# Exaggerating to Break-Even: Reference-Dependent Moral Hazard in Automobile Insurance Claims

Feng Gao† Jaimie W. Lien‡ Jie Zheng±

Current Draft: May 8th, 2021

*Initial Draft: May 2nd, 2016*

**Abstract:**[1]

The effects of asymmetric information are often difficult to detect empirically, such as in insurance settings. We show that allowing for reference-dependent preferences can assist with this empirical challenge. Using detailed auto insurance claims data, we show that policyholders exhibit ex-post moral hazard by exaggerating their damage claims in a manner consistent with reference-dependent preferences. Consistent with our model of reference-dependent claims, we find stronger reference-dependent behavior among policyholders with higher premium levels and low risk policyholders. Also consistent with the model, policyholders are more reference-dependent in their first claims than in subsequent claims, and in the second half of the coverage year compared to the first half. Specifically, the detected reference point is the amount of the original premium paid - in other words, in claiming their accident damages, policyholders seek to claim back the original price paid for their insurance. The pattern does not hold for other potentially salient levels of damages, nor is it consistent with an inclination to merely inflate claims more generally. Since damage claims can only be very precisely manipulated by the policy holder after an accident has occurred, the findings can be attributed primarily to ex-post moral hazard rather than adverse selection or ex-ante moral hazard. Furthermore, the pattern exists only for individually-held policies but not institutionally-held policies, which is most consistent with a reference point in individual utility. Our study illustrates the potential of behavioral economics frameworks in assisting with identification challenges for classical information economics problems.

JEL Codes: D82, D90, G22

Keywords: moral hazard, insurance markets, reference dependence, asymmetric information, automobile insurance

---

[1] † Gao: Department of Finance, School of Economics and Management, Tsinghua University, Beijing, China, gaof@sem.tsinghua.edu.cn ; ‡ Lien: Department of Decision Sciences and Managerial Economics, The Chinese University of Hong Kong, Shatin, Hong Kong, China, jaimie.academic@gmail.com ; ± Zheng: Department of Economics, School of Economics and Management, Tsinghua University, Beijing, China, jie.academic@gmail.com ; We are thankful for comments from Douglas Bernheim, Vincent Crawford, Stefano DellaVigna, Makram El-Shagi, Morten Lau, Zeng Lian, Benson Leung, David K. Levine, Ulrike Malmendier, Seung-Keun Martinez, Enrico Moretti, Shiyuan Pan, Stephen Rassenti, Adam Sanjurjo, Charles Sprenger, Jason Somerville, Justin Sydnor, Xiaoquan Wang, Sarit Weisburd, Zhicheng (Phil) Xu, Keyu Wu, Jipeng Zhang, Zibin Zhang, Xiwei Zhu, and seminar participants at School of Economics - Zhejiang University, International Business School - Beijing Foreign Studies University, School of Economics – Henan University, Institute of Urban Development – Nanjing Audit University; Workshop on Behavioral and Experimental Economics at Southwestern University of Finance and Economics, Inter-departmental Research Workshop at Chinese University of Hong Kong Business School, session participants at the Economic Science Association World Meeting 2017 in San Diego, participants of the 2019 Stanford Institute for Theoretical Economics, Psychology and Economics Workshop, and session participants at the Econometric Society World Meeting 2020 (online). We thank the anonymous insurance company for providing the data for this study. We gratefully acknowledge funding from the National Natural Science Foundation of China, Tsinghua University, The Chinese University of Hong Kong, and Research Grants Council of Hong Kong. All errors are our own.

## 1. Introduction

Moral hazard is a widely acknowledged but difficult to detect problem in most insurance contexts. The key challenge is that the insured hold private information about the behaviors they pursue after becoming insured, which cannot be easily verified by the insurance company or by the researcher. This information asymmetry can create incentives for the insured to engage in various activities leading to higher subsequent insurance claims than they would have in the absence of insurance.

For this reason, empirical evidence on moral hazard in insurance markets is sparse relative to its theoretical importance and welfare implications. Thus far, the literature has largely focused on panel data approaches to detecting information asymmetries such as adverse selection and moral hazard (see Chiappori, Jullien, Salanie and Salanie, 2006). In the case of moral hazard, policyholders who are in higher coverage plans exhibit more risky behavior, and in the case of adverse selection, more risky policyholders select higher coverage plans. Thus in either case, theory predicts a positive correlation between coverage and risky behavior. Additional features of the data may be needed to differentiate these two sources of positive correlation using this standard approach.

In this paper, we provide clear evidence of moral hazard in automobile insurance markets, using policyholder level data from a large insurance company in China. Differing from prior studies on moral hazard in insurance contexts, our method of detection relies on the reference-dependent preferences of the policyholders. We find that policyholders who have been involved in an automotive accident are disproportionately likely to make claims which cover their premium originally paid for the insurance. Since neither the potential selection of policyholders into insurance plans, nor possible (ex-ante) moral hazard in driving behavior alone, can so precisely generate the bunching of claims so closely to the premiums originally paid as we find here, our analysis identifies the "ex-post moral hazard" in insurance claims. Alternative asymmetric information sources such as adverse selection and ex-ante moral hazard are unable to account for this empirical pattern. That is, the distributional patterns of claim timings demonstrate that decisions made by policyholders after the accident drive the bunching of claims around the premium level.

In addition, when examining the claims of individually held insurance policies compared to institutionally held policies, the reference-dependent behavior is present only among policies held at the individual level but not at the institutional level. This provides further support for the notion that the bunching of claims around premium levels is driven by individual preferences, but not any external insurance policy rule, which would apply equally to individually and institutionally held policies.

Our study contributes to the literature which discusses and empirically tests for moral hazard in the automobile insurance context, which has found mixed evidence.[2,3] Among them, Abbring, Chiappori

---

[2] A separate but related literature tests for adverse selections in insurance markets. Cohen (2005) provides evidence for adverse selection in the automobile insurance market, finding that insurance switchers are disproportionately more likely to have poor claims histories. It is also noted in their discussion that a positive correlation between insurance coverage and risk in the automobile insurance market has sometimes been elusive, and has sparked discussion in the literature regarding the best ways to test for asymmetric information (p. 198 – 199).

[3] A broader literature on moral hazard in other insurance markets also finds some degree of supporting evidence, particularly in the health insurance domain. Aron-Dine, Einav, Finkelstein and Cullen (2015) find that health care utilization is responsive to dynamic prices, which in the literature, is interpreted as a type of moral hazard. Wagstaff and Lindelow (2008) find that health insurance in China increases the chance of high subsequent spending, but attribute this to the insurance coverage encouraging individuals to seek health care when they are sick. The RAND Health Insurance Experiment allowed a large

and Pinquet (2003) develop a test of moral hazard apart from adverse selection by using dynamic insurance data. However, using French car insurance data, they find no significant evidence of moral hazard. Dunham (2003) finds that vehicles not driven by their owners depreciate faster, indicative of moral hazard. Using a dynamic structural model on car insurance in Italy, Ceccarini (2007) finds evidence that driving effort may be increasing in the marginal price increase of an accident. Abbring, Chiappori and Zavadil (2008) also examine the dynamics of policyholders' claim incentives and test their model structurally using Dutch insurance data, and find support for moral hazard. Wang, Chung and Tzeng (2008) find support for moral hazard in dynamic automobile insurance coverage and claims data from Taiwan, based on responses to increases in deductibles. Dionne, Michaud and Dahchour (2013) utilize a dynamic model to separately identify moral hazard from adverse selection and learning effects.

While these studies focus on the dynamics of coverage and claims, other studies exploit differences in contracts or policies which alter policyholders' incentives for moral hazard. For example, Schneider (2010) finds evidence of moral hazard in the leasing contracts of New York City taxi drivers. Weisburd (2015) finds that reductions in accident costs due to an employer coverage policy in Israel increases the probability of an accident, and estimates the effect of moral hazard. Ma (2021) finds evidence for moral hazard in Chinese automobile insurance policyholders' increased accident likelihood at the end of the policy period.

While all the above papers focus on ex-ante moral hazard, a smaller number of studies targets the ex-post moral hazard phenomenon, which is the topic of our focus. The insurance literature refers to exploitation of the information asymmetry between policyholder and insurance company which occurs before an accident (ie. ranging from less careful driving to staging of accidents) as ex-ante moral hazard, while exploitation of such information asymmetry after an accident occurs is referred to as ex-post moral hazard (Rowell and Connelly, 2012). As described specifically in Abbring, Chiappori and Zavadil (2007), "Ex-post moral hazard concerns the effects of incentives on claiming actual losses".[4] The most frequently discussed type of ex-post moral hazard is related to fraud or other types of dishonest behavior which can occur after the accident itself. Among the relevant studies on ex-post moral hazard, Cummis and Tennyson (1996), utilize survey data on dishonest insurance behavior, combined with automobile insurance claims frequency, and find support for the existence of moral hazard. Hoyt, Mustard and Powell (2006) examine how state legislation and market conditions contributed to the reduction of automobile insurance fraud in the U.S. from 1988 to 1999.

Compared to these studies, our analysis utilizes a cross-sectional analysis of claims data, identifying misreporting through the claims distribution. One benefit of a cross-sectional approach is a reduced data requirement - individuals may in practice seldom experience changes to their insurance plans over time, which can make dynamic identification approaches difficult to implement in practice. The approach proposed in our study does not require dynamic panel data. Our study also departs from previous moral hazard studies in detection via a specific type of moral hazard behavior, ie. reference-dependent claiming behavior – and thus provides new insight about policyholders' decisions under

---

number of studies on moral hazard in the health insurance domain. Due to space considerations, we omit discussion of them here, but refer to the survey article by Aron-Dine, Einav and Finkelstein (2013) which reviews the main findings.

[4] See also Rowell and Connelly (2012), which provides a thorough historical account and discussion of the term *moral hazard*.

reference-dependent preferences, and tests whether the data largely hold to the predictions generated by an insurance claims model which incorporates reference-dependent preferences.

In identifying the premium as a common reference point for policyholders, our study follows the empirical literature on reference-dependence, which has expanded to provide empirical evidence of reference-dependent behavior in a substantial range of field contexts, including professional basketball (Berger and Pope, 2011), professional golf (Pope and Schweitzer, 2011), high stakes game shows, (Post, Van den Assem, Baltussen and Thaler, 2008), poker gambling (Eil and Lien, 2014), slot machine gambling (Lien and Zheng, 2015), labor supply (Camerer, Babcock, Lowenstein and Thaler, 1997; Fehr and Goette, 2007; Crawford and Meng, 2011), domestic violence (Card and Dahl, 2011), and other important settings. Of particular relevance to our study due to our focus on salient claims ratios, Pope and Simonsohn (2011) find that SAT takers and baseball players modify their behavior disproportionately around round number scores in order to keep their performances above their reference level. Similarly, Allen, Dechow, Pope and Wu (2017) find that marathon runners bunch their completion of the race around round number finishing times, consistent with reference-dependent models. We adopt heavily from the empirical methodology in Allen et al. (2017) in detecting reference-dependent behavior. It also bears similarity in approach to Lien and Zheng (2015) in detecting reference-dependence in a cross-sectional approach by comparing the actual distributions to hypothetical distributions without reference-dependence.[5]

Our study is also related to a recent literature examining individuals' reference-dependent responses to various government policies at intuitive reference points. Rees-Jones (2018) estimates U.S. taxpayers' efforts to reduce their tax liability, and shows evidence that individuals are loss averse in terms of having a positive amount owed on tax day. Jones (2019) finds that homeowners are loss averse to increases in their property tax owed, and are disproportionately more likely to appeal the tax assessment when faced with such increases. In a similar vein but in the online auction domain, Lien, Xu and Zheng (2017) finds that bidders in online penny auctions are disproportionately less likely to bid in subsequent auctions if they have incurred a loss. Our findings in the current paper bear similarity to these studies in terms of finding bunching behavior at a reference point, but differ in that there is no monetary consequence for the reference-dependent behavior detected.[6] Highlighting the differential effects of material and psychological incentives in reference-dependence, Seibold (2020) finds that German citizens' retirement decisions are heavily concentrated at the suggested (statutory) government age levels, in a manner consistent with reference-dependence, rather than based on the discontinuous financial incentives posed by their pension plans.

Insurance decisions have been previously discussed in the theoretical and empirical literature on reference-dependence. On the empirical side, Sydnor (2010) examines homeowners' choices of insurance plans and finds that policyholders select plans which are risk averse to a degree which is difficult to explain without reference-dependent preferences. Barseghyan, Molinari, O'Donoghue and Teitelbaum (2013) study automobile and homeowner insurance choices and find support for probability

[5] These statistical methods also bear resemblance to an experimental economics literature on detecting dishonesty, such as the dice technique proposed by Fischbacher and Follmi-Heusi (2013) and subsequent literature, also later implemented in field experiments such as Hanna and Wang (2015).

[6] In this sense, the findings in our study bear greater similarity to those in Allen et al (2017), in that decision-makers hold a reference point which is purely psychological in nature, but without material consequences for outcomes marginally above or below the reference level. While both types of reference points can be influential, our study highlights the importance of psychologically salient reference points, especially in those situations that lack the corresponding material incentives.

weighting (overweighting of small probabilities) by policyholders. While their main focus is in testing the Koszegi and Rabin (2006) reference-dependence and Gul (1991) disappointment aversion frameworks, they also examine the possibility of ex-post moral hazard. Notably, they do not find evidence for moral hazard in their claims data. One possibility is that since their data are from a U.S. insurance company, the auditing and claims process could be more rigorous than in countries whose insurance markets are still in development. While claiming back the premium paid could be highly attractive irrespective of population studied, administrative procedures might preclude such behavior in some well-developed insurance markets.[7] Therefore, an additional contribution of our study is that it provides an opportune context to understand how insurance consumers view insurance contracts and the claims decision in a less restrained setting.

Theoretically, the endogenously determined expectations-based reference point model of Koszegi and Rabin (2006) allows for an understanding of insurance purchase decisions under reference-dependent preferences. By typically assuming lagged status quo as the reference point, the literature prior to Koszegi and Rabin had difficulty explaining why individuals would purchase insurance in the first place. In the Koszegi and Rabin framework, individuals have rational expectations about their future insurance needs given their behavior, and purchase an insurance plan accordingly. Under this explanation, the premium paid for the insurance plan is not coded by the policy holder as a loss.[8] Some recent studies such as Gneezy, Goette, Sprenger and Zimmermann (2017), Baillon, Bleichrodt and Spinu (2020), and Goette, Harms, and Sprenger (forthcoming) find empirical limitations of expectations-based reference dependence.

Our analysis shows there are at least a sizable number of individuals who hold the lagged status quo as a reference point in the insurance context. One possibility is that given the insurance mandate for car owners, policyholders may still hold lagged status quo as a strong reference point, since the decision to take-up the insurance is not made fully voluntarily. However, the insurance mandate alone cannot explain why some car owners purchase insurance plans with greater coverage and higher premiums. Another possibility is that policyholders might have originally held lagged status quo minus the premium paid as their reference point, but over time adjusted their reference point upward to lagged status quo. Finally, another factor is that in an insurance context, it may be substantially difficult for individuals to have accurate forward-looking expectations about their eventual insurance needs when making the original purchase decision.

Our study is the first to our knowledge, to detect moral hazard in insurance markets by utilizing a reference-dependent approach, and simultaneously, is also one of relatively few studies to confirm the presence of moral hazard in the auto insurance industry, in particular for ex-post moral hazard. We highlight the contribution of reference-dependent utility in enabling us to identify moral hazard behavior, a question which has been long posed as a challenging one in the literature on markets with asymmetric information. Without allowing for the possibility of reference-dependent utility of policyholders, the claims behavior we observe in the data would in itself be difficult to explain. Furthermore and importantly, reference-dependence provides a new way of solving the important empirical question of identifying moral hazard, which can be applied to other settings even under limited panel data or purely cross-sectional data.

---

[7] Barseghyan et al (2013) also note that "…the empirical evidence on moral hazard in auto insurance markets is mixed." Our study provides positive evidence for moral hazard in the automobile insurance setting.

[8] For a summary on the relationship between behavioral models of risk aversion and insurance choices, see O'Donoghue and Somerville (2018).

Finally, our paper contributes to the study of insurance attitudes in China, where the insurance industry is in rapid development, quickly becoming the largest insurance market in the world. Fang (2013) provides an overview of insurance markets in China. On the issue of asymmetric information, Gao and Wang (2011) find evidence of adverse selection in China's auto insurance market through the positive empirical relationship between chosen coverage and claims. Gao, Powers and Wang (2017) decomposes the effects of adverse selection, ex-ante and ex-post moral hazard in China's auto insurance market, finding strong evidence of adverse selection, fairly little effect of ex-ante moral hazard, and some ex-post moral hazard among low-coverage policy holders. Our paper differs from these prior studies in using a reference-dependent approach for detecting moral hazard, as well as in our focus on positively detecting moral hazard patterns across several types of policy holders. In addition, our current study does not provide any definitive evidence for adverse selection or ex-ante moral hazard in this setting, detecting instead the ex-post moral hazard which occurs after an automobile incident.

The remainder of the paper is organized as follows: Section 2 describes the automobile insurance context, data and empirical approach; Section 3 provides a theoretical framework of reference-dependent claims behavior of policyholders which provides the hypotheses for our empirical tests; Section 4 details the results; Section 5 discusses alternative reference point candidates and robustness checks; Section 6 concludes.

## 2. Context and Data

The automobile insurance industry has evolved rapidly in China over recent decades, coinciding with the rise of vehicle ownership in the country, which currently makes it the largest automobile market in the world.

Automobile insurance in China works in largely the same way as in the United States and other countries around the world. Policyholders pay a fee or premium upfront for insurance coverage for a period of time, typically one year. After a collision, the claims process typically works as follows: A law enforcement officer is called to the scene to determine which party is at fault in the accident. The insurance company of the party at fault will be responsible for paying the corresponding damages, and each party involved will make a claim directly to that insurance company for reimbursement.

Official receipts with stamp ("发票") from each party's chosen auto repair shop are needed for the reimbursement process. The policyholder who is at fault is the last party to submit their claim to their insurance company, and typically has access to the information about the claims made by the other parties involved in the accident. The claims in the dataset, which must be accompanied by official receipts, thus represent the sum of reported retail value of repairs made at the auto repair shops of all parties involved.[9]

Our data consist of all auto insurance policies with claims from a single province from one of the largest insurance companies in China, in the years 2008 and 2009. The aforementioned claim procedure

[9] A natural question arises about the exact mechanism for the claim manipulation. While we do not focus on demonstrating the precise mechanism in this study due to the data limitations, we can offer some anecdotal possibilities. For example, the policyholder visits the auto repair shop after the accident, and suppose that the actual value of repairs assessed is lower than the policyholder's desired claim. The policyholder may ask the repair shop to also repair or improve other components of their vehicle which were not damaged in the accident itself, up to the amount of the desired claim. Based on the insurance claim rules and repair shop practices during the time period of our data, we believe that this is the likely mechanism for the claim manipulation.

implies that the policyholders in the data were at least partially held responsible in the incident. We observe the total claim made on a particular policy for a single accident.

Table 1 shows the summary statistics for our data set. Policyholders spanned a range of ages averaging age 40, while the cars themselves were on average just one year old. The value of the cars varied widely. Premiums are determined by a formula used by the insurance company which typically takes into account automobile characteristics; policyholders can select the premium they would like to pay corresponding to their desired coverage. The average premium paid was 4,721 Yuan, while the average claim amount was 1,993 Yuan. Policyholders in the dataset made an average of 2 claims per year.[10] There are a total of 78,102 insurance policies in the data.

**Table 1: Descriptive Statistics**

| | *Average* | *Standard Deviation* | *Minimum* | *Maximum* |
|---|---|---|---|---|
| Policyholder age | 39.58 | 8.55 | 18 | 70 |
| Car age | 0.99 | 1.58 | 0 | 14.17 |
| Car value | 131,144.4 | 120,840.1 | 10,000 | 3,260,000 |
| Total premium | 4,720.95 | 3,056.04 | 785.11 | 70,954.19 |
| Claim amount | 1,993.23 | 8,148.73 | 10 | 434,108 |
| Number of claims | 2.07 | 1.48 | 1 | 18 |
| **Total Policies** | 78,102 | | | |

Rules regarding use of insurance policies vary by company and by time period, but may include dynamic monetary incentive features which do not interfere with our findings. For example, if making zero claims in a policy period, insurance companies typically offer their policyholder a discount in the next policy period. The behavioral result of this policy is typically private settlement of sufficiently small damages between parties in the accident, activity which lies far below the bunching of claims we detect around the premium. On the other extreme, if a policyholder makes claims beyond a certain threshold, their premium could increase the following year. This threshold varies by company but is typically set to over twice to three times the premium amount, and so does not explain our findings.

Finally, the insurance company typically sets a maximum claim amount, which is generally not a binding constraint for most of the accidents and claims that occur in the data.

### 2.1 Ex-ante Moral Hazard

The abnormalities in claim distributions we find here are difficult to reconcile with adverse selection, because adverse selection alone would not deliver such a precise concentration of claims amounts at the premium level. Similarly, ex-ante moral hazard alone cannot be primarily held accountable for our findings either. In other words, even if policyholders were to get into accidents

[10] While an average of 2 claims per year may appear high compared to U.S. claim statistics, one feature of driving conditions in developing countries is a higher frequency of accidents which are often not as severe in nature. For an explanation, see Jung, Lee, Lien and Zheng (2020).

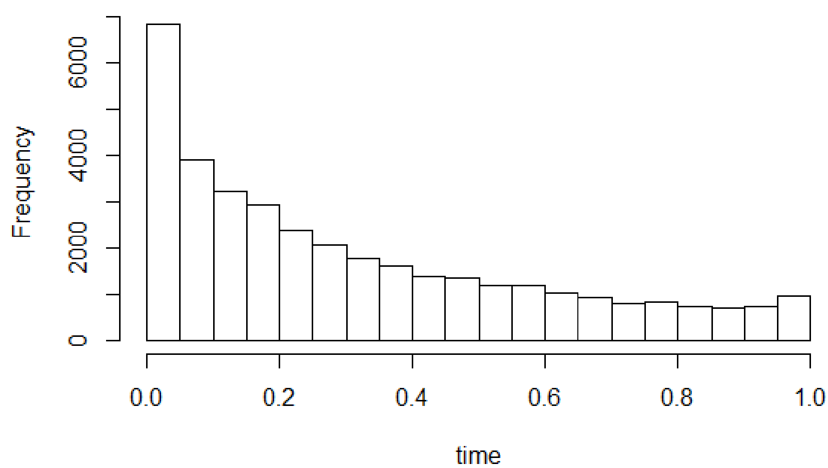
deliberately due to insurance collection motives, it would be difficult to directly generate damages so near to the premium amount without some type of ex-ante moral hazard involved. However, the possibility cannot be ruled out that policyholders may exercise a combination of ex-ante and ex-post moral hazard to generate the claims we observe. In other words, it may be possible that policyholders exercise ex-ante moral hazard in deliberately getting to some type of accident or exercising less caution in their driving behavior, then choosing ex-post moral hazard in claiming the actual damages.

In order to obtain an assessment of to what degree ex-ante moral hazard could be supplementing ex-post moral hazard behavior, we can examine the frequency of accidents over the policy period. Given that making total annual claims beyond a threshold (typically between 2 to 3 times of the premium amount) will result in a higher premium level in the next year, there is an optimal timing for policyholders' ex-ante moral hazard.

Getting into an accident too early in the policy period increases the likelihood to exceeding the threshold. Since any given period carries a risk of accident independently of policyholder behavior (ie. uncontrollable factors such as traffic levels, other drivers' behavior, etc.), we can expect a risk averse policyholder to wait until near the end of the policy period to exhibit ex-ante moral hazard, assuming no prior accidents have occurred for that policyholder thus far. Thus, if ex-ante moral hazard of a deliberate nature is playing a significant role in the market, we should expect to see more first accidents occurring late in the policy period.

Figure 1 shows the actual histogram of first accidents claimed over the policy period, where the horizontal axes normalizes the policy periods across a single zero-to-one scale. The Figure shows that in contrast to the ex-ante moral hazard hypothesis, first accidents tend to occur in the early part of the policy period, and exhibit a steeply downward slope over time until the very end of the policy period at which point the frequency increases slightly.[11] Thus, we infer that ex-ante moral hazard, at least in a deliberate form, is not substantially responsible for the patterns we find in Section 4.

**Figure 1: Timing of First Accidents, Histogram**




[11] A spike in accidents at the end of the policy period is specifically studied in Ma (2021) in the Chinese automobile insurance context, as evidence for ex-ante moral hazard under the sunk cost fallacy. Such a spike does not feature as prominently in our data compared to in the proprietary data in their paper, although such effect may be present in our data as well to a degree. Since our main hypotheses are focused on the ex-post claiming behavior, we do not investigate further into the potential end-of-policy effect.

## 2.2 Methodology

To detect reference-dependence in the post-accident claim amounts, we closely follow the methodologies employed by Allen, Dechow, Pope and Wu (2017) in their analysis of marathon finishing times: First, we utilize the non-parametric McCrary (2008) test, designed to check whether a running variable has been manipulated in a regression discontinuity framework. In our setting, the variable under potential manipulation is the claim amount, or conditional on the premium paid, the claims to premium ratio. Second, we implement the test used in Chetty, Friedman, Olsen and Pistaferri (2011) which fits a hypothetical distribution to tax data in order to detect the excess bunching of income at kinks in the tax schedule.

While the Chetty et al test may be more robust in detecting reference-dependent behavior, as suggested in Allen et al. (2017), our setting bears a slight difference compared to the settings of either marathon running or bunching of income at kinks in a tax schedule: reference-dependent moral hazard predicts an upward manipulation in claims, whereas the marathon and tax settings predict a downward manipulation in finishing times and income earned, respectively. Furthermore, the manipulation in claims is much easier to achieve with precision in the case of auto insurance compared to the case of marathon finishing times, where individual ability poses a constraint on the outcome. For this reason, we conduct both the McCrary test and the Chetty et al test, and we consider the result robust for our purposes only if both tests indicate statistically significant manipulation.

The McCrary test is conducted by comparing the densities on the two sides of the cutoff point, where the densities are estimated using the two-step method provided in McCrary (2008). The first step is to compute histogram using bin size $\hat{b} = 2\hat{\sigma}n^{-1/2}$, where $\hat{\sigma}$ is the sample standard deviation of claims. The second step is local linear smoothing of the histogram to estimate the density, conducted separately for the bins to the right and left of the cutoff point. The bandwidth used in local linear smoothing is based on the automatic bandwidth selection procedure in McCrary (2008).[12]

Similar to the McCrary test, our implementation of the Chetty et al. approach uses 7th order polynomials to directly approximate the entire counterfactual distribution which would occur without reference-dependence.[13] The excess density in the data compared to the estimated counterfactual distribution is calculated, and standard errors are determined by bootstrap. Compared to the McCrary approach, the Chetty et al test focuses on a small window of the running variable where the excess pileup of density is hypothesized to occur. We test several nearby windows around candidate reference points to identify the location of the irregularity in claims, which in turn informs us regarding the location of the reference point.

## 3. Theoretical Framework

We model the claiming decision of an automobile insurance policyholder who has just gotten into an accident. The purpose of the model is to deliver comparative statics results whose predictions we subsequently test using the relevant subsamples of the data in Section 4. In focusing on the ex-post moral hazard motive, we note that only the claiming behavior that occurs after an accident has the potential to be primarily responsible for the bunching of claims we find in the data, as discussed in Section 2.1.

---

[12] For details, we refer the reader to McCrary (2008), p. 705.

[13] We utilize a higher order polynomial than that in the McCrary test in order to better fit the distribution. However, we also test the robustness of the procedure to 6th and 8th order polynomials, and the results are robust to those reported in the paper.

We focus on the decision of the policyholder in making a claim $\tilde{c}$ to the insurance company when an accident of actual damage $c$ occurs, where $c$ can be any number over the interval $(0,+\infty)$. In general, assuming that none of the previous claims made has exceeded the reference point, the policyholder has the following reference-dependent utility function,

$$U(\tilde{c}) = u(w) - \beta |\tilde{c} - c| + \begin{cases} \eta(\min\{\tilde{c},\overline{c}\} - r) & if \quad \tilde{c} \geq r \\ -\lambda\eta(r - \tilde{c}) & if \quad \tilde{c} < r \end{cases}.$$ [14]

where $w$ is the policyholder's wealth level, $r$ is the reference point against which he or she values the final outcome in terms of wealth, $\lambda > 1$ is the loss aversion parameter, $\eta > 0$ measures the marginal effect for the gain loss utility, and $\beta > 0$ is a parameter measuring the cost of misreporting, which can be interpreted as either the intrinsic cost of lying, the penalty for being caught for claims fraud, or both. For simplicity, the misreporting cost is modeled as symmetric around the absolute difference between the claim and damage amount, although the comparative statics predictions on the empirical reference-dependence patterns do not depend on this assumption. As a reasonable assumption, we also assume that the maximum coverage $\overline{c}$ of the insurance policy is greater than $r$. Although we use loss aversion here as the specific form of reference-dependent preference in our model, Allen et al (2017) demonstrates that either reference-dependence without loss aversion, or aspiration utility with a discontinuity at the reference point all generate bunching in cross-sectional data.

The optimal claiming strategy $\tilde{c}^*$ depends on the rank ordering among the parameters $\eta, \beta, \lambda\eta$, as well as the relative sizes of $c$, $\overline{c}$ and $r$, and is described in the following lemma. The proof is relegated to the Appendix.

**Lemma 1:** ***(Claiming Strategy)*** For an individual with parameter profile $(\eta,\lambda,\beta,c,r)$, when $\beta > \lambda\eta$, $\tilde{c}^* = c$; when $\beta < \eta$, $\tilde{c}^* = \begin{cases} c & if \quad c \geq \overline{c} \\ \overline{c} & if \quad c < \overline{c} \end{cases}$; when $\eta < \beta < \lambda\eta$, $\tilde{c}^* = \begin{cases} c & if \quad c \geq r \\ r & if \quad c < r \end{cases}$.

Notice that misreporting of claims occurs in two cases: (1) $\beta < \eta$ while $c < \overline{c}$; (2) $\eta < \beta < \lambda\eta$ while $c < r$. The first case refers to the scenario with a very low cost of misreporting, which results in a misreporting of claims up to the maximum claim allowed. A negligible cost of misreporting is unlikely to be plausible in reality, and furthermore, claiming up to the maximum is not the empirical pattern observed in the data. Thus, from here on we focus on the second scenario under which misreporting occurs, which is a type of reference-dependent behavior with claims inflated to the reference point.

[14] For simplicity, we assume that there is no reference-dependent term in the policyholder's utility function when a previous claim has exceeded the reference point. A relaxation of this assumption will not change the theoretical predictions as long as the reference-dependent effect is significantly smaller when the reference point has already been achieved in an earlier claim. In a more general setup, we also allow for a probability of failed misreporting, as well as curvature in the reference-dependent utility, and the main results hold.

### *3.1 Static Predictions*

Let $\Pr(\tilde{c}^* = r)$ denote the probability that the policyholder reports their reference point (for example, the premium) as their claim. In a static setting where a policyholder follows the claiming strategy described by Lemma 1, we have the following two propositions regarding the likelihood of reference-dependent reporting, $\Pr(\tilde{c}^* = r)$. The proofs are relegated to the Appendix.

**Proposition 1:** ***(Reference point level)*** Policyholders with a higher reference point are more likely to exhibit reference-dependence, that is, $\frac{\partial \Pr(\tilde{c}^* = r)}{\partial r} > 0$.

The reasoning for this result is that the higher the reference point $r$, the higher the likelihood that the actual damage level $c$ is less than $r$, which corresponds to a higher likelihood of the reference-dependent claim behavior as specified in Lemma 1.

**Proposition 2:** ***(Loss Aversion levels)*** Policyholders with greater levels of loss aversion exhibit greater reference-dependent claim behavior, that is, $\frac{\partial \Pr(\tilde{c}^* = r)}{\partial \lambda} > 0$.

Similar to the reasoning in Proposition 1, a higher loss aversion parameter $\lambda$ leads to a greater range of misreporting cost parameters $\beta$ for which reference-dependent claims result under the scenario that $c < r$.

### *3.2 Dynamic Predictions:*

We further examine how a policyholder dynamically decides when to recover the "psychological cost" associated with the reference point over time. Let $\tilde{c}^*_{i,t_i}$ be the policyholder's optimal claiming strategy for the $i$-th accident that occurs in time $t_i$, $i = 1, 2, \cdots$. Suppose that the number of accidents occurring during a policy period follows a Poisson distribution, so that the arrival time of the next accident follows an exponential distribution with probability density function $f(t) = \mu e^{-\mu t}$, where $\mu > 0$ is the rate parameter. The following two propositions result, and the proofs are relegated to the Appendix.

**Proposition 3:** ***(Claim Number)*** Policyholders' first claims exhibit greater reference-dependent bunching compared to subsequently made claims, as long as there is a sufficiently low likelihood of an accident at any given time, that is, $\exists \mu^* > 0$, such that $\forall \mu < \mu^*$, $\forall i > 1$, $\Pr(\tilde{c}^*_{1,t_1} = r) > \Pr(\tilde{c}^*_{i,t_i} = r)$.

The intuition is that as long as the probability of having a subsequent accident is sufficiently small, a reference-dependent policyholder will take the first opportunity to reach their reference level of wealth through the claims process. The resulting hypothesis is that the distribution of first claims exhibits a greater degree of reference-dependent bunching of claims than the distribution of subsequent claims.

**Proposition 4:** ***(Claim Timing)*** Claims $i$ made later in an individual policy period exhibit greater reference-dependent bunching compared to claims $i$ made earlier in the policy period, that is, $\forall i = 1, 2, \cdots$, $t_i' > t_i \Rightarrow \Pr(\tilde{c}^*_{i,t_i'} = r) > \Pr(\tilde{c}^*_{i,t_i} = r)$.

Following similar reasoning as the result of Proposition 3, accidents occurring later in the policy period have a lower likelihood of being followed by another accident within the same policy period. The reference-dependent policyholder thus tends to exhibit a greater tendency to recover their reference level of wealth for a later-occurring accident through the claims process.

## 4. Results

We report the results of our empirical tests, beginning with those results which have theoretical comparative statics hypotheses based on our previously described model framework. Our approach in testing these hypotheses is by means of a sample splitting method, running the tests described in Section 2.2 on each relevant subsample of the policy data to test whether the predictions of the model are realized.

Each set of results consists of the McCrary test and the Chetty et al test. Kernel density plots based on the counterfactual densities estimated in the Chetty et al test are provided merely to show the reference-dependent claims visually. However, please note that the detection of statistically significant excess mass should be drawn from the Chetty et al test tables displayed in the text as well as in the Appendix for alternative candidate reference points, and not from the visualization in the density plots alone. Due to space considerations, these and other supplementary results, graphs and robustness checks are provided in the Appendix.

### 4.1 Static Predictions: Premium Level and Premium Rate

*4.1.1 Premium Level*

We begin with the static results on policyholders' premium level, which is the price the policyholder pays for the insurance plan, and the "premium rate", which measures the insurance company's own assessment of the riskiness of the policyholder. The premium level paid by the policyholder holds special meaning in our findings because our results show that policyholders hold this amount as a reference point during the insurance process. One question is whether this reference-dependent effect is stronger depending on the premium amount.

We find that the desire to recover the premium paid is concentrated primarily among high premium policyholders, consistently with our theoretical Proposition 1. The McCrary test results for premium levels are shown in Table 2A, which divides the premium amounts into above median and below median levels. While policyholders with premiums of 4129 Yuan and below did not show significant manipulation of claims, those with premiums above this amount were significantly reference-dependent around the 1.00 ratio. Table 2C decomposes the effect further by premium quartile using the McCrary test, showing that the effect is indeed concentrated among both the 3rd highest quartile and the highest quartile. The Chetty et al test confirms these results clearly, with the empirical distribution differing from the estimated counterfactual distribution for several bin ranges near the 1.00 ratio among high premium policies, but not among low premium policies.

**Table 2A: McCrary density test, Premium level**

| | *Low Premium (lower 50th percentile, 4129 Yuan and below)* | | | | | | |
|---|---|---|---|---|---|---|---|
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| statistic | 0.0134 | 0.2318 | 0.0714 | -0.0026 | 0.0399 | 0.0773 | 0.1545 |
| *p-value* | *0.9263* | *0.1436* | *0.6442* | *0.9864* | *0.7996* | *0.6288* | *0.3458* |
| | *High Premium (upper 50th percentile, above 4129 yuan)* | | | | | | |
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| statistic | 0.2412 | 0.2395** | 0.4667*** | 0.5121*** | 0.1591 | 0.0508 | -0.0423 |
| *p-value* | *0.1110* | *0.0391* | *0.0042* | *0.0022* | *0.2953* | *0.7246* | *0.7678* |

*significant at 10% level, ** significant at 5% level, *** significant at 10% level

**Table 2B: Chetty et al test, Premium level**

*Low Premium (lower 50th percentile, 4129 Yuan and below)*

| reference point | bins | statistic | std dev | p-value |
|---|---|---|---|---|
| 0.97 | 5 | -0.0727 | 0.1579 | 0.6451 |
| 0.98 | 4 | 0.0942 | 0.1884 | 0.6169 |
| 0.99 | 3 | -0.0400 | 0.1995 | 0.8413 |
| 1.00 | 2 | -0.1607 | 0.2338 | 0.4918 |
| 1.00 | 1 | -0.1589 | 0.2843 | 0.5762 |

*High Premium (upper 50th percentile, above 4129 yuan)*

| reference point | bins | statistic | std dev | p-value |
|---|---|---|---|---|
| 0.97 | 5 | 0.2081 | 0.1482 | 0.1604 |
| 0.98 | 4 | 0.3055* | 0.1705 | 0.0731 |
| 0.99 | 3 | 0.5278*** | 0.2024 | 0.0091 |
| 1.00 | 2 | 0.7756*** | 0.2451 | 0.0016 |
| 1.00 | 1 | 0.9596*** | 0.3052 | 0.0017 |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

For illustrative purposes only, Figure 2 plots the ratio of excess density, the smoothed kernel density of the empirical distribution, divided by the counterfactual density estimated using the Chetty et al test. The excess density near the 1.0 ratio is visually apparent for the high premium subsample, but not for the low premium subsample, reflecting the statistical results in Tables 2A, 2B and 2C. Note that since we conduct the Chetty et al test specifically around candidate reference points (1.00, and other round number ratios in subsequent sections), the statistical significance of the excess density can be viewed only from the tables, but not from the Figures themselves. The tables for alternative candidate reference points 0.90 and 1.10 are provided in the Appendix.

**Table 2C: McCrary density test, Premium level**

*By quartiles, p-values shown in italics*

| | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
|---|---|---|---|---|---|---|---|
| **1st quartile** *(below 3214 yuan)* | 0.0168 | 0.2158 | 0.1467 | 0.0532 | 0.1667 | 0.1106 | 0.1890 |
| *pvalue* | *0.9280* | *0.2895* | *0.4624* | *0.7877* | *0.4171* | *0.5875* | *0.3741* |
| **2nd quartile** *(between 3214 and 4129 yuan)* | 0.1101 | 0.2463 | 0.0068 | -0.0172 | -0.0635 | 0.1410 | 0.2101 |
| *pvalue* | *0.6112* | *0.2837* | *0.9748* | *0.9383* | *0.7652* | *0.5482* | *0.3613* |
| **3rd quartile** *(between 4129 and 5318 yuan)* | 0.3069 | 0.5921** | 0.5101** | 0.5965** | 0.3185 | 0.0113 | -0.0430 |
| *pvalue* | *0.2086* | *0.0300* | *0.0337* | *0.0130* | *0.1896* | *0.9643* | *0.8628* |
| **4th quartile** *(above 5318 yuan)* | 0.2263 | 0.2128 | 0.5105** | 0.5065** | 0.1269 | 0.0300 | -0.1065 |
| *pvalue* | *0.2672* | *0.2742* | *0.0212* | *0.0246* | *0.5325* | *0.8827* | *0.5825* |

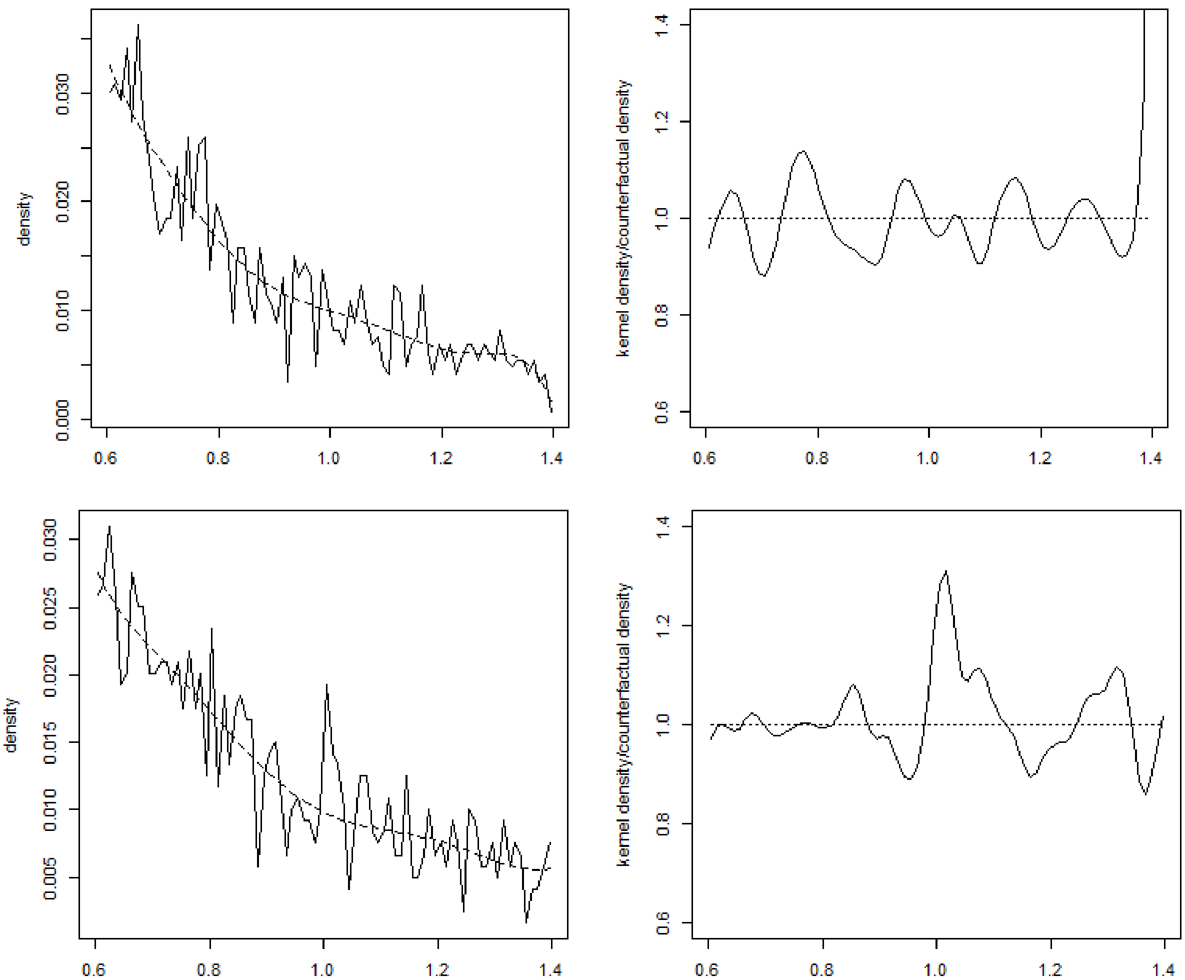
*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Figure 2: Ratio of Excess: Kernel Density / Counterfactual Density**

*Top Panels: Low Premium, Bottom Panels: High Premium*

*(Left Panels: raw frequencies; Right Panels: ratio of excess)*

statistical significance results around candidate reference points provided in Tables 2A, 2B, and Appendix




*4.1.2 Premium Rate*

The second premium-related measure we examine is the "premium rate", which is defined as the premium level of the policyholder divided by the total coverage amount of the policy. The premium rate is a variable used by the insurance company to measure the riskiness of a policy. Since the menu of premium levels and total coverage amounts available to the policyholder is determined by an actuarial formula, once conditioning on driver and automobile characteristics, the premium rate is an indication of the riskiness of the policyholder as assessed by the insurance company. Higher premium rates correspond to higher riskiness of the policyholder – ie. holding the premium amount fixed, the insurance company is only willing to offer a lower total coverage of the policyholder.

The premium rate serves as our best objective assessment of the risk attitudes of a policyholder, through their externally assessed risk-taking behavior. This allows for an opportunity to test Proposition 2, regarding the relationship between reference-dependent claims behavior and risk/loss attitudes.

**Table 3A: McCrary density test, Premium rate**

| | *Low Premium Rate (lower 50th percentile, rate of 0.0153 and below)* | | | | | | |
|---|---|---|---|---|---|---|---|
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| Test statistic | 0.1795 | 0.2519* | 0.2827** | 0.3996*** | 0.1986 | 0.1519 | 0.1357 |
| *p-value* | *0.1658* | *0.0713* | *0.0433* | *0.0051* | *0.1421* | *0.2754* | *0.3257* |
| | *High Premium Rate (upper 50th percentile, rate above 0.0153)* | | | | | | |
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| Test statistic | 0.0959 | 0.2969* | 0.2030 | -0.0037 | 0.0291 | -0.0146 | 0.0042 |
| *p-value* | *0.5495* | *0.0887* | *0.2566* | *0.9819* | *0.8618* | *0.9309* | *0.9800* |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table 3B: Chetty et al test, Premium rate**

*Low Premium Rate (lower 50th percentile, rate of 0.0153 and below)*

| **reference point** | **bins** | **statistic** | **std dev** | **p-value** |
|---|---|---|---|---|
| 0.97 | 5 | 0.0634 | 0.1433 | 0.6584 |
| 0.98 | 4 | 0.1668 | 0.1616 | 0.3020 |
| 0.99 | 3 | 0.2254 | 0.1833 | 0.2187 |
| 1.00 | 2 | 0.5096** | 0.2251 | 0.0236 |
| 1.00 | 1 | 0.7757*** | 0.2923 | 0.0080 |

*High Premium Rate (upper 50th percentile, rate above 0.0153)*

| **reference point** | **bins** | **statistic** | **std dev** | **p-value** |
|---|---|---|---|---|
| 0.97 | 5 | 0.0416 | 0.1396 | 0.7655 |
| 0.98 | 4 | 0.2180 | 0.1631 | 0.1813 |
| 0.99 | 3 | 0.1965 | 0.1813 | 0.2783 |
| 1.00 | 2 | -0.0254 | 0.2005 | 0.8992 |
| 1.00 | 1 | -0.1192 | 0.2608 | 0.6477 |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

Tables 3A and 3B show that manipulation of claims is significant for low premium rate policyholders, but not robustly so for high premium rate policyholders. In other words, policyholders who are judged to be less risky by the insurance company are *also* more reference-dependent around recovering their premiums paid. This is consistent with the convenient and traditional interpretation of utility functions in economics, which typically allows a single utility function to represent incentives across many possible domains. Individuals with higher degrees of reference-dependent loss aversion are correspondingly more risk averse, particularly near their reference point, and are likely to exhibit loss aversion in both the driving and insurance claims domains. Table 3C shows the McCrary test results for the data disaggregated by quartile. The results show that the reference-dependent behavior is driven by the quartile of policyholders directly below the median premium rate.

**Table 3C: McCrary density test, Premium rate**

*By quartiles, p-values shown*

| | 0.97 | 0.98 | 0.99 | 1 | 1.01 | 1.02 | 1.03 |
|---|---|---|---|---|---|---|---|
| **1st quartile** *(below 0.01362)* | 0.1286 | 0.1753 | 0.2163 | 0.2290 | -0.0285 | -0.0724 | -0.1254 |
| *p-value* | *0.4289* | *0.2895* | *0.1988* | *0.1823* | *0.8641* | *0.6651* | *0.4529* |
| **2nd quartile** *(0.01362 to 0.0153)* | 0.2562 | 0.4865** | 0.5757*** | 0.7006*** | 0.6528*** | 0.5508** | 0.5685** |
| *p-value* | *0.2256* | *0.0219* | *0.0057* | *0.0037* | *0.0036* | *0.0147* | *0.0123* |
| **3rd quartile** *(0.0153 to 0.01764)* | 0.0574 | 0.3049 | 0.3729 | -0.0175 | -0.0172 | -0.0843 | 0.0539 |
| *p-value* | *0.7840* | *0.1901* | *0.1381* | *0.9356* | *0.9396* | *0.6905* | *0.8166* |
| **4th quartile** *(above 0.01764)* | 0.2124 | 0.2691 | 0.1057 | 0.0694 | 0.1196 | 0.0766 | -0.0237 |
| *p-value* | *0.3296* | *0.2591* | *0.6415* | *0.7403* | *0.6052* | *0.7471* | *0.9174* |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

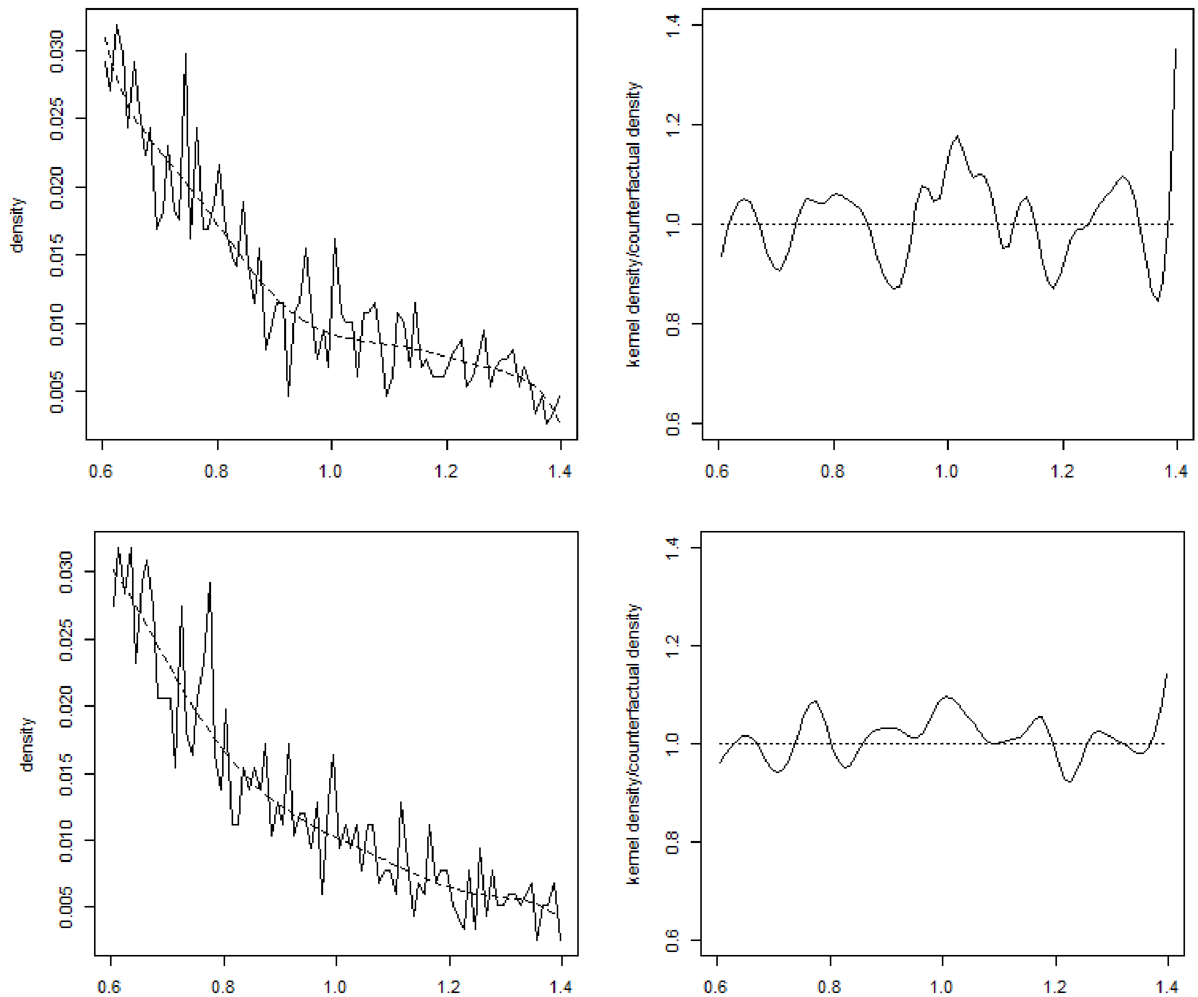
Figure 3 provides a visual display of the above results for low premium rate policyholders and high premium rate policyholders respectively. The increase in the ratio of excess around the 1.00 claims to premium ratio is more pronounced for the low premium rate policyholders, visible from the smoothed density plots in the column on the right.

**Figure 3: Ratio of Excess: Kernel Density / Counterfactual Density**

*Top Panels: Low Premium Rate, Bottom Panels: High Premium Rate*

*(Left Panels: raw frequencies; Right Panels: ratio of excess)*

statistical significance results around candidate reference points provided in Tables 3A, 3B and Appendix




## 4.2 Dynamic Predictions: Claim Number and Timing of Claims

### *4.2.1 Claim Number*

Proposition 3 states that policyholders with reference-dependent preferences over monetary gains and losses incurred throughout the insurance process will be more likely to bunch their claims around their premium level on their first accident compared to for subsequent accidents. The intuition is that a loss averse policyholder with the premium as the reference point will take his or her first possible opportunity to recover the premium amount with certainty, rather than incur the possibility of realized loss if no further opportunity arises to make a claim during the remainder of the premium period.

To test this hypothesis, we divide the sample between first claims made by a policy holder, and repeat claims. For first claims, we test for discontinuities and excess mass around the claims to premium ratio as usual. For repeat claims, in order to control for the claims already made earlier in the policy period, we conduct the tests around the ratio (Claim – Cumulative Prior Claims)/Premium. A reference-dependent policyholder with the premium paid as the reference point will prefer to claim back the premium at the first possible opportunity, but in the event of failing to claim it back on earlier instances, may still attempt to claim it back during subsequent accidents.

**Table 4A: McCrary density test, Claim number**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | *First claims* | | | |
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| Test statistic | 0.1134 | 0.2856** | 0.2397** | 0.2052* | 0.0714 | 0.0238 | 0.0144 |
| *p-value* | *0.3030* | *0.0143* | *0.0499* | *0.0899* | *0.5423* | *0.8379* | *0.9016* |
| | | | | *Repeat claims (second or more)* | | | |
| (Claim-Cumulative Prior Claims)/Premium | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| Test statistic | -0.0529 | -0.0500 | -0.0257 | 0.0923 | 0.0158 | 0.0178 | 0.0325 |
| *p-value* | *0.4440* | *0.4731* | *0.7187* | *0.2028* | *0.8282* | *0.8076* | *0.6615* |

*significant at 10% level, ** significant at 5% level, *** significant at 10% level

Both the McCrary test (Table 4A) and the Chetty et al test (Table 4B) show significant irregularities in the distribution of first claims around the 1.00 claims to premium ratio, while the distribution of the second and subsequent claims do not show any significant irregularities around the relevant ratio. Table 4A shows results for the McCrary test at each possible reference points between 0.97 and 1.03. The distribution of claims to premium ratios displays significant irregularities at 0.98, 0.99 and 1.00, all in close proximity to the 1.00 ratio. Table 2B shows results for the Chetty et al test, where the left-most column shows the candidate reference point, and the next column indicates the number of 0.01 length bins included in the test. The bin from 1.00 to 1.01 shows significant displacement of the distribution compared to the hypothetical estimated distribution.

**Table 4B: Chetty et al test, Claim number**

*First claims*

| candidate reference point | # of bins from reference point tested | test statistic | std error | p-value |
|---|---|---|---|---|
| 0.97 | 5 | 0.0532 | 0.1073 | 0.6204 |
| 0.98 | 4 | 0.1906 | 0.1228 | 0.1206 |
| 0.99 | 3 | 0.2119 | 0.1381 | 0.1249 |
| 1.00 | 2 | 0.2519 | 0.1634 | 0.1231 |
| 1.00 | 1 | 0.3460* | 0.2082 | 0.0967 |

*Repeat claims (second or more)*

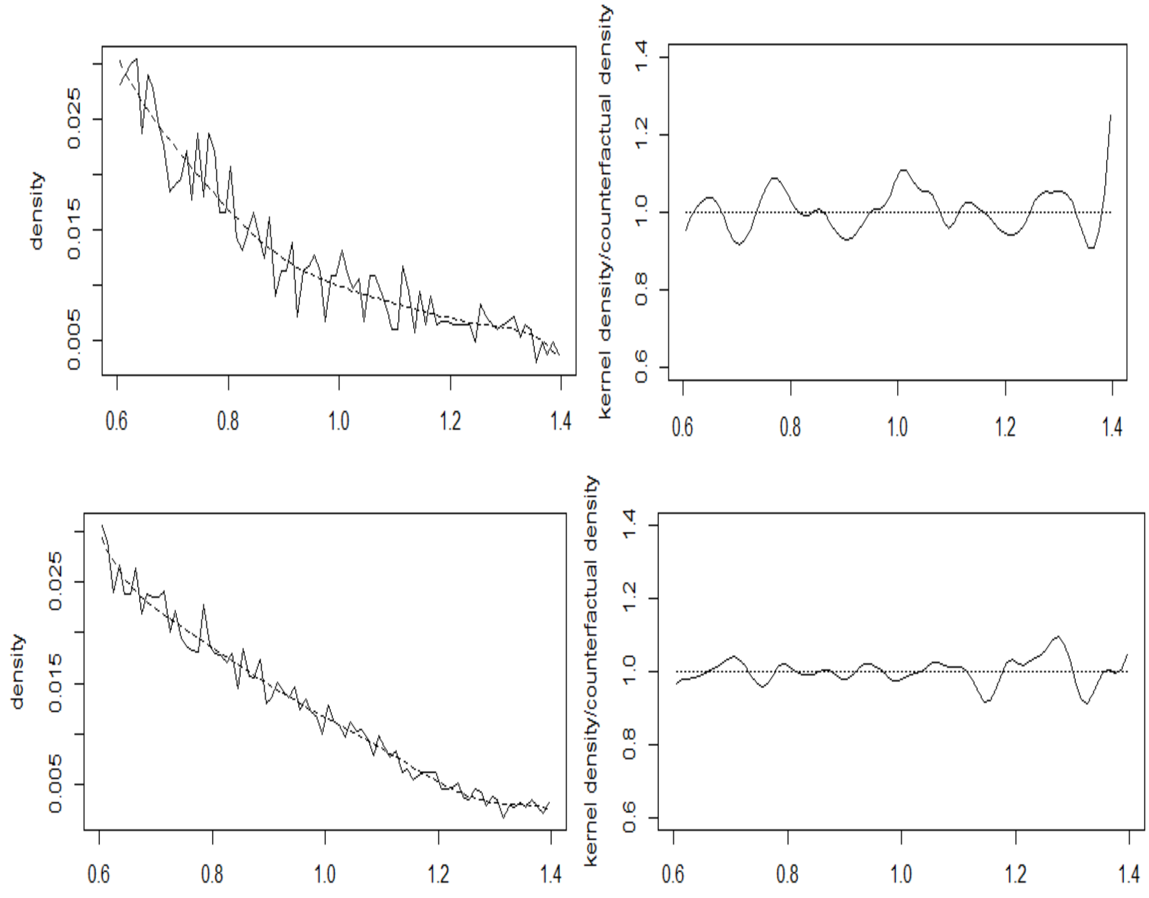
| candidate reference point | # of bins from reference point tested | test statistic | std error | p-value |
|---|---|---|---|---|
| 0.97 | 5 | -0.0199 | 0.0468 | 0.6705 |
| 0.98 | 4 | -0.0240 | 0.0506 | 0.6345 |
| 0.99 | 3 | -0.0170 | 0.0579 | 0.7685 |
| 1.00 | 2 | 0.0669 | 0.0714 | 0.3485 |
| 1.00 | 1 | 0.1290 | 0.0907 | 0.1549 |

*significant at 10% level, ** significant at 5% level, *** significant at 10% level

**Figure 4: Ratio of Excess: Kernel Density / Counterfactual (Chetty et al) Density**

*Top Panel: First Claims, Bottom Panel: Second or Later Claim*

statistical significance results around candidate reference points provided in Tables 4A, 4B, and Appendix




*4.2.2 Claim Timing*

Proposition 4 states that claims made later during the policy period will be more reference-dependent than claims made earlier in the policy period. The intuition can once again be understood in terms of the probability of the policyholder being involved in an accident. Policyholders filing claims earlier in the policy period have a relatively higher likelihood of further accidents occurring before the policy expires, whereas claims filed later in the policy period face this scenario with substantially lower probability. Claims filed later in the policy period therefore give loss averse policyholders greater incentive to reclaim the premium paid under reference-dependent preference compared to claims filed early in the policy period.

Table 5A shows exactly this result using the McCrary test. Greater and more significant discontinuities the distribution occurred near the 1.00 claims to premium ratio in the second half of the policy period compared to the first half of the period. In addition, the Chetty et al test in Table 5B shows that only the [0.99, 1.02] ratio range was statistically different than the estimated distribution, and only in the second half of the policy year. Taking the intersection of significance between the McCrary and Chetty et al tests as statistically robust, we conclude that claims made during the first half of policy periods are unable to detect reference-dependence, while the claims made in the second half of the policy period were robustly reference-dependent around the 1.00 ratio.

**Table 5A: McCrary density test, Claim timing**

| | | | | *First half of policy year* | | | |
|---|---|---|---|---|---|---|---|
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| Test statistic | 0.0379 | 0.1768 | 0.1526 | 0.2408* | 0.0943 | 0.0832 | 0.1494 |
| *p-value* | *0.7660* | *0.1654* | *0.2494* | *0.0725* | *0.4802* | *0.5288* | *0.2567* |
| | | | | *Second half of policy year* | | | |
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| Test statistic | 0.3067* | 0.4664** | 0.4492** | 0.1988 | 0.1091 | 0.0242 | -0.1053 |
| *p-value* | *0.0943* | *0.0147* | *0.0182* | *0.2767* | *0.5466* | *0.8929* | *0.5544* |

*significant at 10% level, ** significant at 5% level, *** significant at 10% level

**Table 5B: Chetty et al test, Claim timing**

*First half of policy year*

| reference point | bins | statistic | std dev | p-value |
|---|---|---|---|---|
| 0.97 | 5 | 0.0383 | 0.1497 | 0.7981 |
| 0.98 | 4 | 0.1499 | 0.1775 | 0.3984 |
| 0.99 | 3 | 0.1418 | 0.1935 | 0.4636 |
| 1.00 | 2 | 0.2980 | 0.2378 | 0.2102 |
| 1.00 | 1 | 0.4550 | 0.3062 | 0.1372 |

*Second half of policy year*

| reference point | bins | statistic | std dev | p-value |
|---|---|---|---|---|
| 0.97 | 5 | 0.0829 | 0.1468 | 0.5724 |
| 0.98 | 4 | 0.2734 | 0.1747 | 0.1176 |
| 0.99 | 3 | 0.3563* | 0.2013 | 0.0768 |
| 1.00 | 2 | 0.1612 | 0.2243 | 0.4725 |
| 1.00 | 1 | 0.1335 | 0.2978 | 0.6540 |

*significant at 10% level, ** significant at 5% level, *** significant at 10% level

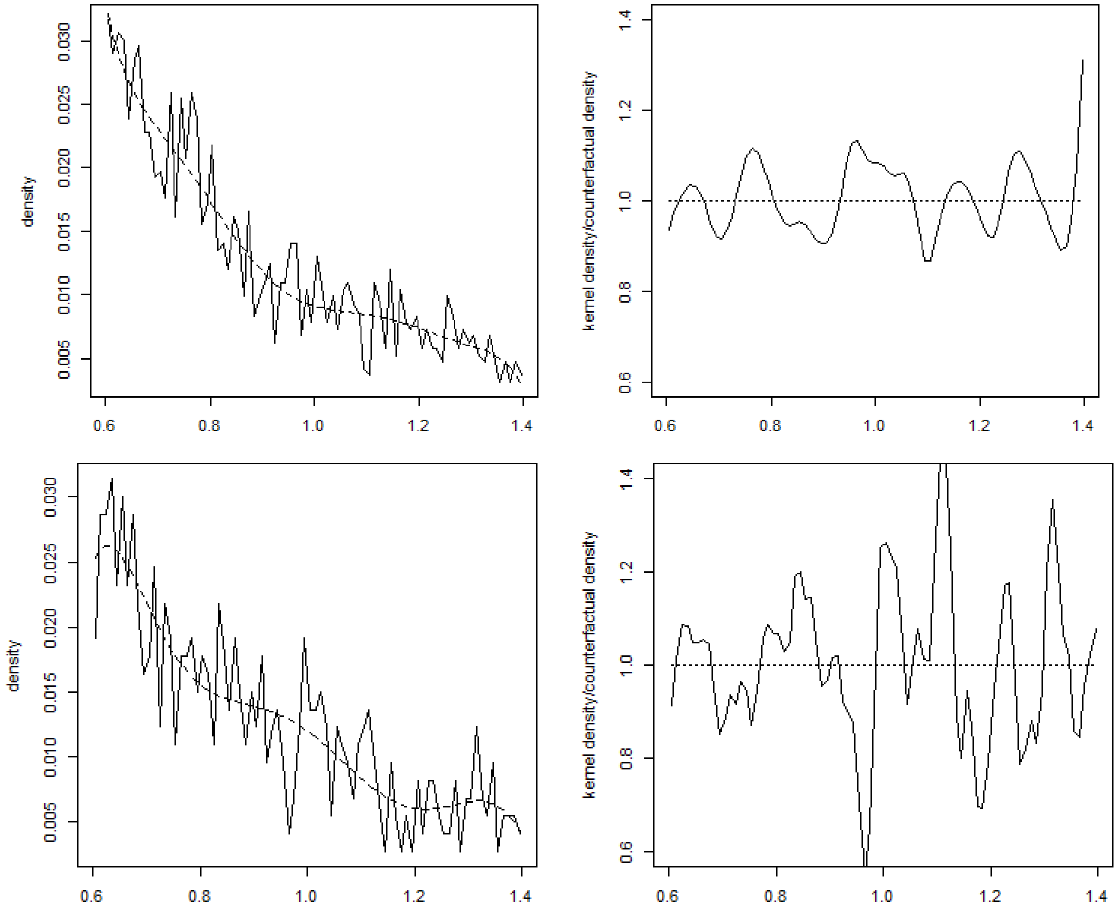
Figure 5 shows for illustrative purposes, the raw frequencies (left panels) and smoothed ratio of excess (right panels) for claims made during the first half of the policy year (top panels), and the claims made in the second half of the policy year (bottom panels). While the statistical significance of the ratio of excess can be obtained only through the Chetty et al test tables, the figures show where the displacements in the distribution are located. Our statistical results around candidate reference points besides 1.0 (ie. 0.8, 0.9, 1.1) shown in the Appendix, demonstrate that the increases or decreases in claims frequencies relative to the counterfactual density are statistically insignificant.

**Figure 5: Ratio of Excess: Kernel Density / Counterfactual Density**

*Top Panels: First half of policy year, Bottom Panels: Second half of policy year*

*(Left Panels: raw frequencies; Right Panels: ratio of excess)*

statistical significance results around candidate reference points provided in Tables 5A, 5B, and Appendix




## 4.3 Original Price and Age of Vehicle

We conduct the same analyses for subsamples of the data based on the original price and age of the vehicle. In terms of the vehicle price, reference-dependence around the premium level is robust for the more expensive half of the cars in the dataset, but not for the less expensive cars. While for the McCrary test (Table 6A), both expensive and inexpensive cars showed some manipulation of claims, in the Chetty test (Table 6B), only the more expensive cars had significant deviations from the hypothetical distribution. This pattern can also be seen from the ratio of excess plots in Figure 6. Table 6C shows the more detailed results by quartiles of car prices, revealing that in fact the most expensive 25% of cars account for the significant manipulation of claims.

**Table 6A: McCrary density test, Price of car**

| | *Inexpensive Car (lower 50$^{th}$ percentile, 104,000 yuan and below)* | | | | | | |
|---|---|---|---|---|---|---|---|
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| Test statistic | 0.2281 | 0.3634** | 0.2524 | 0.2059 | 0.2059 | 0.2134 | 0.2216 |
| *p-value* | *0.1283* | *0.0262* | *0.1022* | *0.1721* | *0.1828* | *0.1759* | *0.1693* |
| | *Expensive Car (upper 50$^{th}$ percentile, above 104,000 yuan)* | | | | | | |
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| Test statistic | 0.1333 | 0.2159 | 0.3588** | 0.4241*** | 0.1072 | -0.0277 | -0.0742 |
| *p-value* | *0.3874* | *0.1511* | *0.0228* | *0.0087* | *0.4808* | *0.8512* | *0.6119* |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table 6B: Chetty et al test, Price of car**

*Inexpensive Car (lower 50$^{th}$ percentile, 104,000 yuan and below)*

| **reference point** | **bins** | **statistic** | **std dev** | **p-value** |
|---|---|---|---|---|
| 0.97 | 5 | -0.0253 | 0.1565 | 0.8717 |
| 0.98 | 4 | 0.1538 | 0.1871 | 0.4111 |
| 0.99 | 3 | 0.0144 | 0.2054 | 0.9442 |
| 1.00 | 2 | -0.1315 | 0.2244 | 0.5579 |
| 1.00 | 1 | -0.1278 | 0.2854 | 0.6544 |

*Expensive Car (upper 50$^{th}$ percentile, above 104,000 yuan)*

| **reference point** | **bins** | **statistic** | **std dev** | **p-value** |
|---|---|---|---|---|
| 0.97 | 5 | 0.1251 | 0.1293 | 0.3330 |
| 0.98 | 4 | 0.2237 | 0.1467 | 0.1272 |
| 0.99 | 3 | 0.4002** | 0.1690 | 0.0179 |
| 1.00 | 2 | 0.6284*** | 0.2034 | 0.0020 |
| 1.00 | 1 | 0.7920*** | 0.2737 | 0.0038 |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

Tables 6C also shows that among cars in the lower half of the price distribution, the lowest price quartile of car policyholders are responsible for the significant manipulations in claims in the more aggregated McCrary test of Table 6A. However, the corresponding Chetty et al test, as shown in Table 6D (which shows only the lowest 25% priced cars, out of space concerns) indicates that there are actually *fewer* claims made at the 1.00 ratio, compared to the estimated distribution. One possible explanation which is consistent with the combination of evidence in Table 6C and 6D is that upward manipulation of claims is substantial for the lowest priced 25% of cars, however they do not center on the 1.00 ratio as the desired claim amount.

**Figure 6: Ratio of Excess: Kernel Density / Counterfactual Density**

*Top Panel: Inexpensive car, Bottom Panel: Expensive car*

*(Left Panels: raw frequencies; Right Panels: ratio of excess)*

statistical significance results around candidate reference points provided in Tables 6A, 6B, and Appendix

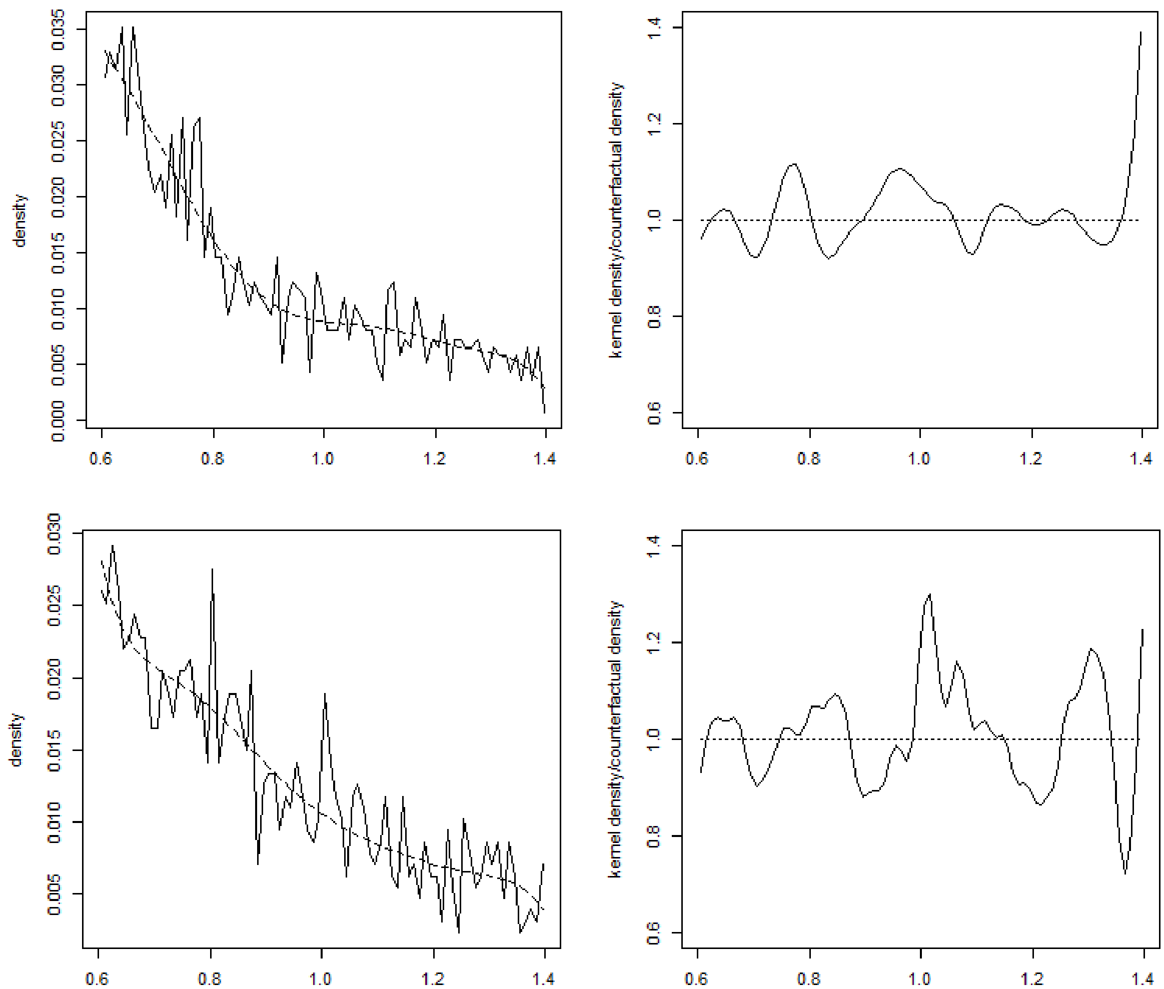


**Table 6C: McCrary density test, Price of car**
*By quartiles, p-values shown*

| | 0.97 | 0.98 | 0.99 | 1.00 | 1.01 | 1.02 | 1.03 |
|---|---|---|---|---|---|---|---|
| **1st quartile** *(77,040 yuan and below)* | 0.2966 | 0.4417** | 0.2641 | 0.2140 | 0.3792* | 0.4077* | 0.5089** |
| *p-value* | *0.1384* | *0.0397* | *0.1981* | *0.2892* | *0.0805* | *0.0632* | *0.0239* |
| **2nd quartile** *(77,040 to 104,000 yuan)* | 0.3143 | 0.3275 | 0.3013 | 0.2291 | 0.0773 | 0.0175 | -0.1284 |
| *p-value* | *0.1398* | *0.1692* | *0.1823* | *0.3049* | *0.7291* | *0.9394* | *0.5908* |
| **3rd quartile** *(104,000 to 145,800 yuan)* | *-0.0712* | *0.3283* | *0.3647* | *0.4095* | *0.0342* | *-0.1664* | *-0.1215* |
| *p-value* | *0.7389* | *0.2036* | *0.1659* | *0.1555* | *0.8947* | *0.4757* | *0.6108* |
| **4th quartile** *(145,800 yuan and above)* | *0.1903* | *0.1433* | *0.2786* | *0.3620*** | *0.0956* | *0.0244* | *-0.0395* |
| *p-value* | *0.2518* | *0.3885* | *0.1163* | *0.0490* | *0.5672* | *0.8834* | *0.8103* |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table 6D: Chetty et al test, Price of car**
*Inexpensive Car (lower 25th percentile, 77,040 yuan and below)*

| reference point | bins | statistic | std dev | p-value |
|---|---|---|---|---|
| 0.97 | 5 | -0.1339 | 0.2012 | 0.5056 |
| 0.98 | 4 | -0.0154 | 0.2368 | 0.9483 |
| 0.99 | 3 | -0.2147 | 0.2448 | 0.3804 |
| 1.00 | 2 | -0.4192 | 0.2718 | 0.1230 |
| 1.00 | 1 | -0.6065** | 0.2943 | 0.0394 |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

Tables 7A and 7B show that reference-dependent claim behavior is concentrated among policyholders with older cars, rather than those with newer cars. Both the McCrary test and Chetty et al test show significant manipulation for cars older than a year old, while neither test shows significant manipulations for cars one year old or below. This result can also be seen visually from Figure 7, which shows a larger deviation from the estimated distribution around 1.00 for the older subset of cars.

This result could be due to the perceived value of the insurance premium paid for new cars, as well as possible mental accounting effects. For new cars, policyholders may be more willing to accept the idea of paying the premium to cover potential risks. For older cars, policyholders may not think the premium paid was worthwhile to them unless they are able to recover the full premium amount should an accident occur. In addition, the purchase of a new car in China is often accompanied by the sale of the insurance directly at the dealership. Thus, buyers may mentally bundle together the price of the new car and the insurance, and may not easily form a reference point around the insurance price alone.

**Table 7A: McCrary density test, Age of car**

*New Car (1 year old and below)*

| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
|---|---|---|---|---|---|---|---|
| Test statistic | -0.1945 | 0.0679 | 0.2298 | 0.2551 | 0.1247 | 0.0688 | 0.0089 |
| *p-value* | *0.3194* | *0.7124* | *0.1877* | *0.1684* | *0.4837* | *0.6930* | *0.9587* |

*Old Car (above 1 year old)*

| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
|---|---|---|---|---|---|---|---|
| Test statistic | 0.1627 | 0.2917** | 0.1959 | 0.1235 | 0.0387 | 0.0124 | 0.0503 |
| *p-value* | *0.2207* | *0.0361* | *0.1500* | *0.3637* | *0.7730* | *0.9271* | *0.7179* |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table 7B: Chetty et al test, Age of car**

*New Car (1 year old and below)*

| reference point | bins | statistic | std dev | p-value |
|---|---|---|---|---|
| 0.97 | 5 | -0.1239 | 0.1403 | 0.3772 |
| 0.98 | 4 | -0.0166 | 0.1620 | 0.9184 |
| 0.99 | 3 | 0.0786 | 0.1955 | 0.6879 |
| 1.00 | 2 | 0.2019 | 0.2332 | 0.3864 |
| 1.00 | 1 | 0.2645 | 0.2934 | 0.3674 |

*Old Car (above 1 year old)*

| reference point | bins | statistic | std dev | p-value |
|---|---|---|---|---|
| 0.97 | 5 | 0.2350 | 0.1546 | 0.1286 |
| 0.98 | 4 | 0.4029** | 0.1783 | 0.0238 |
| 0.99 | 3 | 0.3427* | 0.1874 | 0.0674 |
| 1.00 | 2 | 0.3000 | 0.2198 | 0.1729 |
| 1.00 | 1 | 0.4231 | 0.2869 | 0.1402 |

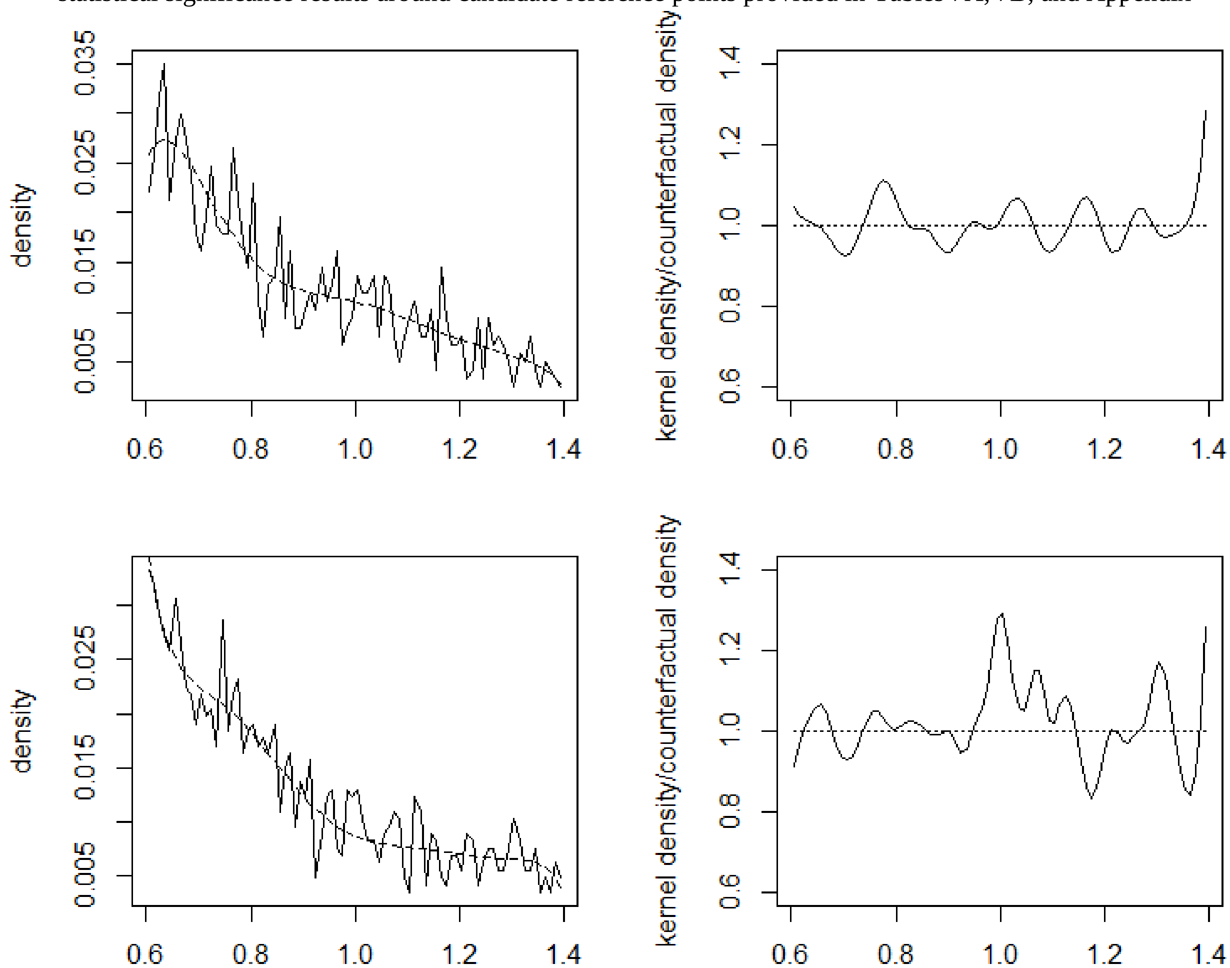
*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Figure 7: Ratio of Excess: Kernel Density / Counterfactual Density**

*Top Panel: New car, Bottom Panel: Old car*

*(Left Panels: raw frequencies; Right Panels: ratio of excess)*

statistical significance results around candidate reference points provided in Tables 7A, 7B, and Appendix




### 4.4 Age and Gender of Policyholder

Conducting our tests on demographic characteristics of the policyholder does not yield the consistency of results apparent in the variables examined in earlier sections. In other words, neither specific age nor gender groups appear to significantly drive the result.

Although the McCrary test indicates that policyholders aged 39 and below are most responsible for reference-dependent claims, the Chetty et al test results (not shown here) for the same age group marginally fail the significance test at the 10% level. The distributional excess calculated via the Chetty et al test are shown in Figure 8, and the pattern of younger drivers being more reference-dependent around the 1.00 ratio can be observed qualitatively. However, we find that these results are less robust than the previous ones, due to the inconsistencies in significance levels between the McCrary and Chetty et al tests.

We also conduct the tests by policyholders' gender. While the McCrary test results suggest that male policyholders are driving the reference-dependent claims, the Chetty et al test shows the opposite, which is that female policyholders are significantly reference-dependent while male policyholders are not. Figure 9 shows the qualitative pattern of excess claims behavior by gender.

**Table 8A: McCrary density test, Age of Policyholder**

*By quartiles, p-values shown*

| | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
|---|---|---|---|---|---|---|---|
| **1st quartile** *(below 33)* | 0.2459 | 0.2538 | 0.3667* | 0.2704 | 0.1765 | 0.1180 | 0.1362 |
| *p-value* | *0.1833* | *0.1797* | *0.0699* | *0.1521* | *0.3496* | *0.5362* | *0.4943* |
| **2nd quartile** *(33 to 39)* | 0.3371* | 0.7015*** | 0.6126*** | 0.5225** | 0.3452 | 0.3081 | 0.0904 |
| *p-value* | *0.0959* | *0.0039* | *0.0083* | *0.0274* | *0.1168* | *0.1781* | *0.6677* |
| **3rd quartile** *(39 to 45)* | -0.1074 | 0.0078 | -0.0704 | -0.0118 | -0.0497 | -0.1927 | -0.0030 |
| *p-value* | *0.5615* | *0.9696* | *0.7172* | *0.9546* | *0.7949* | *0.3247* | *0.9875* |
| **4th quartile** *(45 and above)* | 0.3058 | 0.3100 | 0.2020 | 0.2192 | 0.1576 | 0.1752 | 0.2395 |
| *p-value* | *0.1539* | *0.1647* | *0.3736* | *0.3226* | *0.4876* | *0.4293* | *0.3043* |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Figure 8: Ratio of Excess: Kernel Density / Counterfactual Density**

*Top Panel: Young drivers, Bottom Panel: Older drivers*

*(Left Panels: raw frequencies; Right Panels: ratio of excess)*

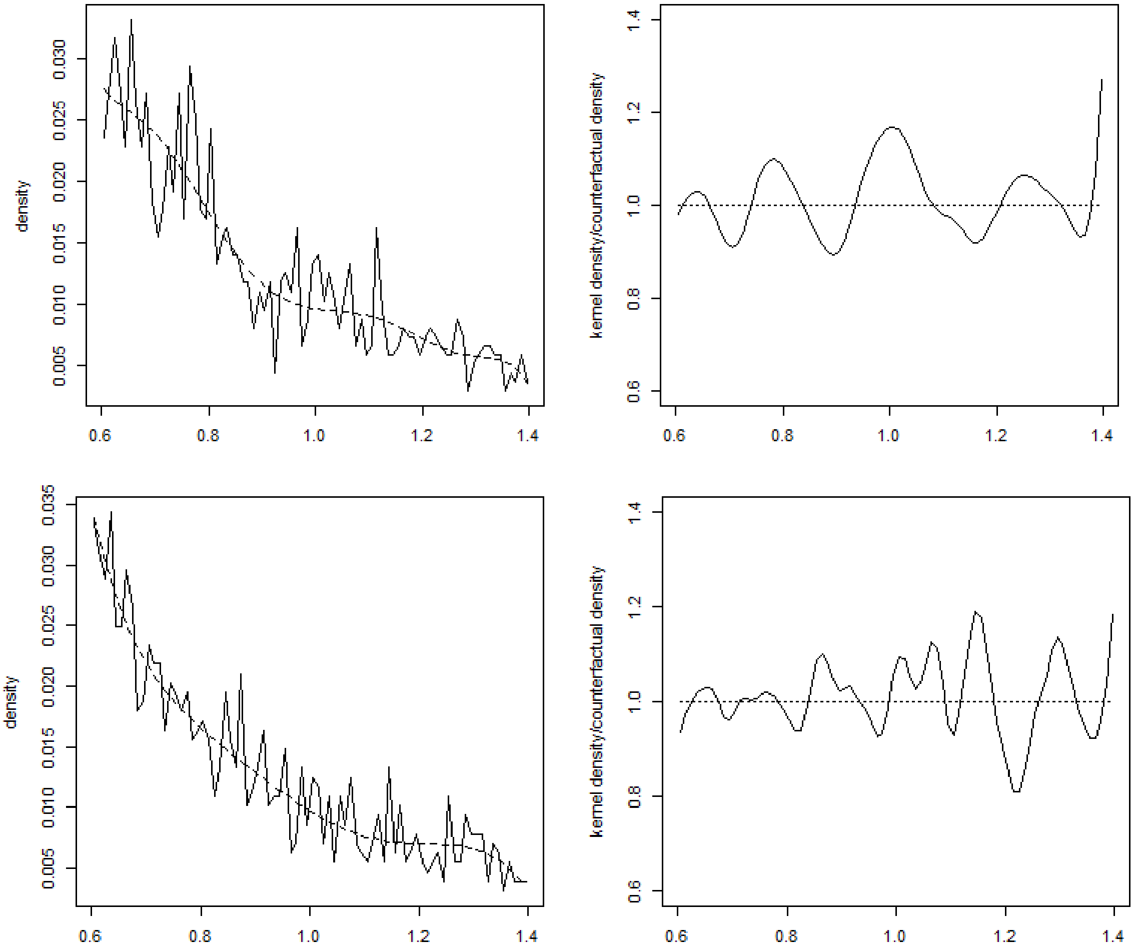



**Table 10: McCrary density test, by Gender**

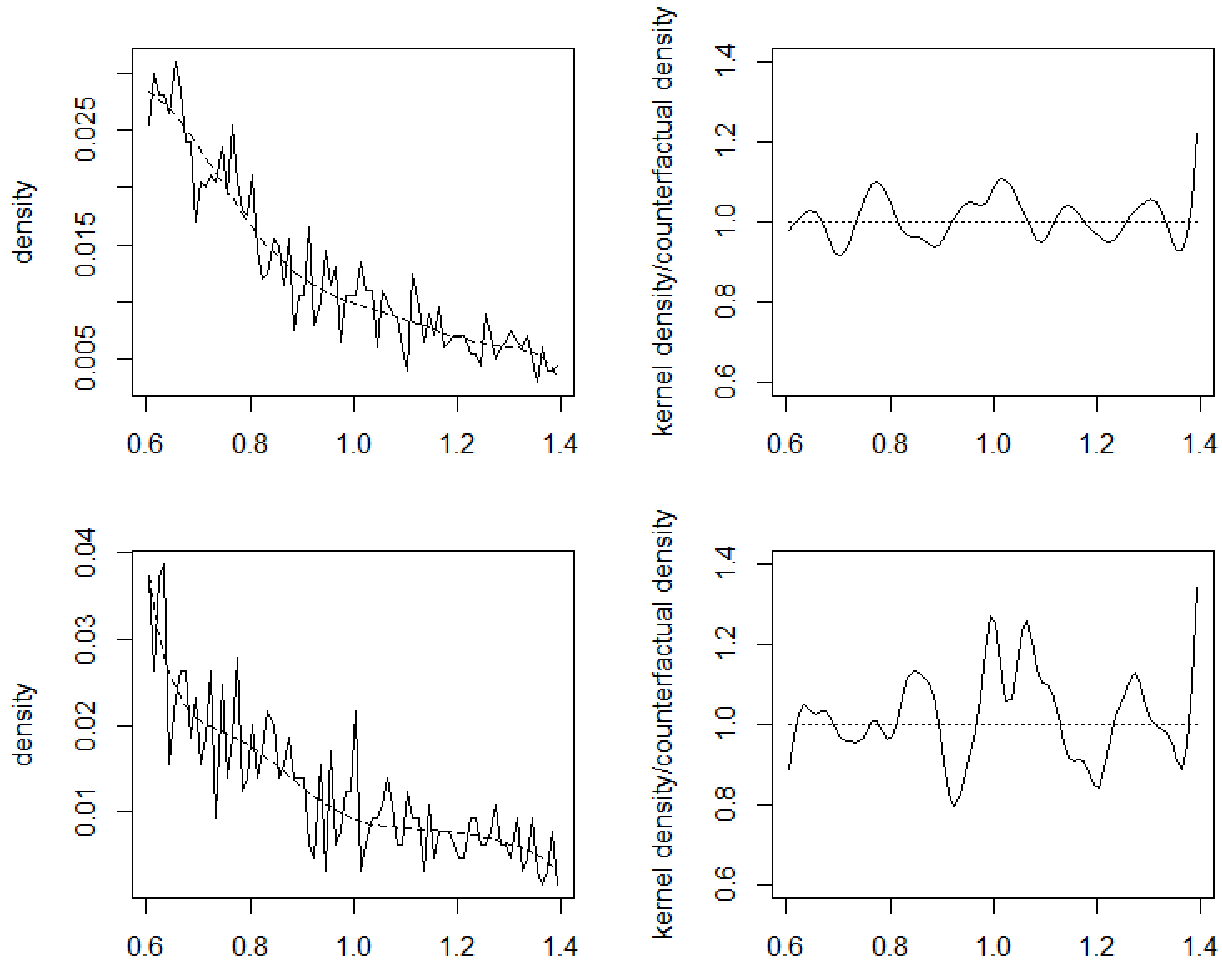
| | | | | *Male Policyholder* | | | |
|---|---|---|---|---|---|---|---|
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| Test statistic | 0.0719 | 0.2471* | 0.2364* | 0.1983 | 0.1914 | 0.0426 | -0.0386 |
| *p-value* | *0.5724* | *0.0589* | *0.0757* | *0.1449* | *0.1495* | *0.7444* | *0.7708* |
| | | | | *Female Policyholder* | | | |
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| Test statistic | 0.2319 | 0.3191 | 0.2564 | 0.2073 | -0.3176 | 0.0472 | 0.1997 |
| *p-value* | *0.2875* | *0.1208* | *0.2051* | *0.3002* | *0.2070* | *0.8390* | *0.3581* |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Figure 9: Ratio of Excess: Kernel Density / Counterfactual Density**

*Top Panel: Male drivers, Bottom Panel: Female drivers*

*(Left Panels: raw frequencies; Right Panels: ratio of excess)*




## 5. Robustness Checks: Other Possible Reference Points

So far our empirical findings focus on testing for a reference point at the 1.0 claims to premium ratio. Noting that both the McCrary and Chetty et al tests utilize small windows of the running variable in order to implement the test, it is beneficial to test around alternative reference point candidates to put our results at the 1.0 ratio into perspective. The results show that the 1.0 ratio is unique among round ratio values across the sample of policyholders, which indicates a significant psychological appeal of wealth levels before the premium was paid, as a reference-point.

### 5.1 Alternative Claims to Premium Ratios

First, we check whether policyholders who have had an accident tend to bunch their claims around other possible reference levels. One possibility is that insurance holders have some positive valuation of the insurance contract, so they are willing to forgo monetary amounts less than this valuation in their claims. For example, upon getting into an accident, a policyholder may be somewhat inclined to claim back the whole premium, but could reconsider and view it as "fair" for the insurance company to keep a percentage of the premium revenue.

Using identical procedures as the analysis described above, we test other possible round number reference points below the 1.00 ratio, such as 0.9 and 0.8. We do not find any consistent evidence of over-concentration of claims around those percentages. We also test the ratio 1.1 as a counterfactual test for the possibility that the desire to inflate claims alone can generate the empirical patterns found at the 1.0 ratio. Similarly, we do not find robust significant overconcentration of claims around this ratio. The results for the 0.9 and 1.1 ratios are provided in Appendix A and B, respectively. Results for other claims to premium rations also lack regular significant patterns. They are omitted due to space concerns, but are available upon request.

One intuitive reason for not detecting manipulation of claims around these alternative claims to premium ratios might be related to the calculation requirement on the part of policyholders of making such claims hypothetically. While claiming back the premium paid only requires recalling the amount paid or looking it up in one's records, claiming back a fraction of the premium requires additional deliberation and calculation of the desired claim amount.

The role of memory and bracketing in establishing the reference point is also supported by the earlier findings on new cars compared to old cars. Recall that we found significant manipulation driven by older cars rather than those cars purchased within the past year. That finding is also consistent with the memory requirement of claiming back the premium. For policyholders who likely purchased their insurance plan together with a new car, it is more difficult to recall the exact price of the insurance policy alone, thus the reference-dependent claim pattern does not appear.

Another likely reason for the lack of results around these alternative ratios is the "coordination" difficulty across policyholders. Recall that our methodology relies on cross-sectional identification. This means that for a reference point to be identified as statistically significant in our data, there must be enough policyholders in the data who hold that ratio as their personal reference point and behave accordingly. It may be much easier for the premium amount to hold popular appeal as a reference point rather than alternative ratio counterparts.

### 5.2 Nominal Claim Amounts

Alternatively, policyholders who have had an accident might hold a nominal monetary amount as a reference point, such as 500 RMB, for example. We test the presence of reference points at nominal levels, in intervals of 50 RMB from the nominal break-even point relative to the premium paid, as well as in 50 RMB intervals from the 500 RMB gain level, and 500 RMB loss level. This includes scenarios that policyholders may want to claim back particular round nominal amounts, as well as scenarios in which policyholders may want to claim back some round number amounts above their premium paid.

Similarly to the results for alternative reference points in the claims to premium ratio, none of these results for nominal claim amounts are robustly significant across the statistical tests. When on occasion, significant results do exist, the typical pattern is that the McCrary test may show significant manipulation, but it is not robust to the Chetty et al test, or alternatively, vice versa. Due to space concerns, we omit these null results, but they are available upon request.

Due to the lack of robustness across the nominal claim amount results, while we infer strong support for the nominal break-even level of claims as a reference point in the previous sections, we do not find any robust support for policyholders' desire to claim back specific nominal amounts besides that of their premium paid, even for salient round number amounts.

We note here that unlike the cases of alternative claims to premium ratios, the claiming of round number monetary amounts does not require any special calculation to be made by the policyholder. In fact, it may require even *less* effort than claiming back the premium, since the policyholder does not need to remember or look up the premium amount. However, similar to the case of the alternative claims to premium ratios, nominal amount reference point candidates require commonality or popularity across policyholders in order to be detected in the data using the current methodology. Our analysis implies that at the least, these nominal amounts were not sufficiently common as policyholder reference points.

Taken together, the robustness checks in Sections 5.1 and 5.2 indicate that the premium level, or in other words the wealth level of the policyholder prior to paying the insurance premium, seems to be a unique common reference point around which policyholders are most substantially inclined to inflate their claims.

### 5.3 Robustness Check: Private versus Organization-held Vehicle Policies

An additional robustness check for the claim abnormalities being due to policyholders' reference-dependent utility and not due to any special policy of the insurance company, is to use corporate and government vehicle insurance policies as a counterfactual. Corporate and government insurance policies generally hold no direct monetary consequences to the individual making the claims, since all insurance premiums and payouts are billed and paid out to the organization. Therefore, if reference-dependent utility functions are the reason for claims bunched around the premium level, we should expect not to observe the patterns in our main findings when restricting the dataset to the set of insurance policies held by firms.

Indeed, none of the patterns we find in the previous section exist among the organization-held insurance policies in our data. The statistical results for Propositions 1 through 4 among organization-held policies are provided in Appendix C. The results do not demonstrate any of the robust statistical significance apparent in the overall sample, let alone following the predictions of the four hypotheses. The absence of significant manipulation of claims around the 1.00 ratio serves as supporting evidence that the manipulations are driven by individual preferences over monetary outcomes, meaning individual policyholder utility, and not other potential alternative explanations related to the insurance company's policies.

## 6. Conclusions

Identifying the effects of asymmetric information in markets is typically a challenging task using naturally occurring field data, and evidence on moral hazard has been particularly sparse. In this paper, we draw upon the reference-dependence and asymmetric information literatures to demonstrate a method of identifying moral hazard when the possibility of reference-dependent policyholders is permitted in the model of policyholder behavior.

We build a model of moral hazard in claims behavior under reference-dependent preferences of policyholders, which carries testable comparative static predictions among subgroups of policyholder profiles. These predictions are confirmed in the data. The first claims made in a policy period, and those claims made chronologically later within a policy period, were drivers of reference-dependent claim behavior. The intuition in the model is that loss averse policyholders would like to take the first opportunity to claim back the premium, and accidents occurring later in the policy period leave less time remaining for additional claim opportunities.

In terms of characteristics of policyholders, we find reference-dependence concentrated among higher premium level plans and policyholders who are evaluated as less risky by the insurance company.

The finding of concentration of reference-dependence among high premium policyholders is consistent with our finding of reference-dependence being concentrated mostly among expensive vehicles. The results on premium rates show that the high premium results do not originate in riskier policyholders – in fact, reference-dependent claims behavior was driven by those policyholders which the insurance company determined to be less risky. Our model is consistent with less risk-taking policyholders being more eager to recover the insurance premium, due to a loss averse utility function which persists across different domains.

Our study has also revealed some insights which are not necessarily related to the domains in the theoretical model prediction, but may still be useful in policy applications. We find that reference-dependent claims are largely driven by policies on expensive cars, and by older cars. If auto expenses are positively correlated with wealth (which we do not observe in the data set), the former phenomenon suggests that wealthier policyholders may be more reference-dependent than less wealthy policyholders. At the same time, we do find some evidence of manipulation among policies on the least expensive cars in the data, but the manipulation does not seem to be reference dependent around the premium amount specifically. The concentration of reference-dependence among cars older than one year could be explained by policyholders' perceptions of the value of insurance on new cars versus old cars. Owners of older cars may perceive that the insurance policy is not as worthwhile, thus having a stronger preference for recovering the premium. This finding also makes sense from a mental bracketing perspective, since buyers of new cars are frequently sold their auto insurance policies together with the car itself.

In addition, our study may be informative regarding how policyholders in a developing insurance market intrinsically value their mandated insurance policies. The results imply that the typical amount of reimbursement that a disproportionate number of policyholders seek is the amount of their premium paid - in other words, they simply want to break even in the entire insurance proposition. As China's insurance market develops further, it remains to be seen whether policyholders remain fixated on the premium as a reference amount to be recovered, or whether other reference points emerge among policyholders over time, as suggested in the reference point updating experiment of Baucells, Weber and Welfens (2011). An additional direction is to elicit how the claims result relative to the premium corresponds to policyholders' reported satisfaction levels, as in Ockenfels, Sliwka and Werner (2015) on workers' bonus payments. A further possibility may be that reference-dependent behavior diminishes altogether with experience as found in List (2003, 2004) and Tong et al (2006).

Over the past decades since the introduction of Prospect Theory (Kahneman and Tversky, 1979) and mainstream incorporation of reference-dependence into economic modeling (Koszegi and Rabin, 2006), studies have successfully identified several sources of solid evidence for loss aversion. Despite the strong evidence for psychological reference points, economists have sometimes viewed reference-dependent utility as a burden on classical theory, which may unnecessarily complicate previously tractable models, among other concerns. However, as we show in this study, reference-dependent utility can be a helpful tool in identifying classical economic problems such as moral hazard. By allowing the possibility of a utility-based reference point which is motivated not by material incentives, but by psychological incentives, a combination of behavioral theory and statistical techniques can successfully detect manipulation of claims where it would be otherwise difficult to do so.

**References:**


Abbring, Jaap H., Pierre-Andre Chiappori, James J. Heckman and Jean Pinquet, "Adverse Selection and Moral Hazard in Insurance: Can Dynamic Data Help to Distinguish?" *Journal of the European Economic Association,* Vol. 1 (2003), No. 2/3, p. 512 – 521.

Abbring, Jaap H., Pierre-Andre Chiappori and Jean Pinquet, "Moral Hazard and Dynamic Insurance Data", *Journal of the European Economic Association*, Vol. 1 (2003), No. 4, p. 767 – 820.

Abbring, Jaap H., Pierre-Andre Chiappori and Tibor Zavadil, "Better Safe than Sorry? Ex Ante and Ex Post Moral Hazard in Dynamic Insurance Data", *Tinbergen Institute Discussion Paper*, No. 08-075/3, 2008.

Allen, Eric J., Patricia M. Dechow, Devin G. Pope and George Wu, "Reference-Dependent Preferences: Evidence from Marathon Runners", *Management Science,* Vol. 63 (2017), No. 6, p. 1657 – 1672.

Aron-Dine, Aviva, Liran Einav and Amy Finkelstein, "The RAND Health Insurance Experiment, Three Decades Later", *Journal of Economic Perspectives*, Vol. 27 (2013), No. 1, p. 197 – 222.

Aron-Dine, Aviva, Liran Einav, Amy Finkelstein and Mark Cullen, "Moral Hazard in Health Insurance: Do Dynamic Incentives Matter?" *Review of Economics and Statistics,* Vol. 97 (2015), No. 4, p. 725 – 741.

Baillon, Aurelien, Han Bleichrodt and Vitalie Spinu, "Searching for the Reference Point", *Management Science*, Vol. 66 (2020), No. 1, p. 93 – 112.

Barseghyan, Levon, Francesca Molinari, Ted O'Donoghue and Joshua C. Teitelbaum, "The Nature of Risk Preferences: Evidence from Insurance Choices", *American Economic Review,* Vol. 103 (2013), No. 6, p. 2499 – 2529.

Baucells, Manel, Martin Weber and Frank Welfens, "Reference Point Formation and Updating", *Management Science*, Vol. 57 (2011), NO. 3, p. 506 – 519.

Camerer, Colin F., Linda Babcock, George Loewenstein, and Richard H. Thaler, "Labor supply of New York City cabdrivers: One day at a time," *The Quarterly Journal of Economics*, Vol. 112 (1997), No. 2, p. 407-441.

Card, David and Gordon B. Dahl, "Family violence and football: The effect of unexpected emotional cues on violent behavior." *The Quarterly Journal of Economics* Vol. 126 (2011), No. 1, p. 103 – 143.

Ceccarini, Olivia, "Does Experience Rating Matter in Reducing Accident Probabilities? A Test for Moral Hazard", *Working Paper*, 2007.

Chetty, Raj, John N. Friedman, Tore Olsen and Luigi Pistaferri, "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records", *Quarterly Journal of Economics*, Vol. 126 (2011), p. 749 – 804.

Chiappori, Pierre-Andre, Bruno Jullien, Bernard Salanie and Francois Salanie, "Asymmetric Information in Insurance: General Testable Implications", *The RAND Journal of Economics*, Vol. 37 (2006), No. 4, p. 783 – 789.

Crawford, Vincent P. and Juanjuan Meng, "New York City Cabdrivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income," American Economic Review, Vol. 101 (2011), No. 5, p. 1912 – 1932.

Cohen, Alma, "Asymmetric Information and Learning: Evidence from the Automobile Insurance Market", *Review of Economics and Statistics*, Vol. 87 (2005), No. 2, p. 197 – 207.

Cummins, David J., and Sharon Tennyson, "Moral Hazard in Insurance Claiming: Evidence from Automobile Insurance", *Journal of Risk and Uncertainty*, Vol. 12 (1996), No. 1, p. 29 – 50.

Dionne, Georges, Pierre-Carl Michaud, and Maki Dahchour, "Separating Moral Hazard from Adverse Selection and Learning in Automobile Insurance: Longitudinal Evidence from France", *Journal of the European Economic Association*, Vol. 11 (2013), Issue 4, p. 897 – 917.

Dunham, Wayne R., "Moral Hazard and the Market for Used Automobiles", *Review of Industrial Organization*, Vol. 23 (2003), p. 65 – 83.

Eil, David and Jaimie W. Lien, "Staying Ahead and Getting Even: Risk Attitudes of Experienced Poker Players", *Games and Economic Behavior*, Vol. 87 (2014), p. 50 – 69.

Fang, Hanming, "Insurance Markets in China", The Oxford Companion to the Economics of China (Oxford University Press), ed. Shenggen Fan, Ravi Kanbur, Shang-Jin Wei and Xiaobo Zhang, 2013.

Fehr, Ernst and Lorenz Goette, "Do workers work more if wages are high? Evidence from a randomized field experiment," *The American Economic Review*, Vol. 97 (2007) No. 1, p. 298-317.

Fischbacher, Urs, and Franziska Follmi-Heusi, "Lies in Disguise – an Experimental Study on Cheating" *Journal of the European Economic Association*, Vol. 11 (2013), p. 525 – 547.

Gao, Feng, Michael R. Powers and Jun Wang, "Decomposing Asymmetric Information in China's Automobile Insurance Market", *Journal of Risk and Insurance*, Vol. 84 (2017), Issue 4, p. 1269 – 1293.

Gao, Feng and Jun Wang, "Testing for Adverse Selection in China's Auto Insurance Market", *Asia-Pacific Journal of Risk and Insurance*, Vol. 5 (2011), Issue 1

Gneezy, Uri, Lorenz Goette, Charles Sprenger and Florian Zimmermann, "The Limits of Expectations-Based Reference Dependence", *Journal of the European Economic Association*, Vol. 15 (2017), No. 4, p. 861 – 876.

Goette, Lorenz, Annette Harms, and Charles Sprenger, "Randomizing Endowments: An Experimental Study of Rational Expectations and Reference-Dependent Preferences", *American Economic Journal, Microeconomics*, forthcoming.

Hanna, Rema and Shing-Yi Wang, "Dishonesty and Selection into Public Service: Evidence from India" *American Economic Journal: Economic Policy*, Vol. 9 (2017), No. 3, p. 262 – 290.

Hoyt, Robert E., David B. Mustard and Lawrence S. Powell, "The Effectiveness of State Legislation in Mitigating Moral Hazard: Evidence from Automobile Insurance", *The Journal of Law and Economics*, Vol. 49 (2006), No. 2, p. 427 – 450.

Jones, Peter, "Loss Aversion and Property Tax Avoidance", *Working Paper*, December 2019.

Jung, Hanjoon Michael, Chul-In Lee, Jaimie W. Lien and Jie Zheng, "Discretionary Red Tape", *Working Paper*, 2020.

Koszegi, Botond, and Rabin, Matthew, "A Model of Reference-Dependent Preferences", *Quarterly Journal of Economics*, Vol. 121 (2006), Issue 4, p. 1133 – 1165.

Lien, Jaimie W. and Jie Zheng, "Deciding When to Quit: Reference-Dependence over Slot Machine Outcomes", *American Economic Review: Papers and Proceedings*, Vol. 105 (2015), p. 336 – 370.

Lien, Jaimie W., Minbo Xu, and Jie Zheng, "What Brings Consumers Back for More? Evidence from Quantifiable Gain and Loss Experiences in Penny Auctions", *Working Paper*, July 2017.

List, John A., "Does Market Experience Eliminate Market Anomalies?" *Quarterly Journal of Economics*, Vol. 118 (2003), No. 1, p. 41 – 71.

List, John A., "Neoclassical Theory versus Prospect Theory: Evidence from the Marketplace", *Econometrica*, Vol 72 (2004), No. 2, p. 615 – 625.

Ma, Chao, “Be Cautious in the Last Month: The Sunk Cost Fallacy Held by Car Insurance Policyholders”, *International Economic Review*, forthcoming.

McCrary, Justin, “Manipulation of the running variable in the regression discontinuity design: A density test”, *Journal of Econometrics*, Vol. 142 (2008), p. 698 – 714.

Ockenfels, Axel, Dirk Sliwka and Peter Werner, “Bonus Payments and Reference Point Violations”, *Management Science*, Vol. 61 (2015), No. 7, p. 1496 – 1513.

O’Donoghue, Ted, and Jason Somerville, “Modeling Risk Aversion in Economics”, *Journal of Economic Perspectives*, Vol. 32, No. 2 (Spring 2018), p. 91 – 114.

Picard, Pierre, “Auditing claims in the insurance market with fraud: The credibility issue”, *Journal of Public Economics*, Vol. 63 (1996), p. 27 – 56.

Pope, Devin G. and Maurice.E. Schweitzer, “Is Tiger Woods loss averse? Persistent bias in the face of experience, competition, and high stakes,” *The American Economic Review*, Vol. 101 (2011), No, 1, p. 129-157.

Pope, Devin and Uri Simonsohn, “Round Numbers as Goals: Evidence from Baseball, SAT Takers, and the Lab”, *Psychological Science*, Vol. 22 (2011), No. 1, p. 71 – 79.

Post, Thierry, M.J. Van den Assem, G. Baltussen, and R.H. Thaler, “Deal or no deal? Decision making under risk in a large-payoff game show,” *The American Economic Review*, Vol. 98 (2008), No. 1, p. 38-71.

Rees-Jones, Alex, “Quantifying Loss-Averse Tax Manipulation”, *Review of Economic Studies*, Vol. 85 (2018), p. 1251 – 1278.

Rowell, David, and Luke B. Connelly, “A History of the Term ‘Moral Hazard’”, *Journal of Risk and Insurance*, Vol. 79 (2012), No. 4, p. 1051 – 1075.

Schneider, Henry Seth, “Moral Hazard in Leasing Contracts: Evidence from the New York City Taxi Industry”, *Journal of Law and Economics*, Vol. 53 (2010), No. 4, p. 783 – 805.

Seibold, Arthur, “Reference Points for Retirement Behavior: Evidence from German Pension Discontinuities”, *Working Paper*, April 2020.

Sydnor, Justin, “(Over)insuring Modest Risks”, *American Economic Journal: Applied Economics*, Vol. 2 (2010), No. 4, p. 177 – 199.

Tong, Lester C. P., Karen J. Ye, Kentaro Asai, Seda Ertac, John A. List, Howard C. Nusbaum and Ali Hortacsu, “Trading Experience Modulates Anterior Insula to Reduce the Endowment Effect”, *Proceedings of the National Academy of Sciences*, Vol. 113 (2016), No. 33, p. 9238 – 9243.

Wagstaff, Adam and Magnus Lindelow, “Can insurance increase financial risk? The curious case of health insurance in China”, *Journal of Health Economics*, Vol. 27 (2008), p. 990 – 1005.

Wang, Jennifer L., Ching-Fan Chung and Larry Y. Tzeng, “An Empirical Analysis of the Effects of Increasing Deductibles on Moral Hazard”, *Journal of Risk and Insurance*, Vol. 75 (2008), No. 3, p. 551 – 566.

Weisburd, Sarit, “Identifying Moral Hazard in Car Insurance Contracts”, *Review of Economics and Statistics*, Vol. 97 (2015), Issue 2, p. 301 – 313.

## Appendix: Proofs

**Lemma 1: (Claiming Strategy)** For an individual with characteristics profile $(\eta,\lambda,\beta,c,r)$, when $\beta>\lambda\eta$, $\tilde{c}^*=c$; when $\beta<\eta$, $\tilde{c}^*=\begin{cases} c & if \quad c\geq\overline{c} \\ \overline{c} & if \quad c<\overline{c}\end{cases}$; when $\eta<\beta<\lambda\eta$, $\tilde{c}^*=\begin{cases} c & if \quad c\geq r \\ r & if \quad c<r\end{cases}$.

Proof:

First note that reporting any claim level that is lower than the actual damage $c$ is strictly dominated by reporting $c$, which would make the reference-dependent term greater and at the same reduce the cost of misreporting to zero. Thus, without loss of generality, we only conduct the analysis where $\tilde{c}^*\geq c$.

When $\beta>\lambda\eta$, $\frac{\partial U(\tilde{c})}{\partial\tilde{c}}\leq\lambda\eta-\beta<0$, so $\tilde{c}^*$ should be set as low as possible, resulting in $\tilde{c}^*=c$.

When $\beta<\eta$, $\frac{\partial U(\tilde{c})}{\partial\tilde{c}}=\begin{cases} -\beta<0 & if \quad \tilde{c}\geq\overline{c}\geq r \\ \eta-\beta>0 & if \quad \overline{c}>\tilde{c}\geq r \\ \lambda\eta-\beta>0 & if \quad \overline{c}>r>\tilde{c}\end{cases}$, so $\tilde{c}^*$ should be set as high as possible if $\tilde{c}<\overline{c}$ and it should be set as low as possible if $\tilde{c}\geq\overline{c}$, indicating $\tilde{c}^*=\begin{cases} c & if \quad c\geq\overline{c} \\ \overline{c} & if \quad c<\overline{c}\end{cases}$.

When $\eta<\beta<\lambda\eta$, $\frac{\partial U(\tilde{c})}{\partial\tilde{c}}=\begin{cases} -\beta<0 & if \quad \tilde{c}\geq\overline{c}\geq r \\ \eta-\beta<0 & if \quad \overline{c}>\tilde{c}\geq r \\ \lambda\eta-\beta>0 & if \quad \overline{c}>r>\tilde{c}\end{cases}$, so $\tilde{c}^*$ should be set as high as possible if $\tilde{c}<r$ and it should be set as low as possible if $\tilde{c}\geq r$, indicating $\tilde{c}^*=\begin{cases} c & if \quad c\geq r \\ r & if \quad c<r\end{cases}$.

**Proposition 1:** ***(Reference Point level)*** Drivers with higher reference point levels exhibit greater reference-dependence, that is, $\frac{\partial \Pr(\tilde{c}^*=r)}{\partial r}>0$.

Proof:

Note that $\Pr(\tilde{c}^*=r)=\Pr(\eta<\beta<\lambda\eta,c<r)$. $\forall r'>r$, we have $\Pr(\eta<\beta<\lambda\eta,c<r')>\Pr(\eta<\beta<\lambda\eta,c<r)$, implying $\Pr(\tilde{c}^*=r')>\Pr(\tilde{c}^*=r)$.

**Proposition 2:** ***(Loss Aversion level)*** Drivers with greater loss aversion levels exhibit greater reference-dependence, that is, $\frac{\partial \Pr(\tilde{c}^* = r)}{\partial \lambda} > 0$.

Proof:

Note that $\Pr(\tilde{c}^* = r) = \Pr(\eta < \beta < \lambda\eta, c < r)$. $\forall \lambda' > \lambda$, we have $\Pr(\eta < \beta < \lambda'\eta, c < r) > \Pr(\eta < \beta < \lambda\eta, c < r)$, implying $\Pr(\tilde{c}^* = r \mid \lambda') > \Pr(\tilde{c}^* = r \mid \lambda)$.

**Proposition 3:** ***(Claim Number)*** A policyholder's first claim exhibits more reference-dependence compared to subsequently made claims, as long as there is a sufficiently low likelihood of an accident at any given time, that is, $\exists \mu^* > 0$, such that $\forall \mu < \mu^*$, $\forall i > 1$, $\Pr(\tilde{c}^*_{1,t_1} = r) > \Pr(\tilde{c}^*_{i,t_i} = r)$.

Proof:

Without loss of generality, consider a policyholder that is involved in accident 1 of damage $c_1$ at time $t_1$ with $\beta \in (\eta, \lambda\eta)$.

If $c_1 \geq r$, the policyholder can simply choose $\tilde{c}^*_{1,t_1} = c_1$ and there is no reference-dependent claim at $t_1$. Since claim 1 has already exceeded the referent point, the reference-dependent term no longer exists for decision making with future claims, hence there is no more reference-dependent claim.

It suffices to show that when $c_1 < r$ a policyholder chooses $\tilde{c}^*_{1,t_1} = r$ for sufficiently low $\mu$.

Suppose $c_1 < r$. The policyholder can either (1) make a reference-dependent claim $\tilde{c}_{1,t_1} = r$ or (2) truthfully report $\tilde{c}_{1,t_1} = c_1$, and it suffices to consider only these two actions as other actions of the policyholder can be easily shown to be strictly dominated by either of the two.

By setting $\tilde{c}_{1,t_1} = r$, for accidents that occur subsequently, there will be no reference-dependent term and hence the policyholder will always make truthful report for subsequent claims. So the overall utility of the policyholder is $U_1(\tilde{c}_{1,t_1} = r) = -\beta(r - c_1)$.

By setting $\tilde{c}_{1,t_1} = c_1$, the policyholder truthfully reports in accident 1 and delays the potential reference-dependent claim to subsequent accidents (if there is any). In the most ideal case, the next accident occurs with damage $c_2$ no less than $\bar{c}$ and the optimal strategy is to truthfully report $\tilde{c}_{2,t_2} = c_2$, resulting in $\max U_2(\tilde{c}_{1,t_1} = c_1) = -\lambda\eta(r - c_1) + \int_{t_1}^{T} \eta(\bar{c} - r) f(t_2) dt_2 = -\lambda\eta(r - c_1) + \eta(\bar{c} - r)(F(T) - F(t_1))$.

Since $f(t)=\mu e^{-\mu t}$, we have $F(T)-F(t_1)=e^{-\mu t_1}-e^{-\mu T}$ increasing in $\mu$ for $\mu\in(0,\frac{\ln T-\ln t_1}{T-t_1})$. It is easy to see that $\frac{\ln T-\ln t_1}{T-t_1}$ is decreasing in $t_1$ and $\lim_{t_1\to T}\frac{\ln T-\ln t_1}{T-t_1}=\frac{1}{T}$. Thus, $\forall\mu<\frac{1}{T}$, $\forall t_1\in[0,T)$, $e^{-\mu t_1}-e^{-\mu T}$ is increasing in $\mu$.

Comparing $U_1(\tilde{c}_{1,t_1}=r)$ with $\max U_2(\tilde{c}_{1,t_1}=c_1)$, we obtain the following result:

$$e^{-\mu t_1}-e^{-\mu T}\le\frac{(\lambda\eta-\beta)(r-c_1)}{\eta(\bar{c}-r)}\Leftrightarrow U_1(\tilde{c}_{1,t_1}=r)\ge\max U_2(\tilde{c}_{1,t_1}=c_1)\cdot$$

Define $\mu^*$ such that $\mu^*=\sup\{\mu\in(0,\frac{1}{T})\,|\,e^{-\mu t_1}-e^{-\mu T}\le\frac{(\lambda\eta-\beta)(r-c_1)}{\eta(\bar{c}-r)}\}$. Obviously $\mu^*\in(0,\frac{1}{T}]$. Therefore, $\forall\mu<\mu^*$, we have $U_1(\tilde{c}_{1,t_1}=r)\ge\max U_2(\tilde{c}_{1,t_1}=c_1)$.

**Proposition 4:** ***(Claim Timing)*** Claim $i$ made later in the individual policy period exhibits more reference-dependence compared to claim $i$ made earlier in the policy period, that is, $\forall i=1,2,\cdots$, $t_i'>t_i$ $\Rightarrow$ $\Pr(\tilde{c}^*_{i,t_i'}=r)>\Pr(\tilde{c}^*_{i,t_i}=r)$.

Proof:

It suffices to show the proposition holds for $i=1$. For $i>1$, the reasoning is exactly the same.

Suppose $t_1'>t_1$ and $c_1<r$. In the following, we show that for the first claim, if the policyholder makes a reference-dependent report at $t_1$, then he or she will do so for $t_1'$.

Based on the proof for Proposition 3, the policyholder's utility for making a reference-dependent report at $t_1$, is $U_1(\tilde{c}_{1,t_1}=r)=-\beta(r-c_1)$. His/her utility for not making a reference-dependent report at $t_1$, is denoted by $U_2(\tilde{c}_{1,t_1}=c_1)$. Note that $U_2(\tilde{c}_{1,t_1}=c_1)$ is a decreasing function of $t_1$, since the potential value of delaying the reference-dependent claim to a later accident diminishes over time. Thus, keeping everything else equal, we have $U_2(\tilde{c}_{1,t_1}=c_1)>U_2(\tilde{c}_{1,t_1'}=c_1)$, while $U_1(\tilde{c}_{1,t_1}=r)=-\beta(r-c_1)=U_1(\tilde{c}_{1,t_1'}=r)$. Therefore, if the policyholder makes a reference-dependent report at $t_1$, implying $U_1(\tilde{c}_{1,t_1}=r)>U_2(\tilde{c}_{1,t_1}=c_1)$, we will have $U_1(\tilde{c}_{1,t_1'}=r)=U_1(\tilde{c}_{1,t_1}=r)>U_2(\tilde{c}_{1,t_1}=c_1)>U_2(\tilde{c}_{1,t_1'}=c_1)$, indicating that the policyholder will also make a reference-dependent report at $t_1'$.

## Appendix A: McCrary and Chetty et al Tests for 0.90 Claims to Premium Ratios

This section reports the results of the McCrary test and Chetty et al test for Hypotheses 1 through 4, for the alternative potential reference point of 0.9 claims to premium ratio. The subsequent tables show that consistency in statistical significance between the McCrary and Chetty et al tests does not occur for any of the hypotheses except the case of first claims. The main reason is that the Chetty et al test around the 0.9 ratio does not attain statistical significance for any of the hypotheses besides that for claim number. In terms of the McCrary test, the results tend to favor a detected discontinuity at the upper end of the ratios considered in this group, in other words 0.93. However, the robustness in terms of excess mass at the 0.9 ratio is not substantiated through the Chetty et al test.

For the case of first claims, while the McCrary test shows a marginally significant discontinuity in the density at the 0.93 ratio, the Chetty et al test for first claims does not show significance at the 0.9 ratio, and shows negative excess mass two and three percentage points below at 0.87 and 0.88. This differs from the pattern we observe around the 1.00 ratio, which is positive excess mass at the 1.00 ratio and ratios slightly below that. Thus, we can reasonably conclude that the 0.9 ratio does not generally display the robust patterns of reference-dependence that are detected at the 1.00 ratio.

**Table A1: McCrary density test, Claim number, around 0.90 ratio**

| | | | | *First claims* | | | |
|---|---|---|---|---|---|---|---|
| Claims to Premium Ratio | **0.87** | **0.88** | **0.89** | **0.90** | **0.91** | **0.92** | **0.93** |
| Test statistic | -0.0517 | -0.1005 | -0.0034 | 0.0438 | 0.0726 | 0.0207 | 0.1899* |
| *p-value* | *0.5938* | *0.2829* | *0.9727* | *0.6625* | *0.4988* | *0.8427* | *0.0884* |
| | | | | *Repeat claims (second or more)* | | | |
| Claims to Premium Ratio | **0.87** | **0.88** | **0.89** | **0.90** | **0.91** | **0.92** | **0.93** |
| Test statistic | -0.0129 | 0.0020 | -0.0860 | 0.0121 | 0.0744 | 0.0414 | 0.0163 |
| *p-value* | *0.8327* | *0.9739* | *0.1654* | *0.8490* | *0.2502* | *0.5275* | *0.8033* |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table A2: Chetty et al test, Claim number, around 0.90 ratio**

*First claims*

| **reference point** | **bins** | **statistic** | **p-value** |
|---|---|---|---|
| 0.87 | 5 | -0.0458* | 0.0865 |
| 0.88 | 4 | -0.1201* | 0.0956 |
| 0.89 | 3 | -0.0217 | 0.1156 |
| 0.90 | 2 | 0.0367 | 0.1404 |

*Repeat claims (second or more)*

| **reference point** | **bins** | **statistic** | **p-value** |
|---|---|---|---|
| 0.87 | 5 | -0.0180** | 0.0429 |
| 0.88 | 4 | -0.0150** | 0.0476 |
| 0.89 | 3 | -0.0700* | 0.0528 |
| 0.90 | 2 | -0.0196* | 0.0663 |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table A3: McCrary density test, Claim timing, around 0.90 ratio**

| | *First half of policy year* | | | | | | |
|---|---|---|---|---|---|---|---|
| Claims to Premium Ratio | **0.87** | **0.88** | **0.89** | **0.90** | **0.91** | **0.92** | **0.93** |
| Test statistic | -0.0383 | -0.1495 | -0.0093 | 0.0805 | 0.1270 | 0.0962 | 0.2954** |
| *p-value* | *0.7307* | *0.1934* | *0.9369* | *0.4931* | *0.2774* | *0.4320* | *0.0196* |
| | *Second half of policy year* | | | | | | |
| Claims to Premium Ratio | **0.87** | **0.88** | **0.89** | **0.90** | **0.91** | **0.92** | **0.93** |
| Test statistic | 0.1062 | 0.0706 | 0.1219 | 0.1061 | 0.0887 | 0.0358 | 0.0495 |
| *p-value* | *0.5192* | *0.6704* | *0.4683* | *0.5238* | *0.5993* | *0.8287* | *0.7698* |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table A4: Chetty et al test, Claim timing, around 0.90 ratio**

| *First half of policy year* | | | |
|---|---|---|---|
| **reference point** | **bins** | **statistic** | **p-value** |
| 0.87 | 5 | -0.0592 | 0.1203 |
| 0.88 | 4 | -0.1637 | 0.1286 |
| 0.89 | 3 | -0.0680 | 0.1621 |
| 0.90 | 2 | 0.0127 | 0.1978 |
| *Second half of policy year* | | | |
| **reference point** | **bins** | **statistic** | **p-value** |
| 0.87 | 5 | -0.0156 | 0.1413 |
| 0.88 | 4 | -0.0205 | 0.1554 |
| 0.89 | 3 | 0.0825 | 0.1832 |
| 0.90 | 2 | 0.0893 | 0.2172 |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table A5: McCrary density test, Premium level, around 0.90 ratio**

| | *Low Premium (lower 50th percentile, 4129 Yuan and below)* | | | | | | |
|---|---|---|---|---|---|---|---|
| Claims to Premium Ratio | **0.87** | **0.88** | **0.89** | **0.90** | **0.91** | **0.92** | **0.93** |
| Test statistic | 0.0764 | -0.0017 | 0.0376 | 0.1080 | 0.2200 | 0.1792 | 0.4942*** |
| *p-value* | *0.5697* | *0.9900* | *0.7793* | *0.4330* | *0.1253* | *0.2080* | *0.0016* |
| | *High Premium (upper 50th percentile, above 4129 yuan)* | | | | | | |
| Claims to Premium Ratio | **0.87** | **0.88** | **0.89** | **0.90** | **0.91** | **0.92** | **0.93** |
| Test statistic | -0.0349 | -0.1516 | 0.0222 | 0.0406 | 0.0568 | -0.0189 | -0.0318 |
| *p-value* | *0.7791* | *0.2401* | *0.8669* | *0.7622* | *0.6597* | *0.8856* | *0.8149* |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table A6: Chetty et al test, Premium level, around 0.90 ratio**

| *Low Premium (lower 50th percentile, 4129 Yuan and below)* | | | |
|---|---|---|---|
| **reference point** | **bins** | **statistic** | **p-value** |
| 0.87 | 5 | -0.0228 | 0.1536 |
| 0.88 | 4 | -0.1035 | 0.1644 |
| 0.89 | 3 | -0.1034 | 0.1837 |
| 0.90 | 2 | -0.0604 | 0.2319 |
| *High Premium (upper 50th percentile, above 4129 yuan)* | | | |
| **reference point** | **bins** | **statistic** | **p-value** |
| 0.87 | 5 | -0.0705 | 0.1123 |
| 0.88 | 4 | -0.1379 | 0.1219 |
| 0.89 | 3 | 0.0700 | 0.1608 |
| 0.90 | 2 | 0.1448 | 0.1948 |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table A7: McCrary density test, Premium rate, around 0.90 ratio**

| | *Low Premium Rate (lower 50th percentile, rate of 0.0153 and below)* | | | | | | |
|---|---|---|---|---|---|---|---|
| Claims to Premium Ratio | **0.87** | **0.88** | **0.89** | **0.90** | **0.91** | **0.92** | **0.93** |
| | -0.0945 | -0.1119 | -0.0054 | 0.0524 | 0.0494 | 0.0705 | 0.2683** |
| *p-value* | *0.4147* | *0.3324* | *0.9630* | *0.6590* | *0.6897* | *0.5691* | *0.0366* |

| | *High Premium Rate (upper 50th percentile, rate above 0.0153)* | | | | | | |
|---|---|---|---|---|---|---|---|
| Claims to Premium Ratio | **0.87** | **0.88** | **0.89** | **0.90** | **0.91** | **0.92** | **0.93** |
| | 0.1155 | 0.0860 | 0.1148 | 0.1252 | 0.1782 | 0.0515 | 0.1208 |
| *p-value* | *0.4271* | *0.5291* | *0.4327* | *0.4026* | *0.2465* | *0.7245* | *0.4114* |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table A8: Chetty et al test, Premium rate, around 0.90 ratio**

*Low Premium Rate (lower 50th percentile, rate of 0.0153 and below)*

| reference point | bins | statistic | p-value |
|---|---|---|---|
| 0.87 | 5 | -0.1228 | 0.1536 |
| 0.88 | 4 | -0.1035 | 0.1644 |
| 0.89 | 3 | -0.1034 | 0.1837 |
| 0.90 | 2 | -0.0604 | 0.2319 |

*High Premium Rate (upper 50th percentile, rate above 0.0153)*

| reference point | bins | statistic | p-value |
|---|---|---|---|
| 0.87 | 5 | -0.0705 | 0.1123 |
| 0.88 | 4 | -0.1379 | 0.1219 |
| 0.89 | 3 | 0.0700 | 0.1608 |
| 0.90 | 2 | 0.1448 | 0.1948 |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

## Appendix B: McCrary and Chetty et al Tests for 1.10 Claims to Premium Ratios

This section reports the results of the McCrary test and Chetty et al test for Hypotheses 1 through 4, for the alternative potential reference point of 1.1 claims to premium ratio. Similar to the result for the 0.9 ratio reported in Appendix A, the Chetty et al test does not attain statistical significance (except in the case of claim numbers, to be discussed below), which indicates that any discontinuity detected by the McCrary test is not robust in terms of excess mass at the reference point candidate.

Similar to the case of the 0.9 ratio, the exception is for the case of claim numbers. However here, contrary to Hypothesis 3, some statistical significance is found across McCrary and Chetty et al tests for repeat claims rather than for first claims. There, the test statistic at the 1.13 ratio for the McCrary test is negative, which is not consistent with the general patterns of significance that we find around the 1.00 ratio. The Chetty et al test for repeat claims shows marginally significant excess mass. Despite the statistical significance of the tests on repeat claims, the pattern does not match the comparative statics result of Hypothesis 3, which predicts that first claims have the more reference dependent claim behavior. Thus, we can altogether conclude that 1.1 is not a convincing common reference point among policyholders.

**Table B1: McCrary density test, Claim number, around 1.10 ratio**

| | | | | *First claims* | | | |
|---|---|---|---|---|---|---|---|
| Claims to Premium Ratio | **1.07** | **1.08** | **1.09** | **1.10** | **1.11** | **1.12** | **1.13** |
| Test statistic | -0.1114 | -0.1273 | -0.1031 | 0.0124 | 0.1485 | -0.0351 | -0.0987 |
| *p-value* | *0.3545* | *0.2829* | *0.3942* | *0.9222* | *0.2609* | *0.7815* | *0.4290* |
| | | | | *Repeat claims (second or more)* | | | |
| Claims to Premium Ratio | **1.07** | **1.08** | **1.09** | **1.10** | **1.11** | **1.12** | **1.13** |
| Test statistic | -0.0363 | -0.0577 | 0.0235 | -0.0765 | -0.1174 | -0.0950 | -0.1783** |
| *p-value* | *0.6429* | *0.4632* | *0.7703* | *0.3460* | *0.1547* | *0.2573* | *0.0359* |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table B2: Chetty et al test, Claim number, around 1.10 ratio**

*First claims*

| reference point | bins | statistic | p-value |
|---|---|---|---|
| 1.07 | 5 | -0.1028* | 0.0942 |
| 1.08 | 4 | -0.1255 | 0.1027 |
| 1.09 | 3 | -0.1054 | 0.1229 |
| 1.10 | 2 | 0.0369 | 0.1546 |

*Repeat claims (second or more)*

| reference point | bins | statistic | p-value |
|---|---|---|---|
| 1.07 | 5 | 0.0358* | 0.0588 |
| 1.08 | 4 | 0.0249* | 0.0649 |
| 1.09 | 3 | 0.0813* | 0.0769 |
| 1.10 | 2 | 0.0160* | 0.0931 |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table B3: McCrary density test, Claim timing, around 1.10 ratio**

*First half of policy year*

| Claims to Premium Ratio | **1.07** | **1.08** | **1.09** | **1.10** | **1.11** | **1.12** | **1.13** |
|---|---|---|---|---|---|---|---|
| Test statistic | 0.0126 | -0.0800 | -0.0578 | 0.0894 | 0.3309** | 0.1691 | 0.0970 |
| *p-value* | *0.9249* | *0.5602* | *0.6707* | *0.5255* | *0.0409* | *0.2629* | *0.5336* |

*Second half of policy year*

| Claims to Premium Ratio | **1.07** | **1.08** | **1.09** | **1.10** | **1.11** | **1.12** | **1.13** |
|---|---|---|---|---|---|---|---|
| Test statistic | -0.1333 | -0.1320 | -0.0690 | -0.1171 | -0.1423 | -0.3855** | -0.4979** |
| *p-value* | *0.4599* | *0.4624* | *0.7119* | *0.5220* | *0.4112* | *0.0390* | *0.0158* |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table B4: Chetty et al test, Claim timing, around 1.10 ratio**

*First half of policy year*

| **reference point** | **bins** | **statistic** | **p-value** |
|---|---|---|---|
| 1.07 | 5 | -0.2365 | 0.1166 |
| 1.08 | 4 | -0.2838 | 0.1284 |
| 1.09 | 3 | -0.3271 | 0.1475 |
| 1.10 | 2 | -0.1767 | 0.1964 |

*Second half of policy year*

| **reference point** | **bins** | **statistic** | **p-value** |
|---|---|---|---|
| 1.07 | 5 | 0.3086 | 0.2304 |
| 1.08 | 4 | 0.3643 | 0.2436 |
| 1.09 | 3 | 0.5949 | 0.2996 |
| 1.10 | 2 | 0.6794 | 0.3375 |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table B5: McCrary density test, Premium level, around 1.10 ratio**

| | *Low Premium (lower 50th percentile, 4129 Yuan and below)* | | | | | | |
|---|---|---|---|---|---|---|---|
| Claims to Premium Ratio | **1.07** | **1.08** | **1.09** | **1.10** | **1.11** | **1.12** | **1.13** |
| Test statistic | -0.1690 | -0.1083 | -0.0786 | 0.1006 | 0.3471* | 0.0835 | -0.1226 |
| *p-value* | *0.2945* | *0.5105* | *0.6389* | *0.5648* | *0.0577* | *0.6374* | *0.4773* |
| | *High Premium (upper 50th percentile, above 4129 yuan)* | | | | | | |
| Claims to Premium Ratio | **1.07** | **1.08** | **1.09** | **1.10** | **1.11** | **1.12** | **1.13** |
| Test statistic | 0.0490 | -0.0658 | -0.0562 | -0.0123 | -0.0041 | -0.1320 | -0.0509 |
| *p-value* | *0.7402* | *0.6616* | *0.7163* | *0.9368* | *0.9792* | *0.4332* | *0.7592* |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table B6: Chetty et al test, Premium level, around 1.10 ratio**

| *Low Premium (lower 50th percentile, 4129 Yuan and below)* | | | |
|---|---|---|---|
| **reference point** | **bins** | **statistic** | **p-value** |
| 1.07 | 5 | -0.2255 | 0.1318 |
| 1.08 | 4 | -0.1891 | 0.1635 |
| 1.09 | 3 | -0.1797 | 0.1823 |
| 1.10 | 2 | -0.0014 | 0.2459 |
| *High Premium (upper 50th percentile, above 4129 yuan)* | | | |
| **reference point** | **bins** | **statistic** | **p-value** |
| 1.07 | 5 | 0.0488 | 0.1638 |
| 1.08 | 4 | -0.0512 | 0.1642 |
| 1.09 | 3 | -0.0195 | 0.1923 |
| 1.10 | 2 | 0.0802 | 0.2364 |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table B7: McCrary density test, Premium rate, around 1.10 ratio**

| | *Low Premium Rate (lower 50$^{th}$ percentile, rate of 0.0153 and below)* | | | | | | |
|---|---|---|---|---|---|---|---|
| Claims to Premium Ratio | **1.07** | **1.08** | **1.09** | **1.10** | **1.11** | **1.12** | **1.13** |
| Test statistic | 0.1049 | 0.0385 | 0.0243 | 0.1316 | 0.2107 | 0.0801 | 0.0357 |
| *p-value* | *0.4532* | *0.7775* | *0.8644* | *0.3728* | *0.1567* | *0.5971* | *0.8085* |
| | *High Premium Rate (upper 50$^{th}$ percentile, rate above 0.0153)* | | | | | | |
| Claims to Premium Ratio | **1.07** | **1.08** | **1.09** | **1.10** | **1.11** | **1.12** | **1.13** |
| Test statistic | -0.1440 | -0.0591 | -0.0315 | -0.0099 | 0.1485 | -0.1213 | -0.1814 |
| *p-value* | *0.3937* | *0.7208* | *0.8543* | *0.9562* | *0.4612* | *0.4988* | *0.3613* |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table B8: Chetty et al test, Premium rate, around 1.10 ratio**

| *Low Premium Rate (lower 50$^{th}$ percentile, rate of 0.0153 and below)* | | | |
|---|---|---|---|
| **reference point** | **bins** | **statistic** | **p-value** |
| 1.07 | 5 | -0.1116 | 0.1327 |
| 1.08 | 4 | -0.2060 | 0.1403 |
| 1.09 | 3 | -0.2153 | 0.1586 |
| 1.10 | 2 | -0.0404 | 0.2062 |
| *High Premium Rate (upper 50$^{th}$ percentile, rate above 0.0153)* | | | |
| **reference point** | **bins** | **statistic** | **p-value** |
| 1.07 | 5 | -0.0913 | 0.1564 |
| 1.08 | 4 | -0.0158 | 0.1769 |
| 1.09 | 3 | 0.0457 | 0.1995 |
| 1.10 | 2 | 0.1415 | 0.2447 |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

## Appendix C: McCrary and Chetty et al Tests for Corporate/Government Vehicles

The results in this section show straightforwardly that for non-individually held insurance policies, the statistical tests are never simultaneously significant across the McCrary and Chetty et al tests for Hypotheses 1 through 4.

**Table C1: McCrary density test, Claim number**

| | | | | *First claims* | | | |
|---|---|---|---|---|---|---|---|
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| Test statistic | 0.2532 | 0.1803 | -0.0967 | -0.0256 | -0.0468 | 0.0263 | 0.1293 |
| *p-value* | *0.2997* | *0.4537* | *0.6813* | *0.9132* | *0.8453* | *0.9129* | *0.6076* |
| | | | | *Repeat claims (second or more)* | | | |
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| Test statistic | -0.0142 | 0.0235 | 0.2002 | 0.2811 | 0.3000 | -0.0922 | -0.0225 |
| *p-value* | *0.9391* | *0.9023* | *0.3310* | *0.1477* | *0.1315* | *0.6453* | *0.9101* |

*significant at 10% level, ** significant at 5% level, *** significant at 10% level

**Table C2: Chetty et al test, Claim number**

*First claims*

| candidate reference point | # of bins from reference point tested | test statistic | std error | p-value |
|---|---|---|---|---|
| 0.97 | 5 | 0.4150 | 0.3883 | 0.2852 |
| 0.98 | 4 | 0.3216 | 0.3811 | 0.3987 |
| 0.99 | 3 | -0.2581 | 0.3363 | 0.4428 |
| 1.00 | 2 | -0.2907 | 0.4107 | 0.4790 |
| 1.00 | 1 | -0.1313 | 0.5456 | 0.8098 |

*Repeat claims (second or more)*

| candidate reference point | # of bins from reference point tested | test statistic | std error | p-value |
|---|---|---|---|---|
| 0.97 | 5 | 0.0585 | 0.1741 | 0.7367 |
| 0.98 | 4 | 0.0626 | 0.1879 | 0.7391 |
| 0.99 | 3 | 0.2955 | 0.2345 | 0.2077 |
| 1.00 | 2 | 0.6381** | 0.2907 | 0.0282 |
| 1.00 | 1 | 0.1002 | 0.3380 | 0.7668 |

*significant at 10% level, ** significant at 5% level, *** significant at 10% level

**Table C3: McCrary density test, Claim timing**

| | | | | *First half of policy year* | | | |
|---|---|---|---|---|---|---|---|
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| Test statistic | 0.3421 | 0.1445 | 0.1343 | 0.0530 | 0.1175 | 0.2909 | 0.3001 |
| *p-value* | *0.2543* | *0.6265* | *0.6488* | *0.8579* | *0.6930* | *0.3373* | *0.3326* |
| | | | | *Second half of policy year* | | | |
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| | 0.2384 | 0.2154 | -0.0977 | 0.1131 | -0.0560 | 0.4235 | 0.0292 |
| *p-value* | *0.5375* | *0.5899* | *0.7985* | *0.7730* | *0.8871* | *0.3249* | *0.9416* |

*significant at 10% level, ** significant at 5% level, *** significant at 10% level

**Table C4: Chetty et al test, Claim timing**

*First half of policy year*

| **reference point** | **bins** | **statistic** | **std dev** | **p-value** |
|---|---|---|---|---|
| 0.97 | 5 | 0.1666 | 0.4117 | 0.6858 |
| 0.98 | 4 | -0.0940 | 0.3782 | 0.8039 |
| 0.99 | 3 | -0.4301 | 0.3952 | 0.2765 |
| 1.00 | 2 | -0.5695 | 0.4904 | 0.2455 |
| 1.00 | 1 | -0.0772 | 0.6497 | 0.9054 |

*Second half of policy year*

| **reference point** | **bins** | **statistic** | **std dev** | **p-value** |
|---|---|---|---|---|
| 0.97 | 5 | 0.9008 | 0.7192 | 0.2104 |
| 0.98 | 4 | 1.2311 | 0.8575 | 0.1511 |
| 0.99 | 3 | 0.0624 | 0.5770 | 0.9138 |
| 1.00 | 2 | 0.2479 | 0.7059 | 0.7254 |
| 1.00 | 1 | -0.2224 | 0.8973 | 0.8043 |

*significant at 10% level, ** significant at 5% level, *** significant at 10% level

**Table C5: McCrary density test, Premium level**

| | Low Premium (lower 50th percentile, 4129 Yuan and below) | | | | | | |
|---|---|---|---|---|---|---|---|
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| Test statistic | 0.5425 | 0.3899 | 0.0696 | 0.0278 | 0.0478 | 0.1926 | 0.1631 |
| *p-value* | *0.1086* | *0.2365* | *0.8244* | *0.9304* | *0.8801* | *0.5576* | *0.6182* |
| | High Premium (upper 50th percentile, above 4129 yuan) | | | | | | |
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| Test statistic | 1.8320*** | N.A. | N.A. | N.A. | 5.2507* | N.A. | N.A. |
| *p-value* | *0.0022* | *N.A.* | *N.A.* | *N.A.* | *0.0956* | *N.A.* | *N.A.* |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table C6: Chetty et al test, Premium level**

*Low Premium (lower 50th percentile, 4129 Yuan and below)*

| **reference point** | **bins** | **statistic** | **std dev** | **p-value** |
|---|---|---|---|---|
| 0.97 | 5 | 0.3419 | 0.4202 | 0.4158 |
| 0.98 | 4 | 0.2345 | 0.4048 | 0.5625 |
| 0.99 | 3 | -0.4238 | 0.3682 | 0.2498 |
| 1.00 | 2 | -0.5647 | 0.4513 | 0.2109 |
| 1.00 | 1 | -0.0687 | 0.6882 | 0.9205 |

*High Premium (upper 50th percentile, above 4129 yuan)*

| **reference point** | **bins** | **statistic** | **std dev** | **p-value** |
|---|---|---|---|---|
| 0.97 | 5 | 0.5358 | 0.6475 | 0.4079 |
| 0.98 | 4 | 0.4698 | 0.6254 | 0.4525 |
| 0.99 | 3 | 0.0413 | 0.5923 | 0.9445 |
| 1.00 | 2 | 0.2218 | 0.7133 | 0.7558 |
| 1.00 | 1 | -0.2342 | 0.8792 | 0.7899 |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table C7: McCrary density test, Premium rate**

| | *Low Premium Rate (lower 50th percentile, rate of 0.0153 and below)* | | | | | | |
|---|---|---|---|---|---|---|---|
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| Test statistic | N.A. | N.A. | 0.8251* | 2.1418*** | 4.5843** | N.A. | N.A. |
| *p-value* | *N.A.* | *N.A* | *0.0589* | *0.0025* | *0.0420* | *N.A.* | *N.A.* |
| | *High Premium Rate (upper 50th percentile, rate above 0.0153)* | | | | | | |
| Claims to Premium Ratio | **0.97** | **0.98** | **0.99** | **1.00** | **1.01** | **1.02** | **1.03** |
| Test statistic | 0.3728 | 0.3519 | -0.0874 | -0.1228 | -0.1467 | -0.0667 | -0.0557 |
| *p-value* | *0.2708* | *0.3030* | *0.7899* | *0.7075* | *0.6597* | *0.8437* | *0.8693* |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level

**Table C8: Chetty et al test, Premium rate**

*Low Premium Rate (lower 50th percentile, rate of 0.0153 and below)*

| reference point | bins | statistic | std dev | p-value |
|---|---|---|---|---|
| 0.97 | 5 | 0.3758 | 0.5884 | 0.5231 |
| 0.98 | 4 | 0.0783 | 0.5449 | 0.8858 |
| 0.99 | 3 | -0.3177 | 0.5288 | 0.5480 |
| 1.00 | 2 | -0.2855 | 0.6498 | 0.6604 |
| 1.00 | 1 | -0.2876 | 0.9043 | 0.7505 |

*High Premium Rate (upper 50th percentile, rate above 0.0153)*

| reference point | bins | statistic | std dev | p-value |
|---|---|---|---|---|
| 0.97 | 5 | 0.4435 | 0.4835 | 0.3589 |
| 0.98 | 4 | 0.5071 | 0.5023 | 0.3127 |
| 0.99 | 3 | -0.2171 | 0.4154 | 0.6012 |
| 1.00 | 2 | -0.2942 | 0.4768 | 0.5372 |
| 1.00 | 1 | -0.0243 | 0.6752 | 0.9713 |

*significant at 10% level, ** significant at 5% level, *** significant at 1% level