000 people in 1947 to 33 per 100,000 in 2003. In addition, increased wealth reduces the prevalence of transients who once hopped trains while traveling to find work. The regressions find that the increase in wealth has the expected negative effect on casualties and casualty rates, with an elasticity close to unity.

The final variable measures a technological change that was found by Mok and Savage (2005) to be particularly effective in improving safety at grade crossings. A 1995 federal rule required increased lighting of trains. The traditional single headlight had to be augmented by two additional lights lower down on the front of the locomotive. These are known as ditch or crossing lights, and provide added illumination of the sides of the track and, what is more important, the triangular pattern provides trespassers with a greater perception of an approaching train's speed and how far it is away. Assuming that locomotives were fitted with these additional lights at a constant rate from the announcement of the rule in September 1995 to the deadline for fitting them in December 1997, the average proportion of locomotives so fitted would be 0.33 in 1996, 0.78 in 1997 and unity from 1998 onwards. Unlike the other variables, this is not expressed in logarithms. The negative binomial regression suggests that installing ditch lights reduced casualties by a statistically significant 13%, whereas the log-linear equation estimates a statistically insignificant 7.5%.

There were two other variables that were tested but were found to be less satisfactory and were dropped from the analysis. The first was the implementation of *Operation Lifesaver* (OL). The public outcry concerning grade crossing safety in the late 1960s led to the formation, starting in Idaho in 1972, of state-based nonprofit organizations to promote education and awareness of railroad-related hazards. While risks at grade crossings have been a primary focus, activities are also directed to trespassing and suicide prevention mainly by means of presentations in schools. The program spread state by state across the nation between 1972 and 1986. A variable was constructed indicating the proportion of the population in a given year who resided in a state in which OL had been established. Unfortunately from an analytical perspective, the growth in OL coincided with the peak period for railroad abandonments. A high correlation between these variables made it impossible to include both in the regressions. Fortunately, an analysis of OL effectiveness is possible for more recent years as data is now available on the magnitude of OL activity in various states. The results of this latter analysis are discussed later in the paper.

With an eye to examining possible trends in the portion of reported casualties who are undocumented suicides, a variable was constructed that measures the national rate of total documented suicides per million population (National Center for Health Statistics, annual). The rate has fluctuated over the years between 97 and 131 per million population. The rate was particularly low between 1951 and 1961, and in the period since 1999. It was particularly high between 1975 and 1978, and again between 1984 and 1988. The correlation between the national suicide rate and the trespassing casualty rate is -0.36, implying a counterintuitive negative relationship. This may also lend support to the notion that undocumented suicides are less of a factor in reported trespasser fatalities in the United States than they are in Europe.

DECOMPOSITION OF TIME-SERIES TRENDS

From 1947 to 2003 the number of annual trespassing casualties fell by 1,594 from 2,490 to 896. The regressions can be used to estimate the contribution of the various causes to the decline. Using the negative binomial regression results, the change in the predicted number of casualties from year t to year t+1 can be decomposed to the following:

$$\begin{aligned}\text{Casualties}_{t+1} - \text{Casualties}_t = {} & \left[\text{Population}_{t+1} - \text{Population}_t\right] e^{a} e^{b_1 \ln(\text{RoadMiles}_t)} \\ & e^{b_2 \ln(\text{Trains}_t)} e^{b_3 \ln(\text{Age15-44}_t)} e^{b_4 \ln(\text{Wealth}_t)} e^{b_5 \text{Ditch Lights}_t} \\ & + \text{Population}_t \; e^{a} \left[e^{b_1 \ln(\text{RoadMiles}_t)} - e^{b_1 \ln(\text{RoadMiles}_{t-1})} \right] e^{b_2 \ln(\text{Trains}_t)} \\ & e^{b_3 \ln(\text{Age15-44}_t)} e^{b_4 \ln(\text{Wealth}_t)} e^{b_5 \text{Ditch Lights}_t} + \ldots + \varepsilon_t - \varepsilon_{t-1}\end{aligned}$$

The equation will also include (in place of the ellipses) similar terms to the first two that involving changes from period t to t+1 for the variables Trains, Age 15-44, Wealth and Ditch Lights. In addition there will be cross-product terms involving every possible combination of the value of variables in period t and changes in variables. There will be 63 terms in total. Of course, most of the cross-product terms will be quite small as they involve the product of two (or more) relatively small changes in the constituent variables. In addition some of the cross-product terms will be positive and some negative, and will tend to cancel each other out.

This decomposition was carried out for each of the annual changes from 1947 to 1948 through 2002 to 2003. The resulting decompositions are shown in Table 4. Over the entire period, the regression predicts that increases in population should have increased casualties by 900 a year. However, this was counteracted by abandonment of parts of the network (reducing casualties by 498), reductions in the number of average daily trains (369), changes in the age distribution of the population (232), installation of ditch lights (136), and increases in wealth which promote less risk-taking behavior (1,133). The cross product terms produce a further reduction of 62 casualties a year. The sum of the error terms, which is to say the changes not explained by the regression, total a net decrease of 68 casualties. As inspection of Figures 4 would suggest, the explanation for the latter is that the actual number of casualties in 2003 was somewhat below the predicted value, whereas in 1947 the actual and predicted numbers are much closer.

Table 4 also breaks down the decomposition into four sub-periods (1947-1960, 1960-1974, 1974-1988 and 1988-2003). The break points were chosen because of observed changes in the trend of one, or more, of the explanatory variables. In interpreting the table, there are a number of notable features. Increases in population and wealth have affected casualties in all four time periods. However, the numerical size of both the population and wealth effects were much larger in the earlier time periods because trespassing risks (as explained by the other variables) were much higher fifty years ago. The reduction in railroad network size had its primary effect in the 1974-1988 period. The aging of the baby boom generation is reflected in the unusually high proportion of people between the ages of 15 and 44 in the period from 1960 to 1988. The reduction in the average number of daily trains in the 1950s has a particularly large negative effect. The reverse was true when train traffic density started to increase rapidly after 1988. In the past fifteen years, the increase in train traffic, coupled with an increase in population, has counteracted other factors that would tend to reduce trespassing. This explains the apparent lack of improvement in casualty numbers in recent times.

PANEL DATA ANALYSIS 1996-2003

A second analysis has the principal objective of investigating the effectiveness of Operation Lifesaver in discouraging trespassing. The analysis makes use of the considerable variation in observed OL activity between states. OL is inherently a state-based organization, and each state has considerable independence in designing and implementing its programs. Indeed, a national office was only established in the late 1980s, and uniform reporting of activities dates from only 1996. This analysis makes use of the OL data to form a panel data set for 45 states for each of the years 1996 to 2003.

Three negative-binomial regressions are conducted. Two analyze total trespasser casualties, while the final one focuses on casualties to child trespassers under the age of 16. Data for the dependent variables are obtained from the FRA's web-based interactive database, which gives a breakdown of casualties by age and by state.

The exposure variable is the annual midyear (July 1) state population count produced by the Population Estimates Program of the U.S. Census Bureau. In the analysis of child trespassers, the count is for children under the age of 15. Note that this is slightly different from the age break used by the FRA, who use a break point of 16. The implication is that a state with twice the population of another state should, all else being equal, have twice as many trespassing casualties.

Other variables mirror those in the time-series analysis. The size of the railroad network in each state has been published annually by the Association of American Railroads (AAR) since 2001. The AAR provided comparable data for earlier years to the author. The correlation between road miles and population is only about 0.6, much lower than one might expect. While road mileage generally increases with population, there are several large yet sparsely populated states with expansive railroad mileage, and some highly populated states on the Eastern Seaboard with comparatively limited mileage.

Disaggregated data on train miles are not available by state. Therefore, one cannot directly calculate average daily trains by year and state. However, a point estimate of the state-by-state distribution of train frequency can be obtained from the FRA's grade crossing inventory database. For each crossing, an estimate is given for the number of daily trains. The most current inventory file for public at-grade crossings was downloaded from the FRA's web site, and the average number of daily trains was calculated for each state. A "state correction factor" was derived by comparing the state average to the national average. This factor varies from 1.72 in Nebraska (72% above the national average) to 0.21 in South Dakota (79% below the national average). The variable representing the average number of daily trains for state i in time period t is calculated by multiply the national annual estimate of average daily trains, used in the time series analysis, by the state correction factor. The average number of trains per day varies from 20 a day in Illinois and Nebraska down to less than four a day in Rhode Island, Maine, New Hampshire and South Dakota.

The Census Bureau's midyear Population Estimates Program data are used to calculate the proportion of the population between the ages of 15 and 44 in each state and for each year. The ratio varies from greater than 0.47 in Alaska, Georgia and Utah down to 0.43 in Pennsylvania, West Virginia, Montana and Arkansas, and 0.41 in Florida. This variable is, of course, not used in the equation analyzing child casualties.

Information on personal income per capita by state is obtained from the annual Census Bureau's *Statistical Abstract of the United States*, and converted to 2003 dollars using the Consumer Price Index. Income varies widely by state from greater than $38,000 in Massachusetts, New Jersey and Connecticut, to less than $24,000 in New Mexico, Arkansas, West Virginia and Mississippi.

Based on the calculations in the time-series analysis, the proportion of locomotives fitted with ditch lights is 0.33 in 1996, 0.78 in 1997 and unity from 1998 onwards. It is not possible to determine whether the rate of installation varied by state.

Descriptive statistics suggest that Arizona and (especially) New Mexico have an unusually high rate of trespassing casualties. Anecdotal evidence suggests that this is partly due to illegal immigrants from Central America. To test this theory, a variable of one minus the proportion of a state's boundaries that border Mexico is constructed. Data on Mexican border length are obtained from Van Zandt (1976), sea coast length from the U.S. Geological Survey, and the boundaries between states from Holmes (1998). The non-Mexican proportion is 0.606 for Texas, 0.737 for Arizona, 0.878 for New Mexico, 0.93 for California and unity for all other states.

In the regressions for all trespassing casualties, the measure of activity by OL is the number of presentations and special training events expressed as a rate per capita. This varies from more than 150 per 100,000 population in Nebraska, Wyoming, North Dakota, Idaho and Montana down to less than 25 per 100,000 population in Delaware, Pennsylvania, New Jersey, New Hampshire, Florida and Michigan. This is felt to be the most reliable and consistently reported measure of activity. Presentations make up 95% of the total. Special training events are also included because it is not clear whether certain activities such as talks to truck drivers, school bus drivers and emergency responders are consistently classified in one category or the other. State coordinators are also required to report the number of people in attendance at these events. Theoretically, this would be a preferable measure. However, in practice there are some anomalies that make the use of this measure questionable. This is particularly the case with presentations to adult groups, such as those conducted at state and county fairs. OL has other activities beside presentations, such as placing public service announcements. Casual reading of the state reports suggests that states that are very active in making presentations are also very active in other activities. So the variable used should be seen as a proxy for the total level of activity.

The number of attendees at presentations in schools is reported much more reliably and consistently. Therefore in the regression for child casualties, the OL variable is the number of attendees at presentations at kindergarten to eighth grade schools (serving children from ages 5 to 13) divided by the state population of under 15s. A high proportion is children attend presentations each year in Rhode Island (10%), Montana (6%), North Dakota (5%) and Illinois (4%). However, in states such as Michigan, North Carolina, Tennessee, Alaska and Nevada less than half of one percent attend presentations each year.

The theoretical maximum number of observations is 392, representing 49 states for each of the eight years from 1996 to 2003 (Hawaii has no railroad service, and the District of Columbia did not establish an OL program until 2002). Unfortunately, some of the observations had to be dropped either because the state OL organization did not submit a report to the national office for a given year, or because, after consultation with the national office of OL, it was felt that the reported data were incomplete or considered unreliable. Generally, volunteer coordinators are reporting on the activities of volunteer presenters, and in some cases the reporting is not very good. Consequently, four states were excluded in their entirety (Arizona, Maryland, Massachusetts, and Virginia), data on 13 annual observations involving 10 states were missing, and 17 years involving nine different states were dropped because the data were questionable or were incomplete. The total usable sample size was therefore 330 out of a possible 392 observations for the regression on total trespassing casualties. For the regression on child casualties, a further six observations had to be dropped because the state coordinator did not report attendance, and one observation (New Hampshire in 1998) was dropped as zero presentations were made, and this variable, as with all the other variables, is expressed in logarithms.

The regression results are shown in Table 5. The first two regressions analyze total trespassing casualties. The first regression does not include the OL variable, whereas the second one does. The final regression is an analysis of child casualties. Overall the explanatory power is very low. Casualties are rather rare and there is considerable random variation both geographically and temporally.

Certain findings from the time-series analysis carry over to the cross-sectional analysis. Increased per capita income is a strong predictor of reduced trespassing risk. The coefficients are similar to those in the time-series analysis, implying an elasticity of -1.0 between income and casualties. This effect also carries over to the analysis of children, suggesting that the income of the household in which a child lives influences the proclivity of the child to trespass.

Increasing railroad road miles is estimated to increase child casualties with an elasticity of 0.45, and total casualties with an elasticity of 0.7. The fact that the elasticity is less than unity is to be expected given that states with the largest amounts of railroad mileage are characterized by operation in rural and mountainous areas with low population densities.

The elasticity of the average daily number of trains with respect to casualties of all ages is 0.14. This is a much lower ratio than was found in the time-series analysis. The elasticity with respect to child casualties was -0.45, providing the counterintuitive result that an increase in average daily trains leads to a decrease in child casualties. Part of the explanation is that a slightly higher proportion of child casualties result for injuries that are sustained from causes other than being struck by a train compared with trespassers in general (30% versus 23%). It may also be the case that children who live in states with frequent train service are more aware of the dangers on the tracks and are less likely to trespass.

The fitting of ditch lights is found to reduce casualties. However, excepting the second model (for all trespassers and including an Operation Lifesaver variable) where it is statistically significant at the 10% level, the evidence is weak.

States bordering Mexico do have a higher incidence of casualties. The effect is quite strong, suggesting that California has a 9% higher casualty count than would be expected given the other variables, New Mexico 17%, Arizona 44% and Texas 82%. However, this relationship is not found for child trespassers, where the opposite result was found. This is not surprising given that most illegal immigrants are adult males searching for work.

A surprising result is that the proportion of the population in a state between the ages of 15 and 44 is, weakly, negatively related to the casualty count. This runs counter to the clear evidence in Table 2 that the risks are much higher for people in this age range. The probable explanation for this counterintuitive result is that there is not much variability in the data between most of the states and over the relatively short period of time considered in this analysis.

The estimated regressions find that a higher level of OL activity is positively related to casualties for both children and for persons of all ages. One would hope that this does not imply that OL activities encourage rather than discourage trespassing. A more likely explanation is that states in which trespassing is prevalent have a greater incentive to promote an active OL program (examples include Arkansas, Illinois, Montana, and North Dakota). Endogeneity of this variable will tend to obscure any analysis of whether increasing OL activity leads to a reduction in trespassing casualties. (Because of the possible endogeneity, Table 5 presents the regression for all trespassers both with and without the OL variable.)

Fortunately, it is possible to circumvent the problem of endogeneity by looking at the considerable year-to-year fluctuations in OL activity in certain states. There are myriad exogenous reasons for these year-to-year variations: the volunteer coordinators change and are replaced by either energetic new people or less-organized people, prolific presenters retire, school districts change their policies regarding presentations in schools, and railroads change their policies regarding allowing employees to make presentations during working hours. There can be as much variability over time in individual states as there is cross-sectional variation between states, and the variable is much more likely to result from exogenous factors unrelated to trends in casualties.

Table 6 shows data for six large states where there were dramatic changes in activity. The analysis focuses on child trespassing because we can be certain that the OL presentations given to kindergarten to eighth grade schoolchildren will be primarily concerned with the risk of trespassing, as opposed to the risks associated with grade crossings which historically have been the major focus of OL. For each of the six states, the table shows annual data from 1996 to 2003 on the number of children in attendance at OL presentations and the number of trespassing casualties under the age of 16.

The periods of interest are shaded. There was a considerable spike in activity in Illinois in 1999 and 2000, in Michigan in 2002 and 2003, and in Texas from 1998 to 2000. In California activity declined over time starting in 1998. There were also declines in activity in Ohio after 2000, in Pennsylvania after 1998, and in Wisconsin after 1999.

The first point to note is that even in large states such as California, the number of annual child casualties can be counted on the fingers of two hands. As a result, considerable year-to-year fluctuations limit the ability to make strong statistical statements. That said, it is tough to observe any relationship between OL activity and the number of child casualties, with the possible exception of the experience in Texas.

CONCLUSIONS

Trespassing on the railroad is a problem that does not seem to be going away. There is a strong epidemiological aspect to it, affecting single males in their 20s and 30s. Consequently the amount of trespassing will be directly related to the demographic trends affecting the size of the population in the target group. The good news for the railroads is that the baby boom generation has aged, lowering the proportion of the population in this high-risk group. The bad news is that the total size of the population is increasing, meaning that the absolute number of 15-44 year olds is still rising.

Evidence from other studies indicates that many victims are under the influence of alcohol, have a history of alcohol abuse, and a low level of educational attainment. Mainstream activities such as public service announcements and Operation Lifesaver presentations may be an ineffective way of communicating the dangers to this segment of the community.

Finally, the good news is that the nation's increasing affluence has reduced the propensity for people to expose themselves to the risks of trespassing on the railroad. However, increases in train traffic and the size of the population conspire to ensure that the annual toll of injuries and fatalities has remained in the 400 to 500 range.

REFERENCES

Arias, E., R.N. Anderson, K. Hsiang-Ching, S.L. Murphy, and K.D. Kochanek (2003). Deaths: Final data for 2001. *National Vital Statistics Reports*, vol. 52(3). Hyattsville, Maryland: National Center for Health Statistics.

Association of American Railroad (annual). *Railroad Facts*. Washington, D.C.: Association of American Railroads.

Centers for Disease Control and Prevention (1999). Injuries among railroad trespassers in Georgia, 1990–1996. *Mortality and Morbidity Weekly Report*, vol. 48(25), pp. 537-541.

Federal Railroad Administration (annual). *Accident/Incident Bulletin*. Washington, D.C.: FRA, U.S. Department of Transportation. (Prior to 1966 was published by the Interstate Commerce Commission under the title *Accident Bulletin*. In 1997 it was renamed *Railroad Safety Statistics*)

Federal Transit Administration (annual). *Transit Safety and Security Statistics and Analysis*. Washington, D.C.: FTA, U.S. Department of Transportation.

Holmes, T.J. (1998). The effects of state policies on the location of industry: Evidence from state borders. *Journal of Political Economy*, vol. 106(4), pp. 667-705. The state boundary data can be found at: http://www.econ.umn.edu/~holmes/data/BorderData.html.

Interstate Commerce Commission (annual to 1994). *Annual Report on the Statistics of Railways in the United States*. Washington, D.C.: Interstate Commerce Commission. (From 1954 renamed *Transport Statistics in the United States*.)

Kochanek, K.D., S.L. Murphy, R.N. Anderson, and C. Scott (2004). Deaths: Final data for 2002. *National Vital Statistics Reports*, vol. 53(5). Hyattsville, Maryland: National Center for Health Statistics.

Mok, Shannon, and Ian Savage (2005). Why has safety improved at rail-highway grade crossings? *Risk Analysis*, vol. 25(4), pp. 867-881.

National Center for Health Statistics (annual). *Vital Statistics of the United States.* Hyattsville, Maryland: National Center for Health Statistics. (Previously published by the Census Bureau.)

National Safety Council (annual). *Injury Facts*. Itasca, Illinois: National Safety Council. (Prior to 1999 this was known as *Accident Facts*).

National Transportation Safety Board. (1978). *Safety Recommendation R-78-42.* Washington D.C.: National Transportation Safety Board.

Pelletier, Andrew (1997). Deaths among railroad trespassers: the role of alcohol in fatal injuries. *Journal of the American Medical Association*, vol. 277(13), pp. 1064-1066.

Prais, S.J., & C.B. Winsten (1954). *Trend Estimators and Serial Correlation*. Chicago: Cowles Commission Discussion Paper 383.

Rail Safety and Standards Board (2005). *Annual Safety Performance Report 2004*. London: Rail Safety and Standards Board.

Savage, Ian (2006). Does public education improve rail-highway crossing safety? Forthcoming in *Accident Analysis and Prevention*.

U.S. Bureau of the Census (annual). *Statistical Abstract of the United States*. Washington D.C.: U.S. Census Bureau.

U.S. Bureau of the Census (2002). *Census 2000 Special Reports, Series CNSR-4; Demographic Trends in the 20th Century*. Washington D.C.: U.S. Census Bureau.

Van Zandt, F.K., (1976). *Boundaries of the United States and the Several States.* Washington, D.C.: U.S. Geological Survey Professional Paper 909.

Table 1: Proportion of 3,628 Trespassing Casualties by Event and Activity 2001-2004

Activity	Event			**Total**
	Struck by On-track Equipment	Slips, Falls, Electric Shock, Crushed, Striking Object	Other	
Walking or Running	29.8%	1.7%	0.5%	**31.9%**
Standing, Bending or Stooping	7.5%	0%	0%	**7.5%**
Sitting	7.4%	0%	0%	**7.4%**
Laying, Lying Down, Sleeping	22.9%	0%	0.9%	**23.7%**
Jumping, Climbing, Crawling, Boarding	5.2%	3.5%	0%	**8.7%**
Driving or Riding or Operating (bicycle, snowmobile etc.)	3.5%	4.8%	0%	**8.3%**
Other activities	0.7%	0%	11.9%	**12.5%**
Total	**76.8%**	**10.0%**	**13.2%**	**100%**

Source: FRA downloadable database (note that data are rounded, so columns and rows may not add up exactly).

Table 2: Distribution of Trespasser Casualties and Casualty Rates by Age

Age Range (years)	% of Total Trespasser Casualties 1999-2001	Annual Rate per million population
0-5	0.7%	0.27
6-10	1.5%	0.67
11-15	5.2%	2.33
16 - 20	10.6%	4.74
21 - 25	14.0%	6.77
26 - 30	11.0%	5.00
31 - 35	9.4%	4.08
36 - 40	11.5%	4.54
41 - 45	9.5%	3.90
46 - 50	6.6%	3.05
51 - 55	3.9%	2.15
56 - 60	1.5%	1.03
61 - 65	1.2%	1.01
66 - 70	1.0%	1.00
71 - 75	0.8%	0.81
76 - 80	0.8%	1.00
>= 81	0.9%	1.00
Not Given	10.0%	

Sources: Trespassing casualties from downloadable FRA database for 1999-2001. Rate calculated as the average annual count 1999-2001 divided by population data from the 2000 U.S. Census.

Table 3: Time Series Regression Results

	Negative Binomial		Prais-Winsten AR(1)	
	Coefficient	t	Coefficient	t
Dependent Variable	Count of Trespassing Casualties		Trespassing Casualties per Head of Population	
Constant	-15.6171	3.78	-13.7662	2.29
United States Population	Exposure		part of Dependent Variable	
Log of Railroad Road Miles	1.0070	4.08	0.9364	2.51
Log of Average Daily Number of Trains	0.9180	8.55	0.8547	6.33
Log of Proportion of Population between Age 15 and 44.	1.1888	4.24	1.2904	3.01
Log of Real Gross Domestic Product per Capita	-0.9633	8.08	-1.0408	6.12
Proportion of Locomotives with Ditch Lights	-0.1360	2.02	-0.0780	0.85
alpha	0.0041	4.35		
Observations	57		57	
Constant-only Log Likelihood	-442.26			
Log Likelihood	-331.37			
rho			0.5555	
Original Durbin-Watson Statistic			0.9318	
Transformed Durbin-Watson Statistic			2.0157	
Pseudo R^2 / Adjusted R^2	0.2508		0.9852	

Table 4: Decomposition of Change in Annual Totals

Totals may not add due to rounding	Sub-Periods				Overall 1947-2003
	1947-1960	1960-1974	1974-1988	1988-2003	
Actual Annual Totals					
Start of period	2,490	1,088	1,004	1,010	**2,490**
End of Period	1,088	1,004	1,010	896	**896**
Change	-1,402	-84	+6	-114	**-1,594**
Changes Explained by Regression					
Increased Population	+415	+184	+135	+170	**+904**
Decreased Road Miles	-68	-84	-218	-128	**-498**
Changes in Average Daily Trains	-806	-1	+65	+374	**-369**
Changes in Proportion of Population between 15 and 44	-338	+110	+109	-114	**-232**
Increased per Capita Real Gross Domestic Product	-436	-373	-181	-142	**-1,133**
Locomotives with Ditch Lights	0	0	0	-136	**-136**
Sum of Cross Product Terms	-15	-9	-27	-11	**-62**
Changes Not Explained by Regression					
	-154	+89	+123	-127	**-68**

Table 5: Panel Regression Results

Dependent Variable	Count of Trespassing Casualties		Count of Trespassing Casualties		Count of Under 16 Trespassing Casualties	
	Coeff.	t	Coeff.	t	Coeff.	t
Constant	-4.2713	2.31	-5.7730	2.70	-2.7169	0.53
State Population	Exposure		Exposure			
State Under 15 Population					Exposure	
Log of Railroad Road Miles	0.0649	1.71	0.0781	1.77	0.4525	3.27
Log of Average Daily Number of Trains	0.1734	2.29	0.1358	1.71	-0.4793	2.38
Log of Proportion of Population between Age 15 and 44.	-1.1274	1.71	-1.3184	1.83	not applicable	
Log of Real Personal Income per Capita	-0.9950	6.38	-0.9099	4.88	-1.2497	2.52
Log of Proportion of Locomotives with Ditch Lights	-0.0238	0.36	-0.1368	1.70	-0.0173	0.09
Log of Proportion of State Boundary not Bordering Mexico	-1.1330	5.19	-1.2934	4.68	1.2053	2.09
Log of Operation Lifesaver Presentations and Special Training per Capita			0.1089	3.25		
Log of Operation Lifesaver K-8 Attendees per Under 15 Population					0.1155	1.87
alpha	0.0903	7.51	0.0985	6.86	0.2300	3.24
Observations	392		330		323	
Constant-only Log Likelihood	-1248.04		-1050.02		-448.71	
Log Likelihood	-1190.99		-1003.29		-437.56	
Pseudo R^2	0.0457		0.0445		0.0249	

Table 6: Analysis of Exogenous Changes in Operation Lifesaver (OL) Activity

	California		Illinois		Michigan		Ohio	
Year	OL K-8 Audience	0-15 Casualties	OL K-8 Audience	0-15 Casualties	OL K-8 Audience	0-15 Casualties	OL K-8 Audience	0-15 Casualties
1996	205,854	9	#N/A	7	3,500	3	35,156	4
1997	221,207	5	43,522	4	3,500	5	50,332	4
1998	146,489	6	55,510	2	#N/A	1	55,933	4
1999	127,715	5	337,238	6	3,000	1	31,184	2
2000	82,060	5	314,615	8	3,500	4	53,935	5
2001	73,567	4	47,902	6	3,700	2	20,120	3
2002	49,891	3	31,139	4	25,821	1	20,882	0
2003	52,178	2	24,818	4	20,179	1	12,816	1

	Pennsylvania		Texas		Wisconsin	
Year	OL K-8 Audience	0-15 Casualties	OL K-8 Audience	0-15 Casualties	OL K-8 Audience	0-15 Casualties
1996	#N/A	3	5,878	9	9,294	2
1997	18,879	14	23,966	4	19,303	3
1998	30,241	9	50,264	4	15,232	6
1999	17,613	1	50,213	1	14,522	0
2000	7,161	5	42,690	5	5,471	3
2001	2,455	4	7,762	5	3,422	1
2002	3,210	4	9,002	9	4,876	1
2003	6,633	3	18,574	8	1,946	2

Figure 1: Annual Casualties by Type

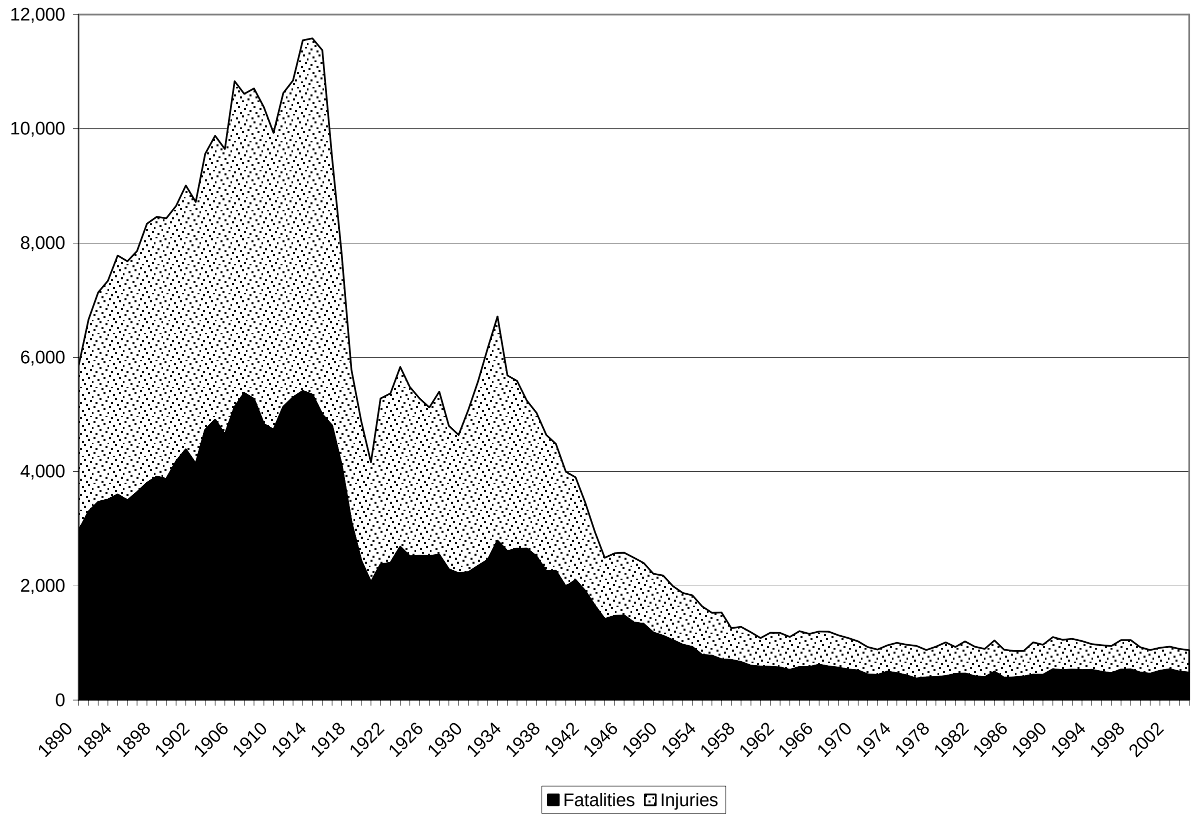

Figure 2: Casualty Rates

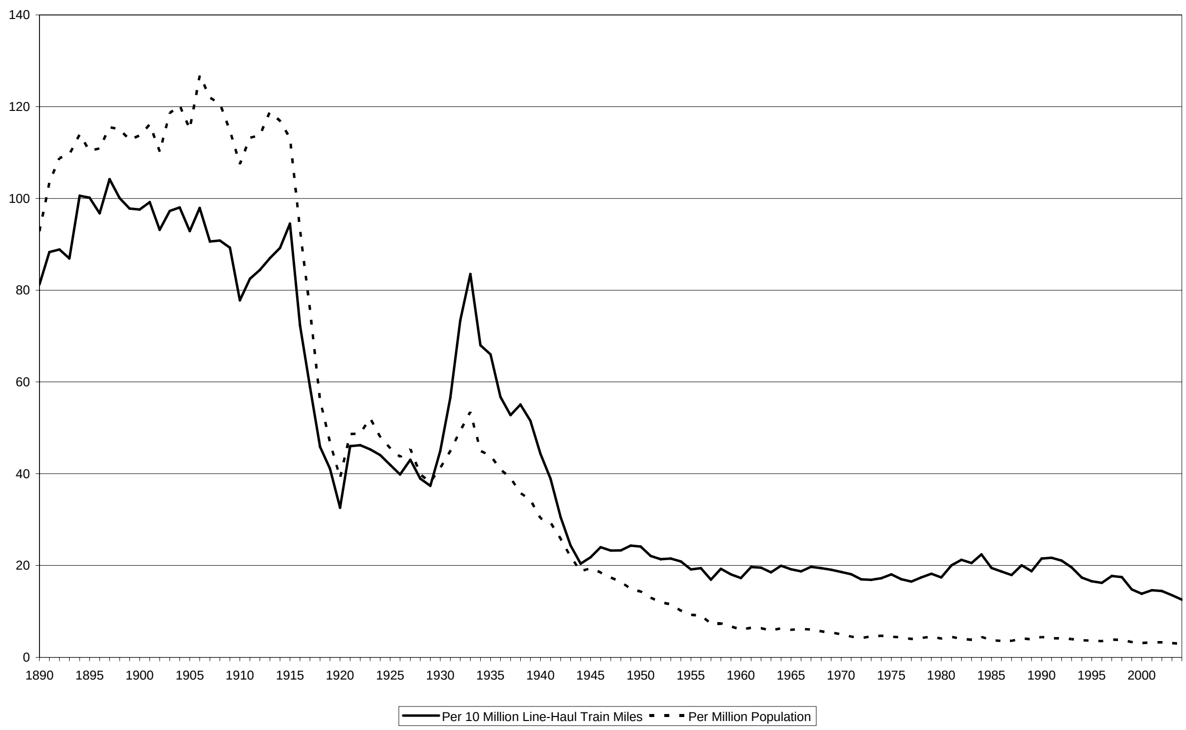

Figure 3: Index of Casualty Rates Since 1947 (1947=100)

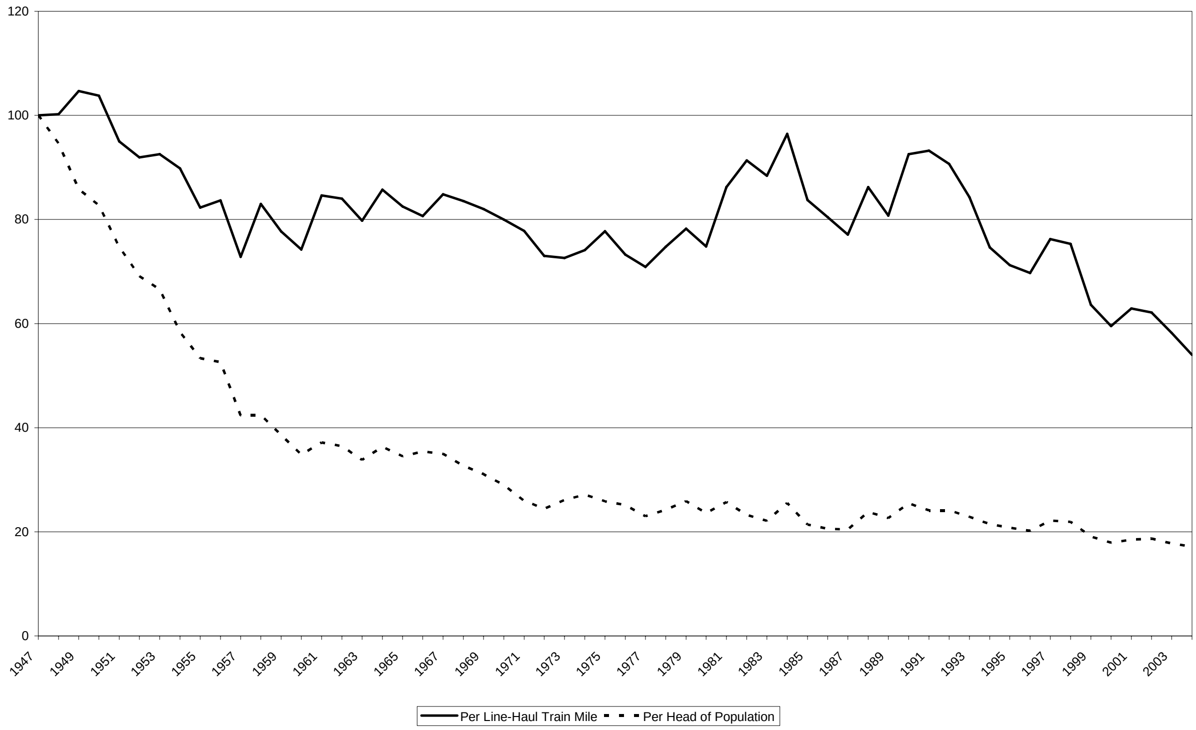

Figure 4: Actual versus Predicted Casualty Rates per Million Population