Exploring the Role of Pseudodeductibles in Auto Insurance Claims Reporting

Dana A. Kerr[1]

Abstract: Many who purchase insurance understand that by reporting covered losses to their insurers they increase the chances of future premium increases or reductions in insurance coverage. If under such circumstances a policyholder decides not to report an otherwise covered loss, the policyholder is effectively displaying the presence of a "pseudodeductible." That is, given a covered loss occurs, a policyholder may define a personal and unobservable threshold that is greater than any stated deductible in the policy below which an insurance claim for the loss will not be reported. The scant amount of empirical research on this topic suffers from a lack of information about losses for which insurance claims were never filed. This research relies on a unique dataset that captures when policyholders choose to forgo insurance claims and why. The findings increase our understanding of the role that pseudodeductibles play in the claims reporting behavior of policyholders. [Key words: deductibles, claims, auto insurance]

INTRODUCTION

In the insurance economics literature, Venezia and Levy (1980) analyze the anecdotally familiar but poorly understood characteristic of the claims filing decision where a policyholder chooses not to report an otherwise covered loss because of the future uncertainties the reporting of the claim creates. They recognize that an individual's optimal insurance claiming strategy is determined by a balance between the benefits of indemnification for loss and the additional costs of paying higher future insurance premiums resulting from the bonus-malus pricing system that exists in the

[1]Dana A. Kerr, PhD, CPCU, ARM is an Assistant Professor of Risk Management and Insurance in the School of Business at the University of Southern Maine. University of Southern Maine, P.O. Box 9300, Portland, ME 04104-9300, Phone: (207) 780-4059, Fax: (207) 780-4662, dkerr@usm.maine.edu

Journal of Insurance Issues, 2012, **35** (1): 44–72.

private insurance market. In this context, the loss suffered must exceed some threshold value before an insurance claim will be pursued. Venezia and Levy develop an economic model based on a multi-period utility function to determine the critical value that a loss must exceed in order for a risk averse decision maker to file an insurance claim despite the related future premium increase.[2] Their theoretical model suggests that the critical threshold loss value defining an individual's optimal claims strategy will decrease with age and increase as risk increases and as a decision maker's discount factor increases.

Venezia (1984) modifies Venezia and Levy (1980) by specifically considering the optimal auto insurance claims strategy given the selection of a deductible by a risk averse driver. The author finds that age does not necessarily influence the critical threshold loss value defining when an insurance claim is (not) filed when the simultaneity of the deductible and claims filing decisions is modeled. Interestingly, Venezia shows how theoretically a claim will be filed notsimply when a given loss amount exceeds the stated deductible, but only when the loss amount is larger than the deductible by some critical amount that accounts for the larger future premium associated with filing a claim for the loss.

Unfortunately, there has been little empirical research conducted to support the theoretical models of Venezia and Levy (1980) and Venezia (1984). A particular problem in researching this behavioral aspect of insurance claims reporting is the dearth of information concerning policyholder losses that occur but for which no insurance claim is ever made. Insurance companies and researchers relying on only the observable elements in closed claims data, for example, are not privy to complete information about the insurance claims decision process. Depending on the nature and design of the study, this lack of information about the full spectrum of losses, both reported and unreported, can create a bias.

Braun, Fader, Bradlow, and Kunreuther (2006) study this claims filing behavior by investigating the pseudodeductible, the term they coin to describe the unstated (and unobservable) threshold amount below which a policyholder may choose to not file an insurance claim despite the fact that the loss amount exceeds any stated deductible in the policy. They use the pseudodeductible concept to explain the curious policyholder decision to "leave money on the table" by not reporting an otherwise covered loss as an insurance claim due to reasons such as the fear of a future premium increase or the loss of coverage. Braun, et al. believe their research on this

[2]Venezia and Levy (1980) do not specifically account for deductibles. As they state, "For ease of exposition we assume that claims are met without deductions" (p. 547).

topic to be the first to attempt to empirically study the behavior prompted by the pseudodeductible.

There are numerous empirical studies that investigate many other aspects of the insurance claims filing decision, but none besides Braun, et al. (2006) have considered the presence of a pseudodeductible and many suffer from the lack of information concerning losses never reported as insurance claims. This information is rarely captured explicitly in the data. For example, much of the prior research studying the insurance claims filing decision by policyholders depends upon proxies to estimate claims filing rates or propensities to file claims.[3] One notable exception is Biddle and Roberts (2003), in which the authors use data about workplace injuries both reported and not reported as workers' compensation claims. However, the authors do not attempt to isolate the effects of any pseudodeductible in the claims reporting decision.

This paper adds to our understanding of insurance pseudodeductibles by empirically examining the decisions made by insurance consumers to forgo reporting losses for reasons that may indicate the presence of pseudodeductibles. The unique opportunity to analyze a dataset that allows us to observe policyholders who have suffered covered losses and yet who have chosen not to file insurance claims is strong motivation for this study. It is an attempt to better understand the extent to which behavioral and psychological factors influence the insurance claims filing process.

The next section of this paper reviews the prior literature concerning behavioral theories that explain the individual decision-making process. A more detailed summary of the pseudodeductible identified by Braun, et al. (2006) and a discussion of how the present paper adds to the Braun, et al. study are then provided. The data used to empirically test the selectivity of policyholders as indicated by their claims filing decisions with respect to a pseudodeductible are described next. The research hypothesis will then be stated and the regression model used will be specified in greater detail. The last two sections will explain the empirical results and offer concluding remarks.

[3]Two such examples are Butler and Worrall (1983) and Cummins and Tennyson (1996). Butler and Worrall model workers' compensation claims filing rates by looking at the number of injury claims filed as a percentage of the number (in thousands) of employees, not as a percentage of the number who have actually experienced a workplace accident. Cummins and Tennyson model the likelihood of filing a third-party bodily injury auto claim by dividing the number of these reported claims by the number of third-party property damage claims, with the latter representing the complete pool of possible bodily injury claims that could have been reported. There is possible measurement error using this proxy in that some of the property damage claimants likely did not suffer bodily injuries.

Review of Individual Decision-Making Literature

With respect to insurance consumers making choices that are prima facie economically irrational, Johnson, Hershey, Meszaros, and Kunreuther (1993) attribute such decision error to possible distortions in how a policyholder evaluates the risk itself, the policy premium, or the benefit paid from the policy. For example, framing can affect the way benefits from an insurance contract are perceived. This will result from what Kahneman and Tversky (1979) call loss aversion, where the negative consequences from a loss are perceived to be greater than the positive consequences associated with a gain of the same size. When alternative decision options are compared against a status quo or default option, the potential losses associated with choosing an alternative different from the status quo may be more heavily weighted in the decision process than the possible gain of making the alternative choice. The result is a specific type of status quo bias that Thaler (1980) calls the endowment effect.[4]

The term status quo bias was first introduced by Samuelson and Zeckhauser (1988), who describe it is a more general classification of the decision biases that are related to but not dependent upon loss aversion, framing, and the endowment effect described above. They suggest that preference for the status quo may be economically rational to the extent there are direct and indirect transaction costs, which they refer to as transition costs, or if there is a measure of uncertainty in the alternative decision choices. Besides the presence of transaction costs that might discourage insurance claiming, policyholders face a new uncertainty with respect to the claims filing decision. What will be the future impact associated with reporting the loss to the insurance company? The post-loss status quo is the insurance policy yet to be affected by the loss. The greater the uncertainty surrounding this potential impact the more likely the policyholder will maintain the status quo by choosing not to report the loss as an insurance claim.

Saumuelson and Zeckhauser (1988) also describe how regret avoidance creates a status quo bias, and it is conceivable that minimizing future regret affects claims filing behavior as well. Kahneman and Tversky (1982) believe that individuals suffer greater regret for the negative consequences of actions taken than the regret they feel for similar negative consequences resulting from no action at all (as cited in Samuelson and Zeckhauser, 1988).

[4]Thaler's (1980) endowment effect describes an individual decision maker who tends to misperceive the gains and losses associated with an alternative choice relative to the status quo. The potential losses associated with an alternative choice will tend to be overestimated and decision makers will therefore be more likely to maintain the status quo scenario.

In the claims filing decision context one can imagine a policyholder, uncertain about the impact on future insurance premiums from reporting a loss, being more inclined not to file the insurance claim (i.e., maintaining the status quo) because of regretting the fallout from the filing decision more than regretting the out-of-pocket loss associated with decision not to file the claim.

Prior research on the subject of individuals giving up otherwise available pecuniary benefits spans several different disciplines other than insurance. The concept of rebate proneness in the marketing field has similarities to the choice by policyholders to file an insurance claim. In finance, there is the question concerning employee enrollment in private pension schemes, especially the failure of some to participate at a level that achieves an employer-matched contribution on behalf of individual employees. To a somewhat lesser extent, there is the issue of take-up rates for government-sponsored insurance and welfare programs.

From a marketing perspective, consumers who purchase products accompanied by rebate offers may choose not to complete the rebate redemption process. Silk (2004) suggests that such "breakage" results from psychological characteristics of consumers, such as procrastination and underestimation of the effort involved in redeeming the rebates (as cited in McCall, Bruneau, Ellis, and Mian, 2009). McCall, Eckrich, and Bruneau (2007) research the idea of consumptive delay as an explanation for not claiming rebate benefits, where some consumers are more inclined to be attracted to products with rebates because they have relatively greater willpower with respect to accepting the delayed rebate payoff (as cited in McCall, et al, 2009). As described in McCall, et al. (2009), the burden of completing the rebate redemption process, including requirements to complete forms, provide written receipts, and remove proof of purchase labels, provides rebate program disincentives (see Tat, Cunningham, and Babakus, 1988). While perceptions of products using rebates and perceptions of the rebates themselves are investigated in the marketing literature, one of the few efforts to understand the demographic profile of consumers most likely to purchase products promoted by rebate offers found that age and sex were not significant factors (McCall, et al., 2009).

The finance literature investigates situations when employees forgo an employer's matching retirement contribution as a result of their individual choices either not to participate in the employer-sponsored retirement plan or to participate at such low contribution levels that they fail to maximize the employer's contribution. Madrian and Shea (2001) provide a thorough discussion of many of the same behavioral economic theories described above that might explain this phenomenon. For example, employee procrastination in making changes to default choices results in the status quo bias defined by Samuelson and Zeckhauser (1988). The direct and indirect transaction costs of voluntarily participating or initiating changes to default 401(k) choices contributes to this procrastination.[5] Madrian and Shea also suggest the presence of an endowment effect and anchoring that can lead to a status quo bias resulting in the observed employee savings behavior.

Forgoing welfare or social insurance benefits to which a person is otherwise entitled is a similar choice as forgoing private insurance indemnification for a loss. Andrade (2002) reviews the primary empirical studies concerning low take-up rates in means-tested welfare programs. He describes how all direct and indirect costs associated with the choice to participate in welfare programs can be viewed as sources of disutility. In a rational, utility-maximizing decision model the net utility of claiming welfare benefits, which includes both the utility and disutility of doing so, is compared to the utility of not claiming the benefits. As the disutility associated with take-up increases due to increasing transaction or psychological costs, it becomes more likely that the net utility from claiming will be less than the utility from not claiming. Based on the empirical evidence, Andrade concludes that low take-up rates are due to psychological costs such as the stigma associated with collecting welfare (see also Moffitt, 1983).

Hernandez, Pudney, and Hancock (2007) estimate the implicit value of the disutility arising from transaction or psychological costs resulting in an individual's decision against collecting government-provided pension benefits in the United Kingdom. In estimating the overall costs to claiming welfare benefits, they find that income per head of household, education, the presence of disability benefits, and home ownership are significantly related to high overall costs of claiming, resulting in a lower propensity of benefit take-up.

Non-means-tested social insurance programs should not have the same level of stigma as means-tested welfare programs. However, take-up rates are still less than 100 percent. Ebenstein and Stange (2010) investigate low take-up rates in the non-means-tested unemployment insurance programs in the United States. Generally, their findings suggest that reducing barriers within the claims process will not necessarily affect program participation.

[5]Madrian and Shea (2001) include as indirect costs the costs of gathering and evaluating information about 401(k) plan design and features and multiple investment options and contribution levels.

The Pseudodeductible

In response to the shortcomings of studying policyholder claims filing behavior without information about losses that simply went unreported, Braun, et al. (2006) use Bayesian inference to better understand how insured yet unreported losses might generate different conclusions than what would be reached by looking only at the observable data. Applying their model to the observable information in homeowner's claims data provided by State Farm Fire and Casualty Company, their findings suggest that policyholders who suffer frequent losses may actually be more selective about which claims they report than policyholders who experience fewer losses.

The State Farm data used by Braun, et al. (2006) contain information about households with homeowner's insurance, including deductible levels and claims activity. Whether a household filed any claims during the 1998 to 2003 research period can be directly observed, but the losses never reported as claims remain unobservable. Therefore, no distinction can be made between households that never suffered a loss and therefore never filed a claim and households who suffered a loss but simply chose not to report it. There is also no way to directly observe whether the reported claims represent all losses that a given household has suffered. The authors rely on Bayesian probability models and the observable data to make inferences about these unknowns.

Braun, et al. (2006) were interested in better understanding why it was observed in the data that nearly a fifth of the households in their sample filed at least one small claim during the research period. Using their Bayesian models, they look at the percentage of households filing at least one "small" claim and find that accounting for the pseudodeductible significantly enhances the ability of their predictive model to replicate the observed likelihood of households filing small claims. This supports their contention that the pseudodeductible and the propensity to file small claims are closely related.

Braun, et al. (2006) also infer from their model that the riskiness of a given household, in terms of likelihood of reporting or filing an insurance claim, may be driven by the size of reported claims in addition to simply looking at the number of claims filed. Relying on observed frequencies alone may not tell the whole story. For example, their research suggests that policyholders who have filed relatively small claims in the past may be exhibiting less selectivity in their claims reporting behavior than policyholders who have filed the same number of claims but of greater severities. Despite the fact that policyholders who have filed small claims may be, in the words of Braun, et al., less risk-prone, they may have a greater propensity to file. Those who have filed larger claims may be less likely to file and may have more unclaimed losses, thereby exhibiting a larger pseudodeductible and a more selective nature than those who filed smaller claims.

The Braun, et al. (2006) study demonstrates the nonstationarity of the pseudodeductible as well. Their results suggest that increasing severities of reported claims within households may be due to an increasing pseudodeductible and a greater selectivity in the claims reporting decision. The authors posit that policyholder fears of future premium increases resulting from previously reported claims might lead them to absorb increasingly larger loss amounts. In essence, having filed a prior claim makes the filing of future claims less likely. Therefore, the recent claim-filer may be a good risk to insure, given that they will be less likely to file future claims (i.e., they will be more selective about which claim to file or report in the future).

The research described in the following sections of this paper furthers our understanding of pseudodeductibles. It improves upon Braun, et al. (2006) in the sense that it utilizes data that capture both claimed and unclaimed losses. The unclaimed losses were unobservable in Braun, et al. It also complements the Braun, et al study by analyzing pseudodeductibles in an auto insurance context. Braun, et al studied homeowner's insurance claims data.

Description of the Data

The primary source of data for this research is from a survey conducted by the Insurance Research Council (Insurance Research Council, 1999), hereafter referred to as the IRC. The survey participants were selected as the result of their responses to an initial random nationwide mailing in which it was indicated that someone in that household had been injured in at least one auto accident during the period between calendar years 1995 and 1998. The term "injured" refers to the bodily injuries of household members and does not include information about physical damage to property or the extent to which a household member injured a third party. These data include some demographic information about each injured household member, detailed information concerning the injuries suffered and how those injuries were medically treated, and information about various sources of compensation for the financial loss associated with the injuries. Of particular interest here, respondents identify when, despite a loss and the presence of auto insurance coverage, they elect not to file an auto insurance claim for reasons indicative of the presence of a pseudodeductible. A respondent could indicate one or more of the following reasons for not filing an auto insurance claim:

(1) My household member was at fault—couldn't collect from other driver.
(2) Other driver was uninsured.
(3) Other driver was relative/friend, didn't want to make a claim against him/her.
(4) No other car was involved.
(5) Amount involved was small, not worth the bother.
(6) Expenses were covered by other benefit sources.
(7) Own car didn't have injury coverage for its passengers.
(8) Could have collected from own auto insurer but didn't want to raise our rates.
(9) Other car hit and ran.
(10) Other reason.

Braun, et al. (2006) describe their concept of a pseudodeductible in terms of a policyholder "leaving money on the table," which is the decision by the policyholder to forgo a loss indemnity payment because its value is less than the policyholder's expected utility derived from not filing the insurance claim. From the list of reasons above, items 3, 5, and 8 are consistent with that definition. The general idea behind this research is to analyze the likelihood of those injured in auto accidents to file auto insurance claims relative to those deciding not to file claims because of a pseudodeductible-related reason. Those choosing not to file because of a pseudodeductible are essentially indicating that the insurance indemnity amount is less than the expected utility associated with not filing the claim. To this end, the presence of at least one of these three reasons is required when analyzing the non-filers. Further, the non-filers are limited to those who did not indicate any other reason for not filing, in order to avoid the possible bias created by other reasons dominating the pseudodeductible reasons.

The survey instructions request that the questions be answered by the household member *most familiar* with the accident and ask that detailed answers to survey questions reflect only the most recent auto accident during the period 1995–1998 if more than one auto accident has occurred. Each survey respondent answers the same set of questions on behalf of a maximum of three household members who have been injured in the same auto accident. The complete set of data captures detailed information about losses suffered by 5,768 individuals, some of whom do not report their losses as insurance claims. This is what makes this dataset unique.

The survey's treatment of each of the 5,768 observations as separate and independent despite the fact that some observations are household members injured in the same auto accident might not present a problem in terms of analyzing economic losses or types of injuries sustained. However, given that this is a study about the claims filing decision process of policyholders, this design may create a bias in the data. For example, if an adult and one minor child were injured in an auto accident and the adult responded to the survey questions for both injured household members, two observations are created that contain information about the related decision to file an insurance claim for each injury.

There is nothing in the data that explicitly links the individual observations together by household. To address the potential bias created when a single household member makes claims filing decisions on behalf of other individual observations throughout the data, clusters of individuals injured as part of the same households were manually created. The strategy is to change the unit of observation from the individual to the household level by collapsing the multiple observations of a single identified cluster into one household-level observation. To begin, all observations were sorted according to the following variables: the type of accident that occurred (a collision, a single-vehicle accident, an accident involving a pedestrian, or other), the month in which the accident occurred, the year in which the accident occurred, the state and city where the accident occurred, the location of the accident (in a central city, a suburb, a medium city, a small town, or a rural environment), and how many household members were injured in the accident.

There are 3,959 out of the original 5,768 observations in the original dataset with only a single injured person where clustering is unnecessary. Of the remaining 1,590 observations with multiple injured household members indicated, 707 fairly obvious clusters of the observations were identified that appeared to represent 707 separate households based on the sort parameters.[6] From the cluster identification process, the original dataset was reduced from 5,768 to 5,549 individual observations, with the 219 missing observations resulting from multiple-injured responses that did not include enough information for clustering purposes.

[6]The 1,590 observations subject to clustering can be broken down as follows: 1,142 indicated having two injured household members; 319 had three injured household members; 98 had four injured household members; 20 had five injured household members; 8 had six injured household members; and 3 had seven injured household members.

The claims filing decision for each family member within an identified cluster was reviewed. There were no clusters in which some members filed an insurance claim and others did not, but there were situations where the reasons given for not filing an auto claim were inconsistent within a given cluster. Of particular importance to this study are the instances where an injured household member did not indicate at least one pseudodeductible-related reason for not reporting a loss. Collapsing these clusters into one household decision-making entity would be difficult and so these clusters were removed from the dataset resulting in the elimination of 17 clusters, or a total of 42 individual observations.

The dependent variable used in the model described later identifies those households that have chosen not to report an otherwise covered loss as an auto insurance claim specifically because of at least one pseudo-deductible-related reason (i.e., fear of rate increase, friend or family was the other driver, or the loss was too small to bother). There were 424 observations removed from the analysis because the reason indicated for not filing an auto insurance claim did not include one of these specific reasons. To avoid the potential bias created by the presence of other reasons in addition to any of the pseudodeductible reasons for not reporting an insurance claim, the dependent variable was further limited to those observations identifying only one or more of the pseudodeductible reasons and no others. This removes another 90 observations from the analysis.

The original idea of a pseudodeductible relates primarily to first-party insurance. Although the same concerns giving rise to the pseudodeductible (e.g., fears of premium increases or non-renewals of coverage) can appear in the decision to report third-party liability claims, there is much more at stake if a third-party loss goes unreported. It is likely, given the possibility of attorneys' fees, non-economic damage awards, and punitive damage awards, that the costs associated with choosing not to report a third-party loss far exceed the costs related to either premium increases or coverage terminations, or both. Similar to Braun, et al. (2006), who analyze the pseudodeductible within a first-party (homeowner's property) insurance context, the focus of this study is first-party auto insurance.

The survey captures information about bodily injuries resulting from multi-vehicle collisions, single-vehicle accidents, and pedestrian accidents. There is no information about physical damage to autos. All of these accident scenarios are included in this research because first-party insurance benefits for bodily injuries existing under Personal Injury Protection (PIP), Medical Payments (MedPay) coverage, or Uninsured/Underinsured Motorists (UM/UIM) coverage can respond under any of these types of accidents. However, to draw direct comparisons between the likelihood of reporting a first-party loss relative to the likelihood of not reporting a first-party loss, observations in which an insurance claim is filed against a third-party's auto insurer, either by itself or in conjunction with a first-party claim, are removed. This limitation to a first-party insurance comparison further reduces the number of observations from 4,993 to 2,112. The idea of a policyholder forgoing insurance indemnity as a result of a pseudodeductible due to a relatively larger expected utility value implies that there is otherwise insurance coverage available for a loss. Observations that indicated no auto insurance coverage existed (N = 4) were therefore removed as well.

Each household member from an identified cluster that was injured and listed as a separate observation was given the same identification number. The clusters were collapsed into the observation for the oldest household member within a given cluster. This means that the multiple-injury households take on the demographic characteristics (e.g., age, gender) of their oldest members. The economic losses reported for the oldest household member was replaced by the sum of the economic losses reported by all other members of a given cluster, in essence becoming a total household economic loss value. This process of collapsing each identified cluster into one single household observation further reduced the dataset from 2,112 to 1,273 observations.

As previously noted, of particular interest here is the response to the survey question indicating one or more pseudodeductible-related reasons, but no others, for choosing not to file a first-party auto insurance claim for an otherwise covered loss. There were 20 non-filers that satisfy this requirement. How the three pseudodeductible-related reasons are combined across the remaining 20 non-filers is summarized in Table 1. In short, there were no respondents indicating "Other driver was friend/relative" as a reason for not filing a claim. Of the remaining two pseudodeductible reasons, only one respondent indicated both the "Fear of premium increase" and the "Minor accident only" reasons together. All other non-filers reported either of these two pseudodeductible reasons exclusively.

The primary advantage of using the data described above is that they allow for a unique analysis of the reasons why a policyholder with an otherwise covered loss would choose not to file an insurance claim. Despite this benefit, however, there are limitations to the dataset. Only bodily injury losses are identified, which could bias the empirical results assuming that some policyholders' decisions are also influenced by the presence of first-party physical damage losses.[7] There is very little personal demographic information available for each household, such as income, wealth, and

[7]Survey respondents may have also suffered first-party physical damage losses in addition to bodily injury, but that information is simply not captured by the IRC survey instrument.

Table 1. Breakdown of Pseudodeductible Reasons Indicated for Choosing Not to File a First-Party Bodily Injury Auto Insurance Claim

Pseudodeductible reasons for not filing	% of all non-filers identifying particular pseudodeductible reason (N=20)	When "Minor accident only" reason given (N=14)	When "Other driver was friend/relative" reason given (N=0)	When "Fear of premium increase" reason given (N=7)
Minor accident only	70.0%	100.0%	0.0%	14.3%
Other driver was friend/relative	0.0	0.0	0.0	0.0
Fear of premium increase	35.0	7.1	0.0	100.0

Note: Data represent the percentages of each sample indicating the stated reason for not filing. Because multiple responses were permitted, column sums may exceed 100%.

occupation and there is no insurance policy information such as policy limits, deductibles, and type of coverage that applies (e.g., UM/UIM, PIP, or Medical Payments).

Research Hypothesis and Methodology

The research hypothesis for this study can be developed from the inferences and conclusions in the Braun, et al. (2006) paper. Their findings suggest that policyholders who have higher claims filing rates also have higher percentage differences between the amounts of the stated policy deductibles and their pseudodeductibles. Conversely, those with lower filing rates have smaller percentage differences. The general implication is that those who have filed more insurance claims in the past will have relatively higher pseudodeductibles and will be more selective about reporting future losses as insurance claims. Their Bayesian models also suggest that those with smaller losses may be more likely to report small claims and those with larger losses may be more likely to absorb the losses. Given these earlier findings, the alternative hypothesis tested in this paper can be stated as follows:

H_a: *A household with a history of previous claims and a relatively more severe loss will be less likely to file an auto insurance claim due to a pseudo-deductible.*

A logistic regression model is used to test this hypothesis. The dependent variable (FILECLAIM) is a binary response equal to one if the household loss was reported as an insurance claim with the respondent's own auto insurer. The variable is equal to zero if at least one of the pseudodeductible-related reasons for not reporting a loss as an auto insurance claim (i.e., the loss was too minor and/or there was a fear of a future premium increase) is indicated. The model identifies factors that affect the propensity to file a first-party insurance claim for bodily injuries relative to not filing because of the presence of a pseudodeductible.

Explanatory Variables

In this study, loss severity is measured by the natural log transformation of total economic damages incurred (LNECONLOSS). Previous studies have found that the severity of the loss incurred by the policyholder has a bearing on the claims reporting decision. This has been empirically supported in the workers' compensation context (see Biddle and Roberts, 2003) and from a third-party moral hazard perspective (see Cummins and Tennyson, 1996), with both studies suggesting that greater losses increase the likelihood of filing insurance claims. However, Braun, et al. (2006) find that because of the presence of a pseudodeductible the likelihood of filing an insurance claim may decrease as the size of the loss increases. The sign of the estimated coefficient on this variable is therefore uncertain.

An important result of the Braun, et al. (2006) study is the finding that a more active claims-filing history may result in a lower propensity to file future claims. The IRC data do not explicitly indicate the reporting of previous insurance claims, but they do include the number of auto accidents that have occurred during the three-year period leading up to the present loss to which the survey questions apply. Dummy variables are used to measure previous auto insurance claiming activity. ONEPRIORACC equals one when there has been just one accident previous to the accident being reported upon and zero when there have been no previous accidents. MULTPRIORACC equals one when at least two prior accidents have occurred and zero otherwise. The category of no previous accidents (NOPRIORACC) is the holdout scenario in the regression analysis. It is anticipated that as the number of previous accidents increases relative to no prior accidents the likelihood of reporting the current accident as an auto claim decreases. This would support the findings in Braun, et al.

A dummy variable (FEMALE) identifying the gender of the injured respondent equals one if a female was injured and zero otherwise. There are also three age dummy variables indicating when the injured person is under age 25 (AGE25), between the ages of 25 and 55 (AGE2555), or age 55 or older (AGE55). These variables may be important factors to the extent that they alter the degree of relative risk aversion (Halek and Eisenhauer, 2001). As an individual becomes relatively more risk averse, there may be an increased willingness to forgo filing an insurance claim if it creates uncertainty that leads to a pseudodeductible, such as a fear of future premium increases. On the other hand, the decision not to report a loss, especially one involving bodily injuries where the precise loss values are difficult to initially predict, could create uncertainty that would discourage the more risk averse from being influenced by a pseudodeductible. In either case, the arguments, and therefore the value of these particular variables, only hold when the person making the decision to report a loss as an insurance claim is the same as the person injured in the accident. As described earlier in the paper, the dataset is designed so that one household member makes the claims filing decision for multiple household members. Recall that to remove the potential data bias this creates, the multiple-injury households (of which there were 1,590 out of the original 5,768 observations) were clustered into common households and each clustered household was collapsed into an individual household observation of the oldest household member. To control somewhat for the effect this clustering and collapsing process may have on the regression results, a dummy variable is included (MULTINJ) that is equal to one when the observation is a combined household version of previously individual observations and zero otherwise.

Soft tissue injuries are relatively more susceptible to moral hazard than other types of injuries because they are difficult to objectively diagnose and treat (Dionne and St-Michel, 1991; Weisberg and Derrig, 1991; Cummins and Tennyson, 1996; and Derrig, Weisberg, and Chen, 1994). Cummins and Tennyson (1996) argue that soft tissue injuries reduce claims filing costs for fraudulent claims thereby increasing the number of claims filed for such injuries. To control more generally for the claims-filing incentives created by the treatment of relatively subjective injuries, a dummy variable (ALTTX) is included that equals one if the respondent's injuries were treated by a chiropractor, some other type of alternative care provider, or a psychologist, and zero otherwise. A positive coefficient is expected.

Control dummy variables indicating the states in which the accidents took place are included in the model. These binary variables, identified by each state's postal abbreviation, capture both observable and unobservable cross-sectional differences created by differences among the states that might influence pseudodeductibles and more generally policyholder claims-filing behavior.[8] Massachusetts is the holdout state.

Dummy variables are also included indicating when the auto law in a given accident state is tort liability (TORT), no-fault with a monetary threshold (MONETARYNF), no-fault with a verbal threshold

(VERBALNF), choice no-fault (CHOICENF), or add-on no-fault (ADDONNF). Drivers in choice no-fault states can choose at the time insurance is purchased to be covered by either tort or no-fault rules with the no-fault option carrying with it the usual tort restrictions.[9] The add-on laws are separately identified because the no-fault options do not place restrictions on the ability to file tort claims. The holdout group in the regression models is the tort law category. Even in tort liability states where Personal Injury Protection (PIP) coverage is not necessary, limited first-party benefits for bodily injuries in personal auto policies exist under Medical Payments coverage. However, given the emphasis on first-party PIP coverage in no-fault states, there may be a stronger effect on the first-party claims decision in no-fault states compared to tort states.

The data collected cover the period 1995 through 1998. Control dummy variables for each accident year are included, with the holdout year being 1998.

Of the 1,273 observations remaining from the original dataset after limiting the data to first-party auto claims and collapsing the individual claimant observations into household observations, a total of 1,100 observations are used in the final analysis. The difference is the result of missing data for particular variables, the removal of other observations that presented conflicting survey information, and removal of particular data points causing computational problems. Variables with missing data include the following: the gender of the household decision maker (N = 14); various accident year identifiers (N = 17); the number of previous accidents (N = 3); the age of the head of household (N = 16), and various state identifiers (N = 13). Accuracy checks on the survey responses showed 93 observations where the respondents collected payments from their own

---

[8]The IRC data identify the states where accidents occurred and not necessarily the states of residency. While much of the variation created by the different states is the result of the accident location, some of the variation may also be attributable to differences among the states of residency. It is possible that an accident occurs in a different state than the residency state, but much of that variation should still be captured because in a large majority of the cases the states are the same. An early investigation of auto accident locations relative to victims' areas of residence (see Durand, 1980) found that nearly 51% of injured auto accident victims were involved in accidents within five miles of their homes and almost 71% were injured within a ten-mile radius. Only about 10% of injured people were involved in accidents more than thirty miles from their homes.

[9]The three choice no-fault states are Kentucky, New Jersey, and Pennsylvania. Owings-Edwards (2004) suggested that Kentucky and New Jersey should be considered de facto no-fault states in any empirical work because roughly 90 percent of the drivers in each state elect the no-fault option. Therefore, Kentucky and New Jersey are categorized as monetary and verbal threshold no-fault states, respectively, leaving Pennsylvania as the only choice no-fault state.

auto insurers while at the same time indicating that no claims were filed with their own auto insurers (and vice versa). Also, 15 observations indicated collecting from other benefit sources (such as workers' compensation, group health insurance, or others) and they did not report a bodily injury claim to their own auto insurers. These observations were removed to avoid any potential bias created by the possibility that these respondents did not pursue a first-party auto insurance claim because of the other benefit sources and simply did not indicate that as a reason for not filing the claim. Finally, two observations, one each from Alaska and Washington DC, caused convergence problems with the logistic regression model and were therefore removed. Summary statistics for all variables appearing in the empirical model are presented in Table 2 with the exception of the states in which the accidents occurred, which can be found summarized in Table A-1 of the Appendix.

Research Model

This research evaluates the propensity to file a first-party auto insurance claim relative to not filing such a claim because of a pseudodeductible. When an otherwise covered claim is not filed due to a pseudodeductible, it identifies a case where a policyholder is willing to give up the indemnity payment because its value does not exceed the policyholder's expected utility associated with not filing the claim. Because the dependent variable in this model has a 0/1 binary outcome, a logistic regression model relying on the maximum likelihood estimation process is used (Greene, 1997). The probability of filing an insurance claim is modeled using the following equation:

$$\begin{aligned} logit(p) = ln\left(\frac{p}{1-p}\right) &= \beta_0 + \beta_1\, LNECONLOSS + \beta_2 ONEPRIORACC \\ &+ \beta_3 MULTPRIORACC + \beta_4 LARGE + \beta_5 FEMALE \\ &+ \beta_6 AGE25 + \beta_7 AGE55 + \beta_8 MULTINJ + \beta_9 ALTTX \\ &+ \beta_{10} MONETARYNF + \beta_{11} VERBALNF + \beta_{12} CHOICENF \\ &+ \beta_{13} ADDONNF + \beta_{14} 1995 + \beta_{15} 1996 + \beta_{16} 1997 \\ &+ \sum_{i=17}^{63} \beta_i\, ACCIDENTSTATE \end{aligned}$$

where p is the probability that a first-party auto insurance claim is filed and $1 - p$ is the probability that no insurance claim is filed due to the presence of a pseudodeductible.

Table 2. Summary Statistics for Variables Used in Analysis (N = 1,100)[a]

Variable	Definition of variable	Mean
FILECLAIM	Binary dependent variable equal to 1 if first-party auto claim filed and 0 if claim not filed due to pseudodeductible.	0.98 (0.13)
LNECONLOSS	Natural log of total economic loss in each household involved in auto accident.	7.63 (1.84)
ONEPRIORACC	Dummy variable indicating previous auto accidents equal to 1 when one previous accident and 0 otherwise.	0.08 (0.27)
MULTPRIORACC	Dummy variable indicating previous auto accidents equal to 1 when at least two previous accidents and 0 otherwise.	0.02 (0.12)
LARGE	Dummy variable indicating size threshold of loss equal to 1 when economic loss greater than or equal to $1,000 and 0 otherwise.	0.67 (0.47)
FEMALE	Dummy variable equal to 1 when gender of oldest injured household member and presumed decision maker is female and 0 otherwise.	0.61 (0.49)
AGE25	Dummy variable equal to 1 when age of oldest injured household member and presumed decision maker is under 25 and 0 otherwise.	0.22 (0.42)
AGE55	Dummy variable equal to 1 when age of oldest injured household member and presumed decision maker is at or over 55 and 0 otherwise.	0.19 (0.39)
MULTINJ	Dummy variable equal to 1 when household had more than one injured member and 0 otherwise.	0.15 (0.36)
ALTTX	Dummy variable equal to 1 if medical treatment involved chiropractor, other alternative treatment, or psychologist and 0 otherwise.	0.31 (0.46)
MONETARYNF	Dummy variable equal to 1 when monetary threshold no-fault state and 0 otherwise.	0.09 (0.29)
VERBALNF	Dummy variable equal to 1 when verbal threshold no-fault state and 0 otherwise.	0.18 (0.39)
CHOICENF	Dummy variable equal to 1 when choice no-fault state and 0 otherwise.	0.13 (0.33)
ADDONNF	Dummy variable equal to 1 when add-on no-fault state and 0 otherwise.	0.17 (0.38)

Table continues

Table 2. continued

Variable	Definition of variable	Mean
1995	Dummy variable equal to 1 when accident year is 1995 and 0 otherwise.	0.26 (0.44)
1996	Dummy variable equal to 1 when accident year is 1996 and 0 otherwise.	0.30 (0.46)
1997	Dummy variable equal to 1 when accident year is 1997 and 0 otherwise.	0.40 (0.49)

[a]Dummy variables for the different states in which accidents occurred are summarized in Table A-1 of the Appendix.
Note: Standard deviations are in parentheses. Minimums and maximums for all dummy variables are, by definition, 0 and 1 respectively. The minimum and maximum values for LNECONLOSS are 1.10 and 13.57, respectively. For all discrete variables, the mean value reflects percentage of total number of observations.

In the sample of closed claims data used by Braun, et al. (2006), 18 percent had filed at least one claim of less than $1,000. The authors perform a posterior predictive check of their Bayesian probability models for the percentage of households with at least one reported insurance claim less than threshold values of $100, $250, $500, or $1,000. For the binary logistic regression model used in this paper, a dummy variable (LARGE) is created that equals one when the economic loss is greater than or equal to a $1,000 claim threshold and zero otherwise. Based upon the findings of Braun, et al., a significant and negative coefficient estimate is expected. This would support the contention that policyholders with large losses are more selective in the claims they file when pseudodeductibles exist.

Empirical Results

The results of the logistic regression of FILECLAIM on all independent variables are reported in Table 3. The adjusted R-square for the model is nearly 62 percent and the Hosmer and Lemeshow (1989) statistic indicates the model is a good fit with the data.

As the magnitude of the total economic loss increases, households do not appear to be more selective in the claims filing decision, as is suggested by Braun, et al. (2006). As households sustain higher economic losses they become more likely to file auto claims than to forgo the claims because of a pseudodeductible. This suggests that increasing loss amounts tend to dominate potential pseudodeductible issues such as future premium increases, making it less likely that policyholders will decide to forgo insurance indemnity payments.

Braun, et al. (2006) indicate that as the previous number of claims increases, policyholders may become more selective in the claims they file and more likely to forgo available indemnity payments. The results reported here suggest that policyholders with one previous loss are significantly more likely to report the current loss as an insurance claim compared to those who have had no previous losses. The way the dependent variable is defined, this result also implies that policyholders with one previous loss are less likely to assert that the presence of a pseudodeductible affected their claims filing decision. This, alone, seems to contradict Braun, et al.

There is no statistical difference between households with no prior accidents and households with two or more prior accidents in terms of the propensity to be influenced by a pseudodeductible. Although not reported in Table 3, a simple change to the model was made to draw a direct comparison between households with only one prior accident and households with two or more prior accidents. The results provide some evidence that households with no previous accidents and households with multiple previous accidents are both more likely to be influenced by pseudodeductibles compared to households with exactly one prior accident.[10] This may be indicative of policyholders initially experiencing a status quo bias but then changing their behaviors as losses accumulate. With no previous accidents, policyholders are reluctant to report the present loss as an insurance claim because of the uncertain future pseudodeductible-related implications of such a decision. Having retained the financial impact of one loss already, those who suffer a second loss (i.e., one accident prior to the present one) are more inclined to report the second loss. For those who are unfortunate enough to have had two or more previous accidents, their claims filing behavior may again become more selective because of updated information about the psuedodeductible-related effects of previously reported losses. This is consistent with the more general finding in Braun, et al. (2006) that policyholders with more extensive claims histories may be paradoxically more appealing as insurance customers because their future claims filing behavior becomes more selective.

The dummy variable indicating total economic losses exceeding the $1,000 large-claim threshold is statistically significant, with the expected

[10]The holdout category was changed from the no prior accident households (NOPRIORACC) to households with two or more prior accidents (MULTPRIORACC). The estimated coefficient on ONEPRIORACC is positive relative to the new holdout category and is significant at the 10% level. The other empirical results reported in Table 3 remained unchanged.

sign. Consistent with Braun, et al. (2006), households with "large" losses are less likely to report a claim because of a pseudodeductible. Despite the result that generally increasing economic loss values reduce the influence of pseudodeductibles (per the LNECONLOSS coefficient estimate), it appears that households are more concerned about the potential effect of large loss amounts on the circumstances giving rise to pseudodeductibles, such as future premium increases.

Gender appears to be a significant determinant in the role that pseudodeductibles play in auto claims filing decisions, with females being less likely than males to file claims and therefore more likely to be influenced by pseudodeductibles. Given that Halek and Eisenhauer (2001) found evidence that female heads of households are relatively more risk averse than males, females may perceive the uncertainty from reasons creating pseudodeductibles, such as future premium increases, as being greater than the uncertainty associated with the loss.

Halek and Eisenhauer (2001) also found, in part, that the youngest and oldest household heads are more risk averse relative to the middle-aged.[11] The results from this research indicate that those younger than 25 years are no more or less likely to report a loss as an insurance claim than middle-aged policyholders between the ages of 25 and 55. However, those ages 55 and older are more likely to file an insurance claim without regard to a pseudodeductible reason than those in the middle-aged category. One explanation for this mixed result may be that the oldest policyholders view the risks from retaining the bodily injury loss as being significantly greater than the risks leading to the development of a pseudodeductible. The oldest policyholders may realize that their age can exacerbate the negative financial consequences from a given bodily injury, and therefore less concern is given to pseudodeductible-related risks.

As expected, the design of a given accident state's auto law significantly affects the decision to file first-party bodily injury auto claims. In particular, households across all no-fault auto regimes were more likely to file first-party bodily injury auto claims relative to households in tort states. With the exception of add-on no-fault states, all no-fault laws were statistically significant at the conventional 5 percent level.

While including individual state dummy variables in the model is the ideal method for controlling for many otherwise unobservable state-specific factors that could influence the claims reporting decision, there were computational problems associated with including these dummy variables in the model largely because (1) the model is a logistic regression model

[11]More precisely, Halek and Eisenhauer (2001) found that the least risk averse heads of households are 40 years old (p. 77).

and (2) the number of non-filer responses for the dependent variable (where FILECLAIM = 0) is very sparse. There are many states in which all of the responses equal one. This leads to what is known as a quasi-complete separation of the data, a failure of the maximum likelihood estimation process to converge and therefore no maximum likelihood estimate for many of the individual state variables. This is an indication that the logistic regression model has come close to perfectly predicting the binary dependent variable.

Ironically, one way to resolve the quasi-complete separation problem present in logistic regressions applied to such data is to change the model so it does not predict the response as perfectly. For example, the individual state dummy variables could be collapsed into a small number of regional variables for which there will be both zeros and ones appearing in each broader category. This relatively simple approach, however, comes at the expense of losing the ability to control for many state-specific factors.

A more appealing approach to address the quasi-complete separation problem is a proposal by Firth (1993) to make certain changes to the gradient vector such that the Newton-Raphson numerical method for maximizing the log-likelihood function defining the logistic regression model can be performed. Using the Firth method allows the maximum likelihood estimation process to converge despite the presence of nearly all individual state dummy variables in the model. The Firth method also corrects for the bias created by the relative lack of variability in the binary responses of the dependent variable (Allison, 2008).

A number of individual states influence the claiming decision in the context of pseudodeductibles. Relative to the holdout state of Massachusetts, the likelihood of policyholders reporting first-party auto bodily injury claims was statistically significant and greater in seventeen states and significant and lower in three states.[12] This suggests that a pseudodeductible is more likely to influence the claims decision in Massachusetts than in seventeen other states but is less likely to be a reason for choosing not to report a loss in three other states. As reported by the Insurance Information Institute, Massachusetts had the tenth highest overall average expenditures for auto insurance in 1998 and was ranked as high as third in 1994 (Insurance Information Institute, 2001).[13] This could partly explain the

[12]Statistical significance is defined at the standard 5 percent level.

[13]The overall average expenditure for auto insurance data is provided by the National Association of Insurance Commissioners and represents what the average consumer actually spends on insurance per vehicle. It does not specifically control for influential factors such as types of coverages purchased, types of autos being insured, per capita income, number of miles driven, and others.

Table 3. Logistic Regression Results with Firth (1993) Bias Correction (N=1,100)[a]

Variable	Description	Expected sign	Estimated coefficient	*p*-value
LNECONLOSS	Natural log of economic loss	+/–	1.1086	<0.0001
ONEPRIORACC	One previous accident in past three years	–	0.7369	0.0057
MULTPRIORACC	Two or more previous accidents in past three years	–	–0.0328	0.9441
LARGE	Economic loss greater than or equal to $1,000	–	–0.3612	0.0303
FEMALE	Gender of accident victim	+/–	–0.2351	0.0432
AGE25	Age of accident victim is under 25	+/–	0.0548	0.6429
AGE55	Age of accident victim is greater than or equal to 55	+/–	1.7327	< 0.0001
MULTINJ	Multiple household members injured	+/–	0.1986	0.3010
ALTTX	Alternative treatment used (chiropractor, alternative treatment, or psychologist)	+	–0.1391	0.2827
MONETARYNF	Monetary threshold no-fault state	+	1.1073	0.0435
VERBALNF	Verbal threshold no-fault state	+	1.6551	0.0019
CHOICENF	Choice no-fault state	+	2.0317	0.0003
ADDONNF	Add-on no-fault state	+/–	0.2678	0.7127
1995	1995 accident year	+/–	0.0348	0.8851
1996	1996 accident year	+/–	–0.1251	0.5960
1997	1997 accident year	+/–	–0.1649	0.4743
AL	Alabama	+/–	1.4159	0.0146
AR	Arkansas	+/–	0.9112	0.1770
AZ	Arizona	+/–	1.4752	0.0181
CA	California	+/–	2.1511	0.0001
CO	Colorado	+/–	0.7876	0.0921
CT	Connecticut	+/–	1.0034	0.1095
DE	Delaware	+/–	0.6012	0.4125
FL	Florida	+/–	–0.1769	0.5687
GA	Georgia	+/–	1.3731	0.0285
ID	Idaho	+/–	0.9076	0.1501

Table continues

Table 3. *continued*

Variable	Description	Expected sign	Estimated coefficient	p-value
IL	Illinois	+/–	1.7096	0.0020
IN	Indiana	+/–	0.9844	0.0742
IA	Iowa	+/–	1.5455	0.0137
KS	Kansas	+/–	0.7058	0.2415
KY	Kentucky	+/–	–0.3522	0.4485
LA	Louisiana	+/–	1.2279	0.0301
MD	Maryland	+/–	1.2082	0.0693
ME	Maine	+/–	–1.6075	0.0408
MI	Michigan	+/–	–0.5729	0.0300
MN	Minnesota	+/–	1.3016	0.0057
MO	Missouri	+/–	1.7950	0.0043
MS	Mississippi	+/–	0.1625	0.7691
MT	Montana	+/–	0.6230	0.3348
NC	North Carolina	+/–	2.8422	<0.0001
ND	North Dakota	+/–	0.2517	0.7007
NE	Nebraska	+/–	1.9893	0.0041
NH	New Hampshire	+/–	0.0000	No est.
NJ	New Jersey	+/–	–0.8919	0.0108
NM	New Mexico	+/–	2.9368	0.0002
NV	Nevada	+/–	0.0000	No est.
NY	New York	+/–	0.0000	No est.
OH	Ohio	+/–	1.4245	0.0089
OK	Oklahoma	+/–	1.0504	0.0648
OR	Oregon	+/–	1.9137	0.0053
PA	Pennsylvania	+/–	0.0000	No est.
RI	Rhode Island	+/–	0.6208	0.3950
SC	South Carolina	+/–	1.0899	0.0713
SD	South Dakota	+/–	–0.1706	0.8086
TN	Tennessee	+/–	1.4087	0.0194
TX	Texas	+/–	1.4486	0.0129
UT	Utah	+/–	0.1582	0.7286
VA	Virginia	+/–	0.8880	0.1376
VT	Vermont	+/–	1.3565	0.0732
WA	Washington	+/–	1.7764	0.0106
WI	Wisconsin	+/–	0.8431	0.2361
WV	West Virginia	+/–	1.2148	0.0666

Table continues

Table 3. *continued*

Variable	Description	Expected sign	Estimated coefficient	*p*-value
WY	Wyoming	+/–	0.8372	0.2157
Max-rescaled R^2 = 0.6253[b]				
Hosmer and Lemeshow Goodness-of-Fit Chi Square Statistic = 14.4650 (8 d.f.)[c]				

[a]Dependent variable (FILECLAIM) equals 1 when first-party auto claim filed and 0 when claim is not filed due only to pseudodeductible reason.
[b]Cox and Snell (1989) developed a generalized pseudo-R^2 goodness-of-fit measure comparable to the traditional R^2 statistic computed in ordinary least squares regressions. However, the pseudo-R^2 has a maximum value less than one. To permit the pseudo-R^2 statistic to reach a value of one, Nagelkerke (1991) suggested rescaling it by the inverse of its maximum value, and this is the max-rescaled R^2.
[c]The Hosmer and Lemeshow statistic (see Hosmer and Lemeshow, 1989) indicates how a given model fits the data being analyzed. Differences between the observed and expected number of occurrences in groupings of the data by percentile are summarized by the Pearson chi-square statistic, which is compared to a chi-square probability distribution with degrees of freedom equal to the number groups created less two. The null hypothesis is that the data fit the specified model well. Therefore, *p*-values in excess of .05 would indicate that the data fit the given model well. The *p*-value associated with the reported statistic of this particular model is 0.0704.

result. The state variables may also be controlling for the variation in other state-specific characteristics including subrogation laws, consumer attitudes towards insurance, degree of urbanization, and others.[14]

SUMMARY AND CONCLUSIONS

Observable insurance claims filing rates in closed claims data might be misleading when used to estimate likelihoods of reporting future losses. The presence of a pseudodeductible may induce policyholders to avoid filing claims for otherwise covered losses. Braun, et al. (2006) use Bayesian probability models to make inferences concerning unobservable claims filing decisions. They find that the presence of a pseudodeductible in

[14]The states of Nevada, New Hampshire, New York, and Pennsylvania were not estimable due to exact linear dependency resulting from combinations with various other states included in the model and the auto law dummy variables. However, the fact that estimates were not determined for these four state variables does not bias the coefficient estimates for the remaining variables in the model.

homeowner's insurance prompted policyholders to be more selective when choosing whether to report larger losses as claims. The policyholders also appear to be more selective about filing homeowner's insurance claims when there have been previous claims reported. The large loss and the previously reported loss characteristics appear to decrease the probability of a given policyholder to file an insurance claim, giving rise to the pseudodeductible.

This paper analyzes similar policyholder behavior but in an auto insurance context. Based on a survey of people injured in auto accidents, it specifically identifies what is typically unobservable in closed claims databases: circumstances where someone has suffered a loss but has chosen not to file an insurance claim. Invaluably, the survey captures the presence of pseudodeductibles in the sense that it identifies when the reason for not filing the insurance claim is because of a fear of future premium increases or because the amount of the loss is too small to bother despite the presence of coverage.

A logistic regression model is used to determine how the likelihood of filing an insurance claim relative to avoiding the claim because of a pseudodeductible is influenced by the primary factors identified by Braun, et al. (2006), namely, the size of the economic loss suffered and the previous loss history of the household. In essence, the probability of injured people to forgo auto insurance indemnity payments given these two characteristics (and other control variables) is modeled.

The empirical analysis confirms the Braun, et al. (2006) findings for the most part but with some differences revealed. As a household's economic losses generally increase, it appears that it is more likely to file an insurance claim. However, households experiencing economic losses exceeding a $1,000 large claim threshold are more selective and are more likely to forgo filing an insurance claim because of a pseudodeductible compared to households with losses below the threshold amount. This is consistent with Braun, et al. Also, somewhat contrary to the findings of Braun, et al., those who have had one prior accident are less likely to be influenced by a pseudodeductible (i.e., they are not more selective) compared to someone with no previous accident history. However, when compared against those with just one previous accident, those with two or more previous losses are more selective and therefore less likely to file a claim for pseudodeductible reasons. So it appears that selectivity in the claims filing decision does increase as prior loss history becomes more extensive.

There are interesting results with some of the control variables as well. In particular, those 55 and older appear to be more likely to file an auto claim (or, less likely to exhibit a pseudodeductible) than those injured who are younger than 55. Females also appear to be more likely to be influenced by the presence of pseudodeductibles.

The insurance claims function is an important element of any insurer's profitability. There is keen interest from both a managerial and a legal perspective about what drives claims filing decisions. What has in the past been difficult to research in this area are some of the behavioral aspects of the decision process itself. Braun, et al. (2006) have hit on one such behavioral aspect, the application of a pseudodeductible, that has been known anecdotally to exist but never rigorously researched until their study. This paper adds to our overall understanding of how pseudodeductibles influence claims filing decisions in a first-party auto context.

REFERENCES

Allison, PD (2008) Convergence Failures in Logistic Regression, *Proceedings of the SAS Global Forum 2008 Conference,* Cary, NC: SAS Institute Inc., available at www2.sas.com/proceedings/forum2008/360-2008.pdf.

Andrade, C (2002) The Economics of Welfare Participation and Welfare Stigma: A Review, *Public Finance and Management*, 2(2): 294–333.

Biddle, J, and K Roberts (2003) Claiming Behavior in Workers' Compensation, *Journal of Risk and Insurance*, 70(4): 759–780.

Braun, M, PS Fader, ET Bradlow, and H Kunreuther (2006) Modeling the "Pseudodeductible" in Insurance Claims Decisions, *Management Science*, 52(8): 1258–1272.

Butler, RJ, and JD Worrall (1983) Workers' Compensation: Benefit and Injury Claims Rates in the Seventies, *Review of Economics and Statistics*, 65: 580–589.

Cox, DR, and EJ Snell (1989) *The Analysis of Binary Data*, Second Edition, London: Chapman and Hall/CRC.

Cummins, JD, and S Tennyson (1996) Moral Hazard in Insurance Claiming: Evidence from Automobile Insurance, *Journal of Risk and Uncertainty*, 12: 29–50.

Derrig, RA, HI Weisberg, and X Chen (1994) Behavioral Factors and Lotteries Under No-Fault with a Monetary Threshold: A Study of Massachusetts Auto Claims, *Journal of Risk and Insurance*, 61(2): 245–275.

Dionne, G, and P St-Michel (1991) Workers' Compensation and Moral Hazard, *Review of Economics and Statistics*, 73(May): 236–244.

Durand, A (1980) An Analysis of Accident Location in Relation to Area of Residence, *Insurance Industry Studies by the All-Industry Research Advisory Council*, Research Report A80-4.

Ebenstein, A, and K Stange (2010) Does Inconvenience Explain Low Take-Up? Evidence from Unemployment Insurance, *Journal of Policy Analysis and Management*, 29(1): 111–136.

Firth, D (1993) Bias Reduction of Maximum Likelihood Estimates, *Biometrika*, 80: 27–38.

Greene, WH (1997) *Econometric Analysis*, Third Edition, Upper Saddle River, NJ: Prentice-Hall.

Halek, M, and JG Eisenhauer (2001) Demography of Risk Aversion, *Journal of Risk and Insurance*, 68(1): 1–24.

Hernandez, M, S Pudney, and R Hancock (2007) The Welfare Cost of Means-Testing: Pensioner Participation in Income Support, *Journal of Applied Econometrics*, 22: 581–598.

Hosmer, DW, and S Lemeshow (1989) *Applied Logistic Regression*, New York, NY: John Wiley and Sons.

Insurance Information Institute (2001) *The Fact Book 2001*, New York, NY: I.I.I..

Insurance Research Council (1999) *1998 Consumer Panel Study of Auto Accident Victims*, Malvern, PA: Insurance Research Council.

Johnson, EJ, J Hershey, J Meszaros, and H Kunreuther (1993) Framing, Probability Distortions, and Insurance Decisions, *Journal of Risk and Uncertainty*, 7: 35–51.

Kahneman, D, and A Tversky (1979) Prospect Theory: An Analysis of Decision Under Risk, *Econometrica*, 47: 263–291.

Kahneman, D, and A Tversky (1982) The Psychology of Preference, *Scientific American*, 246: 160–173.

Madrian, BC, and DF Shea (2001) The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior, *The Quarterly Journal of Economics*, 116(4): 1149–1187.

McCall, M, DW Eckrich, and CL Bruneau (2007) A Preliminary Investigation of Consumptive Delay and Rebate Programs, in Shepard, D (Ed.), *Proceedings of the Association of Marketing Theory and Practice*, 16: 126–130, Panama City, FL: AMTP.

McCall, M, CL Bruneau, AD Ellis, and K Mian (2009) A Framework for Understanding Consumptive Delay, *Journal of Product & Brand Management*, 18(6): 461–467.

Moffitt, R (1983) An Economic Model of Welfare Stigma, *American Economic Review*, 73(5): 1023–1035.

Nagelkerke, NJD (1991) A Note on a General Definition of the Coefficient of Determination, *Biometrika*, 78(3): 691–692.

Owings-Edwards, S (2004) Choice Auto Insurance: The Experience of Kentucky, New Jersey, and Pennsylvania, *Journal of Insurance Regulation*, 23(1): 25–41.

Samuelson, W, and R Zeckhauser (1988) Status Quo Bias in Decision Making, *Journal of Risk and Uncertainty*, 1: 7–59.

Silk, TG (2004) Examining Purchase and Non-Redemption of Mail-in Rebates: The Impact of Offer Variables on Consumers, unpublished doctoral dissertation, University of Florida, Gainesville, FL.

Tat, P, WA Cunningham III, and E Babakus (1988) Consumer Perceptions of Rebates, *Journal of Advertising Research*, 28(4): 45–50.

Thaler, RH (1980) Toward a Positive Theory of Consumer Choice, *Journal of Economic Behavior and Organization*, 1: 39–60.

Tversky, A, and D Kahneman (1974) Judgment Under Uncertainty: Heuristics and Biases, *Science*, 185: 1124–1131.

Venezia, I (1984) Aspects of Optimal Automobile Insurance, *The Journal of Risk and Insurance*, 51(1): 63–79.

Venezia, I, and H Levy (1980) Optimal Claims in Automobile Insurance, *Review of Economic Studies*, 47(3): 539–549.

Weisberg, H, and RA Derrig (1991) Fraud and Automobile Insurance: A Report on Bodily Injury Liability Claims in Massachusetts, *Journal of Insurance Regulation*, 9: 497–541.

APPENDIX

Table A-1. Summary Statistics for Dummy Variables Representing Accident States (N = 1,100)

Variable	Mean	Std. dev.	Variable	Mean	Std. dev.
AK	0.0000	0.00	MT	0.0055	0.07
AL	0.0164	0.13	NC	0.0264	0.16
AR	0.0155	0.12	ND	0.0018	0.04
AZ	0.0127	0.11	NE	0.0055	0.07
CA	0.0873	0.28	NH	0.0018	0.04
CO	0.0273	0.16	NJ	0.0373	0.19
CT	0.0109	0.10	NM	0.0073	0.09
DC	0.0000	0.00	NV	0.0036	0.06
DE	0.0036	0.06	NY	0.0673	0.25
FL	0.0691	0.25	OH	0.0382	0.19
GA	0.0191	0.14	OK	0.0136	0.12
HI	0.0000	0.00	OR	0.0242	0.15
ID	0.0045	0.07	PA	0.0627	0.24
IL	0.0382	0.19	RI	0.0027	0.05
IN	0.0218	0.15	SC	0.0127	0.11
IA	0.0136	0.12	SD	0.0036	0.06
KS	0.0100	0.10	TN	0.0255	0.16
KY	0.0255	0.16	TX	0.0518	0.22
LA	0.0155	0.12	UT	0.0118	0.11
MA	0.0173	0.13	VA	0.0227	0.15
MD	0.0136	0.12	VT	0.0018	0.04
ME	0.0018	0.04	WA	0.0218	0.15
MI	0.0482	0.21	WI	0.0127	0.11
MN	0.0227	0.15	WV	0.0064	0.08
MO	0.0209	0.14	WY	0.0036	0.06
MS	0.0118	0.11			

Note: No observations indicated HI as an accident state. DC and AK were identified in one observation each. However, these observations were necessarily removed in order for the maximum likelihood estimation iterative process to converge. Minimums and maximums for all dummy variables are, by definition, 0 and 1, respectively and therefore are not reported. For all variables, the mean value reflects percentage of total number of observations.